

情报工程

中国科技核心期刊  
中国科技论文统计源期刊

# 情报工程

TECHNOLOGY INTELLIGENCE ENGINEERING



中国科学技术情报学会 | 主办  
中国科学技术信息研究所

2024年 第10卷 第1期

二〇二四年 第十卷 第一期

中国科学技术情报学会 主办  
中国科学技术信息研究所

## 情报工程

QINGBAO GONGCHENG

2024年 第10卷 第1期  
双月刊



主管：中国科学技术协会  
主办：中国科学技术情报学会  
中国科学技术信息研究所

### 编辑部

主编：刘琦岩  
执行主编：姚长青  
副主编：王莉军 刘志辉 桂 婕 董 诚  
责任编辑：孙 瑶 杨 岩 李孟秋 潘 优  
美术编辑：翟芒芒

本刊为中国科技核心期刊（中国科技论文统计源期刊），数字化期刊全文数据库（万方）、中国学术期刊全文数据库（CNKI）、中文科技期刊数据库（维普）、《中国人文社会科学期刊评价报告（AMI）》引文数据库、国家哲学社会科学学术期刊数据库、超星学术期刊“域出版”等收录。

国内统一连续出版物号：CN 10-1263/G3  
国际标准连续出版物号：ISSN 2095-915X  
地 址：北京市海淀区复兴路 15 号（100038）  
电 话：（010）58882458  
E-MAIL：qbgc@istic.ac.cn  
网 址：http://tie.istic.ac.cn/  
印刷单位：北京科信印刷有限公司  
发行范围：公开发行  
编辑出版：科学技术文献出版社有限公司  
广告发布登记：京海工商广登字 20170148 号

如发现印刷、装订等质量问题，请与本刊发行部联系调换。

ISSN 2095-915X



定 价：55.00 元

CN 10-1263/G3  
ISSN 2095-915X



## 技术情报 | TECHNOLOGY INTELLIGENCE

003

全球创新网络中企业间知识扩散模式研究——基于无标度网络理论

Research on the Model of Knowledge Diffusion among Enterprises in Global Innovation Networks: Based on Scale-Free Network Theory

储节旺 施点 杨子豪

017

基于技术挖掘的投资机会分析模型研究

Research on Investment Opportunity Analysis Model Based on Technology Mining

谷威 龚颐雯

028

基于知识网络特征的产业关键共性技术识别模型与实证研究——以新一代信息技术产业为例

Model and Empirical Study on Identification of Industry Key Generic Technologies based on Knowledge Network Characteristics: Evidence from the New Generation of Information Technology Industry

刘静 胥彦玲 张婧

042

中国个人信息保护应用与技术进展研究——基于科学知识图谱视角

Research on the Progress of Personal Information Protection Application and Technology in China——Based on the Perspective of Mapping Knowledge Domains

刘华玲 梁华璧 王希睿

059

长三角地区技术转移与成果转化效率分析——基于三阶段 DEA 模型分析

Analysis of Technology Transfer and Achievement Transformation Efficiency in Yangtze River Delta Region: Based on Three-Stage DEA Model Analysis

戴力新 李红 戚湧

072

计划行为理论视角下微信用户选择性信息规避行为影响因素研究

Influencing Factors of Wechat Users' Selective Information Avoidance Behavior from the Perspective of Planned Behavior Theory

李廷翰

## 情报分析 | INFORMATION ANALYSIS

085

基于 ARIMA、GM (1,1) 模型的高校 ESI 学科发展预测研究

Research on the Prediction of the Development of ESI Disciplines in Universities Based on ARIMA and GM (1,1) Models

柳佳彤 康榆晨 秦丽岩 曹芳

096

基于文献计量学的铁路运输领域论文研究进展及前沿热点

Research Development and Frontier Hotspots of Railway Transportation Based on Bibliometrics

戚小玉

107

国内外顾客互动研究进展与热点分析——基于科学知识图谱视角

Domestic and International Research Progress and Hotspot Analysis of Customer Interaction: Based on the Perspective of Scientific Knowledge Graph

乐承毅 吴嘉炜 周佳辰

# 情报工程

QINGBAO GONGCHENG

2024年第10卷第1期

## 编委会 • EDITORIAL BOARD

### 主任委员：

刘琦岩

### 副主任委员：

乔晓东 潘云涛 孙建军 陆伟 姚长青 朱礼军

### 编委委员（按姓氏汉语拼音为序）：

Ume, Anton (罗马尼亚)	安小米	Coh, Byoung-Youl (韩国)	陈仕吉
陈文广	程永红	迟东训	杜永萍
顾进广	胡春明	Jung, Hanmin (韩国)	黄智生
李广建	刘桂锋	李贺	Meng, Lin (日本)
刘伟丽	刘玉琴	刘忠宝	卢小宾
罗准辰	孟遥	Rana, F. Omer (英国)	漆桂林
钱庆	孙卫	孙新	Mandl, Thomas (德国)
汪雪锋	吴晨生	吴广印	吴小兰
夏昊翔	杨波	杨思洛	杨秀丹
俞立平	翟云	张德政	张家俊
张军	张全	张学福	张巍 (澳大利亚)
张智雄	章成志	真溱	



开放科学  
(资源服务)  
标识码  
(OSID)

# 全球创新网络中企业间知识扩散模式研究 ——基于无标度网络理论

储节旺 施点 杨子豪

安徽大学管理学院 合肥 230601

**摘要:** [目的/意义] 知识经济背景下,全球创新网络成为全球企业间知识流动的平台,通过构建企业间知识扩散模型,研究全球创新网络中企业间知识扩散模式,以促进全球创新网络的发展。[方法/过程] 研究主要基于无标度网络理论,建立有向网络模型并仿真,结合仿真结果以及知识密集型企业与劳动密集型企业的概念,探究了全球创新网络中两种企业知识扩散的模式。[局限] 模型的参数设置较为简单,难以模拟复杂的企业网络结构。[结果/结论] 根据建立的网络模型,分析得到全球创新网络中企业间知识扩散模式主要包括知识创造、知识密集型企业之间的协同创新、知识密集型企业与劳动密集型企业之间的知识扩散三个方面及其具体方式,为两种企业的知识管理提供了建议。

**关键词:** 全球创新网络; 知识扩散; 知识密集型企业; 劳动密集型企业; 无标度网络理论

**中图分类号:** G302

## Research on the Model of Knowledge Diffusion among Enterprises in Global Innovation Networks: Based on Scale-Free Network Theory

CHU Jiewang SHI Dian YANG Zihao

School of Management, Anhui University, Hefei 230601, China

**Abstract:** [Objective/Significance] In the context of the knowledge economy, the global innovation network has become the platform for knowledge flow between global enterprises. This paper constructs an inter-enterprise knowledge diffusion model to study the modes of knowledge diffusion among enterprises in the global innovation network, promoting the development of the global innovation network. [Methods/Processes] The study is primarily based on scale-free network theory. A directed network model is established and simulated. Combining the simulation results with the concepts of knowledge-intensive enterprises and labor-intensive enterprises, the modes of knowledge diffusion between the two types of enterprises in the global innovation network are explored. [Limitations] The parameter settings of the model are relatively simple, making it difficult to simulate complex enterprise network structures. [Results/Conclusions] Based on the established network model, it is analyzed

**基金项目** 国家社科基金一般项目“数智创新生态系统知识生成动力、扩散逻辑与治理机制研究”(23BTQ055)。

**作者简介** 储节旺(1969-),博士,教授,主要研究方向为知识管理;施点(1999-),硕士研究生,主要研究方向为知识管理, E-mail: 1210768172@qq.com; 杨子豪(1998-),硕士研究生,主要研究方向为专利管理。

**引用格式** 储节旺,施点,杨子豪.全球创新网络中企业间知识扩散模式研究——基于无标度网络理论[J].情报工程,2024,10(1):3-16.



that the modes of knowledge diffusion among enterprises in the global innovation network mainly include knowledge creation, collaborative innovation between knowledge-intensive enterprises, and knowledge diffusion between knowledge-intensive and labor-intensive enterprises which provides suggestions for knowledge management of the two types of enterprises.

**Keywords:** Global Innovation Network; Knowledge Diffusion; Knowledge-Intensive Enterprises; Labor-Intensive Enterprises; Scale-free Network Theory

## 引言

在知识经济以及经济全球化的背景下,创新驱动发展已经成为全世界各家企业的基本战略。而在世界范围内的许多行业中,知识创新资源的来源和使用已经从单个企业转移到一群网络化的企业中<sup>[1]</sup>,全球创新网络就是这样一个在国际之间吸收、整合、发散知识创新资源以形成价值的网络形态<sup>[2]</sup>。1991年,Freeman<sup>[3]</sup>首次提出了“创新网络”的概念,这种整合了企业内外部知识创新资源的形式引起了国内外学者的广泛关注。Ernst<sup>[4]</sup>在2009年首次提出全球创新网络(Global innovation network)的概念,并把全球创新网络定义为一种在跨组织、跨区域边界上整合产品研发活动的网络形态。此定义有三层含义,首先指出全球创新网络是一个跨组织跨区域的全球性结构,其组织结构包括但不限于各个国家的企业、科研机构、高校等,其次说明全球创新网络的基础是产品的开发研发与合作,最后点明了全球创新网络是一个网络性结构,符合网络结构的特征。学者们对于全球创新网络的探究包括企业的创新模式<sup>[5-6]</sup>,企业与外部知识的交互行为<sup>[7]</sup>,企业与全球创新网络之间的关系<sup>[8-9]</sup>,以及企业融入全球创新网络的途径<sup>[10-11]</sup>。

而在全球创新网络中,由于企业与各组织之间存在着知识浓度差,各个组织之间会发生明显的知识扩散现象,因此对全球创新网络的

研究便集中在知识扩散的方面。在以往研究中,学者们对于知识扩散的探究主要集中在科研合作网络<sup>[12-14]</sup>,学科交叉<sup>[15]</sup>以及产业创新网络<sup>[16-17]</sup>三个方面。一众学者基于知识扩散的网络性特征,研究了知识扩散与网络结构的关系,例如将知识创新引入规则网络、网络、随机网络以及无标度网络等网络结构,探究在这些网络中知识扩散的规律,并且有学者发现,在无标度网络模型中知识的扩散效率最高<sup>[18]</sup>。但是学者们对于企业之间知识扩散的研究仍有不足。首先,尽管学者们已将企业间知识扩散拓扑成无标度网络模型,但没有引入知识密集型企业的概念,用以定义网络中节点度明显高于其他节点的超级节点。其次,在以往研究中,学者们会通过收集企业的专利数据、创新活动等对企业的知识创新能力进行定性分析,但很少有学者为企业间知识扩散建立数学模型。最后,学者们对无标度网络模型的研究仍有不足,由于知识扩散的有向性,网络结构中节点与节点之间的连接应是有向的,而现有研究大多仅构建了无向的无标度网络模型,并且由于企业之间产生知识联系的概率并不完全依照于网络中的节点度,因此传统的BA模型并不适用于模拟企业间的知识扩散模式。

基于此,本文基于刘浩广等<sup>[19]</sup>的理论,将全球创新网络中企业的拓扑结构抽象为优化过的有向无标度网络模型,即Poisson有向模型。

在全球创新网络中每一个与其他企业发生知识扩散现象的企业都是网络结构中的一个节点，企业之间的知识扩散视为节点之间的连接，并以此构建全球创新网络中企业之间知识扩散的模式。在研究中，本文将探索企业在全局创新网络中与其他企业的知识联系以及知识流动的模式。本文的研究目标主要包括：

(1) 对知识扩散、知识密集型企业、劳动密集型企业的概念进行辨析，并探究企业间知识扩散的特点与性质。

(2) 基于优化后的有向无标度网络理论模型建立企业间知识扩散数学模型并进行仿真，探究部分模型参数对全球创新网络的影响，分析全球创新网络中企业间知识扩散的模式。

(3) 分析知识密集型企业与劳动密集型企业在全局创新网络的知识扩散过程中扮演的角色，为两种类型的企业的知识管理提供策略与思路。

## 1 研究回顾

### 1.1 网络中的知识扩散

知识扩散是有用信息在个人、群体或组织之间的流动，是创新知识在有效机制控制下的传递过程<sup>[14]</sup>。知识扩散是创新思想与技术在人與人之间、组织与组织之间传播的重要社会过程。思想的传播与疾病在人群中的传播在定性上有许多相似之处。受到流行病传播的启发，Bettencourt等<sup>[20]</sup>将发表论文数量或发表作者数量的主题增长视为知识扩散的代理，并引入了流行模型来量化理论和实验物理中的知识传播。由于流行病传播网络与社交网络具有高度的相似性，许多学者利用复杂网络的模型来对知识

扩散的过程进行研究。Zhang等<sup>[21]</sup>研究了四种拓扑结构，找出哪种拓扑结构更适合知识的扩散，并建立了一个新的知识扩散模型，其中考虑了学习与遗忘的机制，网络结构中的个体可以同时扮演传授者与学习者的角色。Guan等<sup>[22]</sup>建立了单个出版物的详细网络，并证明了将扩散网络模型应用于研究传播的可行性。Zhang等<sup>[23]</sup>基于技术专利引用量、科学期刊、科研人员合作等维度，建立了具有空间维度的知识流动网络。Qiao等<sup>[24]</sup>考察了三种不同的知识角色：寻求者、贡献者和中间人。定义了三种类型的选择机制：客观选择机制、基于反馈的选择机制和随机选择机制，并通过对规则网络、随机网络、小世界网络和无标度网络四种典型网络进行基于智能体的知识扩散的过程仿真。Lei等<sup>[25]</sup>着重于两个特征，即网络密度和项目角色，基于代理的建模方法来构建知识扩散的仿真模型并进行了模拟实验。

不难看出，目前学者们对知识扩散模式的探讨与复杂网络的模型紧紧关联，并且研究的方式主要围绕着如何建立更科学、可信的网络模型，来模拟知识扩散的过程。因此本研究将基于复杂网络中的无标度网络模型分析企业间知识扩散的模式演化。但由于企业间的知识扩散具有有向性的特征，且企业间产生知识联系的概率并不完全符合BA无标度网络模型，因此本研究将引入择优连接倾向系数的参数概念，建立有向网络模型，以此刻画全球创新网络中企业间知识扩散模式的特征。

### 1.2 知识密集型企业与劳动密集型企业

知识密集型企业 (Knowledge-Intensive En-

terprises, 简称 KIEs) 是指主要依靠知识、技术、创新等方面的投入来获得利润的企业; 劳动密集型企业 (Labor-Intensive Enterprises, 简称 LIEs) 是指依赖人力和物力资源进行生产的企业。知识密集型企业的员工通常具备高水平的专业技能和知识, 而劳动密集型企业的员工则通常以劳动力为主, 他们的技能和知识水平相对较低。因此学者们对企业知识管理的研究主要集中于知识密集型企业。同时国外也有学者从知识管理、金融、管理理论与实践、社交网络知识共享、开放式创新框架、业务流程管理、四螺旋模型等几个角度探讨了全球知识密集型企业 and 国际技术转移的热点问题<sup>[26]</sup>。

在企业间知识扩散模型中, 知识密集型企业 (KIEs) 与劳动密集型企业 (LIEs) 之间很难定量地区分, 并且两者的概念可能随着演化的过程发生改变, 因此本研究仅将网络模型中节点度明显高于其余节点的企业认定为知识密集型企业。

从知识密集型企业与劳动密集型企业的概念中不难看出, 在全球创新网络中, 知识密集型企业通过产品分工与创新活动等方式, 总是作为知识输出的一方; 而劳动密集型企业由于知识储量不丰富, 通常作为输入知识的一方。研究将基于此, 对全球创新网络的参与者进行定性分析, 并对网络的初始状态进行设定。

## 2 全球创新网络中企业间知识扩散模型构建与仿真

### 2.1 模型构建理论

BA 无标度网络模型是一种常见的复杂网络

模型, 由 Barabási 和 Albert<sup>[27]</sup> 于 1999 年提出。它是一种基于“优先连接”的网络生成算法, 可以用于模拟很多实际现象, 如互联网、社交网络、蛋白质相互作用网络等。BA 模型的一个重要特点是无标度性, 即节点的度数分布呈幂律分布。这意味着只有很少一部分节点的度数很大, 而大多数节点的度数很小。这种特殊的度数分布在很多实际网络中都得到了验证。在 BA 模型被提出之后, 研究人员陆续对其进行了许多研究和拓展。例如, 他们研究了网络的演化规律、拓扑结构和动力学性质等。

也有部分学者将无标度网络理论应用在现实的知识管理领域。有学者通过数值计算, 发现当有一定程度的局部知识存在时, BA 模型的全局属性, 如连通性分布和平均最短路径长度, 具有相当的鲁棒性。与此同时, 其他属性 (如聚类系数和度相关) 则趋近于真实网络的测量值<sup>[28]</sup>。还有学者构建了一个医学知识网络, 并通过研究发现该医学知识网络具有小世界、无标度、分层的特点<sup>[29]</sup>。还有学者从知识转移和知识溢出的角度研究无标度网络在知识管理领域的作用, 并且通过仿真实验得出结论: 无标度网络为知识转移与知识溢出提供了最优路径, 并且可以刺激研发活动, 促进经济增长<sup>[18,30]</sup>。

全球创新网络具有企业数量众多、网络节点较为稠密、网络结构较为稳定等特点, 并且由于其具有一定数量的知识储量丰富、创新活动密集的知识密集型企业, 较为符合无标度网络的拓扑特征, 因此本研究选择全球创新网络作为研究对象, 并采用无标度网络理论对企业间的知识扩散模式进行分析。

## 2.2 企业间知识扩散模型建立

在研究中，全球创新的网络拓扑结构将被视为有向的无标度网络结构，每个节点为企业，节点间的边为企业之间的知识扩散（联合创新活动或产品分工等），边的存在取决于企业之间是否存在知识联系，处于知识联系中的企业有知识输入与输出两种模式。全球创新网络中企业间的知识扩散模式主要受到网络结构与企业决策等多种因素的影响。

本研究主要通过超级节点与普通节点度的演化与分布来体现其网络拓扑结构特性。并引入刘浩广等<sup>[19]</sup>提出的 Poisson 有向无标度网络模型。首先根据无标度特性，全球创新网络企业间知识扩散的初始有向网络  $G(V)$ ，其中  $V$  以节点的形式表示全球创新网络中的所有企业，表示企业之间存在有向的知识联系，即节点间存在有向边联系。例如节点  $a$  与节点  $b$  是网络模型中的两个节点，即全球创新网络中的两个企业，若存在着  $a$  公司向  $b$  公司的知识扩散且没有存在  $b$  公司向  $a$  公司的知识扩散，也就是  $a$  公司存在着知识输出，且  $b$  公司存在着知识输入，即表示为  $e_{ab}=1$  且  $e_{ba}=0$ 。若节点  $a$  作为始点，与  $O_m$  个节点相连，则认为  $O_m$  为节点  $a$  的出度；若节点  $a$  作为终点，与  $I_n$  个节点相连，则认为  $I_n$  为节点  $a$  的入度。

基于此，研究假设全球创新网络中有  $m_0$  个已经产生知识联系的企业，并且网络规模为  $N$ ，即共存在着  $N$  个潜在的全球创新网络中的关联企业 ( $m_0 < N$ )。每过一个时间单位  $T$ ，设置一个企业加入网络。后续的企业加入网络时，产生  $m$  条边，也就是与已在网络中的  $m$  ( $mm_0$ )

个企业产生知识联系，其作为终点的概率为  $P$ ，即知识输入概数，其作为始点的概率为  $1-P$ ，并且其入边数服从二项分布  $B(m, P)$ ，出边数服从二项分布  $B(m, 1-P)$ ，这里的  $P \in (0, 1)$ 。

并且后续加入的企业更倾向于与度值更大的企业产生知识联系，如果新节点是始点，则选择终点  $i$  的概率  $\Pi$  依赖于节点  $i$  的入度  $I_i$ ，即满足关系式：

$$\Pi(I_i) = \frac{P_1 I_i + 1 - P_1}{\sum_j (P_1 I_j + 1 - P_1)} \quad (1)$$

如果新节点是终点则选择始点  $i$  的概率  $\Pi$  依赖于节点  $i$  的出度  $O_i$ ，即满足关系式：

$$\Pi(O_i) = \frac{P_2 O_i + 1 - P_2}{\sum_j (P_2 O_j + 1 - P_2)} \quad (2)$$

其中， $P_j$  是选择网络中已有的节点进行择优连接的概率 ( $P_j$  的取值范围为 0 至 1)，即择优连接倾向系数， $j = 1, 2$ 。当  $P_1 = 0$  时，相当于入度没有择优；当  $P_2 = 0$  时，相当于出度没有择优。

模型主要参数如下表：

表 1 知识扩散模型主要参数

参数	代表意义
$N$	全球创新网络规模
$m_0$	初始网络规模
$m$	新节点加入时产生连边数
$P$	知识输入概数
$P_1$	始点择优连接倾向系数
$P_2$	终点择优连接倾向系数
$s$	初始网络起点集
$t$	初始网络终点集
$k$	网络初始度值
$T$	时间单位



## 2.3 模型仿真

本实验使用 Python 编程语言实现企业间知识扩散模型的构建和演化。通过调用 NumPy 函数库，我们构建了一个三维矩阵来记录网络中的节点与连边情况，以矩阵的形式呈现了网络的结构。同时，我们还利用 Networkx 库构建了网络拓扑模型，以便更好地描述和分析知识扩散过程。此外，我们借助 Pandas 和 Matplotlib 库对数据进行收集、分析和可视化，从而更直观地展示知识扩散的效果。

实验首先进行初始参数设定，用以模拟企业网络的初始状态。通过修改参数并迭代仿真收集各个节点在网络演化过程中的度值数据，分析各个因素对模型演化的影响。以此探究全球创新网络中的两种企业间的知识扩散模式。

### 2.3.1 参数设定

在本项研究中，我们深入探讨了全球创新网络的拓扑结构并进行了深入的仿真研究。为了更准确地捕捉这一复杂网络的动态特性，我们引入了一系列关键参数，其中包括：调节参数  $P$ 、 $P_1$  和  $P_2$ 。这三个参数分别描述新进入的企业作为终点加入网络的概率以及在选择网络中现有节点作为起始点或终点时的择优连接倾向。这些参数为我们提供了一个机会，通过不同的设定来模拟和分析其对整体网络结构的影响。

此外，为了确保模型的稳定性并提供一个可比较的参考框架，我们也设定了几个固定的参数。这包括网络的规模  $N=102$ 、每次仿真的步长  $T=100$  个时间单位、初始网络中的企业数量  $m_0=2$  以及每一个新入驻的企业与已存在企业之间建立的知识联系数量  $m=2$ 。

为了使模型更为直观并确保其从一个明确的状态开始，我们设定网络的初始状态为节点 1 和节点 2 之间存在双向连接，即初始的起点集为  $s=[1,2]$ ，而初始的终点集为  $t=[2,1]$  这种设定不仅简化了模型的起始条件，还为后续的仿真提供了一个明确的基准点。

图 1 展示了在  $P$  设为 0.7，以及  $P_1$  和  $P_2$  都设为 0.5 的条件下，一次仿真的结果，即网络的拓扑结构。而图 2 则进一步展示了在相同的参数条件下，经过 100 次迭代仿真后，各个节点的平均出入度情况。

### 2.3.2 知识输入概数对模型的影响

在图 3 中，通过将择优连接倾向系数  $P_1$  和  $P_2$  均设定为 0.7，并在三种不同的知识输入概数  $P$  值（0.2、0.5 和 0.8）下进行 100 次迭代仿真，我们得到了对知识密集型企业知识交互行为的深入洞察。

首先，图中明确显示，编号靠前的节点（即那些早期加入网络的节点）无论是在知识输入还是输出上，其度数都明显高于后来加入的节点。这在某种程度上反映了“先进者优势”在知识网络中的体现：早期的企业由于其其在网络中的历史地位，能够更广泛地与其他企业进行知识交互，从而积累更多的知识资源和合作关系。

对于知识输入概数  $P$  的影响，实验观察到其在知识输入和输出上都起到了显著的作用。具体而言，当  $P$  值增加时，表示新加入的企业更可能作为知识的生产者。这导致知识密集型企业知识输入上呈现出增长的趋势。这是因为，这些企业作为知识的主要提供者，当网络中有更多的知识接收者时，它们的机会也随之增加，进而提高了其知识的传播和影响力。

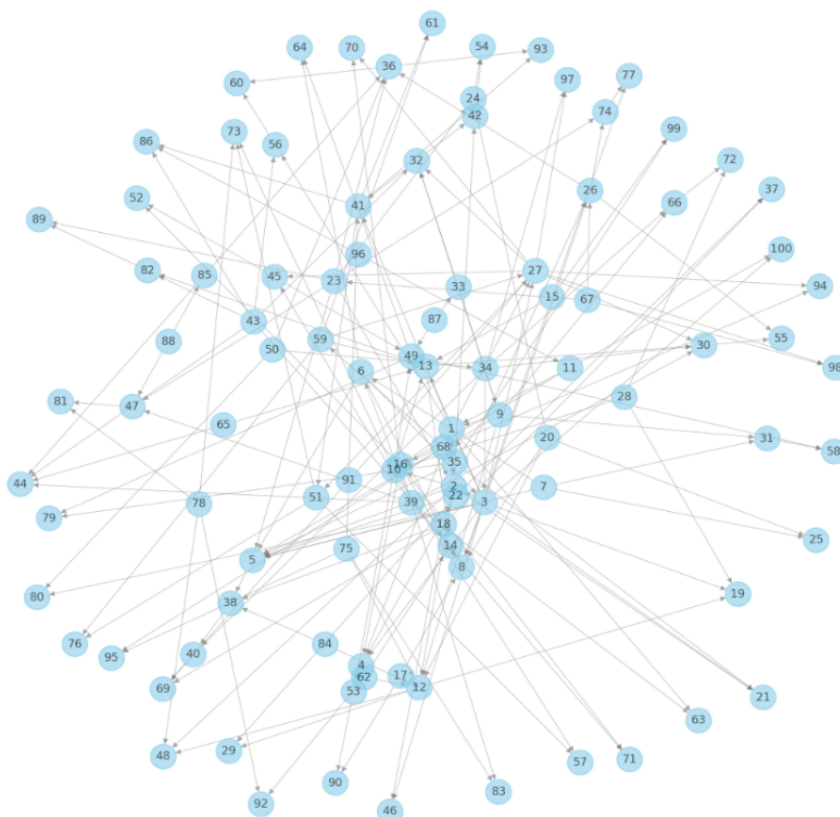


图1 网络结构可视化

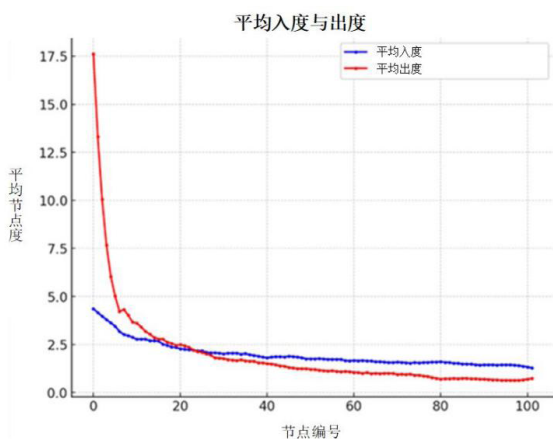


图2 各节点度值情况

然而，从知识输出的角度来看，情况却恰恰相反。随着P值的增加，知识密集型企业的知识输出度数呈现下降趋势。原因在于，较高的P值意味着新加入的企业更倾向于作为知识

的接收者而非提供者。这进一步意味着，知识密集型企业寻找新的知识来源时所面临的选择和机会实际上是在减少的。

### 2.3.3 择优连接倾向系数对模型的影响

图4和图5呈现了始点和终点择优连接倾向系数如何影响模型中各节点的出入度分布。考虑到在真实的企业生态中，新进入的企业通常更多地扮演知识的接受者角色。基于这一观察，实验将知识输入概率P设为0.7，以模拟这一现实情境。

在图4中，固定终点的择优连接倾向系数 $P_2$ 为0.5，并探讨了始点的择优连接倾向系数 $P_1$ 在0.2、0.5和0.8三个不同值下的影响。经过100次的迭代仿真，实验得到了各节点的平均出入度情况。

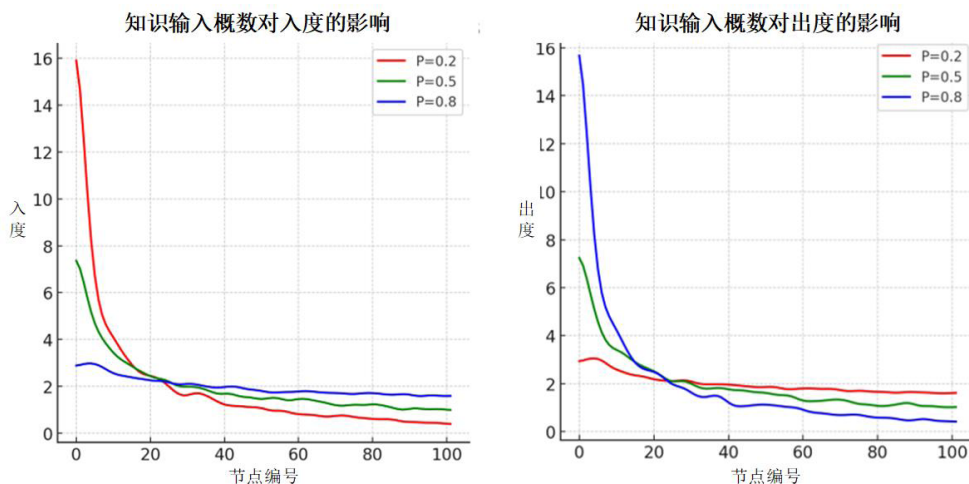


图3 知识输入概数对节点度影响

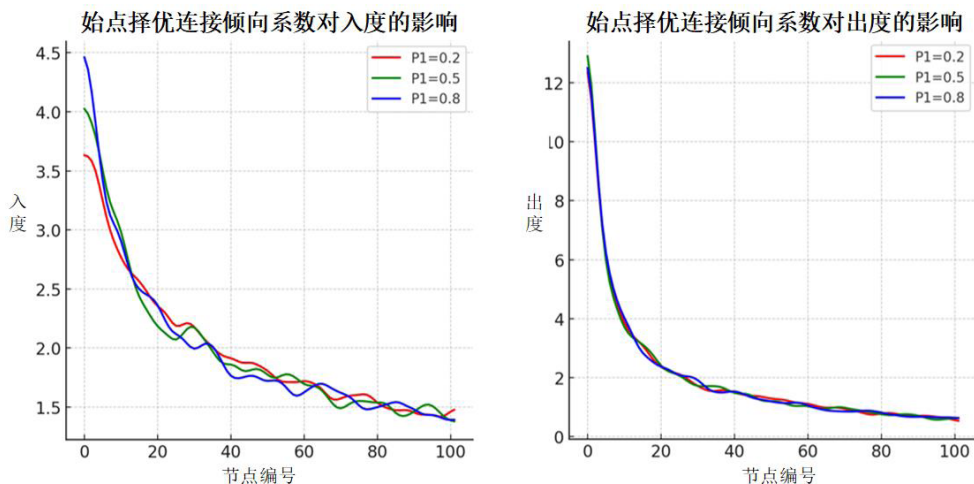


图4 始点择优连接倾向系数对节点度影响

观察结果显示，改变  $P_1$  的值对各节点的出度分布并没有产生显著的影响。这表明在知识传播过程中，知识的提供者（即出度）的选择并不受到  $P_1$  的显著影响。然而，对于入度，情况则明显不同。后期加入的节点由于其在网络中的相对年轻地位，其总度值较小，因此  $P_1$  对其入度的影响也相对较小。但对于较早加入网络的节点，尤其是编号靠前的节点，其入度受  $P_1$  影响显著。随着  $P_1$  的增加，新进入网络的企业在选择知识来源时，更倾向于与那些已经拥有较高知识吸纳能力

的企业建立联系。这导致编号靠前的节点的入度显著增加。

图5为固定始点择优连接倾向系数  $P_1$  为0.5，设置终点择优连接倾向系数  $P_2$  为0.2、0.5、0.8，迭代仿真100次的各节点出入度平均结果。同理，知识接收者（即入度）的选择不受  $P_2$  的显著影响，但对于出度， $P_2$  越大，后续加入网络的企业更倾向从知识输出能力高的企业处获取知识，因此编号靠前的企业的出度越大。而  $P_2$  越小，这种倾向就越小，编号靠前的节点入度也就越小。

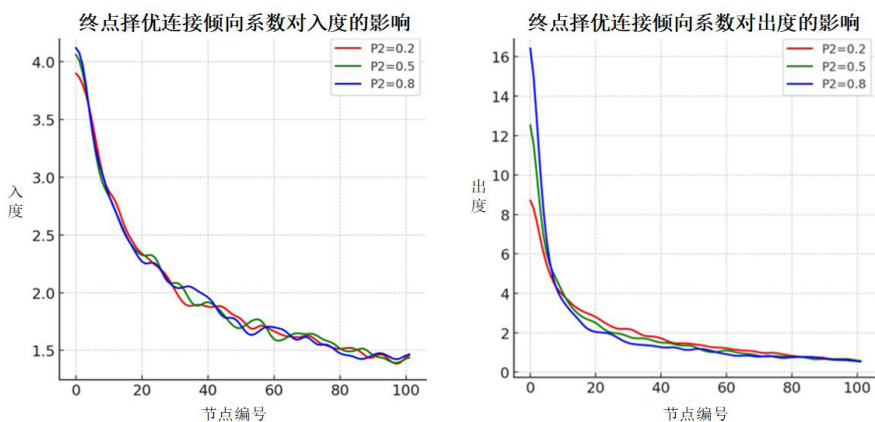


图5 终点择优连接倾向系数对节点度影响

### 2.3.4 仿真结论

在仿真实验中，我们首先进行了部分参数设定并构建了全球创新网络模型的可视化图片，并展示了在该参数设定下各企业知识输入与输出情况。其后探讨了知识输入概率  $P$ 、始点择优连接倾向系数  $P_1$  以及终点择优连接倾向系数  $P_2$  对各企业知识输入与输出的影响。实验发现知识输入概率  $P$  的增加导致新入驻的企业更多地扮演知识的接受者角色，从而刺激知识密集型企业的知识输出。而始点择优连接倾向系数  $P_1$  和终点择优连接倾向系数  $P_2$  的变化则影响了企业与哪些其他企业建立连接，显著影响知识密集型企业的知识输入与输出策略。

在对全球创新网络模型进行深入的研究与分析后，可以观察到企业间知识扩散的两种主要方式：KIEs 间的协同创新和 KIEs 与 LIEs 之间的知识扩散。

KIEs 间的协同创新是这个网络中的一个核心驱动力。这种协同作用主要起源于网络的初始设定。在网络演化的早期阶段，初始的参与者往往具有较高的知识积累和协作能力，使它们逐渐演化为知识的中心节点或所谓的知识密集型企业。这些企业倾向于相互之间建立更紧

密的合作关系，共同探索和创新，从而形成一个强大的协同创新的生态。

随着时间的推移和新企业的不断加入，一个明显的趋势是这些新入驻的企业，尤其是那些知识资源相对较少的 LIEs，更加倾向于与已有的 KIEs 建立联系。这是因为 KIEs 通常具有丰富的知识和经验，而 LIEs 则渴望获取这些知识以增强自身的竞争力。这种机制促进了 KIEs 与 LIEs 之间的知识扩散，确保了知识在整个网络中的广泛传播。

## 3 全球创新网络中企业间的知识扩散模式

### 3.1 知识创造

在本研究中建立的企业间知识扩散模型中，节点直接连接代表了企业之间产生了知识扩散的现象，但全球创新网络中的知识不是凭空产生的，而是通过企业或其他组织（高校、研究所等）进行知识创造而产生的。

企业的知识创造通过两种方式实现：企业内部的知识创造以及跨企业间的知识创造。企业内部的知识创造主要体现在员工与企业



之间显性知识与隐性知识的持续相互作用的过程中<sup>[32]</sup>。而跨企业间的知识创造则是环境的影响促使企业产生知识需求,知识创造的过程就是知识受到环境的影响,在不同组织之间的转移和扩散<sup>[33]</sup>。

在模型中,知识创造的过程体现在模型内与模型外。在模型中,企业之间知识扩散的过程是知识创造的重要方式,在节点的互相连接的过程中,全球创新网络会源源不断地产生新知识。而在模型之外,知识创造的过程包括了企业内部的研发、企业与其他组织之间的知识转移以及其他多种方式。其中以下几种方式均可以促进企业内部的知识创造:技术人员的研发、员工间的知识共享、企业的创新文化以及企业内部组建的创新团队。

### 3.2 KIEs之间的协同创新模式

由于本研究建立的企业间知识扩散网络模型具有明显的无标度性质,在模型演化的过程中会出现一些度值较高的节点,本研究将其定性分类为知识密集型企业。通过分析模型可知,模型中一些超级节点的入度较高,另一部分超级节点的出度较高,即一部分知识密集型企业展现了较高的知识吸收能力,另一部分展现了较高的知识输出能力。模型中KIEs之间的相互连接即为KIEs之间的协同创新,这种模式也是全球创新网络中知识创造的重要方式。

知识密集型企业间的协同创新则是一种基于知识的协作模式,旨在共同开发新产品、新服务或新技术。这种模式的基本理念是,通过将不同的技能、知识和资源整合在一起,企业

可以加速创新过程,并获得更好的创新结果。在这种协同创新模式中,企业之间共享知识和技能,互相学习和支持,以共同实现创新目标。企业可以通过共享知识和技术,节省时间和资源,并加快创新速度。此外,企业间的协作也可以促进不同领域的跨界创新,进一步推动全球创新网络的发展。

知识密集型企业之间的协同创新主要包括以下方式:

(1) 研究合作: KIEs可以通过与其他企业或研究机构合作,共同进行研究和开发项目。这种合作通常是针对某个特定领域或技术开展的,旨在共同解决技术难题,推动技术进步。

(2) 专利授权: KIEs可以通过与其他企业或机构签订专利授权协议,共享自己的专利或获得他人的专利使用权。这种合作有助于企业在技术领域中获得更大的优势,提高创新速度和效率。

(3) 技术许可: KIEs可以通过与其他企业或机构签订技术许可协议,共享自己的技术或获得他人的技术使用权。这种合作有助于企业在技术领域中实现更大的发展和创新。

(4) 供应链合作: KIEs可以通过与供应链上的其他企业或机构合作,共同提高供应链的效率和竞争力。这种合作可以涵盖从原材料采购到产品销售的整个供应链,从而帮助企业降低成本、提高质量和创新能力。

(5) 联合销售: KIEs可以与其他企业或机构联合销售产品或服务。这种合作可以扩大企业的市场规模,提高销售业绩和品牌知名度。

(6) 联合投资: KIEs可以与其他企业或机构共同投资某个项目或企业。这种合作可以

共同分担风险，实现资源共享，共同实现创新和商业目标。

全球创新网络中 KIEs 之间的协同创新模式见图 6。

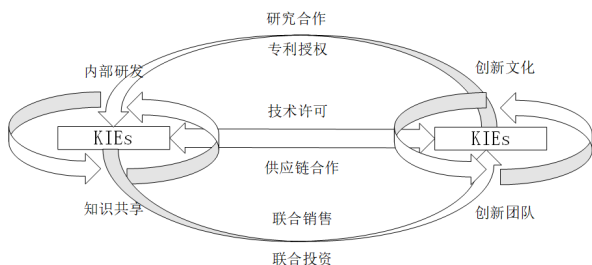


图 6 全球创新网络中 KIEs 之间的协同创新模式

### 3.3 KIEs 与 LIEs 之间的知识扩散模式

在本研究中建立的企业间知识扩散模型中，KIEs 与 LIEs 之间的知识扩散也承担了一部分的知识创造，其扩散的过程体现在新企业加入全球创新网络时与超级节点的连接。根据模型，LIEs 在加入全球创新网络时，更倾向于同 KIEs 产生知识联系，其倾向的强弱取决于始点和终点择优连接倾向系数  $P_1$  和  $P_2$ 。

而 KIEs 与 LIEs 之间的知识扩散模式是一个复杂的过程，涉及技术扩散、知识溢出、知识转移等多个方面：

(1) 技术扩散：KIEs 通常是技术创新的主要源头，LIEs 通常通过购买许可证、技术合作等方式获取新技术，从而提高生产效率和产品质量。同时 KIEs 可以从 LIEs 获取新技术产品或知识产品，以提高创新能力与创新效率。

(2) 知识溢出：知识溢出是指 KIEs 的技术、经验和知识在生产过程中不可避免地传递给 LIEs。例如，KIEs 的员工可能会离开企业，到 LIEs 工作，将他们在原企业获得的技术和经验带到新的工作场所。此外，KIEs 在与 LIEs

的合作中，也会向 LIEs 分享技术和经验，这有助于 LIEs 提高其生产效率和产品质量。

(3) 知识转移：知识转移是指 KIEs 和 LIEs 之间有目的的知识传递的过程。KIEs 可能会主动向 LIEs 提供培训和指导，以帮助它们掌握新的技术和工作方法。另外，KIEs 也可能会为 LIEs 提供技术支持和技术服务，以确保它们在生产过程中顺利地应用新技术。

此外，KIEs 与 LIEs 之间并非是定量区分的。在模型中，当一个 LIEs 在输入或输出大量知识后，其概念会发生改变，即产生了企业转型。LIEs 在转型为 KIEs 时，可以为企业带来诸多好处：提高企业竞争力、提高生产效率、提高产品附加价值、扩大市场份额以及提高员工素质。

KIEs 与 LIEs 之间的知识扩散模式见图 7。

### 3.4 基于 KIEs 和 LIEs 的全球创新网络中企业间知识扩散模式

研究基于知识密集型企业与劳动密集型企业概念来探究全球创新网络中企业间的知识扩散模式。根据本研究建立的模型，企业间知识扩散的模式分为知识创造、KIEs 之间的协同创新、KIEs 与 LIEs 之间的知识扩散三个部分。根据本部分的概述，全球创新网络中企业间知识扩散模式见图 8。

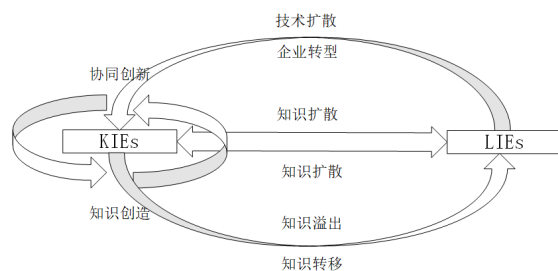


图 7 全球创新网络中 KIEs 与 LIEs 之间的知识扩散模式

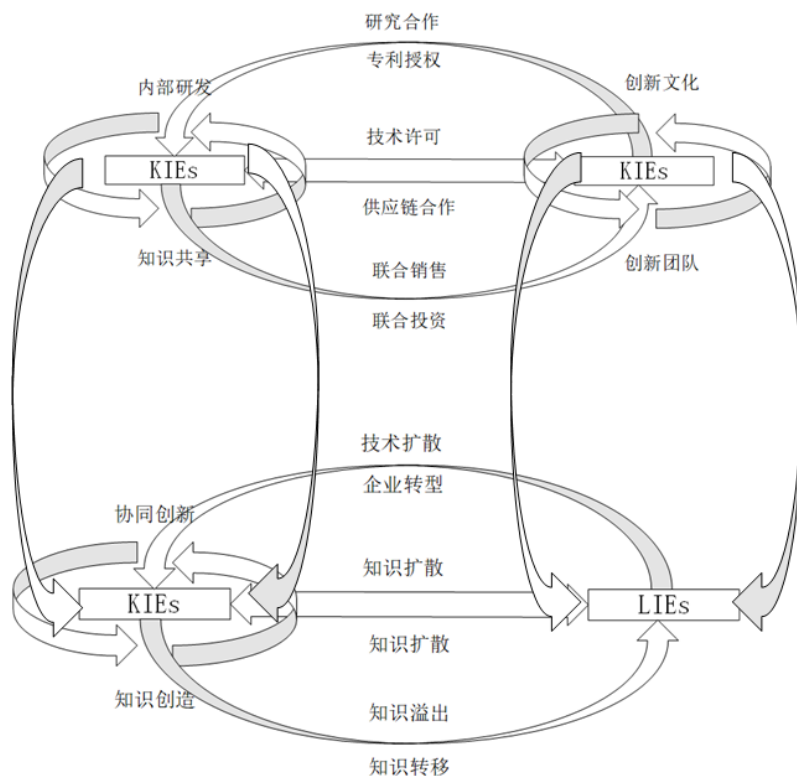


图8 全球创新网络中企业间知识扩散模式

## 4 研究总结

### 4.1 研究结论

企业之间的知识扩散本质上是知识在载体之间的一种流动过程<sup>[34]</sup>，而全球创新网络使这种知识的流动形成了一个系统，任何加入全球创新网络的企业均可输入或输出知识。在以往研究中，学者们通常会建立多种复杂网络模型来探究企业网络中的知识扩散过程，但实验中往往考虑不到知识扩散的有向性，以及企业的策略对网络演化的影响，导致对企业间知识扩散的模式的研究稍显不足。

相较于以往实验，本研究引入了参数知识输入概数  $P$ 、择优连接倾向系数  $P_1$ 、 $P_2$  以及初始网络起点集  $s$ ，终点集  $t$ ，建立了全球创新网

络中企业间知识扩散模型并进行仿真，通过仿真实验发现，当企业加入全球创新网络中时，其决策与  $P$ 、 $P_1$ 、 $P_2$  相关，并且其决策决定了该企业是作为知识输入者或输出者，以及其更倾向于同何种企业产生知识联系。

此外，本研究对模型中的企业进行了定性分类，将产生知识联系较多的企业分类为知识密集型企业，将产生知识联系较少的企业分类为劳动密集型企业。并根据模型发现，全球创新网络中企业间知识扩散模式主要分为知识创造、知识密集型企业之间的协同创新、知识密集型企业与劳动密集型企业之间的知识扩散三个部分。其中知识创造包括企业内部的知识创造以及跨企业间的知识创造，知识密集型企业之间的协同创新主要通过研究合作、专利授权等方式实现，知识密集型企业与劳动密集型企业

业之间的知识扩散包括了技术扩散、知识转移以及知识溢出等方面。并且在全球创新网络的不断演化中,某些劳动密集型企业也可能会朝着知识密集型企业产生企业转型,并创造更高的企业利益。

## 4.2 研究建议

不论是对于知识密集型企业还是劳动密集型企业而言,知识永远是企业收益与发展的最大活力源泉。然而仅靠企业内部的研发,远远不能达到提高企业竞争力的目的,因此促进企业之间的知识传递,积极地参与全球创新网络中的知识流动,便注定成为企业获取知识和收益的最佳选择。

对于知识密集型企业来说,由于很难从劳动密集型企业接受知识,因此需要更加积极地以开展内部研发工作为基础,同时扩展与其他企业的知识交流。例如企业可以开展联合研发项目,共同研究、开发和验证新技术、新产品或新服务。还可以开展专利授权或许可活动,让其他企业使用自己的专利技术,或者购买或交换其他企业的专利技术,并且可以与其他企业共同申请和拥有专利,共同享有专利权利和利益。另外,还可以与供应链上的其他企业共同进行研发、设计、制造、销售和服务等活动。除此之外,知识密集型企业还可以通过出售知识产品的方式来主动输出知识,并以此扩大自身收益。

对于劳动密集型企业来说,最有效的策略是积极地融入全球创新网络以获取知识,尤其是中小型的劳动密集型企业。由于知识密集型企业拥有更高的知识储量或更高的知识输出能力,劳动密集型企业应积极与知识更为密集的

企业产生知识联系,并获取所需要的知识,可以采取参与或投资合作研发项目等方式,以达到促进技术创新和产业升级的目的。此外,劳动密集型企业通常拥有较高的生产技术,此类企业也可以通过技术扩散的方式进行知识输出,并获得企业收益。

## 参考文献

- [1] PERKS H, JEFFERY R. Global network configuration for innovation: a study of international fibre innovation[J]. R&D Management, 2006, 36(1): 67-83.
- [2] BINZ C, TRUFFER B. Global Innovation Systems—A conceptual framework for innovation dynamics in transnational contexts[J]. Research Policy, 2017, 46(7): 1284-1298.
- [3] FREEMAN C. Networks of innovators: a synthesis of research issues[J]. Research Policy, 1991, 20(5): 499-514.
- [4] ERNST D. A New Geography of Knowledge in the Electronics Industry? Asia's Role in Global Innovation Networks[J]. Policy Studies, 2009(54): 1.
- [5] DURMUSOGLU S S. Open innovation: the new imperative for creating and profiting from technology[J]. European Journal of Innovation Management, 2004, 7(4): 325-326.
- [6] 全自强,李鹏翔,杨磊. 二元性创新、全球创新网络与企业逆向创新研究——基于领先企业和后发企业的多案例比较研究[J]. 科技管理研究, 2020, 40(14): 10-19.
- [7] 马琳,吴金希. 全球创新网络相关理论回顾及研究前瞻[J]. 自然辩证法研究, 2011, 27(1): 109-114.
- [8] NEPELSKI D, DE PRATO G. The structure and evolution of ICT global innovation network[J]. Industry and Innovation, 2018, 25(10): 940-965.
- [9] 甘志霞,娜孜木·叶鲁拜,唐文慧. 全球创新网络环境下我国低碳技术创新的模式研究[J]. 产业经济评论, 2015(3): 107-114.
- [10] 亨利·切萨布鲁夫,维姆·范哈佛贝克,乔·韦斯特. 开放创新的新范式[M]. 陈劲,李王芳,谢芳,等,译. 北京:科学出版社, 2010: 129-135.
- [11] 陈志明. 中国企业融入全球创新网络的路径——



- 知识获取与产品内分工整合的视角[J]. 科技管理研究, 2022, 42(14): 1-7.
- [12] OZEL B. 科学合作网络: 土耳其管理学术界的知识传播与碎片化[D]. 伊斯坦布尔: 伊斯坦布尔比尔基大学, 2010.
- [13] 岳增慧, 许海云, 方曙. 基于结构参数的科研合作网络知识扩散建模研究[J]. 情报学报, 2015, 34(5): 471-483.
- [14] 巴志超, 李纲, 朱世伟. 科研合作网络的知识扩散机理研究[J]. 中国图书馆学报, 2016, 42(5): 68-84.
- [15] 王静静, 叶鹰. 国际数字人文研究中的跨学科知识扩散探析[J]. 大学图书馆学报, 2021, 39(2): 45-51, 61.
- [16] 黄玮强, 庄新田, 姚爽. 基于创新合作网络的产业集群知识扩散研究[J]. 管理科学, 2012, 25(2): 13-23.
- [17] 孔晓丹, 张丹. 面向集群创新网络异质企业的知识扩散建模及仿真研究[J]. 运筹与管理, 2020, 29(10): 173-182.
- [18] LIN M, LI N. Scale-free network provides an optimal pattern for knowledge transfer[J]. *Physica A: Statistical Mechanics and its Applications*, 2010, 389(3): 473-480.
- [19] 刘浩广, 蔡绍洪, 张玉强. 无标度网络模型研究进展[J]. 大学物理, 2008(4): 43-47.
- [20] BETTENCOURT L M A, CINTRÓN-ARIAS A, KAISER D I, et al. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models[J]. *Physica A: Statistical Mechanics and its Applications*, 2006, 364: 513-536.
- [21] ZHANG Y, LI X, AZIZ - ALAOUI M A, et al. Knowledge diffusion in complex networks[J]. *Concurrency and Computation: Practice and Experience*, 2017, 29(3): e3791.
- [22] GAO X, GUAN J. Network model of knowledge diffusion[J]. *Scientometrics*, 2012, 90(3): 749-762.
- [23] ZHANG G, LIU L, WEI F. Key nodes mining in the inventor-author knowledge diffusion network[J]. *Scientometrics*, 2019, 118: 721-735.
- [24] QIAO T, SHAN W, ZHANG M, et al. How to facilitate knowledge diffusion in complex networks: The roles of network structure, knowledge role distribution and selection rule[J]. *International Journal of Information Management*, 2019, 47: 152-167.
- [25] XU L, DING R, WANG L. How to facilitate knowledge diffusion in collaborative innovation projects by adjusting network density and project roles[J]. *Scientometrics*, 2022, 127(3): 1353-1379.
- [26] DEL GIUDICE M, CARAYANNIS E G, MAGGIONI V. Global knowledge intensive enterprises and international technology transfer: emerging perspectives from a quadruple helix environment[J]. *The Journal of Technology Transfer*, 2017, 42: 229-235.
- [27] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks[J]. *Science*, 1999, 286(5439): 509-512.
- [28] GOMEZ-GARDENES J, MORENO Y. Local versus global knowledge in the Barabási-Albert scale-free network model[J]. *Physical Review E*, 2004, 69(3): 037103.
- [29] TACHIMORI Y, IWANAGA H, TAHARA T. The networks from medical knowledge and clinical practice have small-world, scale-free, and hierarchical features[J]. *Physica A: Statistical Mechanics and its Applications*, 2013, 392(23): 6084-6089.
- [30] KONNO T. Network effect of knowledge spillover: scale-free networks stimulate R&D activities and accelerate economic growth[J]. *Physica A: Statistical Mechanics and its Applications*, 2016, 458: 157-167.
- [31] BARABÁSI A L. Network science[J]. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2013, 371(1987): 20120375.
- [32] NONAKA I, TOYAMA R, KONNO N. SECI, Ba and leadership: a unified model of dynamic knowledge creation[J]. *Long range Planning*, 2000, 33(1): 5-34.
- [33] MAIER R, SCHMIDT A. Explaining organizational knowledge creation with a knowledge maturing model[J]. *Knowledge Management Research & Practice*, 2015, 13: 361-381.
- [34] ZHAO J, XI X, LI B, et al. Research on radical innovation implementation through knowledge reuse based on knowledge flow: A case study on academic teams[J]. *Information Management*, 2019, 57(8): 103260-103260.



开放科学  
(资源服务)  
标识码  
(OSID)

# 基于技术挖掘的投资机会分析模型研究

谷威<sup>1</sup> 龚颐雯<sup>2</sup>

1. 国家知识产权局专利局初审及流程管理部 北京 100088

2. 北京斯亚技术开发有限公司 北京 100029

**摘要:** [目的/意义] 技术挖掘可以把一些有用的信息,从科技创新信息中有效挖掘出来,基于技术挖掘的投资机会分析模型研究方法是机会分析领域的新尝试。[方法/过程] 将技术挖掘技术融入投资机会分析过程,从热点投资赛道发现、投资机会识别、投资机会评价三个方面揭示了投资机会发现过程。[结果/结论] 基于生物医药领域 2018-2022 年投资数据和生物医药专利数据进行了技术挖掘和投资机会分析进行了实例验证,发现了基因治疗下投资机会,加强技术创新成果和资本之间的联系,为科创投资机会寻找和发现提供了一个新的研究思路。

**关键词:** 技术挖掘; 投资机会分析; 投资机会识别; 投资机会评价

**中图分类号:** G35; F270

## Research on Investment Opportunity Analysis Model Based on Technology Mining

GU Wei<sup>1</sup> GONG Yiwen<sup>2</sup>

1. Preliminary Examination and Flow Management Department of the Patent Office, China National Intellectual Property Administration, Beijing 100088, China;

2. Beijing Siya Technology Development Co., Ltd, Beijing 100029, China

**Abstract:** [Objective/Significance] Technology mining can effectively extract useful information from technological innovation information. The research method of investment opportunity analysis model based on technology mining is a new attempt in the field of opportunity analysis. [Methods/Processes] Integrating technology mining techniques into the process of investment opportunity analysis reveals the process of investment opportunity discovery from three aspects: discovering hot investment tracks, identifying investment opportunities, and evaluating investment opportunities. [Results/Conclusions] Based on investment data and patent data in the biopharmaceutical field from 2018 to 2022, technology mining and investment opportunity analysis were conducted for instance verification. Investment opportunities under gene therapy were discovered, and the connection between technological innovation achievements and capital was strengthened, providing a new research approach for finding and

**作者简介** 谷威 (1978-), 硕士, 主要研究方向为专利审查、专利分析, E-mail: guwei@cnpa.gov.cn; 龚颐雯 (1979-), 硕士, 主要研究方向为金融投资分析、专利分析。

**引用格式** 谷威, 龚颐雯. 基于技术挖掘的投资机会分析模型研究 [J]. 情报工程, 2024, 10(1): 17-27.

discovering scientific and technological innovation investment opportunities.

**Keyword:** Technology Mining ;Investment Opportunity Analysis; Identification of Investment Opportunities; Investment Opportunity Evaluation

## 概述

硬科技概念正在被人们所熟知，具有较高技术门槛和技术壁垒的硬科技越来越受到资本的关注<sup>[1]</sup>。随着科创投资时代的来临，科学技术对投融资事件的影响越来越大，所有投资机构都在努力发现高价值技术赛道，并发现赛道中的高新技术的公司和研发团队。对科技资源进行监测，分析评价科学技术发展的状态和趋势，选择合理优先投资的科技领域、合理配置投资资金、有效开展科创投资的方法和途径已经成为国内投资机构普遍关注的热点。目前的机会分析方法以定量和定性分析相结合为基础，包括基于引文的方法<sup>[2-6]</sup>，基于主题词（关键词）的方法<sup>[7-10]</sup>，引文题的复合方法<sup>[11-14]</sup>，引证分析或聚类分析法<sup>[15-17]</sup>等，以上这些方法只是单一地进行分析，没有针对问题形成规范化的机会分析模型，更没有针对硬科技的投资机会模型进行研究，缺少对投资机会分析的方法体系展开进一步的发展和创新。理论界对于投资机会分析研究的缺乏与实际应用中投资机会分析对于技术创新研究的重要性和紧迫性形成了鲜明的反差，亟待更多的研究者们投入到对投资机会分析的研究和探索工作中来。

针对目前国内外尚未对投资机会发现模型开展系统研究的现状，以提高我国投资机构在进行科技投资决策时对投资机会的把握为思考起点，本文提出一种基于技术挖掘的投资机会

分析模型研究方法，将技术挖掘技术融入“投资赛道发现—投资机会识别—投资机会评价—投资机会决策”这一投资机会发现过程，分析与挖掘投资机会和可能的投资主体，文章最后进行了实证研究，结果也相对比较理想。

## 1 基于技术挖掘的投资机会分析模型研究框架

投资机会分析模型要解决“未来做什么”这一问题，这也正是投资机会分析所要实现的目标。投资机会分析是对特定技术赛道的技术细节和子技术的技术发展情况进行深入地剖析，达到评估和预测技术赛道发展方向、发现有发展潜力的企业投资机会的目的，为投资机构能够更有效地利用有限资源找到技术创新主体的投资活动提供决策依据。因此，通过技术挖掘可以把一些有用的信息，从科技创新信息中有效挖掘出来，不能让有用的信息淹没在如海的数据中，要识别出感兴趣的话题或主题，分解成具体问题，再寻找答案；还要开发出一系列指标，再去找相关数据库和数据来源进行挖掘；利用适当分析工具，比如某些数据分析器，回答之前提出的一系列问题，从而看清技术的发展趋势，进而找到适应发展趋势的投资方向。由此可见，基于技术挖掘的投资机会分析模型可以采用“投资赛道发现—投资机会识别—投资机会评价—投资机会决策”的研究主线，对

投资机会分析的方法展开研究和探讨，主要的研究成果总结如图 1 所示：

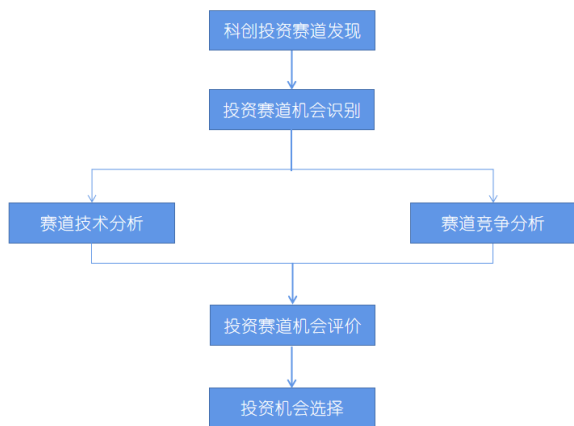


图 1 基于技术挖掘的投资机会分析模型研究框架

### (1) 基于投融资数据实现的热点投资赛道发现模型

由于投资赛道发现是投资机会分析首先要解决的问题，本文首先对基于投融资数据实现的热点投资赛道发现模型展开了研究。针对目前国内外热点投资赛道发现方法单一的问题，对基于多指标分析的热点投资赛道发现方法进行了改进，提出了基于投融资数据 10 个重要指标和逻辑回归分类器算法实现热点投资赛道的训练和预测。

### (2) 基于投资赛道技术分析和竞争分析的投资机会识别系统模型

针对目前投资机会分析领域研究零散，缺乏系统方法体系的问题，本文基于投资赛道技

术分析和竞争分析，提出了投资机会识别系统模型，分别从技术研发、竞争环境两个层面对投资机会进行分析。在技术研发层面，提出了基于技术形态的关键词提取模型；在竞争环境层面上，提出了基于关键词和机构共现分析的竞争环境分析模型。

### (3) 基于技术评价的投资机会评价模型

针对投资机会识别模型得到的技术赛道细分的子技术赛道，在技术要素因子分析基础上，结合专利数据在技术评价中的作用，构建了投资机会评价模型。实现了对子技术赛道的评分和排序，为投资赛道选择奠定基础。

## 2 基于投融资指标体系实现热点投资赛道发现

投资赛道发现研究的重点在于分析投资赛道是否处于现在的热点投资赛道，因此科创投资赛道发现就是基于投融资数据的热点投资赛道选择问题，本文采用了基于多指标分析的热点投资赛道发现方法进行热点投资赛道发现。热点投资赛道发现的核心是基于投资赛道指标进行热点分析，找到热点投资赛道，在投资赛道发现中选择的指标分别是投资次数、投资金额、投资机构数量和获投企业数量，具体指标的含义和应用如下表所示。

表 1 投融资指标体系表

指标名称	含义	应用	权重
投资次数	某一段时间内赛道获得投资的次数	进行赛道的投资数量对比	定量
投资金额	某一段时间内赛道获得的投资金额	进行赛道的投资金额对比	定量
投资机构数量	某一段时间内赛道中投资的投资机构数量	进行赛道的投资机构数量对比	定量
获投企业数量	某一段时间内赛道中投资的获投企业数量	反映赛道的获得投资企业数量对比	定量



### 3 基于投资赛道技术分析和竞争分析的投资机会识别系统模型

#### 3.1 投资赛道技术研发分析

在技术研发分析层面，基于技术形态分析理论（Morphology Analysis，简称MA），本文通过建立技术树，对特定技术领域核心技术科技文献的申请情况进行分析，从而帮助识别和预测未来该技术的技術发展方向。技术形态分析（MA）是对现有和未来技术结构以及新技术发明方向进行分析和预测的一种系统分析方法。技术形态分析引入一种构建和分析某项技术、机构或者社会问题的非量化建模方法。这种方法主要是通过将研究对象分解为几个基本维度来对问题进行建模。因此，在对某特定技术

进行技术研发层面的分析时，引入技术形态分析可以帮助我们通过分析技术结构，对特定技术领域的技术细节和子技术的技术发展情况进行更深入的剖析，达到评估和预测技术发展方向、发现有发展潜力的投资机会的目的，为投资机构能够更有效地利用有限资源开展技术投资活动提供决策依据。

本研究采用的是基于不同字段核心技术关键词与申请时间的关系矩阵或者折线图方法，来帮助快速识别核心技术随时间发展的情况，从而为有效评估核心技术的投资机会提供信息支持。这种相关关系同样是通过核心技术在某段申请时间申请的专利数量来衡量的。在这里，核心技术和申请时间作为实体，由一个节点来表示。通常，不同类型的技术领域可以根据需要通过颜色进行区分。如图2所示。

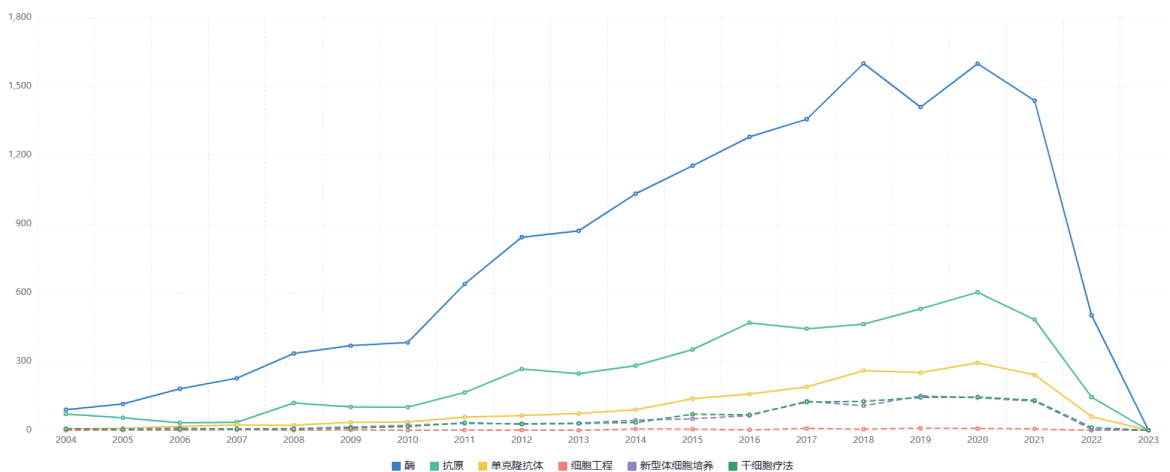


图2 投资赛道技术研发分析

#### 3.2 投资赛道竞争环境分析

技术革新是在硬科技产业中持续胜出的科创企业的核心成长驱动力。因此在技术所处的竞争环境分析中，要在技术研发层面分析的基础上，进一步对有潜力的核心子技术及其重要

的参与者进行分析，同时分析技术所处的整体竞争环境，从而帮助我们进一步了解相应技术都有哪些创新主体在积极参与，各个创新主体的研发兴趣和强项都在哪些方向领域。

研发机构和企业和技术发展的过程中扮演

着将这些有潜力的投资机会转化成产品，并最终投入市场的角色。通过竞争环境分析，可以帮助我们在识别潜力投资机会的基础上，锁定那些在特定技术发展方向具有竞争力的创新主体。从创新主体的角度出发，则是可以发现那些潜在的技术对手，通过竞争环境分析制定相应的技术创新策略。因此，我们通过构建创新主体与技术的关联模型，来帮助创新主体在整个技术领域找准定位，把握优劣势，从而更好地把握投资机会，将其转化成创新动力。

本研究采用的是基于不同字段的关键词与机构共现的关联图方法，来帮助创新主体快速识别技术竞争者，从而为有效评估自身所有的投资机会提供信息支持。技术—创新主体关联图是以网络的方式，来展示技术与创新主体的关联关系。这种关联关系同样是通过创新主体在某个技术领域发表科技文献的数量来衡量的。在这里，技术和创新主体作为实体，由一个节点来表示。节点的大小代表了该实体所包含的科技文献量的多少。数量越多，节点越大。而代表技术的节点与代表创新主体的节点之间的连线代表了它们之间的关联关系。如果创新主体在某技术领域发表了论文或者专利，那么代表它们的两个节点之间则存在连线。线的粗细代表了发表科技文献数量的多少，即关联强度。关联强度越大，线越粗。通常，为了便于观察，技术节点与创新主体节点在技术—创新主体关联图中可以以节点颜色的不同进行区分。同样地，不同类型的创新主体也可以根据需要，通过颜色进行区分。技术—创新主体关联图相对于技术雷达图的优点在于，可以显示更大数量的技术和创新主体之间的关联关系，同时，

还可以根据需要显示和区分不同类型的创新主体。但是，在比较不同创新主体在技术领域的实力以及显示创新主体在技术领域的资源配置上，技术雷达图则相对更加精确。因此，当我们在进行竞争环境分析时，往往需根据需要选用合适的研究工具来帮助我们展示研究结果。

## 4 基于技术评价的投资机会评价模型

单纯从科技文献的角度识别投资机会是不够的，还需要结合其他的影响要素因子对识别的投资机会进行更加全面的评价。在技术要素因子分析基础上，结合专利数据在技术评价中的作用，构建了投资机会评价模型。

基于技术评价的投资机会评价模型的重点，就是技术的专利指标体系的确定。专利指标从大类上可以分为数量指标和质量指标两大类。专利数量指标主要包含专利数量、专利发明人、专利机构数等数量上简单的统计指标，并没有考虑专利所包含的技术质量的差别。而专利质量指标则是通过多种途径代替衡量专利所包含的技术质量水平，如专利引证率、专利当前影响指数、同族专利数等。在本章中，我们选取了如表2所示的六个指标来衡量技术的研发水平。其中，专利数量、专利成长率以及发明人数量指标属于专利数量指标；同族专利数量、当前影响指数以及研发机构数量属于专利质量指标。

投资机会评分公式如下：

$$v = \sum_{i=1..7} w_i f_i \quad (1)$$

其中  $f$  是指标， $w$  是指标权重。

表 2 投资机会评分指标表

指标名称	含义	应用	方法
专利数量	某技术一段时间内专利申请或授权的数量	进行技术领域或研发机构的专利数量对比, 或不同时间段的趋势分析	定量
同族专利数量	某技术具有共同优先权的一组专利数量	反映专利的质量	定量
专利成长率	某技术在某段时间获得的专利数量对比上一阶段的增长幅度	分析技术创新或成果专利化的发展速度	定量
发明人数量	某技术的发明人总数	反映专利或专利组合的发明团队规模	定量
当前影响指数	某产业或企业前五年专利的当年被引次数与平均被引次数的平均值的比值	反映了专利的技术质量以及影响力	定量
研发机构数量	从事某技术研究的机构数量	反映了专利研究的热度	定量

熵权法是一种客观赋权方法。在具体使用过程中, 熵权法根据各个指标的不同变异程度计算出各指标的熵权, 并通过熵权修正各指标的权重, 最后得出较为客观的指标权重。投资机会评价模型采用熵权法作为投资机会评分权重的计算方法。熵权法的基本思路如下: 现有  $m$  个待评项目,  $n$  个评价指标, 形成原始数据矩阵:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix}$$

其中  $r_{ij}$  为第  $j$  个指标下第  $i$  个项目的评价值。

计算各指标值权重的过程为:

(1) 计算第  $j$  个指标下第  $i$  个待评价项目的指标值的比重  $p_{ij}$ :

$$p_{ij} = \frac{r_{ij}}{\sum_{i=1}^m r_{ij}} \quad (2)$$

(2) 计算第  $j$  个指标的熵值  $e_j$ :

$$e_j = -k \sum_{i=1}^m p_{ij} \cdot \ln p_{ij} \quad (3)$$

其中,  $k = \frac{1}{\ln m}$

(3) 计算第  $j$  个指标的熵权  $w_j$ :

$$w_j = \frac{(1 - e_j)}{\sum_{j=1}^n (1 - e_j)} \quad (4)$$

## 5 实验及分析

为验证基于技术挖掘的投资机会分析模型的可行性, 本文选择了生物医药产业投资机会进行了实例验证, 加强技术创新成果和资本之间的联系, 为寻找和发现生物医药产业科创投资机会提供了一个新的研究思路。

### 5.1 生物医药投资赛道分类

确定生物医药产业方向后, 本研究选取了 Wind 数据库中 2018—2022 年投融资数据作为研究对象, 并对融资企业进行生物医药标注, 标注类别是生物医药或非生物医药企业。对标注为生物医药的 985 家企业进行赛道分类, 根据 985 家企业特点, 将赛道分类为上游的生物技术、原材料, 中游的疫苗、血液制品、抗体、细胞治疗、基因治疗、重组蛋白、诊断试剂和核酸药物等分类。

### 5.2 热点投资赛道发现

选用投资次数特征作为热点投资赛道发现指标, 数字如表 3 所示, 可以看到原材料、抗体、基因治疗等 3 个子赛道的投资次数较多。

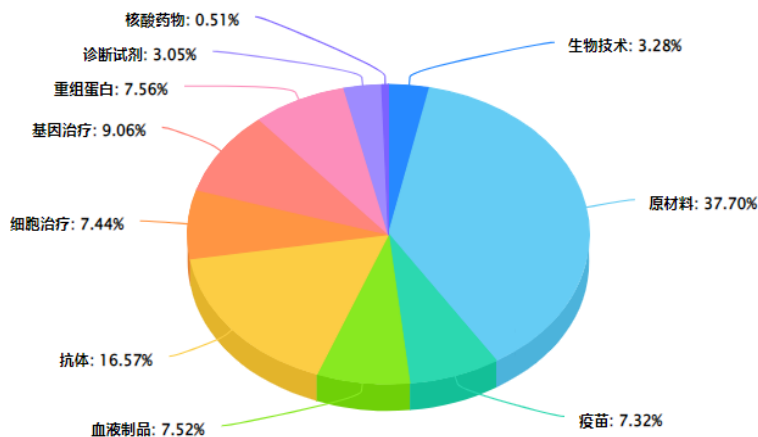


图3 热点投资赛道发现分析

表3 投资次数表

技术	数量
原材料	953
抗体	419
基因治疗	229
重组蛋白	191
血液制品	190
细胞治疗	188
疫苗	185
生物技术	83
诊断试剂	77
核酸药物	13

因此选择原材料、抗体、基因治疗三个赛道作为投机机会分析和识别的投资方向。考虑到原材料赛道过于宽泛，抗体赛道又过于狭窄，因此以基因治疗赛道为例，进行投资机会分析与识别和投资机会评价。

### 5.3 投资机会分析与识别

#### 5.3.1 投资赛道技术研发分析

基因治疗，也称为细胞和基因治疗，是一种利用基因治疗载体将外源的治疗性基因转导至细胞，再通过外源基因的转录和翻译，改变

细胞原有基因表达以治疗疾病的方法。在技术研发分析层面，基于技术形态分析理论，投资赛道技术研发分析首先建立技术树，对特定技术领域核心技术的科技文献申请情况进行分析，从而帮助识别和预测未来该技术的技術发展方向。在投资赛道研发分析方法介绍了通过科技文献数据库识别和分析目标技术机会的方法和步骤。其中重要的一步是基于技术形态的关键词分析。在这里，我们基于基因治疗分析构建专利检索式，在检索数据的基础上构建专利数据库，对基因治疗技术的发展趋势和热点展开研究，并通过分析和比较，识别基因分析的技术前景和投资分析。

在对基因治疗技术的介绍中，提到了该技术的技術构成，主要由病毒载体、非病毒载体等几部分构成。通过关键词识别技术我们获取了技术关键词，在这些关键词中，我们依照关键词提取方法筛选出了与这两个核心技术相对应的技术形态关键词。表4中显示了部分提取的关键词及其与技术之间的对应关系。从表4中我们可以看到根据提取的与病毒载体有关的关键词，腺病毒、痘病毒都可以作为用于基因

治疗的病毒载体。

表 4 基因治疗技术关键词表

序号	技术	关键词
1	病毒载体	单纯疱疹病毒、肠道病毒、伪狂犬病毒、腺病毒、痘病毒、水疱性口炎病毒、麻疹病毒、呼肠孤病毒、新城疫病毒、细小病毒、流感病毒、腮腺炎病毒、人类免疫缺陷病毒、鸡贫血病毒、甲病毒
2	非病毒载体	基因敲除、基因纠正

从表 5 中可以看出腺病毒、流感病毒、痘病毒、伪狂犬病毒、呼肠孤病毒是主要研

究的病毒种类。核心技术—申请时间关系矩阵是以表格的方式，来展示核心技术与申请时间的相关关系。这种相关关系同样是通过核心技术在某段申请时间申请的专利数量来衡量的。在这里，核心技术和申请时间作为实体，由一个节点来表示。节点的大小代表了核心技术所包含的专利申请量的多少。数量越多，节点越大。通常，不同类型的技术领域可以根据需要通过颜色进行区分。腺病毒、痘病毒、HSV-1、细胞治疗也是近年来申请量较为集中的领域。

表 5 投资赛道技术研发分析表

	2012	2013	2014	2015	2016	2017	2018	2019	2020
腺病毒	50	64	50	52	69	78	89	67	97
流感病毒	22	16	14	18	16	11	16	17	16
痘病毒	3	6	14	19	12	24	14	30	21
伪狂犬病毒	1	2	3	10	1	8	2	14	8
呼肠孤病毒	3	8	0	4	0	6	1	2	2
肠道病毒	3	1	2	0	2	6	3	3	6
水疱性口炎病毒	2	0	2	0	0	1	6	2	4
人类免疫缺陷病毒	1	2	0	4	1	1	2	0	0
麻疹病毒	0	2	1	0	0	0	0	0	3
甲病毒	0	0	4	0	0	0	2	0	6
疱疹病毒	0	1	1	0	0	0	0	2	5
新城疫病毒	0	0	2	2	0	0	2	0	1
细小病毒	0	0	0	0	0	3	0	4	0
鸡贫血病毒	0	0	0	0	1	1	0	0	0
腮腺炎病毒	0	0	0	0	1	0	1	1	0

### 5.3.2 投资赛道技术竞争分析

构建创新主体与技术的关联模型，来帮助创新主体在整个技术领域找准定位，把握优劣势，从而更好地把握投资机会，将其转化成创新动力。

除了国家角度的竞争者分析，还需要从研究机构和企业层面分析各机构的技术分布情况。图 4 是腺病毒、痘病毒、单纯疱疹病毒中专利

排名前 20 企业的技术分布图。图中蓝点代表企业，橙点代表技术方向。其中三维生物、恩宝生物、普莱柯生物、元宋生物、本元正阳基因、山东信得、医诚生物、安宇生物、康希诺、康万达、达博、佰芮慷生物与腺病毒有强关联关系；恩宝生物、普莱柯、信得生物在流感病毒的研究比较突出，唯可达生物、功楚生物、温氏食品是痘病毒生产的主要企业。



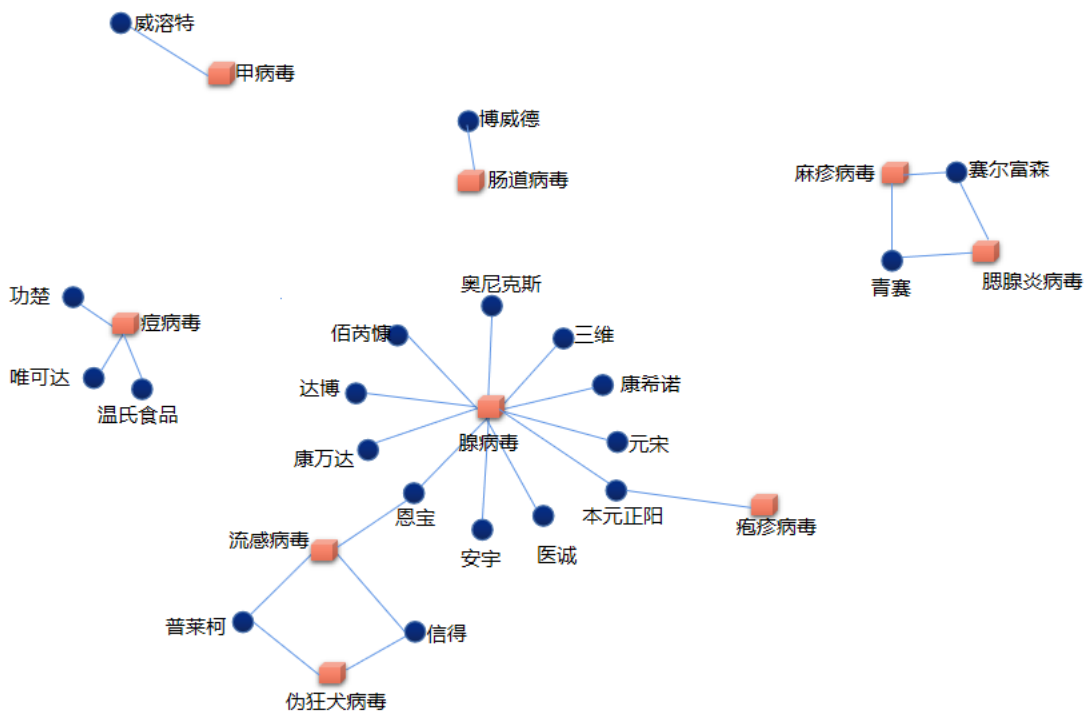


图 4 病毒排名企业技术分布图

### 5.4 投资机会评价

基于基因治疗的专利数据，我们识别出了 5 项具有发展潜力的子技术，分别是腺病毒、流感病毒、痘病毒、伪狂犬病毒和呼肠孤病毒。根据技术研发水平评价指标的定义和计算公式，

我们获得了这 5 项基因治疗子技术的研发水平指标数据，如表 6 所示。在投资机会评价模型中，根据熵权法我们获得了技术评价指标的权重水平（如表 7 所示），最后计算了基因治疗子技术投资机会评价结果如表 8 所示。

表 6 因治疗子技术的研发水平评价指标值

投资机会	专利数量	同族专利数量	专利成长率	发明人数量	研发机构数量	当前影响指数
腺病毒	1180	850	-13.76%	2510	459	1.09
流感病毒	296	194	118.75%	652	113	0.8707
痘病毒	254	185	22.58%	499	119	1.0566
伪狂犬病毒	69	49	0%	260	36	0.8276
呼肠孤病毒	36	26	50%	127	19	0.6452

表 7 投资机会评价指标权重

序号	指标名称	权重	序号	指标名称	权重
1	专利数量	0.19883349	4	发明人数量	0.19455968
2	同族专利数量	0.20341006	5	研发机构数量	0.18890743
3	专利成长率	0.13381174	6	当前影响指数	0.08047759

表 8 基因治疗子技术投资机会评价结果

序号	技术名称	投资机会评分	序号	技术名称	投资机会评分
1	腺病毒	0.86645588	4	伪狂犬病毒	0.07831459
2	流感病毒	0.34580486	5	呼肠孤病毒	0.06625759
3	痘病毒	0.26305361			

从评价的结果来看,腺病毒是基因治疗研究领域目前研发水平最高的子技术形态,其次研发水平较高是流感病毒和痘病毒技术,对伪狂犬病毒和呼肠孤病毒等研究领域研发水平较低。根据 2022 年 7 月 1 日—2023 年 7 月 1 日的生物医药投融资统计数据,腺病毒融资达到 4 件,流感病毒 2 件,痘病毒 1 件,伪狂犬病毒 1 件,呼肠孤病毒 0 件,可见投资机会评价是有效的。

## 6 总结与展望

本文提出一种基于技术挖掘的投资机会分析模型研究方法,形成了投资机会分析的“投资赛道发现—投资机会识别—投资机会评价—投资机会决策”的研究主线和投资机会发现过程,将技术挖掘技术融入投资机会分析过程,主要的创新成果总结如下:(1)基于投融资指标体系实现热点投资赛道发现;(2)基于投资赛道技术分析和竞争分析的投资机会识别系统模型;(3)基于技术评价的投资机会评价模型。基于以上三个模型实现了投资机会发现过程,分析与挖掘投资机会和可能的投资主体。最后,我们基于生物医药领域 2018—2022 年投资数据和生物医药专利数据进行了投资机会分析模型的实例验证,加强技术创新成果和资本之间的联系,为科创投资机会寻找和发现提供了一个

新的研究思路。

对于投资机构的投资行为和企业并购活动,其任务的重点在于结合新的市场需求,找到科技研究成果具有市场潜力的企业。可见投资行为是为了找到被投主体,更是为投资机构推荐被投主体,并对被投主体进行尽职调查,辅助企业在实现自身价值的同时推进产业的发展与变革,因此下一步的工作将在本文研究的模型基础上,进行被投企业的推荐研究。

## 参考文献

- [1] 黄葆春. 基于投融资和专利融合的投资方向发现模型[J]. 情报工程, 2022, 8(2): 109-118.
- [2] NAVONIL M, KORINA K, PAUL F. Exploring the modelling and simulation knowledge base through journal co-citation analysis[J]. *Scientometrics*, 2014, 98(3): 2145-2159.
- [3] TRUJILLO C M, LONG T M. Document co-citation analysis to enhance transdisciplinary research[J]. *Science Advances*, 2018, 4(1): e1701130.
- [4] Blaginin V, Smirnova A, Ergunova O, et al. Russia on world research front of industrial scientific direction[C]//International Conference "Actual Issues of Mechanical Engineering" 2017 (AIME 2017). Atlantis Press, 2017: 113-119.
- [5] 章小童, 阮建海, 引文网络主路径分析法演化脉络及研究现状的文献计量分析[J]. 情报资料工作, 2016(5): 61-66.
- [6] SHIBATA N, KAJIKAWA Y, TAKEDA Y, et al. Detecting emerging research fronts in regenerative

- medicine by the citation network analysis of scientific publications[J]. *Technological Forecasting and Social Change*, 2011, 78(2): 274-282.
- [7] KLEINBERG J. Bursty and hierarchical structure in streams[J]. *Data Mining and Knowledge Discovery*, 2003, 7(4): 373-397
- [8] 洪娜, 张智雄, 乐小虬. 基于决策树的潜在爆发词探测方法 [J]. *情报学报*, 2012, 31(3): 228-241.
- [9] 赵丽梅, 张花. 我国大数据时代数字图书馆研究前沿分析 - 基于共词分析的视角 [J]. *情报科学*, 2019(3): 97-104.
- [10] 关鹏, 王曰芬, 傅柱. 不同语料下基于 LDA 主题模型的科学文献主题抽取效果分析 [J]. *图书情报工作*, 2016(2): 112-121.
- [11] 马腾, 曹吉鸣, 申良法. 知识转移研究演进脉络梳理及前沿热点探析——基于引文分析和共词分析 [J]. *软科学*, 2016(2): 121-125.
- [12] VAN D, BESSELAAR P, HEIMERIKS G. Mapping research topics using word- reference co-occurrences: a method and an exploratory case study[J]. *Scientometrics*, 2006, 68(3): 377-393
- [13] 张艺蔓, 马秀峰, 程结晶. 融合引文内容和全文本引文分析的知识流动研究 [J]. *情报杂志*, 2015(11): 50-54, 49.
- [14] KONG D J, LI M, ZHENG W J. To identify technology frontier for mass- customized production service converged with artificial intelligence based on patent data mining [C]//15th International Conference on Service Systems and Service Management (ICSSSM), 2018: 1-6.
- [15] 张婷, 安嘉璐. 基于专利分析的医学科技重点技术前沿领域的识别研究 [J]. *现代生物医学进展*, 2015, 15(32): 6371-6376.
- [16] 张振刚, 黄洁明, 陈一华. 基于专利计量的人工智能技术前沿识别及趋势分析 [J]. *科技管理研究*, 2018(5): 36-42.
- [17] FUJIMAGARI H, FUJITA K. Detecting research fronts using neural network model for weighted citation network analysis[J]. *Journal of Information Processing*, 2015, 23(6): 753-758.





开放科学  
(资源服务)  
标识码  
(OSID)

# 基于知识网络特征的产业关键共性技术识别模型与实证研究

——以新一代信息技术产业为例

刘静 胥彦玲 张婧

北京市科学技术研究院科技情报研究所 北京 100044

**摘要:** [目的/意义] 构建产业关键共性技术识别模型并进行实证研究,为产业技术创新布局提供参考。[方法/过程] 从技术领域和技术主题两个层面,探索知识网络特征与产业关键共性技术特征的关系:将技术领域共现网络中具有相似性质的节点集合作为产业共性技术领域;根据网络中节点的结构位置,计算技术主题特征指标,最终筛选出产业关键共性技术。[局限] 该模型未考虑不同国家或地区之间技术竞争与布局的差异,后续研究将进一步探讨。[结果/结论] 本文以新一代信息技术产业为例,识别出8个共性技术领域和20个技术主题。本文研究模型从专利文献内容和技术发展规律两个方面综合考量,确保了模型及其结果的科学性、合理性和可靠性。

**关键词:** 产业关键共性技术;知识网络;节点特征;社区发现算法;技术生命周期

**中图分类号:** G35; TP391

## Model and Empirical Study on Identification of Industry Key Generic Technologies based on Knowledge Network Characteristics: Evidence from the New Generation of Information Technology Industry

LIU Jing XU Yanling ZHANG Jing

Institute of Science and Technology Information, Beijing Academy of Science and Technology, Beijing 100044, China

**Abstract:** [Objective/ Significance] Construct the identification model of industry key generic technologies and conduct empirical research. It will be an important basis for governors to formulate suitable policies and strategies. [Methods/Processes] This paper explores the relationship between the characteristics of knowledge network and industry key generic technologies, from both of technical field and technical theme. In the technology field co-occurrence network, the nodes with similar properties are combined to form the industry generic technology field. In the technology theme co-occurrence network, according to the

**基金项目** 北京市科学技术研究院创新培育类项目(1220220491KF0049)。

**作者简介** 刘静(1982-),博士,高级工程师,主要研究方向为科学计量学与科技评价;胥彦玲(1977-),博士,副研究员,主要研究方向为产业技术情报研究;张婧(1985-),硕士,副研究员,主要研究方向为情报理论、方法与技术, E-mail: jinglqq@sina.com。

**引用格式** 刘静,胥彦玲,张婧.基于知识网络特征的产业关键共性技术识别模型与实证研究——以新一代信息技术产业为例[J].情报工程,2024,10(1):28-41.

node's structural position that is the topological structure in the network node, the indicators of technology theme characteristics are calculated, and finally the industry key generic technologies are selected. [Limitations] Yet this paper does not consider the differences of technological competition and layout among different countries or regions, which needs to be further discussed in future research. [Results /Conclusions] Taking the new generation of information technology industry as an example, eight common technology fields are identified, and 20 technology themes are selected as the industry key genetic technologies. This model is comprehensively considered from the content of patent literature and the law of technological development, that ensures the scientificity, rationality and reliability of the model and its results.

**Keywords:** Industry Key Generic Technologies; Knowledge Network; Node Characteristics; Community Discovery Algorithm; Technology Life Cycle

## 引言

产业关键共性技术是制造业创新发展的重要支撑，是促进国家产业转型升级、加快制造强国建设、厚植中国经济发展根基和竞争优势的重要抓手。2018年以来，美国相继针对中兴、华为等中国高科技企业进行无理打压，这些事件凸显了加快我国相关产业关键共性技术突破的重要性与紧迫性。为改变我国关键共性技术受制于人的局面<sup>[1]</sup>，“十四五”规划纲要明确提出“集中力量整合提升一批关键共性技术平台，支持产业共性基础技术研发”。由此，准确识别、遴选产业关键共性技术成为了产业技术创新战略决策中亟待解决的问题。

专利是一种重要的创新产出，尤其是在产业技术研发阶段，发明专利更能体现最核心、最直接的自主创新能力<sup>[2]</sup>。PCT专利作为发明人寻求国际专利保护的一种途径，反映出发明人全球技术布局的意向，其规模和水平代表了申请人以专利为竞争手段开拓国际市场的实力<sup>[3]</sup>。通过分析某一产业或行业内PCT专利文献，能够客观地反映技术创新整体概况和发展态势，从情报学角度为宏观把握产业技术创新态势提供新的研究视角。

本文在梳理产业共性技术、产业关键技术、产业关键共性技术的概念、特征与识别方法的基础上，探索知识网络在产业技术多角度评估中的应用，提出基于PCT专利分析和知识网络特征的产业关键共性技术识别模型，以期为政府对产业关键共性技术的选择和管理提供借鉴。

## 1 相关研究现状述评

### 1.1 产业关键共性技术相关概念与特征

产业关键共性技术、产业共性技术、产业关键技术等相似说法都是在特定的历史背景或社会政治经济条件下产生，它们之间既有区别又有联系。

产业共性技术是对整个行业或产业技术水平、产业质量和生产效率都会发挥迅速的带动作用，具有巨大的经济和社会效益的一类技术<sup>[4]</sup>。共性技术源起于科技政策实践，其概念具有很强的政策性，其内涵受到一个国家经济和技术发展水平的影响<sup>[5]</sup>。从科技管理实践的角度，产业共性技术具有关联性和使能性，它能够应用于多个产业或企业，能够推动多项技术的进步，促进多部门经济社会效益的提高<sup>[6]</sup>。从技术经济学的角度，产业共性技术包含最先应用的科

学知识,能够扩散、溢出到其他技术领域,为后续技术的研发、应用、推广提供基础,技术使用者众多,技术应用范围宽广,因此具有核心性、广泛性和效益性<sup>[7]</sup>。从技术研发进程的角度,产业共性技术作为“竞争前阶段”的技术,衔接了基础研究和面向市场的商业技术开发,还具有准公共品的性质<sup>[8]</sup>。

与产业共性技术一样,产业关键技术也在国家科技规划等政策文件中频繁出现,学术界对其并没有统一定义。从国家层面,国家关键技术是为振兴国家产业,提高国际竞争力,促进经济持续增长,改善人们生活质量,提高国家强盛起决定性作用的技术<sup>[9]</sup>。从产业层面,产业关键技术是指对科技进步有决定性影响,能提高产业竞争力、促进经济和社会持续发展,而且适应科技发展趋势,在一定时间内具有重要商业价值的“瓶颈”技术<sup>[10]</sup>。产业关键技术决定着产业技术领域的发展方向和产品的整体性能<sup>[11]</sup>,关键技术创新能够显著增强产业部门产品的附加值、提升生产效率、优化产品质量,突破产业发展的障碍,打破关键技术瓶颈,从而对产业经济的增长产生决定性的推动作用<sup>[12]</sup>;在未来较长时期内,产业关键技术对国家安全和经济社会发展都会产生重要影响<sup>[13]</sup>。不难看出,产业关键技术具有战略性、前瞻性、重要性、引领性、瓶颈性。

关键共性技术最初出现在科技政策实践中,后逐渐扩展到学术界。国家工信部在《产业关键共性技术发展指南(2011年)》中将产业关键共性技术定义为“能够在多个行业或领域广泛应用,并对整个产业或多个产业产生影响和瓶颈制约的技术”。可见,产业关键共性技术

并不是面向某个企业的技术,而是服务于一个甚至多个产业战略性发展的技术。目前,学术界对产业关键共性技术的概念尚未形成统一的界定。本研究中的产业关键共性技术既包括产业共性技术的特征,又兼具产业关键技术的地位。当某项技术能够在多个领域被普遍应用,实现跨领域的扩散与融合,其研发成果可共享,并对整个产业技术水平、产业质量和生产效率都会产生迅速的提升作用,可以判定该项技术属于产业关键共性技术。因此,本文将从重要性、通用性、基础性、瓶颈性四个特征指标测度技术主题,识别产业关键共性技术。

## 1.2 产业关键共性技术相关识别方法

目前,产业共性技术、产业关键技术、产业关键共性技术识别研究的方法分为定性分析法和定量分析法。

定性分析法主要基于德尔菲调查,例如,袁思达<sup>[14]</sup>应用德尔菲调查识别出氢能源领域的共性技术;然而,德尔菲调查过度依赖专家的技术知识,缺少精准的量化指标而显得过于主观,于晓勇等<sup>[15]</sup>借助决策矩阵改进德尔菲方法,其遴选结果更加接近于真实情况。另外,虞锡君<sup>[16]</sup>在产业集群内关键共性技术选择研究中提出的“三链分析法”“三结合选择方法”和“市场化选择法”也属于定性分析法。

定量分析法多见于学术研究,主要是针对专利数据进行计量学研究、知识网络分析和指标评估。例如,黄鲁成等<sup>[7]</sup>在识别产业共性技术过程中,将技术领域的共性次数作为技术核心评估指标挑选出基础专利技术,若基础专利技术同时具有多领域应用和专利家族逐年增加

等特征，则判定为产业共性技术。栾春娟<sup>[17]</sup>利用德温特创新索引的专利技术特征和分类划分技术领域，计算技术贡献率、技术共现强度等指标，识别出战略性新兴产业共性技术领域；在基于专利数据建立的知识网络中，节点的中介中心度是识别关键技术的指标，度中心指标用于测度核心技术，多重测量中心度作为共性技术识别的指标<sup>[18]</sup>。杨武等<sup>[19]</sup>从文本挖掘、德温特手工代码共现、专利引用三个角度进行评估，逐步识别出锂电池领域共性技术。杨艳萍等<sup>[11]</sup>根据专利文献的共被引聚类形成专利簇，以专利申请数量和被引频次作为专利簇的评价指标，有效地识别出产业层面的关键技术。

综上所述，已有研究表明基于专利分析和知识网络特征是确定产业关键技术、产业共性技术领域的行之有效的方法。本文在前人研究基础上，从产业关键共性技术的本质特性出发，从技术领域和技术主题两个层面，探索知识网络特征与产业关键共性技术之间的关系，建立基于 PCT 专利分析和知识网络特征的产业关键共性技术识别模型，并以新一代信息技术产业为例展开实证研究。

## 2 研究设计

### 2.1 产业关键共性技术识别模型构建

本文构建的产业关键共性技术识别模型主要包括数据获取与预处理、产业共性技术领域识别、技术主题特征指标计算三个阶段：

首先，根据产业链纵向上下游关系或横向并行产业进行单元划分，建立产业技术分解表，依据产业技术分解表检索并下载相关专利数据，形成产业专利技术数据集，绘制 S 曲线，利用专利申请数量随时间推移而变化来缩小专利数据集；其次，基于缩小的数据集构建技术领域共现网络，利用社团发现算法挖掘网络中具有相似性质的节点集合，识别产业共性技术领域；最后，在产业共性技术领域范围内，构建技术主题共现网络，依据网络节点拓扑结构特征，如度及度分布、点权和单位权、介数及其分布特征、节点间距离等，计算技术主题的技术使用频次、技术关联广度、技术关联强度、技术领域中心性等关键技术识别指标，遴选出关键技术主题，见图 1。

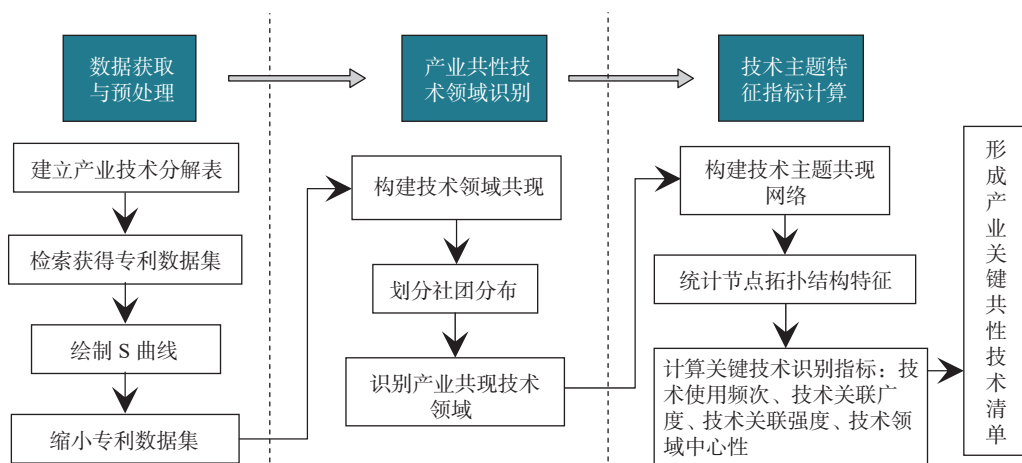


图 1 产业关键共性技术识别模型



## 2.2 数据获取与预处理

本文的全部数据来自德温特世界专利索引数据库 (Derwent World Patents Index, 简称 DWPI)。由字母加数字的形式组成的德温特手工代码 (Derwent Manual Code, DMC) 是 DWPI 特色之一, 便于识别每项发明技术的新颖性及其技术应用领域。本文根据 DMC 为每条数据分别标注技术领域与技术主题, 具体表示为: 以 ANN 为技术领域、ANN-ANN 为技术主题, 其中 A 表示字母, N 表示数字。

根据技术生命周期理论, 一项技术通常要经历萌芽期、成长期、成熟期和衰退期四个阶段<sup>[20]</sup>, 在时间序列下排序专利申请数量形成的曲线呈“S”型, 曲线的不同阶段分别与技术发展阶段相对应<sup>[21]</sup>, 见图 2。本文将以专利累计申请数量作为指标, 绘制 S 曲线, 以曲线的曲率变化为依据缩小检索得到的专利数据集。

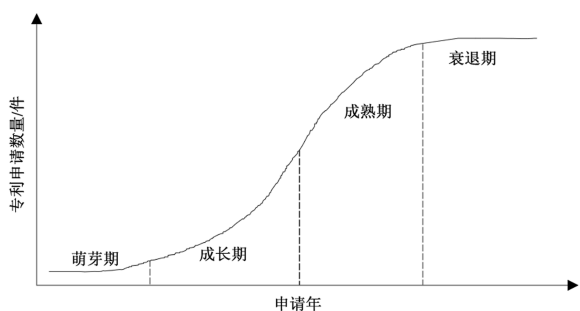


图 2 基于技术生命周期理论的 S 曲线

S 曲线主要分为对称的 Logistic 模型曲线和不对称的 Gompertz 模型曲线, 前者在实践中应用更加广泛。Logistic 模型曲线由 Verhulst<sup>[22]</sup> 提出, 其公式如下:

$$y = \frac{l}{1 + \alpha e^{-\beta T}} \quad (1)$$

其中,  $y$  代表专利累积申请数量,  $\alpha$  代表 S

曲线斜率, 即 S 曲线的成长率,  $\beta$  代表成长曲线中转折点 (Midpoint),  $l$  为成长的饱和水平, 即饱和点 (Saturation), 用于评估专利累积数量的最高值,  $T$  为时间。

## 2.3 产业共性技术领域识别

如果一条专利信息涉及两个或两个以上的技术领域, 即多项技术在该条专利中同时出现, 则视为技术共现。技术共现是关键共性技术识别的基础, 技术共现网络反映了技术/技术领域之间的关联程度。共现网络模型为  $G=(V, E)$ , 是由  $|V|=N$  个节点和  $|E|=M$  条边组成的一个无向加权网络。其中,  $G$  表示节点及其相互共现关系的集合,  $V$  代表节点的集合,  $V = \{v_1, v_2, v_3, \dots\}, v_i \in V$ ,  $E$  代表节点间连线的集合,  $E = \{e_1, e_2, e_3, \dots\}, e_i \in E$ 。网络的邻接矩阵为  $A$ ,  $A=(a_{ij})_{N \times M}$ , 是一个对称矩阵。若  $v_i$  与  $v_j$  直接相连, 代表  $v_i$  与  $v_j$  之间存在共现关系, 则  $a_{ij}=w(i, j)$ , 否则  $a_{ij}=0$ , 其中,  $w(i, j)$  表示  $v_i$  与  $v_j$  相连边的权值, 即共现频数。在技术领域共现网络中, 节点代表专利技术领域, 节点间的连线代表相邻技术领域共同出现, 连线的粗细或长度用于表示相应技术领域之间共现关系的强弱。

知识网络中的节点可以划分入若干社团, 同一社团内部节点间连接稠密, 而各个社团之间的连接相对比较稀疏。知识网络的社团发现实质是对节点的聚类, 基于模块度优化的社团发现算法是目前研究最多的一类算法, 有 GN 算法、CNM 算法和 EO 算法等。模块度 (也称  $Q$  函数) 是指网络中连接社团结构内部节点的边所占的比例与随机网络中连接社团结构内部节点的边所占比例的期望值相减得到的差值<sup>[23]</sup>。

其计算公式为：

$$Q = \frac{1}{2m} \sum_i \sum_j \left[ w_{ij} - \frac{S_i S_j}{2m} \right] \delta(g_i, g_j) \quad (2)$$

其中， $S_i$  表示  $v_i$  的点强度， $m = \frac{1}{2} \sum_i \sum_j w_{ij}$  表示所有边的权重之和， $g_i$  表示  $v_i$  的社区编号，如果  $v_i$  与  $v_j$  属于同一社团，则  $\delta(g_i, g_j)=1$ ，否则  $\delta(g_i, g_j)=0$ 。通常用  $Q$  函数定量地描述社团划分的模块性水平。在技术领域共现网络中，一部分技术领域之间连接稠密，而与其他技术领域之间的连接相对比较稀疏，这些相互连接紧密的技术领域组合起来形成一个整体，视作产业共性技术领域。

#### 2.4 技术主题特征指标计算

依据共现网络模型构建技术主题共现网络，在该网络中，频繁出现的技术主题对不同行业或领域的重要影响力是显而易见的；若某项技术与其他很多的技术共同出现，表明该技术可能是其他多项技术突破的基础；共同出现的伙伴多且相互关联强的技术主题对其他技术的发展具有广泛的渗透性和影响力，该技术主题属于产业技术创新链的重要技术；当一个技术主题处于其他多个技术主题连接路径上，则该技术主题具有控制其他技术“联通”的能力，其在产业技术发展中起着桥梁和纽带的作用。

本文根据产业关键共性技术重点特征，从技术使用频次、技术相关性、技术关联度、技术核心性四个维度，提出了基于技术共现网络的技术主题特征衡量指标：统计专利技术出现的次数，获取技术使用频次，以此评估技术主题“重要性”；分析技术主题的共现关系，以技术关联广度来衡量技术主题“通用性”；技

术共现矩阵与关联强弱分析结合，计算技术关联强度，用来评判技术主题“基础性”；以社会网络中节点中心性判断技术主题地位，量化技术主题中心性，分析技术主题“瓶颈性”。产业关键共性技术是在两个或两个以上共性技术领域特征指标值都比较大的一类技术主题。

(1) 技术使用频次 (Tech-Occurrence)：技术主题出现的频次，用于反映该项技术的重要性。某一项技术主题出现次数多，表明其在不同行业或领域都备受重视。技术使用频次数值越大，代表该项技术越重要。

$$Tech - Occurrence = occ(i) \quad (3)$$

其中， $occ(i)$  表示  $v_i$  出现的频次。

(2) 技术关联广度 (Tech-Correlation)：与技术主题同时出现的其他技术主题的总数(不包括重复共现次数)，即技术共现网络中与  $v_i$  相连的其他节点的个数，用于判断该项技术的通用程度。某一项技术与其他技术节点之间直接联系数量多，表明该技术主题对多个行业或者领域产生了深度影响。技术关联广度数值越大，则该项技术应用越普遍，通用性越强。

$$Tech - Correlation = \sum_{v_i, v_j \in V} e_{ij} \quad (4)$$

其中， $V$  代表技术主题共现网络， $v_i$ 、 $v_j$  代表网络  $V$  中的节点， $e_{ij}$  表示  $v_i$  与  $v_j$  之间的连线。

(3) 技术关联强度 (Tech-Intensity)：技术主题与所有其他同时出现的技术主题相联系的程度(包括重复共现次数)，即技术共现网络中  $v_i$  与其他节点的连线粗细程度，用于衡量该技术主题对其他技术主题的“渗透”能力。某一项技术主题与其他技术主题产生联系程度深，表明该技术主题对其他技术主题的作用力

大,体现其基础性地位。技术关联强度数值越大,则该项技术基础性地位越牢固。

$$Tech - Intensity = \frac{S_i}{k_i} \quad (5)$$

其中,  $S_i$  表示  $v_i$  的点强度, 即与  $v_i$  关联的边权值之和,  $k_i$  表示  $v_i$  的度, 即与  $v_i$  关联的边数量之和。

(4) 技术领域中心性 (Tech-Centrality): 知识网络里所有最短路径中通过  $v_i$  的数量比例, 即节点的介数, 用来刻画节点“桥梁”价值的指标。知识网络的“桥梁”节点具有集成、带动和辐射作用, 技术领域中心性数值越大, 技

术节点在网络中越有用。

$$Tech - Centrality = \sum_{i \neq s \neq t} \frac{n_{st}(i)}{g_{st}} \quad (6)$$

其中,  $n_{st}(i)$  代表最短路径经过  $v_i$  的路径数量,  $g_{st}$  代表连接  $v_s$  和  $v_t$  的最短路径的数量。

这四个指标的量纲不同, 见表 1, 为了消除由此带来的不可公度性, 需进行无量纲化处理, 即基于线性变换的归一化处理, 使结果值映射到 [0-1] 之间。

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (7)$$

其中,  $x_{\max}$  为样本数据的最大值,  $x_{\min}$  为样本数据的最小值,  $\hat{x}$  为样本数据的归一化数值。

表 1 技术主题特征指标

指标名称	计算方法	取值范围	指标意义
技术使用频次	occ(i)	[1, +∞]	反映技术的重要性, 数值越大, 代表该项技术越重要。
技术关联广度	$\sum_{v_i, v_j \in V} e_{ij}$	[0, +∞]	判断该项技术的通用程度, 数值越大, 该技术应用更普遍、通用性更强。
技术关联强度	$\frac{\sum_{v_i, v_j \in V} w_{ij}}{\sum_{v_i, v_j \in V} e_{ij}}$	[1, +∞]	衡量技术的渗透性, 数值越大, 该技术基础性地位越牢固。
技术领域中心性	$\sum_{i \neq s \neq t} \frac{n_{st}(i)}{g_{st}}$	[0, 1]	衡量技术的“桥梁”价值, 数值越大, 该项技术的集成、辐射和带动作用越强。

### 3 实证研究

#### 3.1 新一代信息技术产业的数据获取及预处理

新一代信息技术产业是国民经济的战略性、基础性和先导性产业。世界各国高度重视新一代信息技术产业的发展, 纷纷颁布各类政策框架、发展战略和行动计划, 展开相关部署, 积极培育新一代信息技术产业竞争优势。本文研究以新一代信息技术产业为例, 识别产业共现

技术领域, 计算其中各个技术主题的技术使用频次、技术相关性、技术关联度、技术核心性等特征指标, 从而遴选出新一代信息技术产业关键共性技术, 以期为关键共性技术的准确识别与测度提供参考。

新一代信息技术产业专利检索范围参考 2017 年经济合作与发展组织 (Organization for Economic Co-operation and Development, OECD) 有关 ICT 技术领域专利的划分标准<sup>[24]</sup>, 建立了检索式, 检索日期是 2021 年 7 月 7 日, 检索得

到 PCT 专利 1439056 件，经 DWPI 同族合并，最终获得全球新一代信息技术产业相关 PCT 专利 1244576 件。

依据技术生命周期理论和 Logistic 模型<sup>[25]</sup>，本文研究利用 Loglet Lab 软件模拟新一代信息技术产业专利技术申请趋势，见图 3，大致划分为三个阶段：1987 年以前，专利申请增长率相对较低，曲线平坦；1988—2014 年曲线下凸，相关技术进步速度不断提高；曲线在 2015 年出现拐点，开始变为上凸，技术进步速度开始变小。由于各国专利从申请到公开一般会有 2-3 年的延迟（一般为 18 个月），因此 2019—2021 年数据不全，趋势预测模拟过程不包含 2019—2021 年数据。

从图 3 中可以看出，新一代信息技术产业的 PCT 专利申请数量整体呈波动递增趋势。随着世界经济的快速发展、现代科学技术的进步以及知识产权保护意识的增强，新一代信息技术产业专利申请数量在未来几年将会出现大幅度增长。

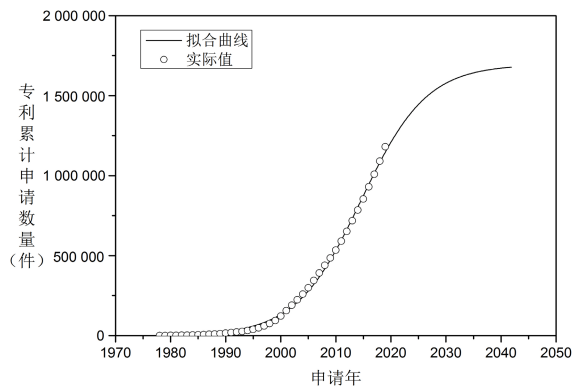


图 3 全球新一代信息技术产业专利申请趋势

### 3.2 新一代信息技术产业共性技术领域识别

基于 S 曲线的拟合结果，选取 2015 以后的专利数据，共计 460207 件专利。根据这些专利数据的技术领域标注，绘制技术领域共现网络，见图 4。

图 4 包含 277 个节点，每个节点代表一种技术领域；节点大小与其出现频次正相关，节点越大表示该技术领域出现次数越多；节点之间连线的宽度代表与其相连的两个技术领域共现强度，连线越宽表示两端的技术领域共现强度越大；同一所属聚类的节点标记相同颜色。

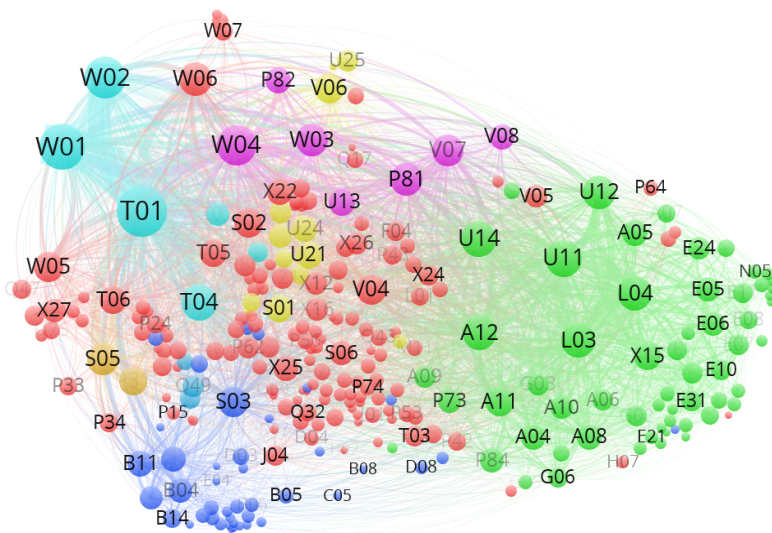


图 4 新一代信息技术产业的技术领域共现网络图



进一步在上述知识网络中进行社区发现算法计算，将划分在同一个社区的技术领域

聚类作为本文识别出的共性技术领域，结果见表 2。

表 2 新一代信息技术产业共性技术领域识别

社区	技术领域	产业共性技术领域
I 社区	W06、W05、V04、S02、T06、T05、X22、X25、S06、X27、X12、X16、V05、P85、X21 等	航空、航海遥控信号相关领域
II 社区	U11、U14、A12、L04、L03、U12、A11、X15、A05、A08、A04、E05、P84、P73、E06 等	半导体存储器、薄膜和混合电路、分立元器件领域
III 社区	S03、B11、B12、B04、D05、B14、P13、D03、P11、C11、B05、C12、C04、B06、B10 等	医药相关科学仪器领域
IV 社区	V06、S01、U21、U24、U22、U23、U25、S04、V03、X13、V01、Q72、J09	电子仪器及其零部件领域
V 社区	W04、W03、P81、V07、U13、P82、V08	音频、视频领域
VI 社区	T01、W01、W02、T04、T07、X23	计算机技术与电话、无线电传输应用领域
VII 社区	Q49、H01、T02	原油、天然气和矿业相关领域
VIII 社区	S05、P31	电子医疗器械领域

由表 2 可知，新一代信息技术产业可分为八个共性技术领域，分别是：（I 社区）航空、航海遥控信号相关领域，（II 社区）半导体存储器、薄膜和混合电路、分立元器件领域，（III 社区）医药相关科学仪器领域，（IV 社区）电子仪器及其零部件领域，（V 社区）音频、视频领域，（VI 社区）计算机技术与电话、无线电传输应用领域，（VII 社区）原油、天然气和矿业相关领域，（VIII 社区）电子医疗器械领域。

### 3.3 新一代信息技术产业的技术主题特征指标计算

针对已识别出的每个共性技术领域，根据上文技术主题指标计算公式，得到技术主题的特征指标得分，选取四个特征指标得分都大于 0.3 的 20 个技术主题，见表 3。

由表 3 可知，注册软件产品（T01-S03）、网络信息传输的应用程序（T01-N01）、网络信息的通讯与监控（T01-N02）、图像处理的数据

处理系统（T01-J10）、电话用户设备（W01-C01）等技术主题在各个共性技术领域的特征指标得分普遍较高，其次是摄像机（W04-M01），行政、商业或信息检索的数据处理系统（T01-J05），工业过程控制的数据处理系统（T01-J07），网络数据信息传输（W01-A06），液晶显示器（U14-K01），程序管理界面、交互等的数据处理系统（T01-J12），用函数或方程求解的数据处理系统（T01-J04）等技术主题，远距离的光学元件、系统、设备（P81-A50），电视接收机和视频显示器（W03-A08），游戏（W04-X02），医疗设备和信息系统的数据处理系统（T01-J06），机电传感器的应用（V06-V04），执行特定程序和系统管理软件的程序控制（T01-F05），无线电系统（W02-C03），与内存相关的数据存储、互联、传输（T01-H01）等技术主题在多个共性技术领域的特征指标得分较高。基于这些技术主题形成新一代信息技术产业的关键共性技术清单，如表 4 所示。

表3 新一代信息技术产业的技术主题特征指标

技术主题	特征指标得分*					所属共性技术领域**
	共性技术领域	技术使用 频次	技术关联 广度	技术关联 强度	技术领域 中心性	
P81-A50	V	0.56	0.98	0.63	0.91	I、II、V、VII
T01-F05	I	0.43	0.71	0.47	0.37	I、II、VI、VII
T01-H01	VII	1.00	0.77	1.00	0.37	III、VII
T01-J04	I	0.39	0.82	0.32	0.60	I、III、IV、VI、VIII、VII
T01-J05	VI	0.65	0.96	0.62	0.84	I、II、III、V、VI、VII、VIII
T01-J06	VIII	0.94	1.00	1.00	1.00	IV、VIII
T01-J07	I	0.45	0.98	0.40	0.97	I、II、III、IV、V、VI、VII
T01-J10	V	1.00	0.85	0.93	0.57	I、II、III、IV、V、VI、VII、VIII
T01-J12	I	0.33	0.77	0.38	0.46	I、II、III、IV、VI、VIII
T01-N01	VI	0.77	0.98	0.83	0.91	I、II、III、IV、V、VI、VII、VIII
T01-N02	VI	0.69	0.98	0.79	0.91	I、II、III、IV、V、VI、VII、VIII
T01-S03	VI	0.97	0.99	1.00	0.94	I、II、III、IV、V、VI、VII、VIII
U14-K01	I	0.44	0.96	0.34	0.84	I、II、III、IV、V、VII
V06-V04	IV	1.00	0.95	0.68	0.81	IV、VIII
W01-A06	VI	1.00	0.93	0.88	0.70	I、II、III、IV、VI、VII
W01-C01	V	0.61	1.00	0.63	1.00	I、II、III、IV、V、VI、VII、VIII
W02-C03	VI	0.46	0.85	0.46	0.55	I、II、III、VI
W03-A08	V	0.40	0.96	0.34	0.91	I、II、III、V
W04-M01	V	0.95	0.99	0.72	0.96	I、II、III、V、VI、VII、VIII
W04-X02	V	0.38	0.92	0.42	0.78	I、II、V、VI

注：\* 针对每个共性领域，计算技术主题的技术特征指标得分，本表格仅展示了1个计算结果。

\*\* 针对每个共性领域，设置技术主题的出现阈值，当某一技术主题出现次数大于出现阈值时，才认为该技术主题属于该共性技术领域。

产业规划和政策文件中有时会列举产业关键技术、产业共性技术的领域范围，比如《“十四五”软件和信息技术服务业发展规划》中列出软件产业的共性技术有软件与系统工程方法、程序设计语言、关键核心算法、数据模型，《物联网新型基础设施建设三年行动计划（2021—2023年）》提到物联网产业的关键共性技术有智能感知、新型短距离通信、高精度定位等，《“十四五”数字经济发展规划》指出关键技术短板包括高端芯片、操作系统、工业软件、核心算法与框架等。注册软件

产品（T01-S03），图像处理的数据处理系统（T01-J10），工业过程控制的数据处理系统（T01-J07），程序管理界面、交互等的数据处理系统（T01-J12），用函数或方程求解的数据处理系统（T01-J04），与内存相关的数据存储、互联、传输（T01-H01）等技术主题都属于这些技术领域范畴。元宇宙被认为是下一代移动互联网雏形，支撑元宇宙的关键技术主要包括扩展现实（XR）、先进无线通信技术（5G/6G）和区块链。美国制定了一系列政策促进元宇宙技术发展，《21世纪就业法案》提出支持区块

表4 新一代信息技术产业关键共性技术清单

技术领域	技术主题
光学元件	远距离的光学元件、系统、设备
数字计算机	执行特定程序和系统管理软件的程序控制
数字计算机	与内存相关的数据存储、互联、传输
数字计算机	用函数或方程求解的数据处理系统
数字计算机	行政、商业或信息检索的数据处理系统
数字计算机	医疗设备和信息系统的数据处理系统
数字计算机	工业过程控制的数据处理系统
数字计算机	图像处理的数据处理系统
数字计算机	程序管理界面、交互等的数据处理系统
数字计算机	网络信息传输的应用程序
数字计算机	网络信息的通讯与监控
数字计算机	注册软件产品
存储器、薄膜和混合电路	液晶显示器
机电传感器和小型机械	机电传感器的应用
电话和数据传输系统	网络数据信息传输
电话和数据传输系统	电话用户设备
广播、无线电和线路传输系统	无线电系统
电视和广播无线电接收器	电视接收机和视频显示器
音频/视频记录与系统	摄像机
音频/视频记录与系统	游戏

链、通信、VR、AR等特定技术领域的研究，《美国创新和竞争法案2021》提出将先进通信技术与沉浸式科技作为关键技术领域，并为其研发、教育和培训、转移和商业化提供资助。我国《虚拟现实与行业应用融合发展行动计划》提出围绕近眼显示、渲染处理、感知交互、网络传输、内容生成、压缩编码和安全可信等核心关键技术取得重要突破。网络信息传输的应用程序

(T01-N01),网络信息的通讯与监控(T01-N02),图像处理的数据处理系统(T01-J10),摄像机(W04-M01),网络数据信息传输(W01-A06),液晶显示器(U14-K01),游戏(W04-X02),与内存相关的数据存储、互联、传输(T01-H01),远距离的光学元件、系统、设备(P81-A50),机电传感器的应用(V06-V04)等技术主题属于这些关键共性技术范畴。此外,美国国家科学技术委员会最新发布的关键和新兴技术清单包含频谱管理技术、光纤链路和光纤技术、陆地/海底电缆、硬件/固件和软件、卫星通信等通讯和网络技术。可见,本文提出的新一代信息技术产业的关键共性技术主题会出现在随后出台的相关产业政策中,从政策层面来看,本文提出的产业关键共性技术识别模型具有一定预测性。但由于专利与政策中关于技术描述词的粒度不同,还需要行业专家做进一步的解读和验证。

### 3.4 我国新一代信息技术产业关键共性技术识别研究

党中央、国务院高度重视新一代信息技术产业发展。2015年以来,我国申请新一代信息技术产业关键共性技术主题相关专利103191件,占全球总量的30.9%,具体情况参见表5。从全球新一代信息技术产业关键共性技术专利申请来看,我国在液晶显示器(U14-K01)主题的专利申请占比最高,其次是网络数据信息传输(W01-A06)、电话用户设备(W01-C01)、无线电系统(W02-C03)等技术主题;而关于医疗设备和信息系统的数据处理系统(T01-J06),游戏(W04-X02),远距离的光学元件、系统、设备(P81-A50)等主题的专利

申请占比偏低。

结合产业发展现状来看，我国在新型显示、信息化建设等方面技术研发优势明显。新型显示产业作为我国重点发展的战略性新兴产业之一，已发展成为新一代信息技术的先导性支柱产业。随着《扩大和升级信息消费三年行动计划(2018—2020年)》《中国光电子器件产业技术发展路线图(2018—2022年)》《超高清视频产业发展行动》《关于2021—2030年支持新型显示产业发展进口税收政策的通知》等政策的陆续出台，涌现出了京东方、深天马、华星光电等一批具有

竞争力的液晶显示面板生产企业。在信息基础设施建设方面，我国建成全球规模最大、技术领先的网络基础设施，光纤网络接入带宽实现从十兆到百兆再到千兆的指数级增长，移动网络实现从“3G突破”到“4G同步”再到“5G引领”的跨越。截至2022年底，我国数据中心机架总规模已超过650万标准机架，在用数据中心算力总规模超180EFLOPS，位居世界第二。此外，卫星遥感、卫星通信广播、北斗导航定位三大系统构成的国家民用空间基础设施体系已初步建成，具备连续稳定的业务服务能力。

表5 我国新一代信息技术产业关键共性技术专利申请统计表

技术主题	全球专利申请数量 (件)	我国专利申请数量 (件)	我国专利占比 (%)
P81-A50 远距离的光学元件、系统、设备	10 522	1 104	10.5
T01-F05 执行特定程序和系统管理软件的程序控制	17 885	4 483	25.1
T01-H01 与内存相关的数据存储、互联、传输	6 688	1 555	23.3
T01-J04 用函数或方程求解的数据处理系统	15 984	4 456	27.9
T01-J05 行政、商业或信息检索的数据处理系统	66 085	15 513	23.5
T01-J06 医疗设备和信息系统的的海理系统	7 983	624	7.8
T01-J07 工业过程控制的数据处理系统	29 434	5 634	19.1
T01-J10 图像处理的数据处理系统	51 173	14 848	29.0
T01-J12 程序管理界面、交互等的海理系统	29 494	6 305	21.4
T01-N01 网络信息传输的应用程序	78 844	23 020	29.2
T01-N02 网络信息的通讯与监控	69 909	18 040	25.8
T01-S03 注册软件产品	98 986	27 538	27.8
U14-K01 液晶显示器	18 055	8 667	48.0
V06-V04 机电传感器的应用	7 225	1 925	26.6
W01-A06 网络数据信息传输	101 791	37 456	36.8
W01-C01 电话用户设备	54 194	18 923	34.9
W02-C03 无线电系统	46 821	14 494	31.0
W03-A08 电视接收机和视频显示器	7 495	2 200	29.4
W04-M01 摄像机	17 833	3 933	22.1
W04-X02 游戏	7 190	716	10.0



我国新一代信息技术产业发展成效显著,5G、云计算、人工智能等新一代信息技术持续向各行各业融合渗透,但目前与实体经济融合发展还仅仅停留在典型、示范、部分制造环节或者部分制造领域,存在数据标准不统一、互操作性差、数据安全风险高等难题,缺乏覆盖全流程的数字化解决方案。以数字医疗产业为例,近年来国家密集发布一系列医疗健康政策,促进医院信息化、远程医疗等智慧医疗发展,但相对于发达国家来说,我国的医疗信息化程度还处于较低的水平,国内从事智慧医疗的企业多数处于商业模式探索阶段。从技术角度,人工智能在医疗健康行业的应用目前总体上仍处于起步阶段,场景应用深度不够,相关技术与产品的研发和创新能力还有待进一步提升;从数据角度,数字化医疗技术涉及大量的个人健康信息,建立安全可靠的数据存储和传输系统是亟待解决的重要问题,还需要制定统一的技术标准和推动系统互联互通,用于提升数字化医疗技术的应用效果。再比如游戏产业,长期以来,社会各界更多关注电子游戏的娱乐属性,而忽视游戏的科技意义。实际上,游戏从诞生起就与前沿科技密不可分,当代游戏产业不仅仅是计算机技术发展的受益者,也是技术发展的孵化器和推进剂。中国科学院研究团队发布的《游戏技术——数实融合进程中的技术新种群》研究报告显示,我国的游戏技术对国内芯片产业、5G高速网络产业和AR/VR产业的科技进步贡献率分别约为14.9%、46.3%和71.6%。但与美日等游戏产业强国相比,我国的游戏技术仍处于并行并跑的发展阶段。随着经济的发展和科技的进步,游戏产业作为数字经济重要组成部分,凭借其蓬

勃的发展态势,有助于实现数字技术与实体经济深度融合,更好地引导新一代信息技术等相关产业的健康可持续发展。

## 4 结论

本文通过辨析产业共性技术、产业关键技术、产业关键共性技术的概念和特征,得出产业关键核心技术的本质特征是重要性、通用性、基础性、瓶颈性。构建了产业共性技术领域的识别方法和关键技术特征指标的测度方法,并选取新一代信息技术产业为例,基于全球PCT专利数据简单划分产业技术发展阶段,在技术领域共现网络中利用社区发现算法识别了产业共性技术领域,进一步根据技术主题共现网络的节点拓扑结构特征遴选出关键技术主题,从而形成最终的产业关键共性技术清单。

本文研究通过两个共现网络的结构特征揭示技术领域之间、技术主题之间潜在的关联关系,从专利文献内容和产业技术发展规律两个方面综合考量,确保了遴选方案的合理性和可靠性。另外,在构建网络之前先从全部专利数据集中筛选出当前阶段的专利数据,这样既能涵盖产业技术创新现状,又能避免与当前产业现状无关的数据参与计算。然而,本文研究成果列出的产业关键共性技术主题清单还需经行业专家解读和验证。在未来研究中,可考虑应用该产业关键共性技术识别体系,研究不同国家或地区之间技术竞争与布局的差异性等问题。

## 参考文献

- [1] 刘坤. 加强攻关掌握更多独门绝技 [N]. 光明日报, 2020-12-28(009).



- [2] 张劲帆, 李汉涯, 何晖. 企业上市与企业创新——基于中国企业专利申请的研究 [J]. 金融研究, 2017(5): 160-175.
- [3] 刘辉锋. 从 PCT 申请和三方专利指标评价中国海外专利申请实力 [J]. 科技和产业, 2017, 17(7): 146-149.
- [4] 李纪珍. 产业共性技术: 概念、分类与制度供给 [J]. 中国科技论坛, 2006(3): 45-47.
- [5] 马名杰. 共性技术的内涵与评判标准 [J]. 新经济导刊, 2004(22): 74-78.
- [6] 许端阳, 徐峰. 产业共性技术的界定及选择方法研究——基于科技计划管理的视角 [J]. 中国软科学, 2010 (4): 73-79.
- [7] 黄鲁成, 张静. 基于专利分析的产业共性技术识别方法研究 [J]. 科学学与科学技术管理, 2014, 35(4): 80-86.
- [8] 薛捷, 张振刚. 国外产业共性技术创新平台建设的经验分析及其对我国的启示 [J]. 科学学与科学技术管理, 2006(12): 87-92.
- [9] 周永春, 李思一. 国家关键技术选择: 新一轮技术优势争夺战 [M]. 北京: 科学技术文献出版社, 1995: 8-9.
- [10] 李建玲, 李纪珍. 产业共性技术与关键技术的比较研究——基于北京市科委资助科研项目的统计 [J]. 技术经济, 2009, 28(6): 11-17.
- [11] 杨艳萍, 董瑜, 韩涛. 基于专利共被引聚类和组合分析的产业关键技术识别方法研究——以作物育种技术为例 [J]. 图书情报工作, 2016, 60(19): 143-148.
- [12] 唐晓华, 景文治, 张英慧. 人工智能赋能下关键技术突破、产业链技术共生与经济“脱虚向实” [J]. 当代经济科学, 2021, 43(5): 44-58.
- [13] 赵彦飞, 王孝炯, 王丽. 国家关键技术选择: 三维综合指数方法研究 [J]. 科学学研究, 2021, 39(6): 1015-1025.
- [14] 袁思达. 技术预见德尔菲调查中共性技术课题识别研究 [J]. 科学学与科学技术管理, 2009, 30(10): 21-26.
- [15] 于晓勇, 尚赞娣, 李金林. 基于技术预见德尔菲调查的共性技术课题选择方法研究 [J]. 数学的实践与认识, 2011, 41(4): 64-68.
- [16] 虞锡君. 产业集群内关键共性技术的选择——以浙江为例 [J]. 科研管理, 2006(1): 80-84.
- [17] 栾春娟. 战略性新兴产业共性技术测度指标实证研究 [J]. 中国科技论坛, 2012(6): 73-77.
- [18] 栾春娟. 网络中心性指标在技术测度中的应用 [J]. 科技进步与对策, 2013, 30(3): 10-13.
- [19] 杨武, 杨大飞. 基于专利数据的产业核心技术识别研究——以 5G 移动通信产业为例 [J]. 情报杂志, 2019, 38(3): 39-45.
- [20] 王山, 谭宗颖. 技术生命周期判断方法研究综述 [J]. 现代情报, 2020, 40(11): 144-153.
- [21] 蔡爽, 黄鲁成. 面向技术战略的专利分析方法述评 [J]. 技术经济, 2008(6): 36-39.
- [22] PELEG M, CORRADINI M G, NORMAND M D. The logistic (Verhulst) model for sigmoid microbial growth curves revisited [J]. Food Research International, 2007, 40(7): 808-818.
- [23] 孙玺菁, 司守奎. 复杂网络算法与应用 [M]. 北京: 国防工业出版社, 2016: 223-225.
- [24] TAKASHI I, MARIAGRAZIA S. ICT: A new taxonomy based on the international patent classification [EB/OL]. (2017-02-18) [2022-12-01]. [https://www.oecd-ilibrary.org/science-and-technology/ict-a-new-taxonomy-based-on-the-international-patent-classification\\_ab16c396-en](https://www.oecd-ilibrary.org/science-and-technology/ict-a-new-taxonomy-based-on-the-international-patent-classification_ab16c396-en)
- [25] 赵莉晓. 基于专利分析的 RFID 技术预测和专利战略研究——从技术生命周期角度 [J]. 科学学与科学技术管理, 2012, 33(11): 24-30.



开放科学  
(资源服务)  
标识码  
(OSID)

# 中国个人信息保护应用与技术进展研究 ——基于科学知识图谱视角

刘华玲 梁华璧 王希睿

上海对外经贸大学 统计与信息学院 上海 201620

**摘要:** [目的/意义] 大数据时代, 公众对个人信息保护的需求日渐多维化, 基于科学知识图谱视角分析近十年中国个人信息保护演化趋势与前沿热点有其必要性。[方法/过程] 基于文献计量、突现检测、科学知识图谱等多种方式展开分析, 以 CSSCI 和 CSCD 刊源分别表征应用与技术维度, 梳理近十年国内个人信息保护研究的双维度发展趋势。[结果/结论] 我国个人信息保护研究文献数量呈现阶段性增长模式; 个人、机构的合作方式以内部学术交流、地域性联系为主, 多学科交叉格局有待形成; 在应用维度上以图情类主导、公共化、数字化为研究趋势; 在技术维度上以计算机类主导、精细化、智能化为研究趋势。全民监督与数据共治成为应对大数据时代个人信息保护新挑战的可行之策。

**关键词:** 个人信息保护; 科学知识图谱; 文献计量; 应用维度; 技术维度; 热点演化

**中图分类号:** G251; G35

## Research on the Progress of Personal Information Protection Application and Technology in China——Based on the Perspective of Mapping Knowledge Domains

LIU Hualing LIANG Huabi WANG Xirui

Department of Statistics and Information, Shanghai University of International Business and Economics, Shanghai 201620, China

**Abstract:** [Purpose/Significance] In the era of big data, the public's demand for personal information protection is becoming increasingly multidimensional. It is necessary to analyze the evolution trend and cutting-edge hotspots of personal information protection in China in the past decade from the perspective of mapping knowledge domains. [Methods/Processes] Bibliometrics method, emergence detection and mapping knowledge domains are used to carry out analysis, while CSSCI and CSCD sources represent the application and technology dimensions respectively. [Results/Conclusions] The growth mode of the number of personal information protection research literature in China is phased stable and generally close to linear growth; The cooperation modes of individuals and institutions are mainly internal academic exchanges and regional connections, therefore

**作者简介** 刘华玲 (1964-), 博士, 教授, 主要研究方向为数据挖掘与欺诈识别; 梁华璧 (1997-), 硕士研究生, 主要研究方向为文献计量与风险管理, E-mail: ml7801113522@163.com; 王希睿 (1999-), 硕士研究生, 主要研究方向为信息管理与知识图谱。

**引用格式** 刘华玲, 梁华璧, 王希睿. 中国个人信息保护应用与技术进展研究——基于科学知识图谱视角 [J]. 情报工程, 2024, 10(1): 42-58.

the interdisciplinary pattern needs to be formed; In the application dimension, it is dominated by library and information science, heading towards popularization and digitization; In the technical dimension, it is dominated by computer science, heading towards refinement and intelligence. National supervision and data co-governance have become feasible strategies to address the new challenges of personal information protection in the era of big data.

**Keywords:** Personal Information Protection; Mapping Knowledge Domains; Bibliometrics; Application Dimension; Technical Dimension; Hotspot Evolution

## 引言

随着“互联网+”与大数据产业的蓬勃发展及其战略地位的上升，个人在充分享受着网络信息高度互联互通带来的服务便利的同时，也在无意识中将个人信息分享给了从事数据收集、分析与推断的调查者或互联网企业。个人作为以几何量级爆炸性增长的信息的贡献者，却在逐渐丧失对信息的掌控权，不仅难以成为信息多次集成再造新价值的受益者，还面临着个人隐私信息被大范围泄露和非法利用的风险。因此，在大数据时代建立完善的个人信息保护机制和体系已成为各界的共识。本文将采用文献计量方法与科学知识图谱可视化技术，对近十年我国相关个人信息保护文献进行系统梳理与总结，并深入分析与挖掘我国个人信息保护研究的前沿热点和动态发展趋势，为我国个人信息保护领域的政策实践与技术研发提供数据参考借鉴。

本文的贡献一方面在于当前从文献计量角度引入定量分析方法专门对个人信息保护研究领域发展概况进行科学阐述的文章比较少<sup>[1-3]</sup>，另一方面则在于文献计量方面的研究大多是基于单一期刊来源或综合多个期刊来源开展关键词检索和计量分析<sup>[4-5]</sup>，而本文在参考谭春辉和

熊梦媛<sup>[6]</sup>在分析数据挖掘领域时将主题内容维度划分为理论类与应用类这一做法的基础上，考虑到CSSCI主要收录管理学、经济学等领域的人文社会科学学术期刊的应用性特点以及CSCD主要收录数学、工程技术等领域的自然科学学术期刊的技术性特点，分别基于CSSCI与CSCD刊源从应用和技术双维度对我国个人信息保护研究的近十年热点演变开展可视化分析，探究两个维度在发文体量、更迭模式、聚合方向等层面的差异与相同之处，以期促进应用与技术维度共建，进而推动个人信息保护研究的双向深入发展。

## 1 数据来源与研究方法

### 1.1 数据来源

在数字主导的信息膨胀时代，隐私与个人信息很难得到绝对的二元分割，看上去不存在私密性的个人信息在大数据技术的深度挖掘下极有可能泄露个人隐私，对个人信息的保护最核心的指向是防范外界对隐私级别信息的侵犯，而维护潜在的隐私利益又能成为个人信息保护体系的建立基础<sup>[7]</sup>。图1展示了个人信息保护与隐私保护的关系演变历程，可见个人信息保护的相关研究离不开隐私保护研究的成果支撑。

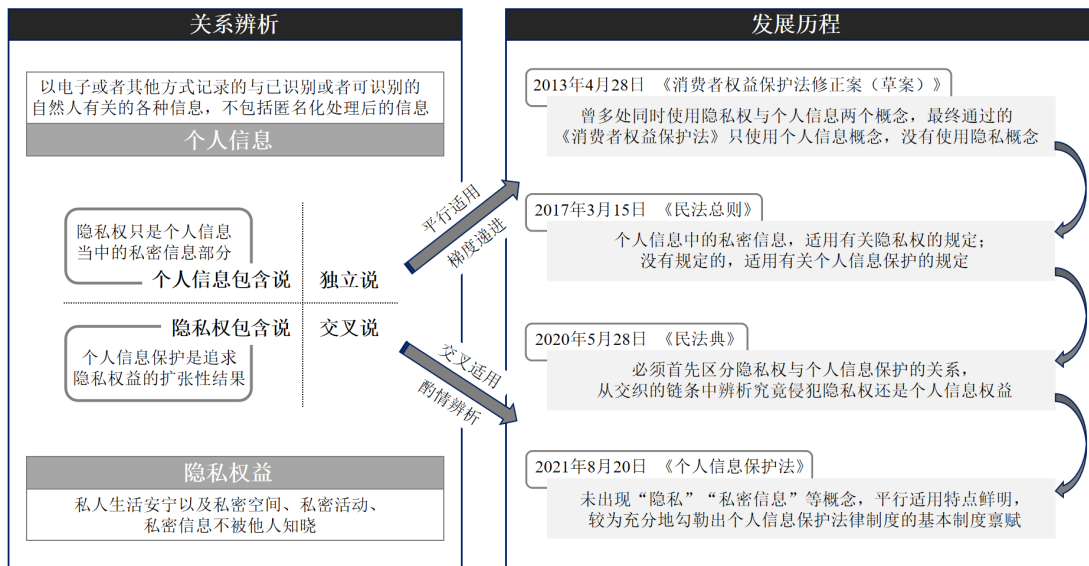


图1 个人信息保护与隐私保护关系演变

为保证原始数据能够全面准确地反映我国当前的研究现状，并兼顾数据分析视角多样性与结果可解释性，本文拟以具有全面的学科覆盖面和完整的学科文献库的中国知网（CNKI）作为基础数据源，以“个人信息保护+隐私保护”为主题、以“2011—2021年”为时间节点、增加以CSSCI、CSCD来源期刊为检索来源的检索条件进行检索，经过筛选后共得到3826篇有效文献。

## 1.2 研究方法

首先，为了解国内个人信息保护研究的时间和空间结构分布特征，本文采用文献计量法分别按不同来源数据库、研究机构、作者统计发文量变化状况，并运用ITGInsight工具绘制研究机构与作者共现网络图来直观刻画其地域分布及合作关系<sup>[8]</sup>。其次，为展示个人信息保护领域基于应用与技术双维度的演进态势和热点分布，利用词频探测技术获取高频关键词列

表，并借助科学知识图谱软件CiteSpace分别绘制来源期刊分布图和细化关键词时间线分布图谱<sup>[9]</sup>，实现研究焦点与未来趋势的可视化分析。

## 2 国内个人信息保护研究文献结构特征分析

### 2.1 时间分布特征

文献数量能够直观反映一个研究领域的受关注程度，本文数据集中包含CSSCI库论文1790篇，CSCD库论文2036篇，总计3826篇文献，分别按CSSCI库发文量、CSCD发文量和总发文量绘制年际变化趋势图，如图2所示。从总发文量来看，我国个人信息保护研究呈现出阶段性平稳、总体接近线性增长的特征，根据前后波动性的差异，可大体划分为启蒙探索阶段（2011—2012年）、一次增长阶段（2013—2016年）和二次增长阶段（2017—2021年）。

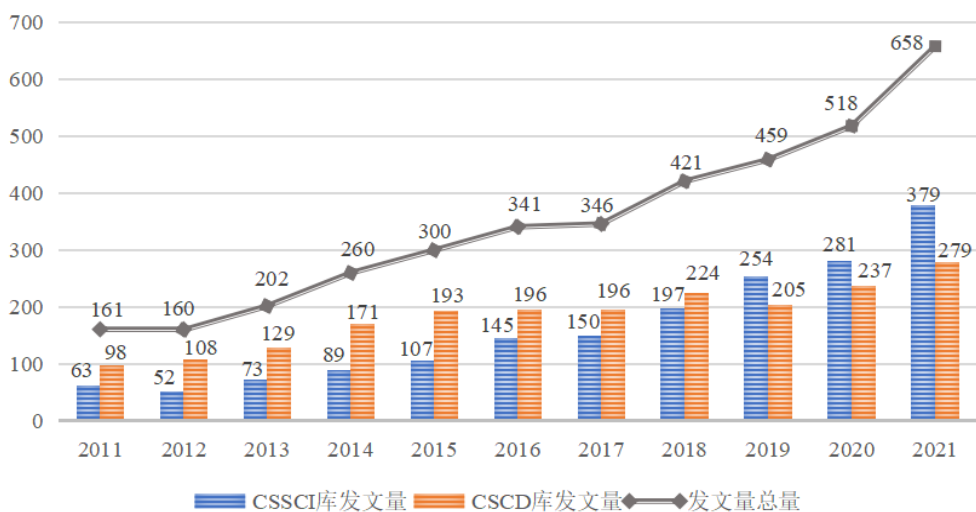


图2 国内个人信息保护研究CSSCI&amp;CSCD期刊发文量年际变化趋势

启蒙探索阶段的年均发文量约为160篇。在一次增长阶段，个人信息保护领域的研究迅速增多，共计发表论文1103篇，年均发文量约为275篇，该阶段发文量增幅为68.81%，一系列法律法规的颁布表明了国家和政府对个人信息保护的重视，进而吸引了各界学者进入该领域从事深入研究。进入二次增长阶段后，共发表学术论文2402篇，年均发文量约为480篇，发文量增幅高达90.17%，2017年个人隐私信息泄露事件频发、个人信息买卖形成非法产业链、企业间不正当数据竞争加剧等严峻现象，都推动着个人信息隐私保护研究的蓬勃发展。

从来源库视角看，CSCD库发文量呈现逐年稳定增长趋势，即使在研究初期也近乎保持着100篇以上的年发文量，可见个人信息保护研究多从技术层面入手，且发文量随时间均衡分布。CSSCI库发文量在近十年内存在指数增长态势，随着个人信息保护技术的更新与完善，学界开始转向研究如何将技术落实为经济社会领域的应用实践。2018年以前，CSCD库年发文量始终高于CSSCI库年

发文量，随后三年CSSCI库发文量实现了逐年超越，成为二次增长阶段个人信息保护文献激增的主要推动力。

## 2.2 空间分布特征

### 2.2.1 研究机构

本文将样本文献进行规范化处理后形成的数据导入ITGInsight，保留每个时间切片内排名前30的研究机构，绘制如图3所示的国内个人信息保护相关研究机构的共现图谱，图谱中每个节点表示一个研究机构，节点大小与该研究机构发文量呈正相关。节点间相连则表明两个机构间存在合作关系，连线的粗细反映了研究机构间合作关系的强弱。

从图3中可以看出，孤立节点大多为各大学的法学院，对于理论辨析、历史沿革等研究往往依赖于过往文献，多依靠机构内部力量推进；计算机科学、信息技术类学院形成的节点关联极为密切，针对算法模型、科学技术等研究更新迭代速度极快，与外部机构保持长期合作交流是把握前沿动态的必然之举。



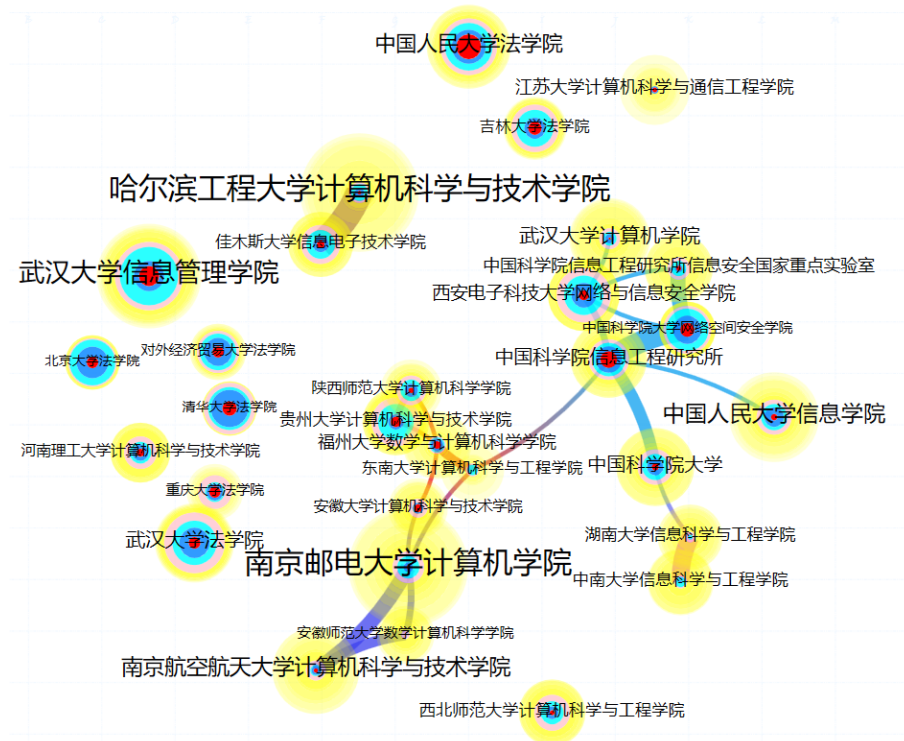


图3 国内个人信息保护研究机构共现图谱

由于大量研究机构以其二、三级下属机构署名发文，为了使文献计量更准确地反映一级研究机构的综合发文水平以及不同研究机构去除内部固有关联后的外部合作关系，本文采用人工统计的方法予以合并，表1展示了国内个人信息保护主题发文量排名前10的研究机构，共发表987篇论文，占研究文献总量的

25.79%，彰显了个人信息保护研究中的核心力量所在。其中，发文量最大的第一梯队是武汉大学（168篇）、电子科技大学（163篇）和中国科学院（154篇），均是第二梯队研究机构发文量的近两倍。

表1 国内个人信息保护主题发文量排名前10的研究机构

机构名称	发文量 (篇)	机构名称	发文量 (篇)
武汉大学	168	中国人民大学	78
电子科技大学	163	南京大学	68
中国科学院	154	南京邮电大学	64
西安电子科技大学	87	哈尔滨工程大学	63
清华大学	81	北京大学	61

研究机构分布呈现出三大特征：第一，研究成果在全国范围分布比较均匀，例如南京、北京、河南、哈尔滨、武汉、贵州、合肥等地都存在节点面积较大的研究机构，即国内各区域对个人信息保护的研究工作都予以足够的重视；第二，高校研究机构的参与程度最高，且以各高校中的法学院、信息学院与计算机学院为主体，这在一定程度上说明当前对于个人信息保护的研究主要还集中于理论层面，而在实际生活中与个人信息接触较多的如法院、图书馆、通讯社等应用型机构对个人信息保护的关

注相对低迷，有可能导致理论研究成果难以落地或者未达到预期效果的问题；第三，研究机构之间的合作具有较强的地域性特征，例如南京邮电大学、南京航空航天大学、安徽大学等长三角地区的高等院校间存在较强的合作关系，中国人民大学、中国科学院信息工程研究所等首都圈的研究机构也形成了较稳定的合作

模式。

### 2.2.2 文献作者

以相同方式绘制我国个人信息保护研究相关作者的共现图谱，图4展现了发文量排名前30的文献作者信息，图谱中每个节点表示一个作者，节点大小与该作者发文量呈正相关。节点间的连线反映不同作者间的合作关系。

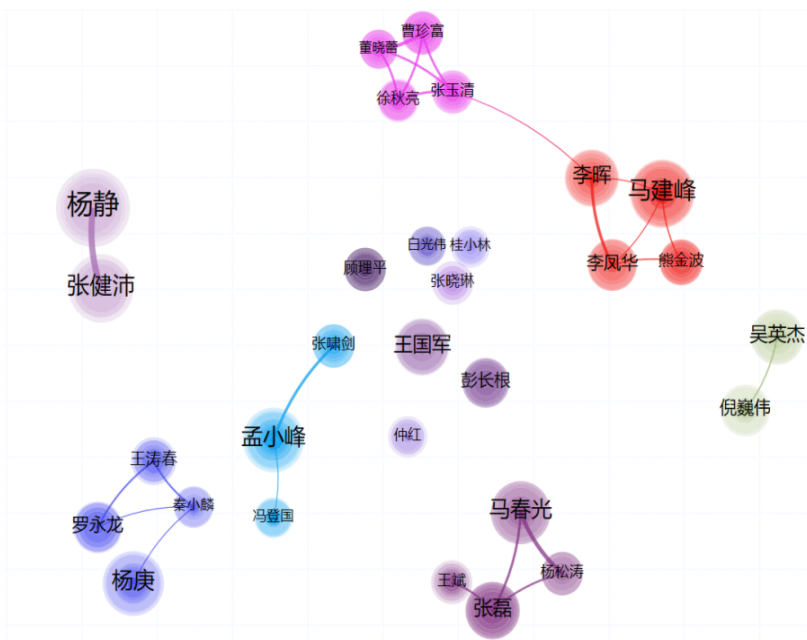


图4 国内个人信息保护研究作者共现图谱

从合作的聚类效应来看，文献作者之间的关系存在内部性、包容性、联系性的特征。一方面可以看出大部分节点之间都存在连线，即高产作者之间普遍建立了不同程度的合作关系，孤立节点中不乏具有一定发文量的文献作者，但其研究联系可能存在于低产作者之间。其次，文献作者共现图中存在4个至少包含4个节点的研究群体，通常以发文量最大者为研究核心，这些研究群体中既包含来自同一研究机构的作者，也包含了所属研究机构在地理位置上跨度较大的作者。

来自西安电子科技大学的马建峰和李晖、来自福建师范大学的熊金波和来自中国科学院信息工程研究所的李风华等人所组成的研究群体，主要涉及加密安全自毁方案<sup>[10]</sup>等创新类研究和云数据确定性删除<sup>[11]</sup>、隐私计算<sup>[12]</sup>、移动群智感知<sup>[13]</sup>等综述类研究；来自佳木斯大学的张磊、杨松涛和王斌与来自山东科技大学的马春光等人所组成的研究群体，研究内容高度集中在基于伪随机置换<sup>[14]</sup>、随机网格<sup>[15]</sup>、匿名区域层级扩张<sup>[16]</sup>等不同技术的位置隐私保护领域。这表明了当前个人信息保护领域的研究群

体在保持了同一研究机构内部紧密学术交流关系的同时,也注重开拓跨省份、跨地区的合作研究。

表2统计了国内个人信息保护主题发文量排名前10的作者,其近十年发表的关于个人信息保护的文章均在20篇以上,这些作者在共现

图谱中均占据了较大的节点位置,是推动我国个人信息保护研究发展的中坚力量。由历年发文量来看,该主题研究的关注度一直处于较高的水平,但排名前10的作者发文量仅占6.38%,表明个人信息保护领域作者基数庞大,高产作者与杰出作者比重有待提升。

表2 国内个人信息保护主题发文量排名前10的作者分布

作者姓名	发文量(篇)	所属单位	作者姓名	发文量(篇)	所属单位
杨静	33	哈尔滨工程大学	马春光	24	山东科技大学
马建峰	27	西安电子科技大学	张磊	21	佳木斯大学
张健沛	27	哈尔滨工程大学	李晖	21	西安电子科技大学
孟小峰	25	中国人民大学	王国军	21	广州大学
杨庚	25	南京邮电大学	吴英杰	20	福州大学

杨静、张健沛和其他学者聚焦研究个性化轨迹隐私保护技术<sup>[17]</sup>与多敏感属性隐私保护方法<sup>[18]</sup>;孟小峰、张啸剑等作者着重探讨大数据管理时代隐私风险管理问题<sup>[19]</sup>以及将差分隐私与自适应网格<sup>[20]</sup>、矩阵分解<sup>[21]</sup>等技术融合开发隐私保护算法的可能性;杨庚、王涛春、秦小麟等作者联合提出传感器网络中的隐私保护方案<sup>[22]</sup>并综合分析联邦学习模式下隐私保护模型<sup>[23]</sup>;王国军与其他学者合作设计移动社交网络中结合跨域代理重加密<sup>[24]</sup>、混淆矩阵变换<sup>[25]</sup>等方法的朋友发现隐私保护机制。

### 3 国内个人信息保护研究应用维度演化分析

#### 3.1 期刊构成

图5呈现了我国个人信息保护研究发文量排名前25的CSSCI来源期刊,其中情报类7种,共计发文283篇;图书馆类7种,共计发文208篇;

法律类5种,共计发文80篇;新闻传播类6种,共计发文77篇。由此可见,CSSCI刊源的个人信息保护文献主要集中发表在图情类期刊,而发文量占据前三的《现代情报》《情报理论与实践》和《情报杂志》均属于情报类期刊。

#### 3.2 关键词突变分析

钟辉新<sup>[26]</sup>指出突现词检测相较于词频分析更易于锁定新兴主题的增长势头,特定领域在长期发展的过程中必然包含了多样化主题的相继突现,而单一突现主题往往伴随着批量含义相近关键词的诞生,共同阐释了领域内某一时间段而非时间点的研究态势。王梦婷<sup>[27]</sup>根据突现检测结果分时间段给出竞争情报领域的研究热点,卢新元等<sup>[28]</sup>借助关键词突变表观察企业知识转移领域在不同阶段的前沿趋势,因此,以核心关键词为导向并从阶段性视角分析个人信息保护领域的演进过程具备可操作性,也是应有之义。

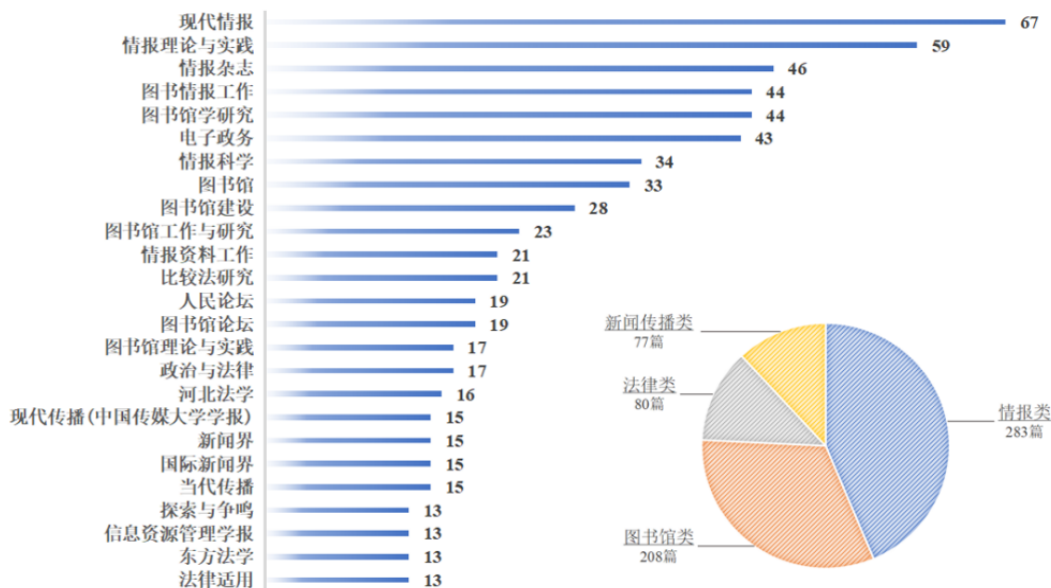


图5 国内个人信息保护研究应用维度期刊分布

本文利用 CiteSpace 提供的词频探测技术,对个人信息保护领域的关键词进行突变分析,表3依次呈现了获取突现词的名称、突现强度、突现开始时间、突现结束时间以及突现持续时间段,所展示的 CSSCI 刊源在个人信息保护研究中最强突现的 26 个关键词基本属于社会科学领域概念,充分反映了 CSSCI 刊源的应用性特征,分别以具有较高的突现度的“图书馆”和“大数据”关键词为阶段核心,将近年来我国个人信息保护研究在应用层面的演进大致分为两个阶段:2011—2015 年和 2016—2021 年。

### 3.2.1 隐私与信息理论萌芽阶段(2011—2015 年)

该阶段的研究主要集中在如何完善法律体系中有关隐私权的界定,以及如何在图书馆等现实信息场所中实现对多样化信息的有序管理和高效传递。以“隐私”为主体的相关高频关键词多达五个,“立法”“法律保护”“伦理

问题”关键词的存在说明了学者对于如何运用法律权威维护个人隐私权益投入了大量的研究资源,杜红原<sup>[29]</sup>阐明隐私权保护的重中之重是对隐私权概念的界定;安宝洋、翁建定<sup>[30]</sup>发现多元学科和部门的协同治理是解决大数据技术带来的网络信息伦理缺失问题的关键。

“图书馆”一词的突现度达 10.41,与“读者”“个性化服务”“个人信用信息”等关键词共同反映了当时图书馆作为线下信息流转的核心在信息服务中起到的重要作用;杨利军和高军<sup>[31]</sup>设计图书馆大数据可视化分析系统框架以期帮助图书馆员有效挖掘海量复杂大数据效能;陈臣和马晓亭<sup>[32]</sup>提出基于小数据的图书馆个性化服务推送模式更有助于响应读者的差异化需求。

### 3.2.2 大数据与法律权益融合阶段(2016—2021 年)

该阶段的研究主要集中探讨在大数据时代

表3 国内个人信息保护研究应用维度突现词

突现词	突现强度	开始年	结束年	持续时间
保护	3.28	2011	2013	
读者	3.21	2011	2014	
实证研究	4.13	2011	2013	
隐私权保护	3.88	2011	2013	
个性化服务	3.75	2011	2014	
云计算	3.65	2011	2014	
隐私权	7.48	2011	2013	
个人信用信息	2.89	2011	2014	
图书馆	10.41	2011	2015	
隐私	4.49	2012	2015	
伦理问题	2.53	2012	2015	
电子商务	3.64	2013	2017	
立法	3.58	2013	2014	
人格尊严	2.67	2013	2015	
法律保护	5.22	2013	2015	
大数据时代	2.80	2014	2018	
大数据	10.16	2015	2016	
被遗忘权	4.76	2015	2018	
美国	4.08	2016	2017	
电子政务	2.76	2016	2019	
网络隐私	2.53	2017	2018	
政府数据	3.17	2017	2018	
高校图书馆	5.11	2017	2019	
数据开放	2.82	2017	2019	
影响因素	3.29	2017	2018	
知情同意	3.57	2019	2021	

如何维护衍生出的新型个人信息保护权利，以及如何在发达的信息网络及时获取有效的信息但又能保证个人隐私不被过度泄露。匡文波<sup>[33]</sup>指出当前的个性化推荐算法对多项个体权益都造成了不同程度的威胁，“被遗忘权”“网络隐私”“知情同意”等最新的高频关键词正是对这类问题的直接反映。蔡培如<sup>[34]</sup>通过欧盟法理实践历史探究以被遗忘权平衡言论自由与知情权利益的可能性；范海潮和顾理平<sup>[35]</sup>认为科

学的知情同意构想能够缓解个人信息保护领域的知情主客体交流鸿沟、执行同意环节失灵等困境。

“大数据”一词的突现度达 10.16，与“数据开放”“电子政务”“网络隐私”等关键词共同描述了数据智能时代的信息对立现象，个人信息保护制度对传统隐私权规范的取代是维护政府数据安全的有益路径探索<sup>[36]</sup>。值得注意的是，在高度开放的网络环境中人们能够便捷



搜寻到大量相关信息，但如何从驳杂的数据集中清洗出最有价值的信息具有较大的难度，反而会陷入一种相对信息匮乏的状态，甚至会主动以个人信息为代价去换取检索过程的简化。

### 3.3 热点演化趋势分析

从图6可以观察到每一主题在整个时间跨度上的演变进程以及由热点事件启发的研究方

向。对于个人隐私保护、信息安全、图书馆管理的研究是贯穿整条时间线的，可见这些主题在个人信息保护研究中处于基础性地位。综合来看，应用维度的个人信息保护领域研究趋势主要呈现为“个人隐私转向公共数据”“人力调查转向数字治理”“信息自决转向政府责任”等宏观化特征，个人信息保护逐渐成为全民所需、全民参与、全民监督的一项公共事业。

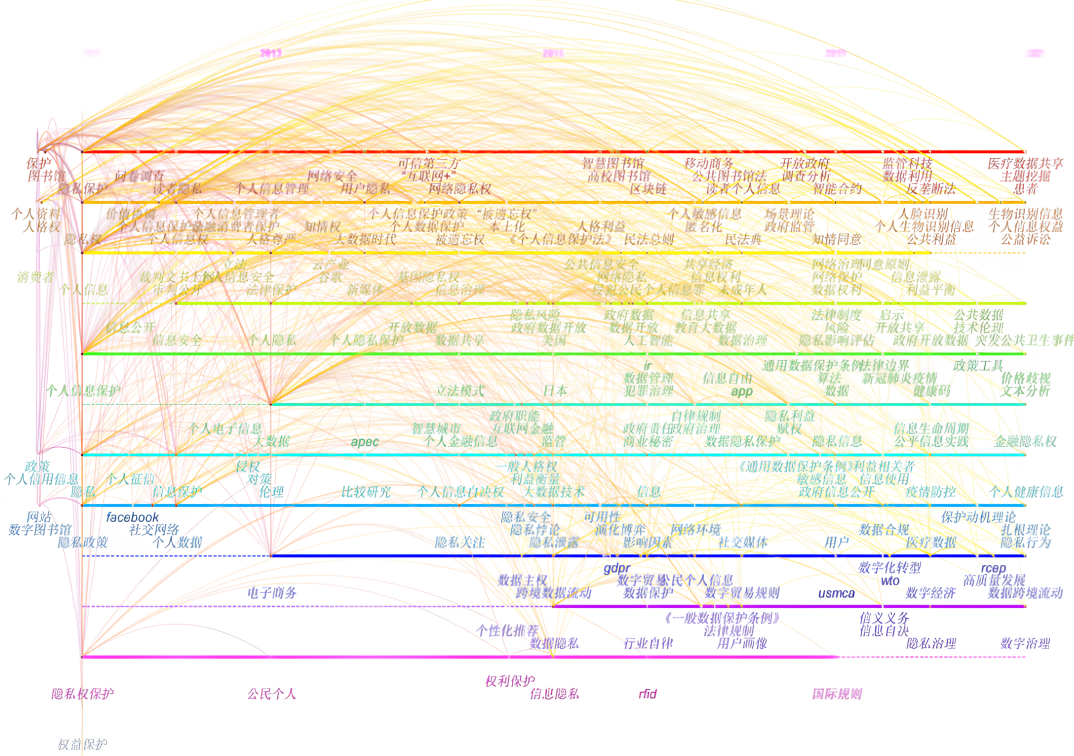


图6 国内个人信息保护研究应用维度关键词时间线分布图谱

新冠肺炎疫情暴发以来，政府对于公民个人信息大范围、多频次的收集和使用引发了媒体和社会公众对个人信息安全担忧，同时也掀起了化解由突发公共卫生事件造成的个人信息安全威胁的研究浪潮。苏今<sup>[37]</sup>通过涉疫信息的场景控制和权限设定分五阶段实现后疫情时代个人信息保护；占南<sup>[38]</sup>从三类个人信息保护关联主体视角出发构建面向隐私保护设计的全生命周期保护框架；沈伟伟<sup>[39]</sup>提出从紧急状态

到日常状态的“恢复机制”理念并将其拓展到健康码记忆敏感信息的规制实践中。

医疗机构借助大数据技术以电子化的形式记录和存储了海量患者健康医疗信息，相关数据泄露将危害社会公共利益和国家安全，因此做好患者隐私和个人信息保护始终是研究焦点与难点。现行医疗隐私规则使得个人隐私保护机制生效范围有限，识别隐私利益的群体维度有助于获取统计意义上的健康信

息<sup>[40]</sup>；郭建<sup>[41]</sup>指出有必要建立一套伦理治理原则以规避健康医疗大数据应用中的伦理风险；相丽玲和陈琬珠<sup>[42]</sup>研究发现个人健康医疗信息保护逐步形成技术、管理、法律三位一体的跨领域保护模式。

## 4 国内个人信息保护研究技术维度演化分析

### 4.1 期刊构成

图 7 呈现了我国个人信息保护研究发文量

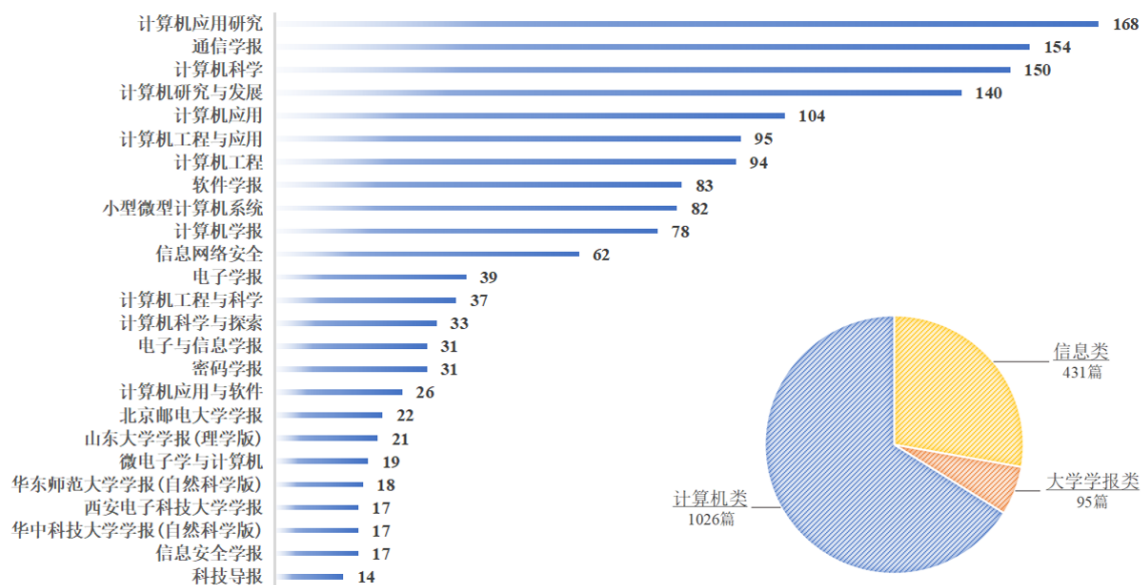


图 7 国内个人信息保护研究技术维度期刊分布

### 4.2 关键词突变分析

表 4 展示了 CSCD 期刊源在个人信息保护研究中最强突现的 37 个关键词，基本属于核心科技领域的概念，充分反映了 CSCD 刊源的技术特征，其中“数据发布”“无线传感器网络”“差分隐私”和“区块链”具有较高的突现度，分别以这些关键词为阶段核

排名前 25 的 CSCD 来源期刊，其中计算机类 12 种，共计发文 1026 篇；信息类 8 种，共计发文 431 篇；大学学报类 5 种，共计发文 95 篇。研究发现，25 种期刊仅占期刊总量的 4.96%，其 1552 篇的发文量占据文献总量的 40.56%，而 CSSCI 刊源文献仅占 16.94%。由此可见，国内学者更倾向于从技术维度研究个人信息保护问题，且高度集中于计算机领域，排名前五中的《计算机应用研究》《计算机科学》《计算机研究与发展》和《计算机应用》期刊发文量均超过 100 篇。

心，将近年来我国个人信息保护研究在技术层面的演进大致分为三个阶段：2011—2012 年、2013—2017 年和 2018—2021 年。此外，结合突现词数量容易发现，个人信息保护领域在技术维度的研究相较应用维度具有更快的更新速度、更多样化的研究对象以及更具波动性的未来趋势。

表4 国内个人信息保护研究技术维度突现词

突现词	突现强度	开始年	结束年	持续时间
关联规则	3.2752	2011	2013	
l-多样性	2.7388	2011	2013	
多敏感属性	3.4594	2011	2014	
泛化	2.7435	2011	2012	
聚类	4.1314	2011	2014	
有损连接	3.1074	2011	2013	
数据发布	7.6127	2011	2012	
数据挖掘	3.1659	2011	2013	
社会网络	4.2134	2011	2016	
数据融合	5.3095	2013	2014	
无线传感器网	8.1881	2013	2014	
数据聚合	2.4937	2013	2016	
数据聚集	2.7981	2013	2014	
完整性验证	2.8056	2013	2017	
大数据	2.7611	2014	2016	
匿名	2.8138	2014	2017	
轨迹数据发布	2.5775	2014	2015	
范围查询	3.5398	2014	2017	
移动互联网	2.5775	2014	2015	
位置隐私	2.8328	2015	2016	
移动社交网络	4.0756	2015	2017	
隐私	3.3081	2015	2016	
位置服务	4.2111	2015	2016	
k匿名	3.0055	2015	2017	
隐私度量	2.8051	2015	2018	
属性基加密	3.5362	2017	2018	
轨迹数据	2.6797	2017	2018	
轨迹隐私	3.9837	2017	2018	
群智感知	2.5718	2018	2021	
移动群智感知	2.6686	2018	2021	
属性加密	2.8592	2018	2019	
智能电网	3.9256	2019	2021	
激励机制	3.4076	2019	2021	
区块链	30.902	2019	2021	
差分隐私	10.3813	2019	2021	
数据共享	3.858	2019	2021	
零知识证明	2.8032	2019	2021	

#### 4.2.1 基础算法与隐私规则的探索优化阶段 (2011—2012年)

该阶段的研究主要集中于对传统个人信息保护规则的优化更新和对基础个人信息保护算法的功能提升，“数据发布”作为该阶段突现度最高的关键词则表明，此时个人信息保护的研究热点依然围绕着数据本身，但一味追求数据的完全模糊化处理难以满足不同应用领域对匿名发布数据的质量需求，如何借助个性化隐私匿名技术平衡个性化服务质量和隐私保护效果受到广泛关注<sup>[43]</sup>。

刘彬等<sup>[44]</sup>在 if-then 算法设计中融入兴趣度、规则左件和逐步移项的思想，隐藏敏感信息的同时有效控制规则丢失率；刘峰等<sup>[45]</sup>借助安全多方计算和随机干扰矩阵缩减数据开销，高效解决半诚实模型下的隐私保护问题；王波和杨静<sup>[46]</sup>利用个性化扩展 L-多样性逆聚类算法构建隐私匿名模型，以满足不同用户个性化的隐私保护需求；徐勇等<sup>[47]</sup>提出的考虑敏感属性权重的数据发布算法能够迎合不同应用领域的隐私保护意图和数据质量要求。

#### 4.2.2 交互网络覆盖与隐私轨迹加密实践阶段 (2013—2017年)

随着智能手机的普及和一系列自带定位功能的社交软件开发，位置信息的泄露风险急剧增加，世界高度互联互通使得行为个体往往会无意识被纳入多个社交关系网络，个人隐私数据更容易被窃取。因此，该阶段的研究更为注重如何紧贴数字时代的信息共享特性，研发能够迅速投入实践的个人信息保护通用技术和隐私攻击抵御机制，而非长期停留在实验精度和效度的数据模拟阶段。此时共现关键词数量出现激增，说明个人

信息保护研究已经具有一定的热度。

以高频关键词“无线传感器网络”为核心，“移动互联网”“范围查询”“移动社交网络”等均体现了隐私数据的大范围散布，数据聚合对原始感知数据的篡改在一定程度上增加了隐私信息加密的难度。丁超等<sup>[48]</sup>基于隐私同态和聚合消息验证码技术提出的可恢复数据聚合方案能够兼顾数据隐私性与完整性；陈燕俐等<sup>[49]</sup>采用加法同态加密和同态消息认证码打造的轻量级安全数据融合保护方案便于处理多源异构数据场景；而这一时期诸如桶技术、前缀成员验证、保序加密等隐私保护范围查询处理技术尚未较好地实现隐私性、完整性、高效性和精确性四者之间的均衡<sup>[50]</sup>。

在隐私轨迹发布过程中如何既保证个体本身的敏感位置不泄露，并防止攻击者通过轨迹行为推导出其他敏感信息，也是当时专家学者热切关注的问题。杨静等<sup>[51]</sup>根据用户轨迹匿名的等级差异构造规模可变的个性化轨迹图模型以改善轨迹数据可用性低下的表现；张博闻等<sup>[52]</sup>通过构造隐私保护 Trie 树并根据 (k,p) 敏感轨迹进行剪枝重构来实现可信第三方隐私保护模块功能。

#### 4.2.3 社群感知和差分隐私智能化应用阶段 (2018—2021年)

该阶段中“区块链”的突现度高达 30.90，区块链技术所具备的难以篡改、智能合约、网状直接协作机制等优势令其实现了对中心记账体系的颠覆，但性能及扩展性不足、数据隐私与访问控制机制不成熟和治理机制不完善等缺陷都限制着区块链产业的规范化发展，“同态加密”“属性加密”“零知识证明”等均反



映了当前研究者为解决区块链个人信息保护问题所做出的努力。田有亮等<sup>[53]</sup>基于属性加密和策略更新算法确保区块链场景下的交易隐私动态保护和溯源信息动态共享同步开展；许重建和李险峰<sup>[54]</sup>在蚂蚁区块链平台上验证其双重加密方案能以字段级别细粒度保护隐私数据并高效完成全链路操作；李莉等<sup>[55]</sup>综合群签名、隐私地址协议、零知识证明等多项技术制作了一类分布式可监管隐私保护方案。

“差分隐私”同样是该阶段突现度较高的关键词，通过向原始数据添加随机噪声以达到降低数据敏感度的目的，但在同等数据量的条件下，发布数据的有效比例会降低，故学界尝试构建不同隐私约束来规避数据恢复时的复杂计算，应用于位置、感知、电表等隐私数据进

而追求高智能化个人信息保护水平。杨丽和陈思光<sup>[56]</sup>使用云雾协作的多级聚合模型节省智能电表通信开销，并借助同态加密技术实现隐私数据的轻量级认证；李卓等<sup>[57]</sup>考虑移动群智感知用户提交数据特点，分别从属性联合与属性独立视角出发设计本地差分隐私保护类算法。

### 4.3 热点演化趋势分析

图8表明通过主题聚类可以将技术维度的个人信息保护研究划分为八大主题。各主题间并非泾渭分明，而是相互渗透、相互演化、相互支撑，同时观察到研究早期各主题均有较多的分支，随着时间的推移，分布图谱的密度有所下降，细分研究方向主要集中于“差分隐私”“云计算”“区块链”和“基于位置的服务”四个主题。

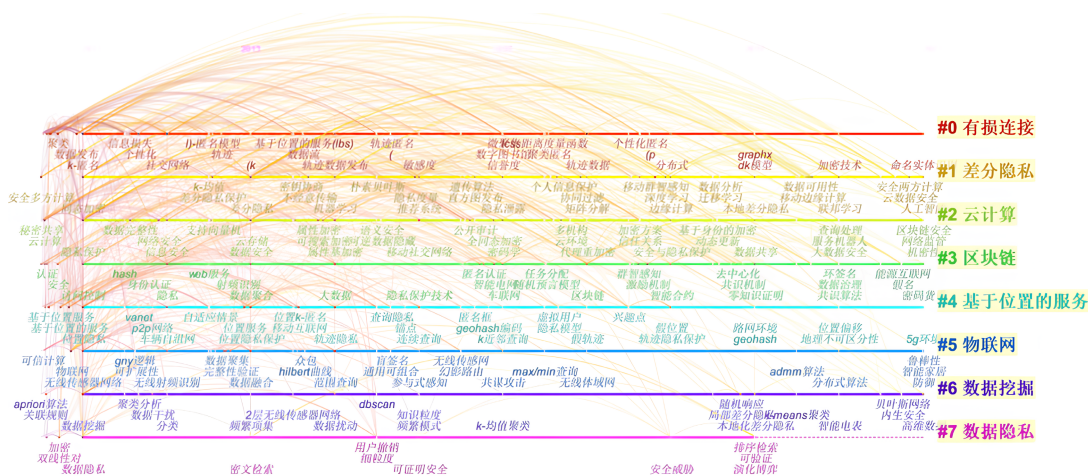


图8 国内个人信息保护研究技术维度关键词时间线分布图谱

联邦学习是近两年来取得迅猛发展的研究热点之一，区别于传统隐私保护方法的原理，联邦学习不会出现数据和模型的传递，通过加密机制的参数交换方式保护用户数据信息的安全，常与区块链相结合用以抵御隐私攻击。方晨等<sup>[58]</sup>引入区块链的共识和激励机制促进设备参与联邦训练，辅以自适应差分方法防范边缘

节点的隐私泄露可能；高胜等<sup>[59]</sup>以指数机制采样梯度贡献值高的隐私点，通过双因子调整机制实现全局模型更新，增强区块链体系下异步联邦学习方案的隐私保护效率。

随着第五代移动通讯技术（5G）的到来，数据传播速度和共享模式必将发生深刻变革，在方便用户长距离通讯和利用数据资源的同时，



个人数据的进一步透明化使个人信息保护将面临诸多未知的新挑战。冯登国等<sup>[60]</sup>将5G安全框架归纳为接入安全、网络安全、用户安全等七大层面,指出认证方案、密钥架构、切片隔离等是5G环境的关键安全技术。姜海洋等<sup>[61]</sup>在5G定位背景下通过降维处理、定位耦合和对称加密传输提升移动用户隐私保护能力。

## 5 结语与展望

本文借助CiteSpace和ITGInsight科学知识图谱软件,选取2011—2021年国内与个人信息保护研究主题相关的CSSCI和CSCD来源期刊文献,首先以文献计量手段和机构作者共现图谱识别个人信息保护的时空分布特征,其次分别从应用与技术维度可视化分析个人信息保护的热点演化趋势,得出结论如下:

第一,国内个人信息保护文献数量呈现出阶段性平稳、总体接近线性增长的时间分布特征,近年来应用类文章增长态势强于技术类文章。该领域发文在2013和2017年分别出现了显著的激增现象,随着《个人信息保护法》的颁布,我国个人信息保护研究必然出现更大的创新和突破。国内个人信息保护领域呈现出研究机构间的合作具有地域性和文献作者间的合作存在内部性的空间分布特征。总体来看,需要推动应用型研究机构积极参与到个人信息保护的研究当中,与理论型研究机构达成优势互补的合作,并注重增强不同研究群体跨省份、跨地区的科研联系。

第二,基于研究内容,可以将应用维度的研究区间分为隐私与信息理论萌芽阶段、大数据与法律权益融合阶段,将技术维度分为基础

算法与隐私规则的探索优化阶段、交互网络覆盖与隐私轨迹加密实践阶段、社群感知和差分隐私智能化应用阶段。基于研究规模,技术维度排名前25的期刊发文体量是应用维度的两倍有余,在单一类型期刊上的集中趋势更强,突现词的平均持续时间更短,迭代也更为频繁。应用维度的热点演化随时间而进一步发散,技术维度的热点演化随时间逐渐聚焦于各自的核心要义,可见应用维度倾向于研究成果多点下探,技术维度倾向于研究成果单点泛化。

第三,中国个人信息保护具象于应用场景的研究呈现出公共化、数字化、宏观化的特征,研究主阵地坐落于情报、新闻传播与法律领域。信息爆炸性增长、信息主体纷繁交互、传播渠道日新月异致使个人信息泄露风险升级,进一步催生个人信息保护意识与隐私保护法律权益的诞生。然而随着个性化推荐算法等技术的不断普及,信息茧房与相对信息匮乏激化了情报价值与自我维权的对立,理论圈层的独立研究并不能从根本上平衡技术发展与信息保护的矛盾。

第四,中国个人信息保护具象于技术路线的研究呈现出专一化、精细化、智能化的特征,研究主阵地坐落于计算机和电子信息领域。早期匿名、脱敏和随机干扰等非密码技术兼顾了个性化信息保护需求与信息服务质量,同时以安全多方计算技术为代表的密码技术及其与其他密码技术的组合也被广泛运用于数据发布、共享和计算;随着位置轨迹与社交网络类信息的泛滥,学者通过改造图、树模型数据结构引入位置、社交网络外部信息以进行隐私攻击抵御机制与轨迹行为加密相关研究,多源异构的数据源提升了数据融合难度,主要使用全同态

加密技术解决数据不匹配、不均衡问题；当区块链的概念被提出，分布式个人信息保护方案与差分隐私算法更为适配，零知识证明具有数据最小化、关键信息隐私化的优势，在区块链方面也得到了广泛的应用，同时以联邦学习为代表的信息保护技术也在隐私 AI 领域被广泛研究，在保护个体核心数据资产的同时降低业务成本与风险<sup>[62]</sup>。从基于数据本身的加密处理到同态加密、零知识证明等信息保护技术应用范围的拓展，再到感知不同的应用场景反哺优化技术主体，实际上是一个被动满足需求到主动感知拓展的过程。

个人信息保护事业的发展离不开应用与技术双维度的融合支撑，应用无法架空，技术终须落地，“事后惩罚”不如“事前监控”，以理论创新为更多具体的信息保护需求场景寻求更加合适和先进的技术解决方案，以应用场景的特殊性与真实性为技术发展引入人为知识的指导，从而推动个人信息保护研究领域内容不断丰富、完善。

## 参考文献

- [1] 孟小峰, 王雷霞, 刘俊旭. 人工智能时代的数据隐私、垄断与公平 [J]. 大数据, 2020, 6(1): 35-46.
- [2] 熊金波, 王敏桑, 田有亮, 等. 面向云数据的隐私度量研究进展 [J]. 软件学报, 2018, 29(7): 1963-1980.
- [3] 方滨兴, 贾焰, 李爱平, 等. 大数据隐私保护技术综述 [J]. 大数据, 2016, 2(1): 1-18.
- [4] 沈光. 文献计量视角下我国生态文明研究的热点与动态——基于 6240 篇 CSSCI 刊源文献信息的分析 [J]. 南京工业大学学报 (社会科学版), 2020, 19(5): 78-89, 116.
- [5] 刘雅琦, 贾利芳. 知识图谱视角下的个人信息保护热点及前沿研究 [J]. 科技管理研究, 2019, 39(18): 189-198.
- [6] 谭春辉, 熊梦媛. 基于 LDA 模型的国内外数据挖掘研究热点主题演化对比分析 [J]. 情报科学, 2021, 39(4): 174-185.
- [7] 李忠夏. 数字时代隐私权的宪法建构 [J]. 华东政法大学学报, 2021, 24(3): 42-54.
- [8] 陈超美, 陈悦, 侯剑华, 等. CiteSpace II: 科学文献中新趋势与新动态的识别与可视化 [J]. 情报学报, 2009, 28(3): 401-421.
- [9] 刘玉琴, 汪雪峰, 雷孝平. 科研关系构建与可视化系统设计与实现 [J]. 图书情报工作, 2015, 59(8): 103-110, 125.
- [10] 熊金波, 姚志强, 马建峰, 等. 面向网络内容隐私的基于身份加密的安全自毁方案 [J]. 计算机学报, 2014, 37(1): 139-150.
- [11] 熊金波, 李风华, 王彦超, 等. 基于密码学的云数据确定性删除研究进展 [J]. 通信学报, 2016, 37(8): 167-184.
- [12] Li F, Li H, Niu B, et al. Privacy computing: concept, computing framework, and future development trends [J]. Engineering, 2019, 5(6): 1179-1192.
- [13] 熊金波, 毕仁万, 田有亮, 等. 移动群智感知安全与隐私: 模型、进展与趋势 [J]. 计算机学报, 2021, 44(9): 1949-1966.
- [14] 周长利, 田晖, 马春光, 等. 路网环境下基于伪随机置换的 LBS 隐私保护方法研究 [J]. 通信学报, 2017, 38(6): 19-29.
- [15] 杨松涛, 王慧强, 马春光. 基于随机网格的位置隐私保护方法 [J]. 系统工程与电子技术, 2018, 40(2): 422-426.
- [16] 张磊, 马春光, 印桂生. 匿名区域层级扩张的位置隐私保护方法 [J]. 系统工程与电子技术, 2021, 43(2): 561-566.
- [17] 杨静, 张冰, 张健沛, 等. 基于图划分的个性化轨迹隐私保护方法 [J]. 通信学报, 2015, 36(3): 5-15.
- [18] 谢静, 张健沛, 杨静, 等. 基于属性相关性划分的多敏感属性隐私保护方法 [J]. 电子学报, 2014, 42(9): 1718-1723.
- [19] 孟小峰, 张啸剑. 大数据隐私管理 [J]. 计算机研究与发展, 2015, 52(2): 265-281.
- [20] 张啸剑, 金凯忠, 孟小峰. 基于自适应网格的隐私空间分割方法 [J]. 计算机研究与发展, 2018, 55(6): 1143-1156.
- [21] 张啸剑, 付聪聪, 孟小峰. 结合矩阵分解与差分隐私的人脸图像发布 [J]. 中国图象图形学报, 2020, 25(4): 655-668.
- [22] 王涛春, 秦小麟, 张吉, 等. 传感器网络中基于路线的隐私保护数据聚集算法 [J]. 电子学报, 2017, 45(6): 1334-1341.
- [23] 杨庚, 王周生. 联邦学习中的隐私保护研究进展 [J]. 南京邮电大学学报 (自然科学版), 2020, 40(5): 204-214.
- [24] 罗恩韬, 王国军, 陈淑红, 等. 移动社交网络中跨域代理重加密朋友发现隐私保护方案研究 [J]. 通

- 信学报, 2017, 38(10): 81-93.
- [25] 罗恩韬, 王国军, 刘琴, 等. 移动社交网络中矩阵混淆加密交友隐私保护策略 [J]. 软件学报, 2019, 30(12): 3798-3814.
- [26] 钟辉新. 新兴趋势探测研究综述 [J]. 现代情报, 2017, 37(12): 162-167.
- [27] 王梦婷. 基于突变检测的主题突变分析研究 [J]. 情报科学, 2016, 34(12): 36-39.
- [28] 卢新元, 张恒, 王馨悦, 等. 基于科学计量学的国内企业知识转移研究热点和前沿分析 [J]. 情报科学, 2019, 37(3): 169-176.
- [29] 杜红原. 论隐私权概念的界定 [J]. 内蒙古社会科学 (汉文版), 2014, 35(6): 105-109.
- [30] 安宝洋, 翁建定. 大数据时代网络信息的伦理缺失及应对策略 [J]. 自然辩证法研究, 2015, 31(12): 42-46.
- [31] 杨利军, 高军. 图书馆个性化服务中的大数据可视化分析与应用研究 [J]. 现代情报, 2015, 35(7): 68-72.
- [32] 陈臣, 马晓亭. 基于小数据的图书馆个性化推送服务与服务质量保证研究 [J]. 情报理论与实践, 2015, 38(10): 95-99.
- [33] 匡文波. 对个性化算法推荐技术的伦理反思 [J]. 上海师范大学学报 (哲学社会科学版), 2021, 50(5): 14-23.
- [34] 蔡培如. 被遗忘权制度的反思与再建构 [J]. 清华法学, 2019, 13(5): 168-185.
- [35] 范海潮, 顾理平. 探寻平衡之道: 隐私保护中知情同意原则的实践困境与修正 [J]. 新闻与传播研究, 2021, 28(2): 70-85, 127-128.
- [36] 商希雪, 韩海庭. 政府数据开放中个人信息保护路径研究 [J]. 电子政务, 2021(6): 113-124.
- [37] 苏今. 后疫情时代个人涉疫信息的控制特点及其路径修正——以隐私场景理论为视角 [J]. 情报杂志, 2021, 40(9): 124-132, 123.
- [38] 占南. 重大疫情防控中的个人信息保护研究——基于隐私保护设计理论 [J]. 现代情报, 2021, 41(1): 101-110.
- [39] 沈伟伟. 论数字紧急状态的恢复机制——以新冠疫情防控为例 [J]. 清华法学, 2021, 15(2): 121-142.
- [40] 刘士国, 熊静文. 健康医疗大数据中隐私利益的群体维度 [J]. 法学论坛, 2019, 34(3): 125-135.
- [41] 郭建. 健康医疗大数据应用中的伦理问题及其治理思考 [J]. 自然辩证法研究, 2020, 36(3): 85-90.
- [42] 相丽玲, 陈琬珠. 个人健康医疗信息保护的研究进展与未来趋势 [J]. 情报科学, 2020, 38(6): 170-177.
- [43] 王波, 杨静. 数据发布中的个性化隐私匿名技术研究 [J]. 计算机科学, 2012, 39(4): 168-171, 200.
- [44] 刘彬, 孟凡荣, 汪楚娇. 基于兴趣度的隐私保护关联规则挖掘算法 [J]. 计算机工程与设计, 2011, 32(6): 2124-2128.
- [45] 刘峰, 薛安荣, 王伟. 一种隐私保护关联规则挖掘的混合算法 [J]. 计算机应用研究, 2012, 29(3): 1107-1110.
- [46] 王波, 杨静. 一种基于逆聚类的个性化隐私匿名方法 [J]. 电子学报, 2012, 40(5): 883-890.
- [47] 徐勇, 秦小麟, 杨一涛, 等. 一种考虑属性权重的隐私保护数据发布方法 [J]. 计算机研究与发展, 2012, 49(5): 913-924.
- [48] 丁超, 杨立君, 吴蒙. 一种同时保障隐私性与完整性的无线传感器网络可恢复数据聚合方案 [J]. 电子与信息学报, 2015, 37(12): 2808-2814.
- [49] 陈燕俐, 张乾, 许建, 等. 无线传感器网络多应用场景下的安全数据融合方案 [J]. 计算机科学, 2017, 44(9): 162-167.
- [50] 张晓莹, 董蕾, 陈红. 无线传感器网络隐私保护范围查询处理技术 [J]. 华东师范大学学报 (自然科学版), 2015(5): 1-13.
- [51] 杨静, 张冰, 张健沛, 等. 基于图划分的个性化轨迹隐私保护方法 [J]. 通信学报, 2015, 36(3): 5-15.
- [52] 张博闻, 陈晶, 杜瑞颖. 一种个性化移动社交网络轨迹隐私保护方案 [J]. 计算机应用研究, 2017, 34(3): 871-874.
- [53] 田有亮, 杨科迪, 王缙, 等. 基于属性加密的区块链数据溯源算法 [J]. 通信学报, 2019, 40(11): 101-111.
- [54] 许重建, 李险峰. 区块链交易数据隐私保护方法 [J]. 计算机科学, 2020, 47(3): 281-286.
- [55] 李莉, 杜慧娜, 李涛. 基于群签名与属性加密的区块链可监管隐私保护方案 [J]. 计算机工程, 2022, 48(6): 132-138.
- [56] 杨丽, 陈思光. 雾辅助的轻量级隐私保护数据多级聚合研究 [J]. 小型微型计算机系统, 2020, 41(6): 1224-1230.
- [57] 李卓, 宋子晖, 沈鑫, 等. 边缘计算支持下的移动群智感知本地差分隐私保护机制 [J]. 计算机应用, 2021, 41(9): 2678-2686.
- [58] 方晨, 郭渊博, 王一丰, 等. 基于区块链和联邦学习的边缘计算隐私保护方法 [J]. 通信学报, 2021, 42(11): 28-40.
- [59] 高胜, 袁丽萍, 朱建明, 等. 一种基于区块链的隐私保护异步联邦学习 [J]. 中国科学: 信息科学, 2021, 51(10): 1755-1774.
- [60] 冯登国, 徐静, 兰晓. 5G 移动通信网络安全研究 [J]. 软件学报, 2018, 29(6): 1813-1825.
- [61] 姜海洋, 曾剑秋, 韩可, 等. 5G 环境下移动用户位置隐私保护方法研究 [J]. 北京理工大学学报, 2021, 41(1): 84-92.
- [62] 霍炜, 郁昱, 杨糠, 等. 隐私保护计算密码技术研究进展与应用 [J]. 中国科学: 信息科学, 2023, 53(9): 1688-1733.



开放科学  
(资源服务)  
标识码  
(OSID)

# 长三角地区技术转移与成果转化效率分析

## ——基于三阶段 DEA 模型分析

戴力新<sup>1</sup> 李红<sup>1</sup> 戚湧<sup>2</sup>

1 江苏省科技资源统筹服务中心 南京 210004;  
2. 南京理工大学知识产权学院 南京 210094

**摘要:** [目的/意义] 研究技术转移与成果转化效率顺应时代发展要求, 对于加强技术创新政策优化配置, 催化科技成果真正转化为现实的生产力具有重要的理论意义和现实意义。[方法/过程] 利用包括随机前沿分析在内的三阶段 DEA 方法, 以长三角地区的江苏、上海、浙江和安徽四个省市作为研究对象, 对其 2015—2021 年的技术转移与成果转化效率进行评价, 并对比分析剔除环境因素和随机因素前后的效率值。[局限] 研究样本较少, 未考虑科技镇长团成员数等可能影响技术转移与成果转化的其他因素。[结果/结论] 结果发现长三角四省市的技术转移与成果转化效率整体较高, 其中服务环境、科技政策环境、经济发展环境和随机扰动是影响技术转移与成果转化效率的重要因素, 通过第三阶段 DEA 分析发现, 四省市的效率均值有所降低。根据上述研究结论, 从合理规划技术转移与成果转化资源投入、健全科技成果转化机制等方面提出对策建议。

**关键词:** 长三角; 三阶段 DEA; 技术转移; 成果转化

**中图分类号:** F124.3; G35

## Analysis of Technology Transfer and Achievement Transformation Efficiency in Yangtze River Delta Region: Based on Three-Stage DEA Model Analysis

DAI Lixin<sup>1</sup> LI Hong<sup>1</sup> QI Yong<sup>2</sup>

1. Jiangsu Province Science and Technology Resources Coordination Service Center, Nanjing 210004, China;  
2. School of Intellectual Property, Nanjing University of Science and Technology, Nanjing 210094, China

**Abstract:** [Objective/ Significance] Research on the efficiency of technology transfer and achievement transformation conforms

**基金项目** 国家自然科学基金面上项目“基于专利分析的智能网联汽车技术创新网络治理研究”(71974096)。

**作者简介** 戴力新(1966-), 学士, 副研究员, 主要研究方向为技术转移与成果转化, E-mail: njust\_qy@163.com; 李红(1982-), 硕士, 副研究员, 主要研究方向为科技创业服务、技术转移与成果转化; 戚湧(1970-), 博士, 教授, 主要研究方向为创新管理、知识产权管理、战略性新兴产业等。

**引用格式** 戴力新, 李红, 戚湧. 长三角地区技术转移与成果转化效率分析——基于三阶段 DEA 模型分析[J]. 情报工程, 2024, 10(1): 59-71.



to the development requirements, and has important theoretical significance and practical significance for strengthening the optimal allocation of technological innovation policies and catalysing the real transformation of scientific and technological achievements into real productive forces. [Methods/Processes] Using a three-stage DEA method including stochastic frontier analysis, the article takes four provinces and cities in the Yangtze River Delta region, namely Jiangsu, Shanghai, Zhejiang and Anhui, as research objects. It evaluates the efficiency of technology transfer and achievement transformation from 2015 to 2021, and compares and analyses the efficiency values before and after removing environmental factors and stochastic factors. [Limitations] The study sample is small and does not take into account other factors such as the number of the Science and Technology Mayor's Corps members that may affect technology transfer and achievement transformation. [Results /Conclusions] The efficiency of technology transfer and achievement transformation of the 4 provinces and cities in the Yangtze River Delta is high overall. Among them, service environment, science and technology policy environment, economic development environment and stochastic perturbation are important factors affecting the efficiency of technology transfer and achievement transformation. Through the third stage of DEA analysis, it was found that the mean value of the efficiency of the 4 provinces and cities had decreased. Based on the above research conclusions, countermeasures and suggestions are put forward in terms of rational planning resource inputs for technology transfer and achievement transformation, and improving the mechanism for transferring and transforming scientific and technological achievements.

**Keywords:** Yangtze River Delta; Three-stage DEA; Technology Transfer; Achievement transformation

## 引言

党的二十大报告指出，要不断提高科技成果转化和产业化水平。实现技术转移与成果转化是促进科技与经济结合、实现创新驱动发展的重要方式，是贯彻落实习近平总书记科技创新思想、实施创新驱动发展战略的一项重大举措，科学技术成果只有同经济社会发展需求、市场需求相结合，才能实现其创新价值。科技成果转化通常指的是将科技成果进行开发、应用和推广，进而把科学技术转化为经济效益。技术转移与成果转化主要包含了四个环节，即研发、小试、中试、产业化。然而，技术转移与成果转化所固有的特质也会使得技术流动的过程产生不可避免的问题。例如，相关技术的开发者和需求者存在信息不对称情形，相较于开发者而言，需求者所能获得的信息较少，他

们无法精准判别所获得的技术收益以及收益的真实性。若技术开发者过度披露信息，则在一定程度上会导致技术泄密，进而损害开发者自身利益。二者之间的信息博弈会轻而易举地降低技术转移与成果转化的效率和效果。<sup>[1]</sup>再者，在技术转移和成果转化的过程中，为了有效促进技术的创新研发，会重点关注开发者权益的保护。但这也可能会抑制技术的普及推广，从而出现技术垄断与议价垄断行为，最终导致技术转移与成果转化无法有效开展。

已有研究表明，技术转移与成果转化能够有效提升地区的创新能力和创新绩效，激发地区创新活力<sup>[2-3]</sup>。本土技术转移可以提高我国省际高新技术产业创新效率，并且与国外技术引入之间存在着互补效应<sup>[4]</sup>。长三角等地区是技术转移和成果转化的核心地带，这些地区拥有着各自独特的资源、区位、产业以及政策优势，

在建设国际科创中心的过程中存在差异<sup>[5]</sup>。王俊松和颜燕<sup>[6]</sup>研究发现北京、上海和深圳三地在技术结构和关联度、复杂性以及演化路径等方面存在着显著的异质性，需要据此制定针对性的技术转移与成果转化的发展策略；方力等<sup>[7]</sup>通过对京沪深科技创新情况进行评价，发现京沪深科技创新发展模式存在明显差异。因此，开展技术转移与成果转化的评价研究能够顺应时代发展要求，对于加强技术创新政策优化配置，提升技术转移与成果转化的效率，催化科技成果真正转化为现实的生产力具有重要的理论意义和现实意义。

## 1 研究现状

目前，国内关于技术转移与成果转化的研究主要集中于技术转移方向、技术转移效率评价以及技术转移与成果转化的影响因素等方面。

从技术转移方向来看，首先是一个地区向另一个地区的转移。地区发展水平不同时，其技术转移与成果转化的水平会存在巨大差异。在跨国公司中，管理授权的方式可以提高其技术转移与成果转化的水平，并且也能帮助公司以合适的方式转让相关技术<sup>[8]</sup>。顾高翔和王铮<sup>[9]</sup>研究发现发达国家在降低碳排放技术转移上发挥出了巨大作用，但发展中国家的碳排放能力和研发技术创新能力还有待提升，因此发达国家与发展中国家之间需要进行低碳技术转移，进而达到降低碳排放的目的。其次是行业或者产业之间的技术转移与成果转化。技术发展水平较高的产业或行业通常会向技术发展水平较低的产业或行业进行有效转移。Uusitalo 和 La-

vikka<sup>[10]</sup>以建筑公司为研究样本，基于满足顾客要求的目标，探讨样本中的公司如何参与技术转移，这为其他相关行业的公司挺进建筑市场提供了借鉴意义。孙理军等<sup>[11]</sup>基于中国的中成药制造业的相关数据，分析了中成药制造企业目前的技术转移和技术创新的实际情形，为相关行业的技术转移和成果转化提供了理论借鉴。

关于技术转移效率评价研究，部分学者以地区作为整体来测算技术转移与成果转化的效率，因为若仅以企业或者高校为研究主体，可能会使得技术转移和成果转化效率高于整体的实际效率<sup>[12-13]</sup>。具体地，Osabutey 和 Jin<sup>[14]</sup>以撒哈拉以南非洲的发展中国家为样本，研究发现产业机构、教育成果和较低的企业拥塞性等相关因素能有效促进撒哈拉以南非洲发展中国家的技术转移与成果转化。王七萍和易凌峰<sup>[15]</sup>利用 DEA 分析方法对安徽省各个地级市的工业技术转移效率进行评价，研究发现超效率模型在评价过程中具备更强的科学性；冯华和单丽曼<sup>[16]</sup>利用 Malmquist 指数与 DEA 相结合的方法对中国技术转移效率进行评价，发现整体上中国技术转移效率偏低，并且超过一半的省区属于低纯技术效率、低规模效率情形；肖国华等<sup>[17]</sup>在技术转移效率的评价过程中发现研发存在投入冗余的问题，杜传忠等<sup>[18]</sup>通过研究测算发现，泛长三角地区的技术转移绩效较高，北京和上海等一线龙头城市的技术转移效率最为突出。

从技术转移与成果转化的影响因素来看，宏观和微观层面均能对技术转移与成果转化产生显著影响。在宏观层面，地区经济发展水平、政策支撑力度等能促进技术转移与成果转化<sup>[5]</sup>，而财政转移支付对西部地区的技术转移与成果

转化有着显著的积极效应<sup>[19]</sup>，并且高技术产业的聚集程度及其与市场的联系紧密度也能够对地区技术转移和成果转化效率产生重要影响<sup>[20]</sup>。在微观层面，知识产权在很大程度上也能够对技术转移和成果转化产生重要影响，企业的合作伙伴数量也是能够影响到技术转移与成果转化的重要因素<sup>[21-22]</sup>。

总体来看，技术转移与成果转化是影响一个国家或地区创新能力的核心条件之一，在科技成果充分地发挥社会和经济效用的过程中扮演着重要角色。但目前研究表明中国整体的技术转移与成果转化效率还不够高，地区之间存在较大差异，且有所起伏。在区域技术转移与成果转化的相关分析评价中，鲜有全面考虑地区之间的横向比较以及单一地区的纵向比较研究。基于此，本文以长三角地区的江苏、上海、浙江和安徽四地为研究对象，利用三阶段 DEA 模型对各地区的技术转移与成果转化效率进行评价，以期构建科学的合理的技术转移与成果转化研究框架，并为优化相关政策提供借鉴。

## 2 指标设计与研究方法

### 2.1 评价指标体系构建

技术转移与转化效率通常包含多个投入和产出变量，需要从投入和产出两个方面来选择相应的指标。根据以往研究，投入指标多包含人力投入和资金投入两方面，而产出指标大多可分为专利产出和经济产出。例如，投入类指标通常有科技活动人员、研发成果应用支出经费、R&D 成果应用、科技活动经费、R&D 人员全时当量等<sup>[23-25]</sup>。产出类指标有专利

授权量、技术合同成交额、新产品销售收入、发明专利授权数等<sup>[26-27]</sup>。本文根据技术转移与成果转化可能会涉及的关键因素，结合以往学者研究，将投入部分分为财政资源投入和人力资源投入，产出部分分为创新成果产出与产出转化效益。具体来讲：

**投入变量选择：**根据国家统计局于 2019 年发布的《研究与试验发展（R&D）投入统计规范（试行）》，科技经费在我国主要可以分为三大类，即研究与试验发展（R&D）经费、R&D 成果应用经费和科技服务经费，而财政科技支出是保障基础研究和关键核心技术攻关等资金需求的重要支撑。基于数据的可获得性，本文将研究与试验发展（R&D）经费和财政科技支出作为财政资源投入变量。由于研究与实验发展人员以及科技活动人员在创新过程中发挥着重要作用，在技术转移与成果转化过程中不容忽视，因此，本文将其作为人力资源投入变量。

**产出变量选择：**专利授权数、有效发明专利数以及有效发明专利密度从专利层面来展现出科技成果的多少<sup>[16, 28]</sup>，论文的产出从学术上体现了科研创新能力的强弱<sup>[29]</sup>，科技成果奖则从获奖层面来对技术转移与成果转化的效果进行度量。综合上述考虑，本文将有效发明专利密度、科技论文数量、专利授权数和省部级科技成果奖数量作为创新成果产出变量。不仅如此，技术合同成交额和项目数量是反映科技成果转化的重要指标<sup>[30]</sup>，而高新技术企业有着知识密集、技术密集性特征，能持续进行研究开发与技术成果转化。因此，产出转化效益部分包含技术合同成交额、技术合同成交项数、技

术合同成交额占当地 GDP 比值以及高新技术产业企业数等相关变量。

环境变量选择：本文将可能影响技术转移与成果转化的环境因素分为服务环境、科技政策环境以及经济发展环境三个方面，服务环境是指对技术转移与成果转化提供相关服务的因素，科技政策环境是能够对技术转移与成果转

化提供政策支撑的因素，而经济发展环境和地区的经济发展程度密切相关。结合以往学者的研究以及相关数据的可获得性，选取了省级以上科技企业孵化器数量、科研机构数、技术转移与成果转化法规和激励政策办法出台数量以及当地 GDP 总量等作为环境变量。

相关变量的具体指标定义及衡量方式见表 1。

表 1 技术转移与成果转化效率评价的指标说明

类别	维度	具体指标	计算方法
技术转移与成果转化资源投入	财政资源投入	I1: 研究与试验发展经费支出占该地区国内生产总值 (GDP) 比重	研究与试验发展经费支出 / 该地区国内生产总值 (%)
		I2: 财政科技支出占财政支出比重	科技支出 / 财政总支出 (%)
	人力资源投入	I3: 研究与实验发展人员数量	研究与实验发展人员累计数量 (人)
		I4: 从事科技活动人数	从事科技活动累计人数 (人)
技术转移与成果转化绩效产出	创新成果产出	O1: 有效发明专利密度	每万人所拥有的各地区每年有效发明专利累计数 (件)
		O2: 科技论文数量	中国科技人员在国内外发表论文累计数量 (篇)
		O3: 专利授权数	专利授权累计数 (件)
		O4: 省部级科技成果奖数量	省部级科技成果奖累计数 (个)
	产出转化效益	O5: 技术合同成交额	技术合同成交额 (亿元)
		O6: 技术合同成交项数	技术合同成交项数 (项)
		O7: 技术合同成交额占比	该年份该地区技术合同成交金额占当地 GDP 比值 (%)
		O8: 高新技术产业企业数	高新技术产业企业累计数 (个)
技术转移与成果转化环境因素	服务环境	E1: 省级以上科技企业孵化器数量	省级以上科技企业孵化器累计数 (个)
		E2: 科研机构数	科研机构累计数 (个)
	科技政策环境	E3: 技术转移与成果转化法规和激励政策办法出台数量	技术转移与成果转化法规和激励政策办法出台累计数 (个)
		经济发展环境	E4: 当地 GDP

## 2.2 样本选择与数据来源

本文以长三角地区的江苏省、上海市、浙江省和安徽省四个地区作为研究对象，基于四个省 2015—2021 年的投入产出数据及其所在地的环境变量数据，分析长三角地区的技术转移与成果转化效率以及动态发展趋势。由于中国的各类年鉴是对上一年度数据的总结，因此本

文研究数据主要来源于 2016—2022 年的《中国统计年鉴》《中国火炬统计年鉴》以及《江苏统计年鉴》等相关城市的统计年鉴，部分数据如技术转移与成果转化法规和激励政策办法出台数量，则通过查询地方政府官网经手工整理计算得出，而其他少数缺失数据则通过插值法计算获得。



## 2.3 研究方法

DEA最初是由Charnes等<sup>[31]</sup>提出的衡量决策单元投入产出效率的一种计算效率值的非参数模型,在经过后续学者的改进和完善后发展了多种类型的DEA模型。传统DEA模型中,最终得出的效率值可能会被内部管理无效率、外部环境和随机误差项三种因素所干扰,而三阶段DEA则能有效避免外部环境和随机因素的干扰<sup>[32]</sup>。因此本文引入三阶段DEA模型来消除外部环境与随机误差项对效率评价单元的影响,以便更加客观、准确地评估各决策单元效率。三阶段DEA模型的基本原理主要包含三个阶段:

第一阶段:运用传统的DEA模型测度决策单元的效率值,并求得投入和产出的松弛值,对于本项目而言,将采用投入导向下对偶形式的DEA-BCC模型:

$$\begin{aligned} & \min \theta \\ & \text{s.t.} \begin{cases} \sum_{i=1}^n \lambda_i x_{ij} + s^- = \theta x_{0j} \\ \sum_{i=1}^n \lambda_i y_{ir} - s^+ = y_{0r} \\ \sum_{i=1}^n \lambda_i = 1 \\ \lambda_i \geq 0; s^+ \geq 0; s^- \geq 0 \end{cases} \quad (1) \end{aligned}$$

其中,  $i=1, 2, \dots, n$ ;  $j=1, 2, \dots, m$ ;  $r=1, 2, \dots, s$ ;  $n$ 为决策单元的个数,  $m$ 和  $s$ 分别为投入与产出变量的个数,  $x_{ij}$ 为第  $i$ 个决策单元的第  $j$ 个投入要素,  $y_{ir}$ 为第  $i$ 个决策单元的第  $r$ 个产出要素,  $\theta$ 为决策单元的有效值。BCC模型计算出来的效率值为综合技术效率值(TE),进一步分解为规模效率(SE)和纯技术效率值(PTE)的乘积,即  $TE = SE \times PTE$ 。

第二阶段:将第一阶段得到的4个投入变量的松弛变量作为被解释变量,将标准化之后的环境变量作为解释变量,构建SFA模型,分析影响技术转移与成果转化效率的环境因素。SFA回归可以将所有决策单元调整到相同的外部环境中,能够避免环境因素和随机因素对效率测度的影响,得到的效率值能够直接反映管理水平的高低。

第三阶段:调整后的DEA模型。将调整后的投入数据代替原始投入数据,再次运用DEA-BCC模型进行技术转移与成果转化的效率评估,最后得到剔除环境因素变量以及随机误差项之后各个地区的技术转移与成果转化效率值。

## 3 实证分析

### 3.1 第一阶段DEA分析

采用DEA2.1软件测量出第一阶段长三角四个地区的技术转移与成果转化效率,如表2所示。从时间趋势上看,江苏的技术转移与成果转化综合效率整体上呈现上升趋势,尽管在2016年开始下降,但第二年开始上升,且2019—2021年的综合效率值均为1;上海在样本期间内,除了2020年的综合效率值为0.952以外,其他年份均为1,整体来看上海的技术转移与成果转化效率较高;浙江的综合效率值在2019年开始下降,随后开始增加,到2021年效率值为1;安徽的综合效率值在2016年开始下降,随后几年逐年递增。从投入冗余的角度来看,浙江和上海的投入冗余较少,存在明显的稀缺性,相应地增加资源供给和优化资源

表 2 第一阶段 DEA 分析结果

省份	年份	技术效率	规模效率	综合效率	规模报酬	投入冗余			
						I1	I2	I3	I4
江苏	2015	0.983	0.97	0.954	irs	0	0	42.459	24.925
	2016	0.936	0.82	0.767	irs	0	0	0	16.972
	2017	0.974	0.957	0.932	irs	0	0	13230.09	19.992
	2018	0.99	0.962	0.952	irs	0	0	0	3.623
	2019	1	1	1	-	0	0	0	0
	2020	1	1	1	-	0	0	0	0
	2021	1	1	1	-	0	0	0	0
上海	2015	1	1	1	-	0	0	0	0
	2016	1	1	1	-	0	0	0	0
	2017	1	1	1	-	0	0	0	0
	2018	1	1	1	-	0	0	0	0
	2019	1	1	1	-	0	0	0	0
	2020	0.975	0.977	0.952	irs	0.056	0.004	0	0
	2021	1	1	1	-	0	0	0	0
浙江	2015	1	1	1	-	0	0	0	0
	2016	1	1	1	-	0	0	0	0
	2017	1	1	1	-	0	0	0	0
	2018	1	1	1	-	0	0	0	0
	2019	0.952	0.995	0.947	irs	0	0.003	4435.925	4.286
	2020	0.995	0.989	0.984	drs	0	0	7264.662	0
	2021	1	1	1	-	0	0	0	0
安徽	2015	1	0.756	0.756	irs	0	0	18218.71	1.483
	2016	1	0.694	0.694	irs	0	0.005	16512.23	0
	2017	0.991	0.785	0.778	irs	0	0.012	18466.15	0
	2018	1	0.879	0.879	irs	0	0.019	33872.69	0
	2019	1	1	1	-	0	0	0	0
	2020	1	1	1	-	0	0	0	0
	2021	1	1	1	-	0	0	0	0

注：“irs”“-”“drs”分别表示规模报酬递增、不变和递减，下同。

配置是提高地区技术转移与成果转化效率的关键所在。江苏和安徽的投入冗余较为明显，说明两地在财政资源配置和人力资源管理等方面存在较大的改进空间。整体而言，在 2015—2021 年的 7 年间，上海的技术转移与成果转化效率均值最高，为 0.99314，其效率值在 0.95~1 之间浮动；浙江位居第二，技术转移与成果转化效率均值为 0.99013，在 2019 年处于规模报

酬递增，2020 年规模报酬递减；江苏位于第三，效率均值为 0.94357；安徽的技术转移与成果转化效率均值最低，为 0.87243，低于 0.9。

由于第一阶段 DEA 分析存在外部环境因素和随机扰动因素对技术转移与成果转化效率的影响，当前的效率值无法反映出真实的地区技术转移与成果转化效率情况，接下来将进行 SFA 回归分析。

### 3.2 第二阶段SFA实证分析

第二阶段的 SFA 回归分析中，将第一阶段的投入变量（松弛值）作为被解释变量，将服务环境、科技政策环境和经济发展环境等四个环境变量作为解释变量，采用 Frontier 4.1 软件进行回归分析，第二阶段的 SFA 回归结果如表 3 所示。从表 3 可以看出，研究与实验发展人

员数量（I3）和从事科技活动人数（I4）投入冗余变量的  $\gamma$  在 1% 的显著性水平下显著，且 LR 的单边似然比检验在 1% 水平下显著，表明四种环境因素对 I3 和 I4 两种投入松弛变量均有显著的影响，这也进一步说明第二阶段的 SFA 回归具有合理性和必要性。然而，财政资源投入的两个投入冗余变量的  $\gamma$  值不存在显著性，且四种环境变量的系数均未达到显著性要求。

表 3 SFA 回归结果

变量	I1	I2	I3	I4
常数	-0.002 (1.000)	0.0048 (1.000)	11349.2680*** (947.1459)	-8.7597*** (3.1735)
省级以上科技企业孵化器数量 (E1)	-0.021 (1.000)	-0.0010 (1.000)	3193.1674*** (178.0207)	7.7984*** (11.5859)
科研机构数 (E2)	0.013 (1.000)	-0.0029 (1.000)	-7520.2069*** (515.5729)	21.5993*** (6.6277)
技术转移与成果转化法规和激励政策办法出台数量 (E3)	0.003 (1.000)	0.00821 (1.000)	15709.3720*** (118.7589)	2.7254*** (4.8623)
当地 GDP (E4)	0.016 (1.000)	-0.0069 (1.000)	-18402.8520*** (319.0943)	12.6636*** (9.3496)
$\delta^2$	0.0001	0.00001	55487594.00***	301.6593**
$\gamma$	0.05	0.05000	0.1828***	0.9603***
LR	89.128774	116.682	1.61425***	7.2075***

注：括号内为标准误；\*\*\* 表示在 1% 置信水平下显著。

从四种环境变量的系数可看出，服务环境中的省级以上科技企业孵化器数量对 I3 和 I4 的投入松弛变量回归系数均为正，科研机构数对 I3 存在负向作用，对 I4 呈正向促进作用；技术转移与成果转化法规和激励政策办法出台数量对两种投入松弛变量回归系数均为正；当地 GDP 总额与 I3 存在负相关关系，与 I4 存在正向关系。具体分析如下所示：

(1) 服务环境。服务环境中的省级以上科技企业孵化器数量对 I3 和 I4 的回归系数分

别为 3193.1674、7.7984，在 1% 水平上显著，这表明科技企业孵化器能够促进地区技术转移与成果转化效率水平的提升。科技企业孵化器能够为科技型企业提供相关的服务与支持，可以在一定程度上降低创业者的创业风险以及创业所带来的巨额成本，从而提高创业的成功率，来有效促进技术转移与成果转化。服务环境中的科研机构数对 I3 和 I4 的回归系数分别为 -7520.2069 和 21.5993，均在 1% 的显著性水平上显著，表明科研机构数增多会减少研究与

实验发展人员变量的投入冗余,有利于提高地区技术转移与成果转化的效率,但是会增加科技活动人员投入冗余,这不利于提高技术转移与成果转化的效率。从理论上来说,科研机构是地区科技创新成果重要供给源,能够为地区的产业结构升级提供重要技术支撑。随着研究的不断深入,一方面加强了各研究机构之间的协同合作,促使相关的技术和科研成果转化为生产力,进而提高技术转移与成果转化的效率;另一方面,科研机构需要大量的高素质高技能人才,但是由于教育运行机制和人才培养模式的固有特点,培养出的人才在掌握的技能上可能会相似,进而会增加相关科技活动人员的冗余,抑制地区技术转移与成果转化效率的提升。

(2) 科技政策环境。科技政策环境中,技术转移与成果转化法规和激励政策办法出台数量对 I3 和 I4 的回归系数分别为 15709.3720 和 2.7254,均在 1% 的水平上显著,说明技术转移与成果转化法规和激励政策办法出台的数量增多,会加剧研究与实验发展人员数量和科技活动人员的投入冗余现象。尽管技术转移与成果转化的相关法规和激励办法数量增多,可能会激励相关人员和机构投身于技术转移和成果转化的工作中来,但是由此也会造成相关的工作人员过多的情形,在一定程度上造成人员投入的冗余现象,从而不利于地区的技术转移与成果转化效率的提高。

(3) 经济发展环境。该环境变量主要为各地区的 GDP 总量,其与 I3 和 I4 的回归系数分别为 -18402.8520 和 12.6636,在 1% 的显著性水平上显著,说明经济发展水平会减少研究与实验发展人员投入冗余,增加科技活动人员投入

冗余。各地的经济发展水平较高时,能够为研究与试验人员提供充足经费,这有利于技术转移与成果转化效率提升。但是地区经济发展水平越高,对人才的集聚吸纳能力也会越强,可以间接地吸引更多的科技创新活动人员的投入,过多的科技活动人员投入则会造成投入冗余现象,反而会抑制技术转移与成果转化效率提升。

### 3.3 第三阶段调整后的DEA模型分析

在第二阶段利用 SFA 剔除了外部环境因素及相关的随机干扰后,将调整后的投入变量值替换原始投入数据,保持产出数据不变,再次将调整后的投入数据与原始的产出数据导入 DEAP 2.1 软件,并且利用 BCC 模型进行效率评估,最终得出了排除环境因素和随机误差的技术转移与成果转化综合效率、纯技术效率和规模效率,相关结果见表 4。从表 4 可看出,在消除环境因素和随机扰动的影响之后,调整后的规模效率、技术效率以及综合效率均发生了变化,不同地区的效率值变化不同。2015-2021 年间,江苏省调整后的技术效率和综合效率的均值有所下降,而规模效率的均值有所上升;上海市的技术效率均值存在上升情形,而规模效率和综合效率均值有所下降;浙江省和安徽省三个效率的值存在统一上升或下降的趋势,例如,浙江省三种效率均值在调整后均有所上升,而安徽省三种效率均值在调整后均有所下降。这也间接表明各地区的技术转移与成果转化效率受服务环境、科技政策和经济发展环境以及随机误差的影响较大,样本中的各地区在上述三种环境因素以及随机误差的干预下存在着效率虚高/偏低的现象。



表 4 第三阶段技术转移与成果转化效率分析

省份	年份	技术效率	规模效率	综合效率	规模报酬	投入冗余			
						I1	I2	I3	I4
江苏	2015	0.974	0.974	0.949	irs	0	0	0	1.302
	2016	0.922	0.83	0.765	irs	0	0	0	3.497
	2017	0.975	0.946	0.922	irs	0	0	3456.618	0.709
	2018	0.988	0.962	0.951	irs	0	0	0	0
	2019	1	1	1	-	0	0	0	0
	2020	1	1	1	-	0	0	0	0
	2021	1	1	1	-	0	0	0	0
上海	2015	1	1	1	-	0	0	0	0
	2016	1	1	1	-	0	0	0	0
	2017	1	0.997	0.997	irs	0.173	0.012	0	0
	2018	1	0.997	0.997	irs	0	0.032	2585.559	0
	2019	1	1	1	-	0	0	0	0
	2020	0.989	0.965	0.954	irs	0	0.001	67.968	0
	2021	1	1	1	-	0	0	0	0
浙江	2015	1	1	1	-	0	0	0	0
	2016	1	1	1	-	0	0	0	0
	2017	1	1	1	-	0	0	0	0
	2018	1	1	1	-	0	0	0	0
	2019	0.956	0.992	0.948	irs	0	0	0	0.008
	2020	0.996	0.994	0.99	drs	0	0	0	0
	2021	1	1	1	-	0	0	0	0
安徽	2015	1	0.703	0.703	irs	0	0	9003.07	3.579
	2016	0.994	0.638	0.634	irs	0	0.002	7360.939	0
	2017	0.986	0.713	0.703	irs	0	0.001	7277.494	0
	2018	1	0.77	0.77	irs	0	0	9786.305	0
	2019	1	1	1	-	0	0	0	0
	2020	1	1	1	-	0	0	0	0
	2021	1	1	1	-	0	0	0	0

调整后的四个典型区域的技术转移与成果转化综合效率变化趋势如图 1 和图 2 所示。在调整各个投入变量之后，各个地区的技术转移与成果转化效率值的大小有着明显变化，但整体来看，各地区的变化趋势差异不显著。例如，江苏省技术转移与成果转化综合效率

整体趋势变化不大，均在 2019 年后上升至 1，调整前，2016 年的综合效率值大于 0.7，而调整后却小于 0.7。这充分说明这个地区在进行技术转移和成果转化时需要兼顾可能存在的外部影响因素，尽量减少外部因素所带来的不利影响。

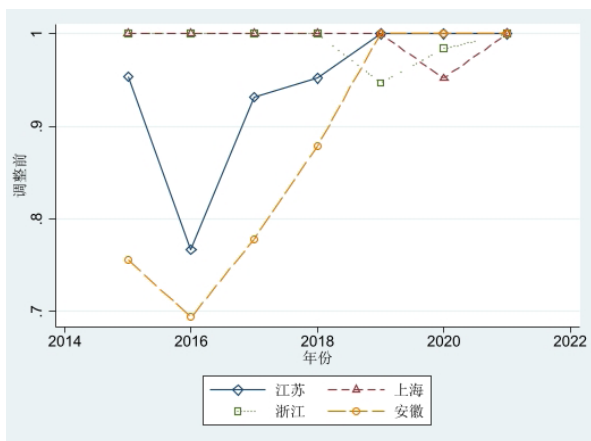


图1 调整前地区技术转移与成果转化综合效率变化趋势

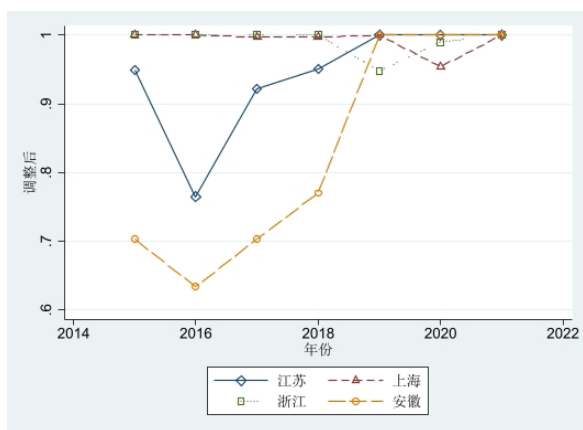


图2 调整后地区技术转移与成果转化综合效率变化趋势

## 4 研究结论及对策建议

### 4.1 研究结论

本文利用三阶段 DEA 模型对 2015—2021 年长三角地区四个省市的技术转移与成果转化效率进行测度，比较了不同省市之间的技术转移与成果转化效率差异，并且分析了科技企业孵化器数量、科研机构数量、技术转移与成果转化法规和激励政策办法出台数量以及当地 GDP 总量对技术转移与成果转化效率的影响，结果发现：第一阶段 DEA 分析表明长三角四个省市的技术转移与成果转化效率整体较高，其中服务环境、科技

政策环境、经济发展环境和随机扰动是影响技术转移与成果转化效率的重要因素。第二阶段 SFA 检验结果显示各环境因素仅对人力资源投入的两个投入松弛变量具有显著影响，其中，省级以上科技企业孵化器数量、技术转移与成果转化法规和激励政策办法出台数量系数显著为正，说明这两个环境因素抑制了技术转移与成果转化效率的提升，可能的原因在于科技企业孵化器数量已达到市场上的最佳供应水平，相关激励政策办法可能会导致科研人员投入过多的情形，这样会使得技术转移与成果转化效率并无明显增长空间。科研机构数和 GDP 总产值两个环境变量有利于降低研究与实验发展人员投入冗余，但会增加科技活动人员投入冗余。第三阶段在提出上述的三种环境因素之后，再次进行 DEA 分析，发现长三角地区的四个省市的技术转移与成果转化效率均值明显降低，但增减趋势大体一致，主要是由于长三角各省市若仅依靠自身的技术转移与成果转化的服务环境、相关激励政策办法以及地区经济发展水平，则无法有效提高技术转移与成果转化的效率，因此需要不断提高技术转移与成果转化的管理质量和重视程度，从而保持领先优势。

### 4.2 对策建议

(1) 合理规划资源投入。在研发经费和财政资源上，改进研发经费审查管理机制，以避免研发投入在投入前和投入中的浪费；探索新的评价体系，以评估政府在技术创新、技术转移中的支持力度及发展程度。除了科研经费投入等硬性指标，还可以考虑采用效率指标来激励政府更好地分配资源，并避免资源浪费和政府懒政行为。同时，可采取税收优惠等措施来

提高研发效率和财政资源的利用效率,从而更好地满足市场需求和创造商业机会。关于研发及科技活动人员,在引进人才时,应科学合理地确定数量,避免盲目扩张,科学评估人才缺口,以便为关键领域提供支持。同时,还应当分析现有研发人员的数量和分布状况,以判断是否存在过度密集的研究领域。若存在上述情形,应当建立人才引导机制,以便引导人才向周边地区流动,以避免人才浪费。

(2) 健全高校科技成果转化机制。针对当前长三角地区各省市目前技术转移与科技成果转化存在的问题,可适当完善高校、科研院所等分类考核评价机制,把科技成果转化、产学研协同创新作为重要考核指标。引导高校院所建立技术转移机构,建立职务科技成果披露制度,完善技术转移服务人员的职称评定、收入分配等制度。探索科技经纪人试点工作,组织各地高校、科研院所等研究机构面向全国选派科技经纪人,促进地区的技术转移与科技成果的产业化。

(3) 激发技术转移与成果转化的积极性。加强绩效评估、质量考核和第三方监督等手段,以激发各机构的研发积极性,提升科技成果转化能力,增加企业的创新收益。保障生产商和经营者的合法权益,促进技术创新整合和再造,推动产业转型升级,促进产业创新进步。鼓励各地企业打破单向合作方式,建立资金支持系统,采取市场化和协同化方式提高技术转移和成果转化的竞争优势。通过提高政策透明度、减轻税费负担、完善市场服务等有效手段培育新兴产业,促进区域产业转型升级和经济稳定增长,推动区域全面发展,形成可持续发展的

技术产业生态链。

### 4.3 研究局限及展望

本文仅以上海、江苏、浙江和安徽为样本进行研究,研究的个体较少,在未来关于技术转移与成果转化效率评价的研究中,可以长三角地区城市群、珠三角地区城市群等样本,扩充样本量,以便横向或者纵向对比各城市(群)之间的技术转移与成果转化效率差异。此外,由于数据的可得性等原因,可能影响技术转移与成果转化效率的其他指标因素我们并未考虑。譬如,各地区的技术经理人数量、科技镇长团成员数、科技招商以及省科技创新券服务企业数等,在后续技术转移与成果转化效率的研究中,可以考虑上述指标。

## 参考文献

- [1] 余辉,马超,张羽帆. 技术转移相关概念辨析与研究现状 [J]. 中国高校科技, 2022(10): 84-91.
- [2] JONATHAN E, KORTUM S. Trade in ideas Patenting and productivity in the OECD[J]. Journal of international Economics, 1996, 40(3-4): 251-278.
- [3] 侯健敏,党兴华. 研发合作及技术转移影响区域创新能力路径研究 [J]. 科学学与科学技术管理, 2010, 31(9): 56-61.
- [4] 冯宗宪,丁梦. 本土技术转移对省际高新技术产业创新效率影响的实证分析 [J]. 统计与决策, 2018, 34(22): 108-111.
- [5] 刘承良,管明明,段德忠. 中国城际技术转移网络的空间格局及影响因素 [J]. 地理学报, 2018, 73(8): 1462-1477.
- [6] 王俊松,颜燕. 复杂度、关联度与城市技术演化路径——基于北京、上海、深圳的对比分析 [J]. 地理科学进展, 2022, 41(4): 554-566.
- [7] 方力,张士运,王健. 京沪深科技创新综合效应比较评价研究 [J]. 世界科技研究与发展, 2020, 42(2): 192-205.

- [8] CHEN H C. Entry mode, technology transfer and management delegation of FDI[J]. *International Review of Economics & Finance*, 2018(54): 232-243.
- [9] 顾高翔, 王铮. 《巴黎协定》背景下国际低碳技术转移的碳减排研究[J]. *中国软科学*, 2018(12): 8-16.
- [10] UUSITALO P, LAVIKKA R. Technology transfer in the construction industry[J]. *The Journal of Technology Transfer*, 2021, 46(5): 1291-1320.
- [11] 孙理军, 张曼婷, 周国华. 战略性新兴产业技术转移、内部研发能力与技术创新绩效关系研究——以中成药制造业为例[J]. *科技管理研究*, 2020, 40(5): 68-76.
- [12] 吕荣杰, 贾芸菲, 张义明. 我国省份技术转移效率评价——基于高校、企业比较的视角[J]. *科技管理研究*, 2018, 38(12): 86-91.
- [13] 张凤, 汤奥灵, 沈映春, 等. 基于 DEA 模型的京沪深三地技术转移效率研究[J/OL]. *北京航空航天大学学报(社会科学版)*: 1-10[2023-03-17]. DOI: 10.13766/j.bhsk.1008-2204.2022.0987.
- [14] OSABUTEY E L C, JIN Z. Factors influencing technology and knowledge transfer: Configurational recipes for Sub-Saharan Africa[J]. *Journal of Business Research*, 2016, 69(11): 5390-5395.
- [15] 王七萍, 易凌峰. 安徽省工业技术转移效率评价研究[J]. *华东经济管理*, 2014, 28(10): 29-31, 52.
- [16] 冯华, 单丽曼. 中国技术转移效率评价研究——基于 Malmquist 指数和 Bootstrap-DEA 的实证分析[J]. *学习与实践*, 2016(11): 14-22.
- [17] 肖国华, 杨云秀, 王江琦. 四螺旋参与度对技术转移及其效率的影响研究[J]. *科技进步与对策*, 2016, 33(4): 7-11.
- [18] 杜传忠, 冯晶, 张咪. 中国三大经济圈技术转移绩效评价研究[J]. *财经问题研究*, 2017(7): 95-101.
- [19] 毛军, 李子豪, 石信秋. 财政转移支付对区域技术转移的影响研究[J]. *科研管理*, 2021, 42(1): 88-99.
- [20] 胡欣悦, 任紫娟, 汤勇力. 我国重点高校技术转移效率变化的影响因素研究——基于面板随机前沿分析方法[J]. *技术经济*, 2020, 39(7): 200-208.
- [21] HAYASHI D. Knowledge flow in low-carbon technology transfer: A case of India's wind power industry[J]. *Energy Policy*, 2018(123): 104-116.
- [22] RAVILIOUS G E, CHOUDHRY V, HOWIESON S V, et al. An Analysis of Factors Affecting Federal Laboratory Technology Transfer Transactional Efficiency[J]. *Research-Technology Management*, 2021, 64(3): 20-30.
- [23] 彭峰, 周银珍, 李燕萍. 中国高技术行业的技术转移与效率差异[J]. *求索*, 2016(3): 92-96.
- [24] 高擎, 何枫, 吕泉. 区域环境、科研投入要素与我国重点高校技术转移效率研究[J]. *中国高教研究*, 2020(1): 78-82, 108.
- [25] 钟卫, 陈宝明. 中国高校科技成果转化绩效评价研究[J]. *中国科技论坛*, 2018, 264(4): 41-49.
- [26] 杨剑, 夏慧良. 中部与长三角地区国家级高新区科技创新效率评价——基于三阶段数据包络分析[J]. *科技管理研究*, 2022, 42(16): 70-77.
- [27] 肖兴志, 徐信龙. 区域创新要素的配置和结构失衡: 研究进展、分析框架与优化策略[J]. *科研管理*, 2019, 40(10): 1-13.
- [28] 王方, 李华. 基于 DEA 的中国区域技术转移效率评价[J]. *科研管理*, 2013, 34(S1): 153-160.
- [29] 许晓冬, 秦续天. 高技术产业绿色生态创新效率评价与影响因素研究——基于三阶段 DEA 模型[J]. *现代管理科学*, 2022(5): 50-58.
- [30] 周俊亭, 席彦群, 周媛媛. 区域技术市场、政府扶持与科技创新[J]. *中国软科学*, 2021, 371(11): 80-90.
- [31] CHARNES A, COOPER W W, RHODES E. Measuring the efficiency of decision-making units[J]. *European Journal of Operational Research*, 1979, 3(4): 339-339.
- [32] FRIED H O, LOVELL C, SCHMIDT S S, et al. Accounting for environmental effects and statistical noise in data envelopment analysis[J]. *Journal of Productivity Analysis*, 2002, 17(1/2): 157-174.





开放科学  
(资源服务)  
标识码  
(OSID)

# 计划行为理论视角下微信用户选择性信息规避行为影响因素研究

李廷翰

广州航海学院图书馆 广州 510725

**摘要:** [目的/意义] 探索微信用户选择性信息规避行为的影响因素, 为社交媒体平台设置优化、用户自我调节及成瘾预防法规的完善提供实践启示。[方法/过程] 基于计划行为理论, 分析与归纳微信用户选择性信息规避行为特征, 利用 PLS-SEM 结构方程模型方法对其影响路径进行实证检验。[局限] 未考虑信息类型、推送频率、知识差异及其他外生激励因素的影响。[结果/结论] 用户卷入度显著正向影响社会网络联结与心理所有权; 社会网络联结显著正向影响心理所有权与信息规避规范; 信息规避规范、心理所有权、自我控制能力分别显著正向影响信息规避意向; 信息规避意向显著正向影响信息规避行为。

**关键词:** 信息规避; 卷入度; 社会网络联结; 心理所有权; 计划行为理论

**中图分类号:** G252.0; G35

## Influencing Factors of Wechat Users' Selective Information Avoidance Behavior from the Perspective of Planned Behavior Theory

LI Tinghan

Library, Guangzhou Maritime University, Guangzhou 510725, China

**Abstract:** [Purpose/Significance] Exploring the influencing factors of WeChat users' selective information avoidance behavior, providing practical insights for optimizing social media platform settings, improving user self-regulation, and improving addiction prevention regulations. [Methods/Processes] Based on the theory of planned behavior, analyze and summarize the characteristics of WeChat users' selective information avoidance behavior, and through the PLS-SEM structural equation modeling method to empirically test its impact path. [Limitations] The impact of information type, push frequency, knowledge differences, and other exogenous motivational factors were not considered. [Results/Conclusions] Users' involvement significantly positively affects social network bond and psychological ownership; Social network bond significantly positively affects psychological ownership and information avoidance norms; Information avoidance norms, psychological ownership, and self-control ability significantly positively affect information avoidance intention; Information avoidance intention has a significant positive impact

**基金项目** 广东省教育科学“十三五”规划 2020 年度研究项目“心理契约视角下粤港澳大湾区科研人员知识转移影响因素及促进机制研究”(2020GXJK171)。

**作者简介** 李廷翰(1983-), 博士, 馆员, 主要研究方向为信息行为、知识管理, E-mail: li\_tinghan@163.com。

**引用格式** 李廷翰. 计划行为理论视角下微信用户选择性信息规避行为影响因素研究[J]. 情报工程, 2024, 10(1): 72-84.

on information avoidance behavior.

**Keywords:** Information Avoidance; Involvement; Social Network Bond; Psychological Ownership; Planned Behavior Theory

## 引言

据统计,全球社交媒体用户数量已经突破47亿人<sup>[1]</sup>,随着用户好友增加、使用时长增长及虚拟社交场景的丰富,出现了社交媒体用户的消极使用行为,其中包括:社交媒体倦怠、中辍<sup>[2]</sup>及不持续使用<sup>[3]</sup>等。在真实工作与生活情景中,因好友信息交互、支付及防疫等需要,用户不会因社交媒体存在一些缺陷而彻底停用,但却可能略看甚至规避一些不相关的信息<sup>[4]</sup>。因此有研究将信息规避归入社交媒体消极使用行为,会产生负面效应,如:减弱与好友的情感维系,影响社交媒体广告收益<sup>[5]</sup>等。但其也具有积极作用,如:有利于用户进行信息筛选,减少非必要社交互动,降低信息过载及社交过载<sup>[6]</sup>,也有利于社交媒体平台中重要信息与知识的传播。因此迫切需要探索社交媒体用户选择性信息规避意向及行为的影响因素。本文选择使用率较高的社交媒体平台-微信app,基于计划行为理论框架,通过量化研究方法,验证微信用户信息规避行为影响因素,既能从感知、信念及能效视角扩充用户信息规避意向的归因,又能为优化平台交互设置,细化社交媒体成瘾的预防法规提供实践启示。

## 1 文献回顾

### 1.1 信息规避行为

信息规避是指个体有意阻止或延迟获取可

利用的但不需要的信息的行为<sup>[7]</sup>,其主要表现为:个人针对信息的主动躲避、疏忽、偏颇理解、遗忘及自我设障等。已有研究关注不同种类信息的规避行为:一是公共风险信息,如近年在Covid-19病毒的全球传播环境下<sup>[8]</sup>,公众对疫情新闻<sup>[9]</sup>、国外民众对恐怖主义信息的规避<sup>[10]</sup>等;二是个人健康信息,例如针对肿瘤、儿童自闭症等各类疾病的风险及筛查信息<sup>[11]</sup>,以及食品中的卡路里含量<sup>[12]</sup>、抗生素耐药性<sup>[13]</sup>及电子烟危害等日常消费品健康信息<sup>[14]</sup>的规避;三是机构或平台提供的服务信息,如针对乡村公共文化服务平台<sup>[15]</sup>朋友圈、网络医疗众筹信息的规避<sup>[16]</sup>。此外还有针对特定人群,如中老年群体健康信息规避<sup>[17]</sup>的研究等。

用户信息规避行为的影响因素主要可分为:

①用户的感知因素。研究证实个体对信息处理能力不足引致的信息过载感知会正向影响信息规避意愿或行为<sup>[18]</sup>;除直接作用外,信息过载感知也能通过个体心理不适情绪及健康威胁感知等变量间接作用于信息规避行为的产生<sup>[8]</sup>;与之相似,信息处理的自我效能感、对肿瘤的宿命感知(如:是否可以通过个人努力来阻止患上肿瘤的信念)都可与个体采取信息搜寻或规避行为相关<sup>[11]</sup>。②用户内心的情绪状况。焦虑被认为是重要因素之一,但其对信息规避的影响作用存在差异,如美国用户对患肿瘤疾病的焦虑与肿瘤信息规避负相关,中国用户反而会因患病焦虑产生逃避接收相关疾病信息的行为<sup>[19]</sup>;Siebenhaar等<sup>[20]</sup>证明悲痛情绪会逐渐增

加人们对疫情信息的规避意愿，从而降低防疫措施的遵守程度；而外部环境与个人信念因素则通过导致用户产生焦虑、悲伤及认知失调感，正向影响信息规避的发生<sup>[21]</sup>；③用户间存在的个体特征差异，如社交范围较窄的用户较容易受信息过载影响，增加疾病信息规避机会<sup>[22]</sup>；而疾病检测结果、家族病史与用户对糖尿病和结肠肿瘤信息的规避态度密切相关<sup>[23]</sup>；在没有患肿瘤疾病的人群中，受教育水平、信息来源信任感皆显著反向影响肿瘤疾病的信息规避意愿<sup>[24]</sup>。

## 1.2 社交媒体信息规避

在社交媒体领域，信息规避被认为是应对过载的一种应对方式，也是用户因时间、知识或兴趣缺乏而产生的忽略或规避好友信息的行为<sup>[7]</sup>。目前针对社交媒体场景下的信息规避归因探索包括：Dai等<sup>[25]</sup>认为感知信息过载及其引起的疲劳、挫败及不满等情绪反应，是用户信息规避意向产生的归因；Guo等<sup>[7]</sup>则通过压力源—负担—结果（SSO）理论框架，证实感知信息过载、信息不相关、社交过载，会导致用户产生社交网络疲劳，最终产生信息规避行为结果，此外还考虑了时间压力的调节效应及用户好友数量的控制变量作用等。

已有研究提供了相关理论基础及实践尝试，但仍存在有待探索的领域：一是讨论了信息过载及社交过载带来的负面影响，仍亟待证明主动“选择性”规避行为的正向能效因素；二是与社交媒体不持续使用研究相似，验证了个体焦虑、疲劳或认知失调等消极情绪对信息规避的影响作用，但有必要关注可能存在的积极情

绪感知效应；三是证实了个体信念与认知层面的应激反应，仍需借鉴社会资本、消费者行为等相关理论，扩充社会群体因素产生的影响。

## 2 模型构建及假设提出

我们认为社交媒体用户选择性信息规避是指：社交媒体用户因时间、能力、兴趣或感知价值不足，对平台信息有选择性地忽略、逃避、遗忘或延迟接收的行为。值得注意的是，用户信息规避的对象除好友、群及空间信息外，还应包括平台或第三方推送的公众号、小程序、广告或其他信息等。

### 2.1 计划行为理论

作为社会认知理论的一部分，计划行为理论认为，行为意向是最终开展行为的最佳预测变量，当预测到行为结果有利（态度），感知到社会压力的推动（主观规范），以及感觉到有能力去完成该行为（知觉行为控制）时，人们就会产生某种特定行为的强烈意向<sup>[26]</sup>。本研究设定心理所有权为行为态度因素，信息规避规范为主观规范，自我控制能力为知觉行为控制，探索三者对信息规避意向及行为的预测作用，同时考虑卷入度与社会网络联结二者的外部背景因素的影响，具备理论的契合性与科学性。

### 2.2 背景因素：卷入度与社会网络联结

卷入度源于消费者行为理论，是个人基于内在需求、价值和兴趣对产品的相关程度的感知<sup>[27]</sup>。社会网络联结则是社会成员及群体之间所存在的关系纽带，其由成员间的交互时长、

情感强度、亲密度（相互信任）及互惠行为共同组成<sup>[28]</sup>。而卷入度在一定程度上反映了用户对微信的使用强度，其体现在在线情景下，用户与其他好友及群组之间交互的意愿及行为倾向<sup>[29]</sup>。已有研究证实了社区支持农业消费者的社交媒体参与度正向影响社会联结<sup>[30]</sup>，此外夏少昂等<sup>[31]</sup>认为微信卷入度与社会资本存在正向相关关系，因社会资本包含通过社交网络和人际关系可获得的情感支持、信息或信任，由此可证卷入度也对社会网络产生影响。综上，可作假设如下：

假设 1：微信卷入度正向影响社会网络联结强度。

### 2.3 行为态度：心理所有权

心理所有权是个人感觉拥有物质或非物质目标（或目标一部分）的一种心理状态，该理论的核心是对目标的占有感及心理联系<sup>[32]</sup>。当个体对目标充分了解，自我与目标相融合时，心理所有权便会得到增强<sup>[33]</sup>。由上文可知，卷入度是用户对产品需求和兴趣的一种反映，因此用户对微信的卷入度越高，个人需求与功能之间的匹配度与相关度越高，对其心理所有权就越强。此外，人际关系所形成的亲密度与心理所有权正向相关<sup>[34]</sup>，当顾客与企业服务人员建立密切的关系时，顾客会提高对该企业的心理所有权<sup>[35]</sup>，而亲密度正是社会网络联结程度的反映。由此可见，好友之间的联结越紧密，用户对所在社交平台的认同度越高，对其的心理所有权就越强。最后，微信作为虚拟的空间，用户会产生占有感和私有感。社交媒体用户容易对平台原生广告信息产生反感及厌恶，是由

于广告流信息“侵犯”了用户的社交媒体空间（不管是空间上还是注意力上），用户会有意识关闭或忽略广告信息，进行选择性的规避<sup>[5]</sup>。同理，用户也会产生忽略或屏蔽一些微信好友、朋友圈及公众号信息的意愿，以此消除因这些信息入侵带来的不快感知，保障对社交平台的心理私有性。综上，可作假设如下：

假设 2：微信卷入度正向影响用户心理所有权。

假设 3：社会网络联结强度正向影响用户心理所有权。

假设 4：用户心理所有权正向影响信息规避意向。

### 2.4 主观规范：信息规避规范

规范的形成是受社会因素影响的结果<sup>[36]</sup>。社会规范可分为群体规范及个人感知规范<sup>[37]</sup>，前者是在社会成员互动基础上形成的共同行为准则，后者则是个体对群体规范的理解与感知<sup>[38]</sup>，即：不同个体虽在性格与知识水平等方面存在差异，但选择信息的标准均需建立在对群体信息规范的感知之上。随着社会网络联结感知的增强，个体会越来越重视及认同群体信息规避的信念。此外，信息规避规范可分为感知描述性规范及感知命令性规范两种，其中感知描述性规范负向影响信息搜寻，并会降低搜寻频次，而命令性规范则直接对信息规避产生正向影响。综上所述可知：用户信息规避规范对信息规避意向产生正向预测作用。因此作出如下假设：

假设 5：社会网络联结正向影响用户信息规避规范。

假设 6：用户信息规避规范正向影响信息



规避意向。

## 2.5 知觉行为控制:自我控制能力

自我控制能力是指在长期价值目标与短暂欲望之间的冲突过程中,个人对思想、感觉与行动的自发管理与控制能力<sup>[39]</sup>。研究表明自我控制能力在个人所受到压力对网络成瘾的正向影响中发挥部分中介效应作用<sup>[40]</sup>,而用户感知控制可通过行为控制,对逃离倾向进行调节,也可通过决策控制,发生忽略、潜水、回避等倦怠行为<sup>[41]</sup>。综上可知用户自我控制能力越高,对社交媒体的使用时间与强度的调节控制能力越强。考虑到微信在学习、工作、支付情景下的不可替代性,在不停用的前提下,自我控制能力强的用户往往选择忽略一些其认为不重要的信息,如屏蔽一些好友的朋友圈,有意

不看一些广告、公众号消息等。而且,当个体感觉到行为控制产生时,会为其将要发生的行为提供直接预测的因素<sup>[42]</sup>,这也佐证了自我控制能力对微信信息规避行为的正向影响。此外,Dunwoody等<sup>[43]</sup>证实用户对信息搜寻的态度会负向影响信息规避行为,因此,与信息搜寻相反,信息规避意向会对规避行为产生正向影响作用。

综上,可进行如下假设:

假设7:用户自我控制能力正向影响微信信息规避意向。

假设8:用户自我控制能力正向影响微信信息规避行为。

假设9:用户信息规避意向正向影响规避行为。

综上,本研究的理论假设模型如图1所示。

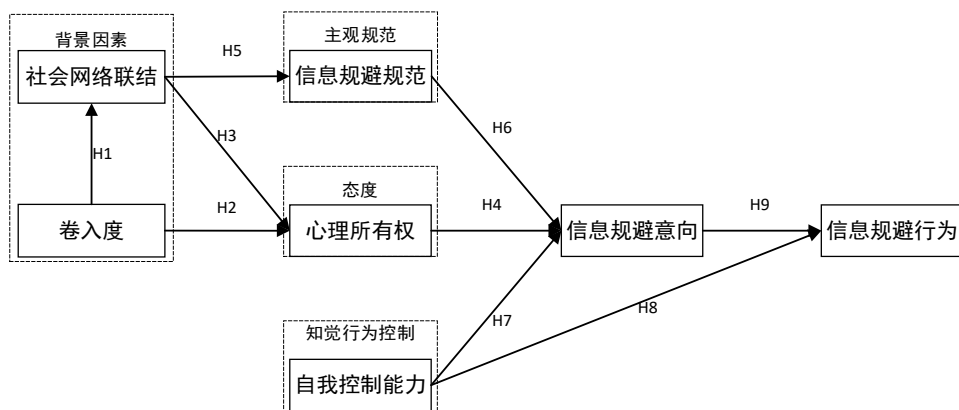


图1 理论假设模型

## 3 研究设计

### 3.1 测量工具

为保证问卷的信度,本研究采用国内外研究中使用的成熟量表进行改编,其中,针对英文量表,我们严格进行了双向互译,并按照

Brislin程序进行校验<sup>[44]</sup>。我们将30个题项进行预测试,删除因子载荷0.7以下的题项2项,如:

“我把微信看成是其所拥有的”“我觉得微信有价值”等,正式生成28个题项的“微信用户信息规避调查问卷”。正式问卷的预测试结果Cronach's a系数均大于0.6,组合信度CR系数

均大于 0.7, 因子载荷均大于 0.6, 平均方差提取量 AVE 均不小于 0.5, 达到预测信效度的要求,

可进行正式测试。正式量表的题项、来源、显变量指标与因子载荷如表 1 所示。

表 1 正式量表的潜在变量、来源、显变量指标与因子载荷

潜在变量	显变量指标	因子载荷
卷入度 (Zaichkowsky 等 <sup>[45]</sup> )	JR1 我觉得微信不重要	0.887
	JR2 我觉得微信有趣	0.896
	JR3 我觉得微信有意义	0.884
	JR4 我觉得微信没有吸引力	0.864
	JR5 我觉得微信使人兴奋	0.905
社会网络联结 (Chen 等 <sup>[46]</sup> )	LJ1 我和微信好友个性比较相似	0.835
	LJ2 我和微信好友能信任彼此	0.875
	LJ3 我和微信好友有相似的兴趣和爱好	0.903
	LJ4 我无法考虑到微信好友的兴趣	0.883
	LJ5 我和微信好友能彼此理解	0.881
信息规避规范 (Link 等 <sup>[24]</sup> 、Gao 等 <sup>[47]</sup> 、Hu 等 <sup>[48]</sup> )	GF1 认识的大部分人觉得可以规避一些微信信息	0.915
	GF2 认识的大部分人规避一些微信信息	0.948
	GF3 我的好友觉得我应规避一些微信信息	0.917
	GF4 我周围的人曾规避过一些微信信息	0.938
心理所有权 (Van Dyne 等 <sup>[49]</sup> )	XL1 我对微信有极高程度的个人所有权	0.897
	XL2 感觉微信属于我	0.916
	XL3 我对微信有很强的归属感	0.794
	XL4 大部分微信用户会认为微信不属于他们个人	0.853
自我控制能力 (Tangney 等 <sup>[50]</sup> )	KZ1 即使感到有趣, 我也不会做对自己不好的事情	0.749
	KZ2 我能集中注意力做事	0.917
	KZ3 我能有效完成长期目标	0.907
	KZ4 即使有时候娱乐放松, 我也能照常完成工作	0.882
信息规避意向 (Dai 等 <sup>[25]</sup> )	YX1 我将选择性地从微信接收信息	0.814
	YX2 我将拒绝接收一些微信信息	0.920
	YX3 我打算规避一些来自在线好友的微信信息	0.949
信息规避行为 (Guo 等 <sup>[7]</sup> )	XW1 我有意忽略一些微信信息	0.935
	XW2 我有意不关注一些微信信息	0.920
	XW3 我下滑页面来规避一些微信信息	0.898
	XW4 我采用技术手段来规避一些微信信息	0.858

### 3.2 研究对象

本研究选择微信用户为调查对象, 以纸质

问卷和电子问卷相结合的方式正式发放, 要求被访者匿名进行填写, 并向其说明问卷数据仅

用于学术研究，不会涉及个人隐私搜集和商业利用。本研究在广东、四川和天津三地共发放问卷 423 份，回收 411 份，其中有效问卷为 381 份。有效样本的人口统计学特征如表 2 所示。

表 2 有效样本的人口统计学特征 (N=381)

统计项	选项	频次	百分比 (%)
性别	男	172	45.14%
	女	209	54.86%
年龄	17 岁及以下	11	2.89%
	18—30 岁	183	48.03%
	31—40 岁	96	25.20%
	41—50 岁	68	17.85%
	51 岁及以上	23	6.04%
学历	高中、中职及以下	59	15.49%
	大专	79	20.73%
	本科	202	53.02%
	硕士研究生	36	9.45%
	博士研究生	5	1.31%
微信使用历史	1 年以内	9	2.36%
	1—3 年	11	2.89%
	3—5 年	141	37.01%
	5 年以上	220	57.74%
微信日均使用时长	1 小时以内	19	4.99%
	1—2 小时	25	6.56%
	2—3 小时	84	22.05%
	3—4 小时	76	19.95%
	5 小时及以上	177	46.46%

表 4 AVE 平方根与潜变量间相关系数

潜在变量	GF	YX	XW	JR	SY	LJ	KZ
信息规避规范 (GF)	0.930						
信息规避意向 (YX)	0.601	0.896					
信息规避行为 (XW)	0.645	0.897	0.903				
卷入度 (JR)	0.719	0.695	0.644	0.887			
心理所有权 (SY)	0.593	0.642	0.590	0.769	0.866		
社会网络联结 (LJ)	0.565	0.481	0.457	0.652	0.620	0.876	
自我控制能力 (KZ)	0.517	0.540	0.478	0.623	0.651	0.637	0.866

## 4 数据分析

### 4.1 信度与效度分析

偏最小二乘法结构方程模型 (PLS-SEM) 算法可用于解释变异，即因果变量之间是否显著相关，适合进行理论模型的探索及验证模型主成分间的预测与解释关系<sup>[5]</sup>，符合分析需求。因此本研究主要采用 Smartpls 3.0 软件进行分析。

对模型进行信度与效度分析，结果如表 3、表 4 所示，模型中各潜变量 Cronach's Alpha 均大于 0.6，CR 值均大于 0.7，因此本模型具备较高的信度。此外，因子载荷均大于 0.7，AVE 均大于 0.5，因此本研究具备较好效度。

表 3 模型的指标体系

潜在变量	Cronbach's Alpha	Composite Reliability (CR)	Average Variance Extracted (AVE)
信息规避规范	0.948	0.962	0.864
信息规避意向	0.875	0.924	0.803
信息规避行为	0.925	0.946	0.816
卷入度	0.932	0.949	0.787
心理所有权	0.888	0.923	0.750
社会网络联结强度	0.924	0.943	0.767
自我控制	0.887	0.923	0.751

采用 Bootstrapping( 拔靴法 ) 进行重复抽样。Bootstrapping 是通过重复随机抽样的方式, 替代原样本而生成新样本进行假设验证的方法<sup>[52]</sup>。为保证稳定性, 本研究设置抽样次数为 5000 次, 得到模型变量间影响的 t 值与置信区间 p 值, 如表 5 所示。此外, Henseler<sup>[53]</sup> 认为标准化均方根残差值 (SRMR) 需小于 0.08, 本模型数值为 0.052, 说明其具有较好拟合度。

最后, 为了避免共同方差的出现, 将数据导入 SPSS26 中进行 Harman 单因子检验分析, 结果发现特征根大于 1 的因子有 8 个, 大于 1 个; 最大因子的方差解释度为 30.712%, 低于 40%, 因此本研究不存在严重的共同方法偏差。

## 4.2 模型假设检验

根据表 5 的模型假设检验结果可知, 卷入度对社会网络联结 ( $\beta=0.652^{***}$ ,

$T=7.771$ )、卷入度对心理所有权 ( $\beta=0.635^{***}$ ,  $T=8.564$ )、心理所有权对信息规避意向 ( $\beta=0.366^{***}$ ,  $T=5.170$ )、社会网络联结对信息规避规范 ( $\beta=0.565^{***}$ ,  $T=6.280$ )、信息规避规范对信息规避意向 ( $\beta=0.312^{***}$ ,  $T=5.419$ )、信息规避意向对信息规避行为 ( $\beta=0.903^{***}$ ,  $T=33.048$ ) 产生显著正向影响 ( $P < 0.001$ ), 因此假设 1、假设 2、假设 4、假设 5、假设 6、假设 9 成立; 社会网络链接对心理所有权 ( $\beta=0.207^{**}$ ,  $T=3.166$ ) 产生显著正向影响 ( $P < 0.01$ ), 自我控制能力对信息规避意向 ( $\beta=0.141^*$ ,  $T=2.054$ ) 产生显著正向影响 ( $P < 0.05$ ), 因此假设 3、假设 7 成立; 此外, 自我控制能力对信息规避行为的影响 ( $\beta=-0.010$ ,  $T=0.287$ ) 不显著, 因此假设 8 不成立。SEM-PLS 路径分析结果如表 5、图 2 所示。

表 5 假设检验结果

原假设	模型构架间关系	路径系数	T 值	结论
假设 1	卷入度 -> 社会网络联结	0.652***	7.771	支持
假设 2	卷入度 -> 心理所有权	0.635***	8.564	支持
假设 3	社会网络联结 -> 心理所有权	0.207**	3.166	支持
假设 4	心理所有权 -> 信息规避意向	0.366***	5.170	支持
假设 5	社会网络联结 -> 信息规避规范	0.565***	6.280	支持
假设 6	信息规避规范 -> 信息规避意向	0.312***	5.419	支持
假设 7	自我控制能力 -> 信息规避意向	0.141*	2.054	支持
假设 8	自我控制能力 -> 信息规避行为	-0.010	0.287	不支持
假设 9	信息规避意向 -> 信息规避行为	0.903***	33.048	支持

注: \* 代表  $P < 0.05$ , \*\* 代表  $P < 0.01$ , \*\*\* 代表  $P < 0.001$  (\*  $t=1.96$ , \*\*  $t=2.58$ , \*\*\*  $t=3.29$ )

## 4.3 讨论

首先, 假设 H1、H2、H3 与 H5 的成立, 验证了背景因素的影响作用。用户对微信的卷入度越高, 越能增加对好友的情感维系意愿, 最终将提高微信产品在心理上的占有感, 产生

及累积对不相关、不重要、没兴趣获取的信息的排除效应, 进而提高针对特定信息的信息规避意向; 因主观规范是对群体规范的理解与感知, 故个人与社群好友间的联系与互动更为频密时, 其信息规避规范感知也会相应得到提高。



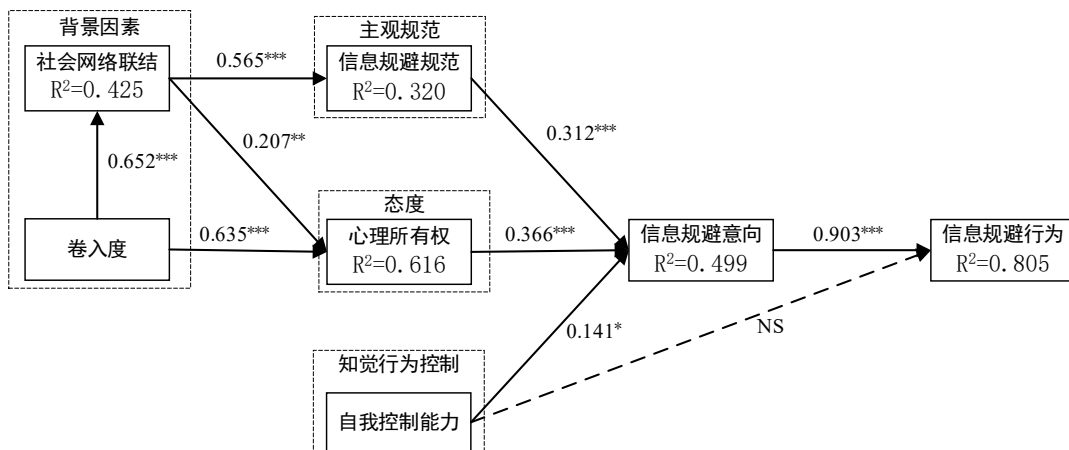


图2 SEM-PLS 分析结果

其次,假设 H4 与 H6 的成立,证实心理所有权、信息规避规范及自我控制能力是用户信息规避意向的归因。心理所有权对信息规避意向的正向预测结果,验证了 Niu 等<sup>[5]</sup>针对广告信息回避的研究,而且本文进一步将研究场景拓宽至好友、群、公众号等其他信息的规避上;与从众行为会在短期内被改变或扭转不同<sup>[54]</sup>,主观规范是一定时期内采取行为所受到的社会压力或看法,其具有恒定性的特征,故用户主观信息规避规范对信息规避意向可产生稳定的预测作用。

再次,假设 H7 成立验证自我控制能力对信息规避意向产生正向影响作用,这与季忠洋等<sup>[41]</sup>对社交媒体倦怠的研究结果相似,但我们进一步证实将规避的对象,从社交媒体平台本身推移到特定信息之上。假设 H8 不成立则反映了自我控制能力对信息规避意向与行为的影响有别,经探索,我们认为可能的原因是个体面临其他重要任务、目的,因此无法抗拒社交媒体使用所导致的自我控制失败<sup>[55]</sup>,或与社会资本驱动下的用户媒体信息错失焦虑感知产生

影响有关。

最后,假设 H9 成立,具有较高的路径系数证实了用户信息规避意向对行为具有强预测作用。现有文献或只关注规避意向,或仅探索规避行为的原因,本文对二者进行比较,并验证了意愿对行为的重要影响。

## 5 结论与展望

本文以微信用户为研究对象,基于计划行为理论框架,将信息规避意愿及行为作为结果变量,从外部因素、用户感知与效能等角度拓展与完善影响因素及路径,丰富信息规避行为的相关研究。采用量化研究方法,通过模型构建及验证,弥补现有研究视角的不足。本文取得的结论如下:

首先,提出社交媒体用户选择性信息规避的含义。与针对特定种类信息的信息规避不同,社交媒体信息规避具有场景性及针对性,其行为的表现形式也具有特殊性。因此,该含义的提出拓展了信息规避行为的研究范围,即:从

以往研究提出的信息种类差异，延伸至对用户的社会资本、社交媒体粘性、社会的规范制约、个体占有及控制能力感知等重要因素的关注上。

其次，扩充感知、信念及能效视角下的信息规避归因。与已有用户消极情绪、过载等负向能效感知不同，本文通过信念感知及正面能效视角，提出并证实心理所有权、信息规避规范及自我控制能力等信息规避意向的归因及规避意向对行为开展的预测作用。具体来说，已有研究认为烦心的各类内置广告，以及信息过多无法处理产生过载感是造成社交媒体用户放弃接收信息的归因，但本文证实了社会整体价值判定，以及用户所处的工作学习生活群体的价值取向，将逐渐形成并影响其对信息的规避准则及态度；而用户对微信等社交媒体创造的虚拟空间作为其内心占有程度、干涉及控制意识，影响其对不相关、不重要、不感兴趣或负面的信息进行选择规避的意向；而自我控制能力则驱动用户对一些即便感兴趣但对工作生活没帮助的“无效”且“浪费时间”的社交媒体信息进行主动抵御，所以自我控制力强的用户往往拥有较高效率及专注力，不易被无关信息打扰。尽管如此，无关信息或广告信息的标题、语言过于吸引，以及现代人际交往重视下社交媒体信息错失焦虑的加剧，均可导致用户本应采取的规避行为失效，其也成为自我控制能力对规避行为影响不显著的原因。

再次，借鉴消费者行为及社会资本等理论进行背景因素创新。卷入度与社会联结皆源于消费者行为研究领域，本文创新地将微信产品视为用户内在需要和价值兴趣的客体（特定产品卷入度）<sup>[56]</sup>，证实其对社会资本理论角度下

社会网络联结感知的显著影响。也证明了微信好友间的联结程度对主观信息规避规范的预测作用，以及卷入度与社会网络联结共同显著影响用户的微信心理所有权。具体来说，用户社会资本，即社交关系维系意识越强，越重视对重要信息的接收及对无关信息的拒绝，其与社交媒体产品的粘性（包括社交网络成瘾）共同提升微信虚拟空间的个人归属感。而对社交媒体平台的卷入或粘性，也会促进用户重视网络虚拟空间中对社会资本的培养与维护。

最后，从平台、用户及社会层面提出实践启示。对社交媒体平台来说，应给予用户更高的信息呈现与交互形式设置，让卷入度高、重视社交互动的用户能有效规避或延迟接收不相关的信息，如通过用户新增好友的群组，智能归类是否关注其朋友圈信息；让用户通过免费或付费等多样形式来选择一定时段内屏蔽广告信息，可提高用户评价及平台使用粘性；提升平台的形象，如可努力为特定人群降低平台的娱乐社交属性，打造或定制为工作学习交流场所，也会提升社会对社交媒体平台的正面评价。对用户来说，需提高自身的数字素养，提升信息筛选与识别能力，提高自控能力，更好地进行时间及行为管控，避免偶遇过多的无效社交及娱乐信息，影响工作和学习。提升广告及无关信息的识别能力，掌握屏蔽或拒收的技巧；最后，从社会宏观层面出发，应细化及完善网络成瘾预防的法律法规，如讨论是否应将青少年防沉迷管理条例延伸至社交媒体、短视频媒体的使用上。出台更严厉的法规或指引，让用户可自主选择消除广告或无关信息的侵扰，让社交媒体平台真正成为用户“拥有”并可“掌控”

的交流工具。

后续研究应考量用户选择性信息规避是否受到信息类型、推送频率、知识差异及其他外生激励因素的影响；此外，本文调查样本的分布以本科及大专以上的中青年群体为主，缺乏对未成年及中老年等两极群体的充足数据，日后可针对性地设计补充调查，探索群体间可能存在的差异。

### 参考文献

- [1] SIMON KEMP. Digital 2023: Global Overview Report[EB/OL]. (2023-01-26) [2024-01-30]. <https://datareportal.com/reports/digital-2023-global-overview-report>.
- [2] 甘春梅, 林晶晶, 肖晨. 社交网络用户间歇性中辍行为的影响因素研究 [J]. 情报理论与实践, 2021, 44(1): 118-123.
- [3] 袁静, 郭玲玉. 在线健康社区用户非持续使用行为影响因素分析 [J]. 现代情报, 2022, 42(2): 81-93.
- [4] SOROYA S H, FAROOQ A, MAHMOOD K, et al. From information seeking to information avoidance: Understanding the health information behavior during a global health crisis[J]. *Information Processing & Management*, 2021, 58(2): 102440.
- [5] NIU X, WANG X, LIU Z. When I feel invaded, I will avoid it: The effect of advertising invasiveness on consumers' avoidance of social media advertising[J]. *Journal of Retailing and Consumer Services*, 2021, 58: 102320.
- [6] SWEENEY K, MELNYK D, MILLER W, et al. Information avoidance: Who, what, when, and why[J]. *Review of General Psychology*, 2010, 14(4): 340-353.
- [7] GUO Y, LU Z, KUANG H, et al. Information avoidance behavior on social network sites: Information irrelevance, overload, and the moderating role of time pressure[J]. *International Journal of Information Management*, 2020, 52: 102067.
- [8] 陈琼, 宋士杰, 赵翔翔. 突发公共卫生事件中信息过载对用户信息规避行为的影响: 基于 COVID-19 信息疫情的实证研究 [J]. 情报资料工作, 2020, 41(3): 76-88.
- [9] DE BRUIN K, dE HAAN Y, VLIENGENTHART R, et al. News avoidance during the COVID-19 crisis: Understanding information overload[J]. *Digital Journalism*, 2021, 9(9): 1286-1302.
- [10] LEE J, KIM Y. How terrorism cues affect attitude polarization over undocumented immigrants via negative emotions and information avoidance: A terror management theory perspective[J]. *The Social Science Journal*, 2021: 1-16.
- [11] YU L, ZHENG F, XIONG J, et al. Relationship of patient-centered communication and cancer risk information avoidance: A social cognitive perspective[J]. *Patient Education and Counseling*, 2021, 104(9): 2371-2377.
- [12] WOOLLEY K, RISEN J L. Hiding from the truth: when and how cover enables information avoidance[J]. *Journal of Consumer Research*, 2021, 47(5): 675-697.
- [13] MEERZA S I A, BROOKS K R, GUSTAFSON C R, et al. Information avoidance behavior: Does ignorance keep us uninformed about antimicrobial resistance?[J]. *Food Policy*, 2021: 102067.
- [14] YANG Q, HERBERT N, YANG S, et al. The role of information avoidance in managing uncertainty from conflicting recommendations about electronic cigarettes[J]. *Communication Monographs*, 2021, 88(3): 263-285.
- [15] 王猛, 蒋琳萍, 郑建明, 等. 乡村公共数字文化服务中用户信息规避行为的影响机理 [J]. 图书馆论坛, 2022, 42(5): 107-117.
- [16] 彭国超, 梁欣婷, 张冰倩. 网络医疗众筹信息规避行为影响因素研究 [J]. 情报资料工作, 2021, 42(4): 72-80.
- [17] 彭丽徽, 张芊慧. 中老年人健康信息规避行为影响因素与关联路径研究——基于扎根理论的探索性分析 [J]. 图书馆学研究, 2022(1): 77-86, 76.
- [18] CHAE J. Who avoids cancer information? Examining a psychological process leading to cancer information avoidance[J]. *Journal of health communication*, 2016,

- 21(7): 837-844.
- [19] LU L, LIU J, YUAN Y C. Cultural differences in cancer information acquisition: cancer risk perceptions, fatalistic beliefs, and worry as predictors of cancer information seeking and avoidance in the US and China[J]. *Health Communication*, 2021: 1-10.
- [20] SIEBENHAAR K U, KOTHER A K, ALPERS G W. Dealing with the COVID-19 infodemic: Distress by information, information avoidance, and compliance with preventive measures[J]. *Frontiers in Psychology*, 2020, 11: 2981.
- [21] SONG S, YAO X, WEN N. What motivates Chinese consumers to avoid information about the COVID-19 pandemic?: The perspective of the stimulus-organism-response model[J]. *Information Processing & Management*, 2021, 58(1): 102407.
- [22] CHAE J, LEE C J, KIM K. Prevalence, predictors, and psychosocial mechanism of cancer information avoidance: Findings from a national survey of US adults[J]. *Health Communication*, 2020, 35(3): 322-330.
- [23] OROM H, SCHOFIELD E, KIVINIEMI M T, et al. Agency beliefs are associated with lower health information avoidance[J]. *Health Education Journal*, 2021, 80(3): 272-286.
- [24] LINK E, BAUMANN E. Explaining cancer information avoidance comparing people with and without cancer experience in the family[J]. *Psycho-Oncology*, 2022, 31(3): 442-449.
- [25] DAI B, ALI A, WANG H. Exploring information avoidance intention of social media users: A cognition-affect-conation perspective[J]. *Internet Research*, 2020, 30(5): 1455-1478.
- [26] AJZEN I. The theory of planned behavior[J]. *Organizational Behavior & Human Decision Processes*, 1991, 50(2): 179-211.
- [27] ZAICHKOWSKY J L. Measuring the involvement construct[J]. *Journal of Consumer Research*, 1985, 12(3): 341-352.
- [28] GRANOVETTER M S. The strength of weak ties[J]. *American Journal of Sociology*, 1973, 78(6): 1360-1380.
- [29] BRODIE R J, ILIC A, JURIC B, et al. Consumer engagement in a virtual brand community: An exploratory analysis[J]. *Journal of Business Research*, 2013, 66(1): 105-114.
- [30] 陈卫平. 社区支持农业(CSA)消费者对生产者信任的建立: 消费者社交媒体参与的作用[J]. *中国农村经济*, 2015(6): 33-46.
- [31] 夏少昂, 周晓虹. 微信卷入度与社会资本[J]. *江海学刊*, 2020(2): 121-128.
- [32] PIERCE J L, KOSTOVA T, DIRKS K T. Toward a Theory of Psychological Ownership in Organizations[J]. *The Academy of Management Review*, 2001, 26(2): 298-310.
- [33] 朱沅, 刘舒颖. 心理所有权前沿研究述评[J]. *管理学报*, 2011, 8(5): 784-790.
- [34] ZHU H, CHEN C C, LI X, et al. From personal relationship to psychological ownership: The importance of manager-owner relationship closeness in family businesses[J]. *Management and Organization Review*, 2013, 9(2): 295-318.
- [35] 赵建彬, 景奉杰. 基于心理所有权的网络嵌入对在线品牌社群公民行为的影响研究[J]. *管理学报*, 2015, 12(2): 276-283.
- [36] RIMAL R N, LAPINSKI M K. A re-explication of social norms, ten years later[J]. *Communication Theory*, 2015, 25(4): 393-409.
- [37] LAPINSKI M K, RIMAL R N. An explication of social norms[J]. *Communication Theory*, 2005, 15(2): 127-147.
- [38] QU Y, SAFFER A J, AUSTIN L. What drives people away from COVID-19 information?: Uncovering the influences of personal networks on information avoidance[J]. *Health Communication*, 2021: 1-12.
- [39] DUCKWORTH A L, TAXER J L, ESKREIS-WINKLER L, et al. Self-control and academic achievement[J]. *Annual Review of Psychology*, 2019, 70: 373-399.
- [40] SONG W J, PARK J W. The influence of stress on internet addiction: Mediating effects of self-control and mindfulness[J]. *International Journal of Mental Health and Addiction*, 2019, 17(4): 1063-1075.
- [41] 季忠洋, 李北伟, 朱婧祎, 等. 情感体验和感知控



- 制双重视角下社交媒体用户倦怠行为机理研究[J]. 情报理论与实践, 2019, 42(4): 129-135.
- [42] LINK E. Information avoidance during health crises: Predictors of avoiding information about the COVID-19 pandemic among german news consumers[J]. *Information Processing & Management*, 2021, 58(6): 102714.
- [43] DUNWODDY S, GRIFFIN R J. Risk information seeking and processing model[J]. *The SAGE Handbook of Risk Communication*, 2015: 102-116.
- [44] BRISLIN R W. Translation and content analysis of oral and written materials[J]. *Methodology*, 1980: 389-444.
- [45] ZAICHKOWSKY J L. The personal involvement inventory: Reduction, revision, and application to advertising[J]. *Journal of Advertising*, 1994, 23(4): 59-70.
- [46] CHEN X P, PENG S. Guanxi dynamics: Shifts in the closeness of ties between Chinese coworkers[J]. *Management and Organization Review*, 2008, 4(1): 63-80.
- [47] GAO L, WANG S, LI J, et al. Application of the extended theory of planned behavior to understand individual's energy saving behavior in workplaces[J]. *Resources, Conservation and Recycling*, 2017, 127: 107-113.
- [48] HU X, WU N, CHEN N. Young people's behavioral intentions towards low-carbon travel: Extending the theory of planned behavior[J]. *International Journal of Environmental Research and Public Health*, 2021, 18(5): 2327.
- [49] VAN DYNE L, PIERCE J L. Psychological ownership and feelings of possession: Three field studies predicting employee attitudes and organizational citizenship behavior[J]. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 2004, 25(4): 439-459.
- [50] TANGNEY J P, BAUMEISTER R F, BOONE A L. High self-control predicts good adjustment, less pathology, better grades, and interpersonal success[J]. *Journal of Personality*, 2004, 72(2): 271-324.
- [51] GEFEN D, RIGDON E E, STRAUB D. An Update and Extension to SEM Guidelines for Administrative and Social Science Research[J]. *Mis Quarterly*, 2011, 35(2): III-XII.
- [52] HAIR J F, RINGLE C M, SARSTEDT M. PLS-SEM: Indeed a silver bullet[J]. *Journal of Marketing theory and Practice*, 2011, 19(2): 139-152.
- [53] HENSELER J, HUBONA G, RAY P A. Using PLS path modeling in new technology research: updated guidelines[J]. *Industrial Management & Data Systems*, 2016, 116(1): 2-20.
- [54] 曹芬芳, 张晋朝, 李力, 等. 施引文献视角下国外从众行为研究述评[J]. *情报科学*, 2022, 40(3): 183-192.
- [55] 陈必忠, 郑雪, 孙晓军. 社交媒体自我控制失败影响硕士生主观幸福感的心理机制[J]. *心理发展与教育*, 2021, 37(6): 882-888.
- [56] 李琪, 任小静. 矛盾性追加评论对感知有用性的影响效应研究[J]. *管理科学*, 2017, 30(4): 139-150.



开放科学  
(资源服务)  
标识码  
(OSID)

# 基于 ARIMA、GM (1,1) 模型的高校 ESI 学科发展预测研究

柳佳彤<sup>1</sup> 康榆晨<sup>1</sup> 秦丽岩<sup>1</sup> 曹芳<sup>2</sup>

- 新疆医科大学公共卫生学院 乌鲁木齐 830017;
- 新疆医科大学图书馆 乌鲁木齐 830017

**摘要:** [目的/意义] 学科建设是高校提升教育质量的关键环节,对科学研究起着重要的支撑作用。采用数学统计建模探索一种科学有效的方法,实现潜力学科入围 ESI 前 1% 的时间预测,对于机构学科发展规划有着重要指导意义。[方法/过程] 基于 ESI 数据库,获取目标机构 4 个潜力学科的被引频次和 ESI 入围阈值,建立时间序列并创建预测模型:先引入转换系数来去除不同数据库的差异,使其可比,然后分别拟合 GM (1,1) 模型、ARIMA 模型,预测目标学术机构学科被引频次和 ESI 入围阈值,找到目标机构学科被引频次赶上 ESI 入围阈值的时间,即预测的入围时间。通过采用平均绝对百分比误差 (MAPE)、平均绝对误差 (MAE) 和均方根误差 (RMSE) 对模型的拟合预测效果进行评估和比较,根据 MAPE、MAE 和 RMSE 三个指标来评价模型拟合及预测效果,以此为学校的学科建设及长远发展规划提供参考依据。[局限] 本研究仅局限于目标机构 4 个学科的数据,尚需获取其他机构、更多学科的数据进行模型预测效果验证。[结果/结论] ARIMA 模型的拟合效果和预测效果优于 GM (1,1) 模型。目标机构的生物学与生物化学学科可能于近期入围 ESI 前 1%;免疫学科有入围 ESI 前 1% 学科的潜力,但入围时间可能会稍微滞后;分子生物学与遗传学和神经科学与行为学学科,离入围还有较大差距。

**关键词:** ESI; Incites; 潜力学科; 灰色模型; ARIMA 模型

**中图分类号:** G35; TP391

## Research on the Prediction of the Development of ESI Disciplines in Universities Based on ARIMA and GM (1,1) Models

LIU Jiatong<sup>1</sup> KANG Yuchen<sup>1</sup> QIN Liyan<sup>1</sup> CAO Fang<sup>2</sup>

- Public Health of Xinjiang Medical University, Urumqi 830017, China;
- Xinjiang Medical University Library, Urumqi 830017, China

**Abstract:** [Objective/Significance] Subject construction is a key aspect for universities to enhance the quality of education

**作者简介** 柳佳彤 (2000-), 本科生, 主要研究方向为预防医学; 康榆晨 (2001-), 本科生, 主要研究方向为预防医学; 秦丽岩 (1990-), 通讯作者, 硕士, 讲师, 主要研究方向为流行病学; 曹芳 (1985-), 硕士, 副研究馆员, 主要研究方向为图书竞争情报分析、数据挖掘, E-mail: Joanna567@xjmu.edu.cn。

**引用格式** 柳佳彤, 康榆晨, 秦丽岩, 等. 基于 ARIMA、GM (1,1) 模型的高校 ESI 学科发展预测研究 [J]. 情报工程, 2024, 10(1): 85-95.

and plays an important role in supporting scientific research. This article adopts mathematical statistical modeling to explore a scientifically effective method for predicting the time it takes for a potential subject to enter the top 1% in ESI rankings. This has significant guidance implications for institutional subject development planning. [Methods/Processes] Based on the ESI database, this paper obtains the citation frequency and ESI shortlisting threshold of the four potential disciplines of the target institution, establishes the time series, and creates a prediction model: first introduce the conversion coefficient to remove the differences between different databases and make them comparable, and then fit the GM(1,1) model and ARIMA model respectively to predict the citation frequency and ESI shortlisting threshold of the target academic institution, and find the time when the citation frequency of the subject of the target institution catches up with the ESI shortlisting threshold, that is, the predicted shortlisting time. By using mean absolute percentage error (MAPE), mean absolute error (MAE) and root mean square error (RMSE) to evaluate and compare the fitting and prediction effect of the model, the model fitting and prediction effect were evaluated according to the three indicators of MAPE, MAE and RMSE, so as to provide a reference basis for the discipline construction and long-term development planning of the school. [Limitations] The limitation of the study is data from only four disciplines in the target institution. Additional data from other institutions and more disciplines are needed to validate the predictive performance of the model. [Results /Conclusions] The fitting effect and prediction effect of ARIMA model are better than those of GM(1,1) model. The biology and biochemistry disciplines of the target institution will be in the top 1% of ESI in the coming months; Immunology has the potential to be shortlisted in the top 1% of ESI, but the shortlisting time may be slightly delayed; The disciplines of molecular biology and genetics and neuroscience and behavior are still far from being shortlisted.

**Keywords:** ESI; Incites; Potential Discipline; Gray Model; ARIMA Model

## 引言

学科建设是各大高校走内涵式发展道路的重要基础，亦是各高校从量的扩张到质的提升的关键因素。“双一流”建设是国家重大战略决策，为我国学科建设指明了新的方向，代表着我们国家为推动一流大学、一流学科建设的坚定决心。2022年1月，党中央、国务院发布了《关于深入推进世界一流大学和一流学科建设的若干意见》，突出强调要培养一流人才、服务国家战略需求、争创世界一流的导向<sup>[1]</sup>。

近年来，广大发展中国家将基本科学指标数据库（ESI）作为评价不同国家、不同地区、不同高校及不同科研机构学术能力及影响力的衡量指标。ESI根据世界各国或研究单位不同分类的学科，将所有学科分为22个种类。ESI

数据库每两个月更新一次，统计的数据范围是近10年的论文总被引频次，在我国ESI评价体系已被教育部门和各级评价部门广泛用以高校绩效的评价。由于ESI数据库仅提供机构入围学科的发文量和被引频次等信息，因此预测未入围学科晋级ESI前1%的时间是高校学科建设工作的重要目标。

灰色系统理论是由邓聚龙<sup>[2]</sup>在1982年创建的解决信息不完全和不确定性问题的学习和建模方法。灰色预测模型通过对系统内部灰色因素的分析 and 建模，对未来趋势进行预测，由于灰色预测模型具有简单易用、适用时序短、数据样本量较小等优点，被广泛应用于各个领域，如经济、环境、医疗等，为决策者提供了重要的参考依据。ARIMA模型其全称是自回归移动平均模型，属于统计模型中最常见的一

种, 是利用时间序列多个历史时刻对应的值预测未来时刻对应的值的一种方法。朱文佳等<sup>[3]</sup>用 ARIMA 模型拟合目标机构 ESI 被引频次预测值的时间序列, 用时间序列模型拟合 ESI 入围阈值时间序列, 以复旦大学经济与商学为例进行实证研究, 结果预测模型在两个时间序列上都有较高的拟合度, 得出的入围时间预测值可信度较高。本文在研究灰色预测模型 GM (1, 1) 及 ARIMA 模型的基础上, 通过实证分析, 对目标机构 4 个学科进行 ESI 前 1% 排名入围时间进行预测, 以期助力该校进一步提升学科建设质量。将笔者所在机构视为目标机构, 研究目标机构的学科建设及发展。目标机构现有医学、理学、工学、管理学、法学等学科门类, 其中有一门学科进入 ESI 全球排名前 2.63%, 还有一门学科位列 ESI 全球排名前 1%。4 个学科入选“十四五”重点学科建设序列。

## 1 研究方法

### 1.1 GM(1,1) 模型

这是一维一次求导的灰色预测模型, 即自变量对自身进行预测分析<sup>[4]</sup>。虽然可以选择各种类型的灰色模型, 但由于计算效率高, 以往的研究大多集中在预测 GM (1,1) 模型<sup>[5]</sup>。

已知参考数据列为:  $X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)\}$ , 做一阶累加生成序列 (AGO):

$X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)\}$ , 其中

$$X^{(1)}(k) = \sum_{i=1}^k X^{(0)}(i) \quad (1)$$

求累加序列的均值:  $Z^{(1)}(k) = \{0.5X^{(1)}(k) + 0.5X^{(1)}(k-1)\}$ ,  $k=2,3, \dots, n$ , 则 GM (1,1) 对应的白化方程为:

$$\frac{dX^{(1)}}{dt} + \alpha X^{(1)}(t) = \mu \quad (2)$$

其中  $\alpha$  为发展灰数,  $\mu$  为内生控制灰数, 利用最小二乘法得:

$$\begin{bmatrix} \alpha \\ \mu \end{bmatrix} = (B^T B)^{-1} B^T Y \quad (3)$$

其中,  $B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(n) & 1 \end{bmatrix}$ ,  $Y = \begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \vdots \\ X^{(0)}(n) \end{bmatrix}$ , 可

得到预测方程:

$$\hat{X}^{(1)}(k+1) = (X^{(0)}(1) - \frac{\mu}{\alpha})e^{-\alpha k} + \frac{\mu}{\alpha} \quad (4)$$

$k=1,2, \dots, n-1$ ,

原始预测值为:

$$\hat{X}^{(0)}(k+1) = \hat{X}^{(1)}(k+1) - \hat{X}^{(1)}(k) \quad (5)$$

$k=1,2, \dots, n-1$ ,

模型的检验:  $e = X^{(0)}(k) - \hat{X}^{(0)}(k)$ ,

$$S_1 = \sqrt{\sum_{i=0}^k (X^{(0)}(i+1) - \hat{X}^{(0)}(i))^2 / k-1}, \quad S_2 =$$

$$\sqrt{\sum_{i=0}^k (e^{(0)}(i+1) - \bar{e})^2 / k-1}$$

计算后验差比值、小误差概率:

$$C = \frac{S_2}{S_1} \quad p = P\{|e(k) - \bar{e}| < 0.6745S_1\} \quad (6)$$

GM (1,1) 模型精度等级评价见表 1 所示。

表 1 GM(1,1) 模型精度等级表

模型精度等级	p	c
1 级 (好)	$p \geq 0.95$	$c \leq 0.35$
2 级 (合格)	$0.80 \leq p < 0.95$	$0.35 \leq c < 0.5$
3 级 (勉强)	$0.70 \leq p < 0.80$	$0.5 < c \leq 0.65$
4 级 (不合格)	$p < 0.70$	$c > 0.65$



## 1.2 ARIMA模型

ARIMA 是一种著名的时间序列预测方法，它强调分析时间序列数据的随机性和概率性，被广泛用于预测传染病的出现和演变。该模型无需考虑各种影响因素，可以将这些综合效应纳入时间变量中，只需分析历史数据即可实现建模和定量预测，在医疗卫生领域具有广阔的应用前景<sup>[6]</sup>。其模型核心部分为自回归（AR）、差分（d）和移动平均（MA）。当模型为平稳序列时，时序问题可用 ARMA 模型进行预测。

### （1）ARMA（p,q）模型

假定时间序列中包括自回归与移动平均两部分，ARMA 模型就可以表达为：

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (7)$$

其中，自回归模型 AR（p）为：

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t \quad (8)$$

移动平均模型 MA（q）为：

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (9)$$

### （2）ARIMA（p,d,q）模型

假设数据一般模型为：

$$Y_t = \mu_t + X_t \quad (10)$$

其中  $\mu_t$  是时变均值， $X_t$  是零均值平稳序列。

如果一个时间序列  $\{Y_t\}$  的 d 次差分  $W_t = \nabla^d Y_t$  是一个平稳的 ARMA 过程，则称  $\{Y_t\}$  为差分自回归移动平均模型。

ARIMA（p,d,q）模型<sup>[7]</sup>，其中 AR 是自回归，p 为自回归项；d 是需要对数据进行差分的阶数；MA 为移动平均，q 为移动平均项数。建模基本步骤：（1）验证数据的平稳性；（2）模型识别与定阶；（3）参数估计；（4）预测并确定最优模型<sup>[8-9]</sup>。建模流程见图 1 所示。

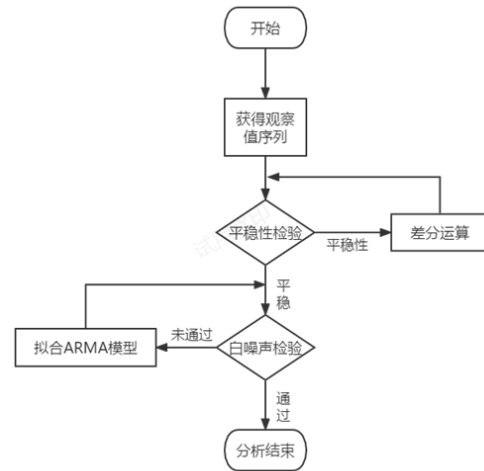


图 1 建模流程图

模型的检验流程如下：

### （1）单位根检验

单位根检验有许多方法，较为常用的是 ADF 检验，是 DF 检验的扩展，其应用的是在建模前对数据进行假设检验来判断数据是否为平稳数据，其中 H0：数据不平稳，H1：数据平稳，检验水准为 0.05。若  $P > 0.05$ ，则存在单位根，表明数据不平稳，需要进行差分处理使之平稳；反之，则数据平稳，可以进行后续建模。

### （2）白噪声检验

对数据进行白噪声检验，若为白噪声，则过去的行为对未来没有影响，不能进行预测，反之则可进行预测。常见的白噪声检验为 Ljung-Box 检验。其中 H0：数据为白噪声序列，H1：数据不为白噪声序列，检验水准为 0.05。经检验，上述数据并非白噪声序列，可以对其进行建模。

### （3）模型识别与定阶

对平稳后的时序绘制自相关图（ACF）、偏自相关图（PACF），ARIMA 模型的定阶原则如表 2 所示。

表 2 ARIMA 模型的定阶原则

模型	ACF	PACF
AR(p)	拖尾	p 阶截尾
MA(q)	q 阶截尾	拖尾
ARMA(p,q)	拖尾	拖尾

但仅自相关图 (ACF)、偏自相关图 (PACF) 不能较为精确进行定阶, 故我们基于 AIC 准则和 BIC 准则, 参考各 ARIMA 模型的 AIC 函数值和 BIC 函数值选出最优的模型。

赤池信息量 (AIC):  $AIC = -2\ln(L) + 2k$ 。其中,  $k$  表示模型参数个数,  $L$  表示似然函数。贝叶斯信息量 (BIC):  $BIC = -2\ln(L) + k\ln(n)$ 。其中,  $k$  表示模型参数个数,  $n$  表示样本数量,  $L$  表示似然函数。AIC、BIC 值越小的模型效果越好。

## 2 实证分析

### 2.1 数据来源

本研究基于 ESI 数据库, 数据来自团队定期采集的历史数据, 本研究使用 2019—2023 年更新的 23 期数据作为历史样本。ESI 数据库最近一次更新为 2023 年 5 月 11 日, 本期 ESI 统计数据覆盖时间范围为 2013 年 1 月 1 日至 2023 年 2 月 28 日, 整 10 年。每年 5 月 ESI 都会剔除最早一年的数据, 数据覆盖的时间范围在 10-11 年间波动。一般年底与年初更新的数据覆盖范围都接近 11 年, 本研究将每年年底作为时间序列的关键时间节点, 因此为保持一致, 将 InCites 与 WoS 采集的数据时间范围也设定为近 11 年。

### 2.2 时间序列的建立

ESI 数据库每年更新 6 次, 最新一次是

2023 年数据的第 3 次更新。将最新一期的 ESI 被引频次作为转换系数的计算依据, 各学科 23 个历史时间节点的入围机构最低被引频次形成的时间序列, 将作为预测未来入围阈值的依据。

WoS 数据库不提供 ESI 学科分类, 要通过 InCites 数据库确定 WoS 数据库论文的 ESI 学科分类。以 2022 年底为例, 先从 InCites 数据库中获取目标机构在目标 ESI 学科 2012—2022 年发表的论文集合, 然后在 WoS 论文合集中检索, 使用 WoS 的引文分析功能, 即可得到近 11 年的 WoS 总被引频次, 再根据转换系数得到 2022 年底时的近 11 年 ESI 总被引频次的估计值。

同理, 可得到截至 2022 年底的近 11 年 ESI 总被引频次的估计值, 从而预测目标机构 2023 年底时的近 11 年 ESI 总被引频次, 再与 2023 年底 ESI 入围机构最低被引频次的预测值进行比较, 如果后者小于前者, 该时间就是 ESI 前 1% 学科预测的入围时间。

朱文佳等学者的研究<sup>[3]</sup>表明, 只要有一定数量的样本, 使用转换系数或者修正因子的平均值, 来推测未入围机构的 ESI 数据, 是具有一定准确性的。笔者选择了修正因子的平均值进行了实证研究。

修正因子的计算方法如下:

$$ESI / WOS \text{修正因子}_i \text{平均值} = \text{Average} \left( \frac{ESI \text{中第 } i \text{ 个学科某个机构的被引频次}}{WOS \text{第 } i \text{ 个学科某机构的被引频次}} \right) \quad (11)$$

基于以上原理, 我们选择目标机构的四个最有潜力的学科, 分别计算其转换系数, 并根据计算所得的转换系数预估目标机构的 ESI 总被引频次的估计值, 得到所选四个学科的 23 期学科阈值的时间序列 (见表 3)。

表3 4个学科ESI入围机构最低被引频次(入围阈值X篇)

更新时间	分子生物学与遗传学	免疫学	神经科学与行为学	生物学与生物化学
2019-09	14174	5141	6581	6538
2019-11	14628	5253	6818	6728
2020-01	14716	5327	6904	6747
2020-03	14681	5419	6795	6823
2020-05	14132	5149	6426	6316
2020-07	14275	5204	6456	6348
2020-09	14621	5281	6545	6441
2020-11	14743	5357	6618	6584
2021-01	14817	5401	6715	6603
2021-03	14990	5492	6793	6769
2021-05	14208	5201	6480	6397
2021-07	14704	5368	6527	6602
2021-09	14615	5462	6650	6694
2021-11	14957	5552	6856	6890
2022-01	15100	5586	6986	6895
2022-03	15205	5668	7080	6986
2022-05	14830	5417	6668	6624
2022-07	15085	5428	6807	6673
2022-09	14849	5525	6889	6812
2022-11	14918	5631	6813	6910
2023-01	15227	5679	6951	7065
2023-03	15437	5699	7059	7126
2023-05	13496	5431	6914	6732

## 2.3 数据处理与预测

### 2.3.1 灰色模型预测结果

研究过程中,通过GM(1,1)模型建立,利用MATLAB,对样本数据进行计算分析及预测。GM(1,1)模型预测相关参数指标见表4。

(1)建立预测公式:

基于以上相关参数指标可以建立相关学科的预测方程。

基于ESI的预测方程为:

表4 基于GM(1,1)模型预测相关参数指标

学科		$\alpha$	$\mu$	c	p
分子生物学与遗传学	ESI	0.0266	15966	0.8071	0.8
	Wos	-0.0042	7870	0.1729	1
免疫学	ESI	0.0102	5781.9	0.7484	0.8
	Wos	-0.1168	1896.4	0.33	1
神经科学与行为学	ESI	-0.0059	6812.5	0.8323	0.8
	Wos	-0.0527	1329.8	0.2074	1
生物学与生物化学	ESI	0.0067	7097.4	0.863	0.8
	Wos	-0.0737	3988	0.2552	1

分子生物学与遗传学:

$$X^{(1)}(k+1) = -585376.56e^{-0.0266k} + 600225.56 \quad (12)$$

其中,  $k=1,2, \dots, n-1$ 。

免疫学:

$$X^{(1)}(k+1) = -561327.94e^{-0.0102k} + 566852.94 \quad (13)$$

其中,  $k=1,2, \dots, n-1$ 。

神经科学与行为学:

$$X^{(1)}(k+1) = 1161550.02e^{0.0059k} - 1154661.02 \quad (14)$$

其中,  $k=1,2, \dots, n-1$ 。

生物学与生物化学:

$$X^{(1)}(k+1) = -1052501.43e^{-0.0067k} + 1059313.43 \quad (15)$$

其中,  $k=1,2, \dots, n-1$ 。

基于WOS的预测方程为:

分子生物学与遗传学:

$$X^{(1)}(k+1) = 1880063.52e^{0.0042k} - 1873809.52 \quad (16)$$

其中,  $k=1,2, \dots, n-1$ 。

免疫学:

$$X^{(1)}(k+1) = 18097.30e^{0.1168k} - 16236.30 \quad (17)$$

其中,  $k=1,2, \dots, n-1$ 。

神经科学与行为学:

$$X^{(1)}(k+1) = 26375.40e^{0.0527k} - 25233.40 \quad (18)$$

其中,  $k=1,2, \dots, n-1$ 。

生物学与生物化学:

$$X^{(1)}(k+1) = 57789.26e^{0.0737k} - 54111.26 \quad (19)$$

其中,  $k=1,2, \dots, n-1$ 。

(2) 预测结果

$$\hat{X}^{(0)}(k+1) = \hat{X}^{(1)}(k+1) - \hat{X}^{(1)}(k) \quad (20)$$

通过上述公式可以计算出原始序列的预测

值, 结果见表 5、表 6。

表 5 目标机构 4 个学科 ESI 入围机构最低被引频次预测结果及残差(灰色模型)

时间	分子生物学与遗传学		免疫学		神经科学与行为学		生物学与生物化学	
	预测值	残差	预测值	残差	预测值	残差	预测值	残差
2022-09	14849	0	5525	0	6889	0	6812	0
2022-11	15364	-446	5696.3	-65.3	6873.2	-60.2	7028.2	-118.2
2023-01	14960	267	5638.3	40.7	6913.7	37.3	6981.4	83.6
2023-03	14567	870	5581	118	6954.5	104.5	6934.8	191.2
2023-05	14184	-688	5524.2	-93.2	6995.5	-81.5	6888.5	-156.5
2023-07	13811		5468		7036.7		6842.6	
2023-09	13448		5412.4		7078.2		6797	
2023-11	13095		5357.4		7120		6751.6	
2024-01	12750		5302.9		7162		6706.6	

表 6 目标机构 4 个学科 ESI 被引频次预估值预测结果及残差(灰色模型)

时间	分子生物学与遗传学		免疫学		神经科学与行为学		生物学与生物化学	
	预测值	残差	预测值	残差	预测值	残差	预测值	残差
2008-2018	6254	0	1861	0	1142	0	3678	0
2009-2019	7912.1	-147.1	2242.2	-212.2	1427.3	-13.3	4420.1	-203.1
2010-2020	7945.1	203.9	2520	192	1504.6	-18.6	4758	187
2011-2021	7978.2	31.8	2832.2	168.8	1586	73	5121.8	187.2
2012-2022	8011.5	-88.5	3183.1	-146.1	1671.9	-39.9	5513.4	-165.4
2013-2023	8044.9		3577.4		1762.3		5935	
2014-2024	8078.4		4020.6		1857.7		6388.7	

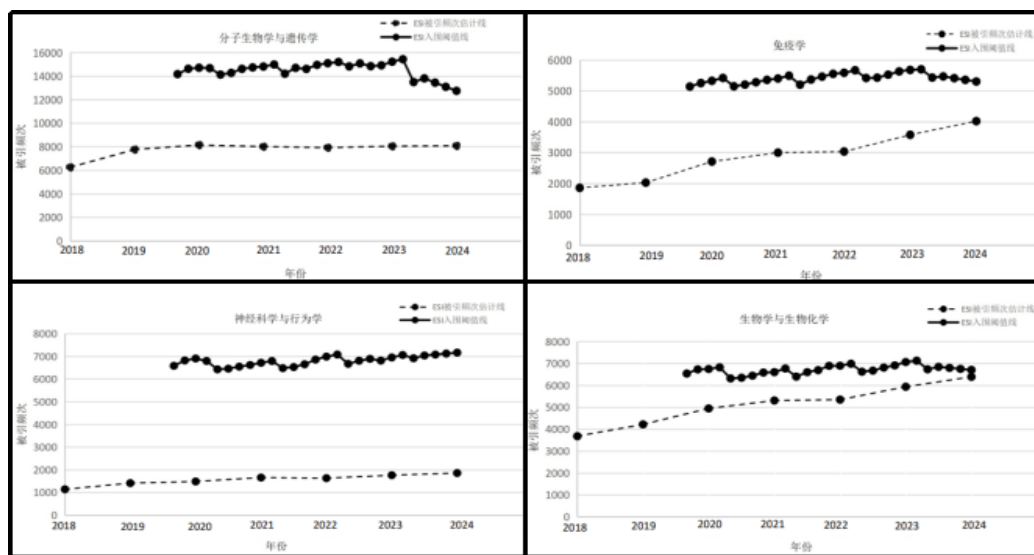


图 2 目标机构 4 个学科 ESI 前 1% 入围时间预测模型(灰色模型)



### 2.3.2 ARIMA模型预测结果

我们分别选取2019年9月到2023年5月的ESI数据和2002年到2023年的ESI被

引频次估计值，分别选取后三组数据测试，其余为训练集，进行模型建立。预测结果见表7、表8。

表7 目标机构4个学科ESI入围机构最低被引频次预测结果及残差(ARIMA模型)

时间	分子生物学与遗传学		免疫学		神经科学与行为学		生物学与生物化学	
	预测值	残差	预测值	残差	预测值	残差	预测值	残差
2022-07	15336.2	-251.2	5514.8	-86.8	6704.8	102.2	6744.0	-71.0
2022-09	14986.3	-137.3	5504.0	21.0	6909.7	-20.7	6811.2	0.8
2022-11	15166.6	-248.6	5608.0	23.0	7015.4	-202.4	6967.6	-57.6
2023-01	15022.7	204.3	5673.3	5.7	6923.2	27.8	6927.4	137.6
2023-03	15089.4	347.6	5753.8	-54.8	7007.7	51.3	7033.4	92.6
2023-05	14676.1	-1180.1	5475.8	-44.8	6654.9	259.1	6651.9	80.1
2023-07	14892.8		5544.5		6738.9		6780.3	
2023-09	14618.6		5626.4		6845.4		6869.1	
2023-11	14649.3		5713.0		6938.2		7010.4	
2024-01	14715.7		5751.4		7060.3		6999.1	

表8 目标机构4个学科ESI被引频次预估预测结果及残差(ARIMA模型)

时间	分子生物学与遗传学		免疫学		神经科学与行为学		生物学与生物化学	
	预测值	残差	预测值	残差	预测值	残差	预测值	残差
2018年	7128.3	-874.1	1714.2	147.2	1227.8	-85.4	3781.7	-104.0
2019年	6922.8	841.7	2195.0	-165.0	1287.0	126.7	4130.9	85.9
2020年	9274.8	-1126.0	2198.5	513.7	1580.2	-94.6	4755.9	189.0
2021年	8533.1	-523.2	3394.6	-393.3	1690.9	-31.4	5673.0	-363.6
2022年	7871.0	51.9	3290.2	-253.3	1765.1	-133.0	5673.9	-325.6
2023年	7732.1		3579.2		1916.4		6038.4	
2024年	7593.2		3868.2		2037.2		6403.0	

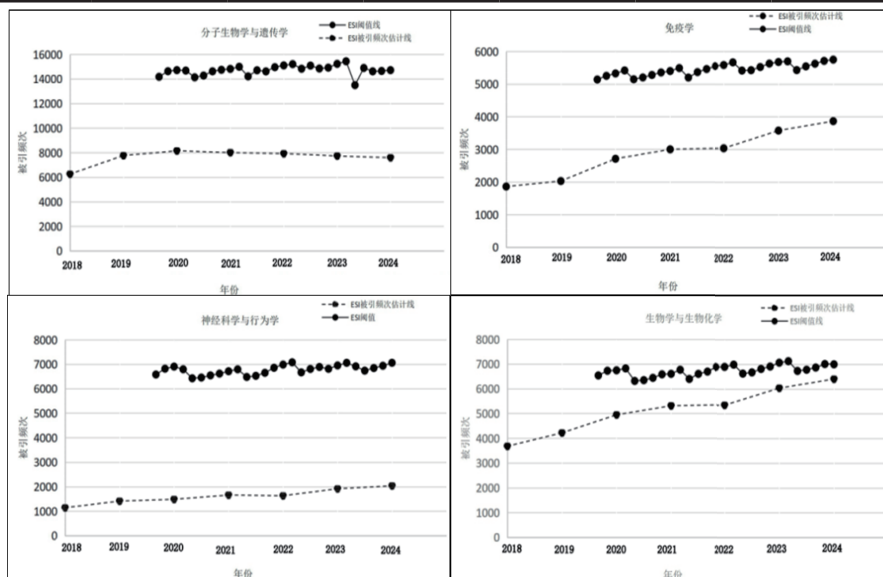


图3 目标机构4个学科ESI前1%入围时间预测模型(ARIMA模型)

## 2.4 模型评价

对于上述两种模型采用平均绝对百分比误差 (MAPE)、平均绝对误差 (MAE) 和均方根误差 (RMSE) 三个指标来评价模型拟合及预测效果, 其中 MAPE、RMSE 越小, 模型的预测精度越高<sup>[10-12]</sup>。计算公式如下:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (21)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (22)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (23)$$

通过对上述三种模型预测的精度比较, 从

MAPE 来看, 拟合效果 ARIMA 优于 GM(1, 1), MAE、RMSE 指标拟合效果与 MAPE 基本一致。综合比较来看 ARIMA 模型拟合效果和预测效果相对于 GM (1,1) 模型较好, 不同模型拟合及预测效果见表 9、表 10。

此外, 基于上述三种模型对 ESI 阈值和 ESI 被引频次估计值的预测, 我们发现, 目标机构的生物学与生物化学学科会于未来几个月入围 ESI 前 1%; 免疫学科有入围 ESI 前 1% 学科潜力, 但入围时间可能会稍微滞后; 分子生物学与遗传学和神经科学与行为学学科, 离入围还有较大差距。

表 9 3 种模型对 ESI 阈值拟合及预测效果

学科	模型	拟合值			预测值		
		MAPE	MAE	RMSE	MAPE	MAE	RMSE
分子生物学与遗传学	GM (1, 1)	1.58%	237.67	300.11	5.37%	779	784.3
	ARIMA	1.03%	153.04	213.25	4.11%	577.33	720.02
免疫学	GM (1, 1)	0.63%	35.33	44.42	1.89%	105.6	106.33
	ARIMA	0.33%	17.79	31.4	0.63%	35.06	40.96
神经科学与行为学	GM (1, 1)	0.47%	32.5	40.89	1.33%	93	93.71
	ARIMA	0.73%	49.83	76.91	1.62%	112.72	153.32
生物学与生物化学	GM (1, 1)	0.96%	67.27	83.59	2.50%	173.85	174.71
	ARIMA	0.53%	35.17	50.91	1.48%	103.44	106.35

表 10 3 种模型对 ESI 被引频次估计值拟合及预测效果

学科	模型	拟合值			预测值		
		MAPE	MAE	RMSE	MAPE	MAE	RMSE
分子生物学与遗传学	GM (1, 1)	1.47%	117	145.16	0.76%	60.15	66.5
	ARIMA	7.59%	463.83	596.52	0.65%	51.86	51.86
免疫学	GM (1, 1)	5.84%	134.73	165.22	5.22%	157.45	157.86
	ARIMA	11.71%	184.21	239.32	8.34%	253.3	253.3
神经科学与行为学	GM (1, 1)	0.73%	10.63	13.2	3.42%	56.45	58.83
	ARIMA	7.00%	60.58	75.06	8.00%	132.96	132.96
生物学与生物化学	GM (1, 1)	2.87%	130.03	159.39	3.31%	176.3	176.64
	ARIMA	6.82%	172.7	219.58	6.00%	325.62	325.62

### 3 总结与建议

在“双一流”建设背景下,学科建设对高校发展至关重要。如何通过学科建设赋能高校发展无疑是我国高校值得探究的重要事项<sup>[13]</sup>。本研究基于ESI采集的历史数据,发现ESI、InCites与WoS数据库之间被引频次存在差异,我们引入转换系数来去除不同数据库的差异,使其可比。其次,我们基于采集的历史数据和从不同数据库收集到的数据,结合时间序列数据的时序性特征,分别采用GM(1,1)模型、ARIMA模型拟合,预测目标机构潜力学科入围ESI前1%排名的时间。以目标机构的生物学与生物化学、免疫学、分子生物学与遗传学、神经科学与行为学四个学科为例进行实证研究,采用绝对百分比误差(APE)、平均绝对百分比误差(MAPE)和均方根误差(RMSE)比较模型拟合预测效果,根据APE、MAPE和RMSE最小原则选择最优预测模型。通过本研究提出以下建议。

#### 3.1 深入了解自身基础,强化潜势学科培育

目标机构需要明确学校发展定位,深入了解自身基础。夯实基础,加强调查研究,充分结合学校情况,分析各学科的优势与不足,找出优势学科与潜势学科,强化潜势学科培育。重视生物学与生物化学学科和免疫学科的培育,提升学科质量,增强科研产出能力。

#### 3.2 营造良好学科环境,保持优势学科特色

高校应积极加强学科科研工作,完善学科科研创新管理机制,营造良好的科研环境。优

势学科是学科生态系统生长发育的基石,高校应该充分利用特色学科的优势。对于进入ESI全球排名前2.63%,以及位列ESI全球排名前1%的优势学科,继续提升学科优势,带动相关学科发展,能够更好地营造良好学科环境。同时要遵循学科发展规律,保持学科发展的灵活性、成长性等特色<sup>[14]</sup>。

#### 3.3 增加学科建设投入,加强师资队伍建设

增加学科建设投入经费,提高高校自主获取经费的能力,促使其通过多渠道获取经费,争取更多中央财政资金和地方财政资金,经费投入向国家战略需求的学科适度倾斜。在人才培养增加投入,加强对高层次人才和团队的建设,给学者提供准确及时的信息,以提高学者的科学研究效率。

总之,基于对ESI数据库的利用和发掘,能够促进高校高质量内涵式发展。各高校应该更加重视学科建设,为世界一流高校的建设强基固本。

### 参考文献

- [1] 周建国,王教志,陈丽,等. “双一流”建设背景下提升研究生学位授予质量的路径探索——以温州医科大学为例[J]. 温州医科大学学报, 2023, 53(4): 340-344.
- [2] JU-LONG D. Control problems of grey systems[J]. Systems and Control Letters, 1982, 1(5): 288-294.
- [3] 朱文佳,朱莉. 基于时间序列分析法的ESI前1%学科入围时间预测模型[J]. 情报理论与实践, 2019, 42(10): 137-145.
- [4] 肖玲. 基于灰色预测模型的B2B电子商务交易规模预测研究[D]. 武汉: 华中师范大学, 2022.
- [5] WEI W, WANG G, TAO X, et al. Time series prediction for the epidemic trends of monkeypox using the ARIMA, exponential smoothing, GM (1,

- 1) and LSTM deep learning methods[J]. *Journal of General Virology*, 2023, 104(4): 001839.
- [6] MALKI Z, ATLAM ES, EWIS A, et al. ARIMA models for predicting the end of COVID-19 pandemic and the risk of second rebound[J]. *Neural Computing & Applications*, 2021, 33(7): 2929-2948.
- [7] 李田田. 基于时间序列分析与神经网络模型的股票价格预测 [D]. 大连: 东北财经大学, 2021.
- [8] 王燕. 应用时间序列分析 [M]. 北京: 中国人民大学出版社, 2022: 134-140.
- [9] 杜亚萍, 魏骅, 陶群山. 基于 ARIMA 和 GM(1, 1) 模型的中药材价格指数预测研究 [J]. *广东药科大学学报*, 2022, 38(5): 53-58.
- [10] 刘天, 姚梦雷, 黄继贵, 等. 组合预测模型在丙型肝炎病毒性肝炎发病率预测中的应用 [J]. *中国疫苗和免疫*, 2018, 24(6): 674-679.
- [11] 孙义, 周陇陇. 基于 Python 的金融时间序列 ARIMA 模型教学 [J]. *现代信息技术*, 2021, 5(10): 192-195.
- [12] WEI W, JIANG J, GAO L, et al. A new hybrid model using an autoregressive integrated moving average and a general-ized regression neural network for the incidence of tuberculosis in Heng County, China[J]. *The American Journal of Tropical Medicine and Hygiene*, 2017, 97(3): 799-805.
- [13] 巫芯宇, 商润泽. 以学科建设赋能我国高校发展研究——基于对 43 所高校一流学科建设方案的共现频谱分析 [J]. *西南师范大学学报 (自然科学版)*, 2023, 48(5): 102-110.
- [14] 荆林波, 杨佳乐. 哲学社会科学学科建设与人才培养: 成绩、问题及建议 [J]. *北京大学学报 (哲学社会科学版)*, 2022, 59(5): 150-158.





开放科学  
(资源服务)  
标识码  
(OSID)

# 基于文献计量学的铁路运输领域论文研究进展及前沿热点

戚小玉

中国铁道科学研究院集团有限公司 电子计算技术研究所 北京 100081

**摘要:** [目的/意义] 分析并跟踪铁路运输领域研究趋势可为铁路运输领域学术研究提供参考和依据。[方法/过程] 基于文献计量学分析铁路领域论文研究发展及前沿热点, 结合中图分类, 采集知网数据库中有基金支撑的核心期刊论文数据, 对发文总量、发文期刊来源、发文学科与主题、支撑基金等数据进行了综合统计分析与预测, 借助 CiteSpace 可视分析工具对研究热点高频关键词、关键词共现图谱、研究热点聚类进行分析与阐述, 利用 Kleinberg 突发检测算法对研究前沿进行识别与追踪并对不同类型研究机构的阶段性优势进行综合分析。[局限] 本文仅就知网数据进行了分析, 在样本选择方面存在一定局限性, 后续需要进一步扩展。[结果/结论] 文章结合多维度的分析结果总结了铁路运输领域科研发展特点, 在各类基金支撑、科研热点跟踪、产业成果转化等方面提出建议, 具有一定参考价值。

**关键词:** 文献计量学; 可视分析; 共现图谱; 铁路运输; 前沿热点

**中图分类号:** G35

## Research Development and Frontier Hotspots of Railway Transportation Based on Bibliometrics

QI Xiaoyu

Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China

**Abstract:** [Objective/Significance] Analyze and track research trends in the field of railway transportation can provide reference and basis for academic research in the field of railway transportation. [Methods/Processes] This paper analyzes the research development and frontier hot spots of papers in the field of railway based on bibliometrics. In combination with the classification of the Chinese Library, this paper collects the data of core journals supported by funds in the CNKI database, and carries out comprehensive statistical analysis and prediction of the total number of publications, the source of publications, the subjects and topics of publications, and the supporting funds. With the help of the CiteSpace visual analysis tool, the paper analyzes and elaborates the high-frequency keywords, keyword co-occurrence graph, and clustering of research hotspots. The Kleinberg burst detection algorithm is used to identify and track the research frontier, and the phased advantages of different types of research

**作者简介** 戚小玉 (1985-), 硕士, 助理研究员, 主要研究方向为数据可视化与可视分析, E-mail: qi\_xy@139.com。

**引用格式** 戚小玉. 基于文献计量学的铁路运输领域论文研究进展及前沿热点 [J]. 情报工程, 2024, 10(1): 96-106.

institutions are comprehensively analyzed. [Limitations] This paper only analyzes the data from CNKI, and there are certain limitations in sample selection, which need to be further expanded in the future. [Results/Conclusions] The article combines multidimensional analysis results to summarize the characteristics of scientific research development in the field of railway transportation, and puts forward suggestions in various fund support, research hotspot tracking, and industrial achievement transformation, which provides certain reference value.

**Keywords:** Bibliometrics; Visual Analysis; Co-occurrence Graph; Railway Transportation; Frontier Hotspots

## 引言

自 2008 年第一条时速 350 公里高速铁路——京津城际铁路开通运营以来,经过 15 年的发展,截至 2022 年底,全国铁路营业里程已达 15.5 万公里,其中高速铁路 4.2 万公里<sup>[1]</sup>,是世界上运营里程最长、在建规模最大、运营场景最丰富、商业运行时速最高<sup>[2]</sup>的高速铁路网,处于国际领先地位。

科技论文是科研成果的载体,展示相应领域的研究进展及热点趋势<sup>[3]</sup>,在中国高铁诞生至今的 15 年里,各类国家级、省部级基金项目为铁路行业持续发展提供了有力支撑,产出了大量高质量论文,发表于核心期刊。文献计量

学,是以文献体系和文献计量特征为研究对象,采用数学、统计学等计量方法,评价和预测科学技术的现状与发展趋势的学科<sup>[4]</sup>。因此,选取铁路运输领域近 15 年有基金项目支撑并发表于核心期刊的论文进行文献计量分析与可视化展现,详细分析与梳理铁路运输领域相关论文研究发展及前沿热点,总结铁路运输领域发展脉络及演化进程,为铁路运输领域科研工作提供参考,为科研管理工作提供决策依据。

本文的分析基于知网数据库,筛选自 2008 年 1 月—2022 年 12 月,中图分类号以 U2 开头的,含 U2 大类下所有子类,即铁路运输领域有基金项目支撑并发表于北大核心期刊的论文,检索框如图 1 所示。



The screenshot shows the CNKI search interface. The search criteria are as follows:

- Search term: CLC%='U2'
- Filters:  仅看有全文,  包含资讯,  网络首发,  增强出版,  基金文献,  中英文扩展,  同义词扩展
- Time range: 出版年度: 2008 -- 2022, 更新时间: 不限
- Source categories:  全部期刊,  SCI,  EI,  北大核心,  CSSCI,  CSCD,  AMI
- Buttons: 重置条件, 检索, 结果中检索

图 1 铁路运输领域有基金项目支撑并发表于北大核心期刊的论文检索框

经过数据清洗之后得到有效数据 30985 条,形成样本数据集。分析工具主要选用 Tableau 与

CiteSpace 以及工具中集成的算法模型。

## 1 发文数据综合分析

利用知网搜索页面的可视化分析功能,对近15年有基金项目支撑并发表于核心期刊的论文数据进行可视化概览并下载发文学科、发文主题等统计数据,结合样本数据集,从发文总量、发文期刊、发文学科、发文主题以及支撑基金等维度进行综合分析。

### 1.1 发文总量

从2008—2022年铁路运输领域核心期刊论文发文总量来看,整体呈上升趋势,2022年略有回落。但这些发文中,有基金支撑的论文占比呈逐年上升状态,在2022年发文总量由6064回落到5585的情况下,有基金支撑的论文占比由62.76%增长到64.83%,如图2所示,

体现了各类基金在铁路运输领域的科研工作中有着关键导向作用,为高质量核心期刊论文的产出提供了重要支撑。

同时,选用指数平滑算法基于前面15年的历史数据对未来1年的数据值进行短期预测。指数平滑算法可自动追踪数据变化,不断调整对数据序列中蕴含的短期趋势信息的估计,得出较为准确的短期预测结果<sup>[5]</sup>。调用Tableau集成的指数平滑算法模型得到2023年铁路运输领域核心期刊发文总量预测值为6266,有基金支撑的论文占比预测值为68.93%,这两个预测值均高于前面15年的历史最高值,基于历史数据得出的未来预期向好,说明高质量核心论文产出数量逐年增加的趋势较为稳定,各类基金在论文成果产出中持续发挥积极作用。

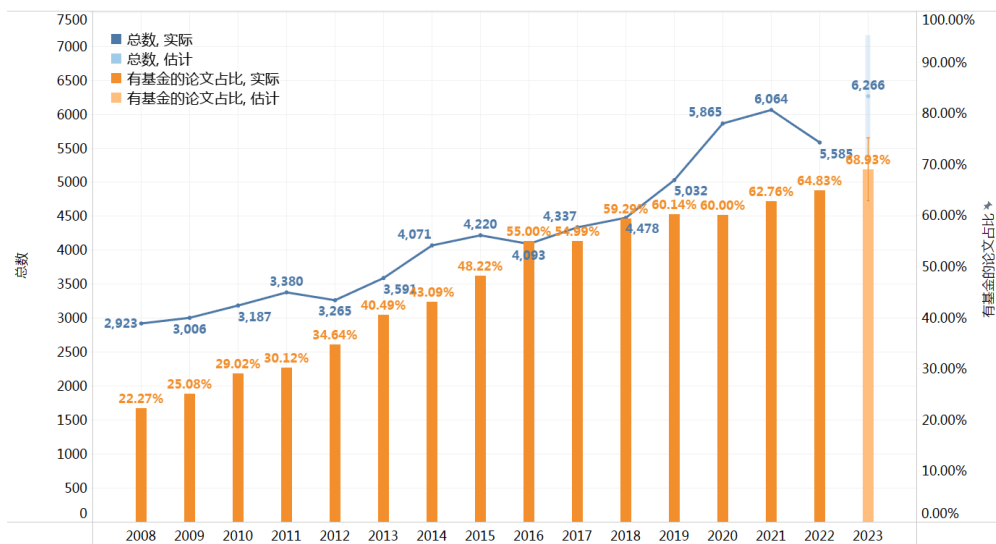


图2 近15年铁路运输领域核心期刊论文发文总量及有基金支撑的论文占比(加2023年预测值)

为确保预测值相对准确,对算法质量指标进行评估,上述预测属于时间序列预测,本文选用平均绝对百分误差MAPE作为预测算法的质量指标。一般认为当MAPE值低于10%时,

预测精度较高。发文总量与有基金的论文占比两个系列指数平滑模型的MAPE值分别为6.5%和7%,如图3所示,均小于10%,说明上述预测结果准确度较高。





## 1.4 支撑基金

从样本数据集中支撑论文的基金项目数据来看,由国家自然科学基金项目支撑的核心论文数量最多,为13376篇,远远多于其他各类

基金项目,如图6所示。这与国家自然科学基金聚焦基础、前沿、人才,注重创新团队和学科交叉的理念一致,是国家自然科学基金在铁路运输领域培育源头创新能力、支持基础科学研究的重要体现<sup>[6]</sup>。

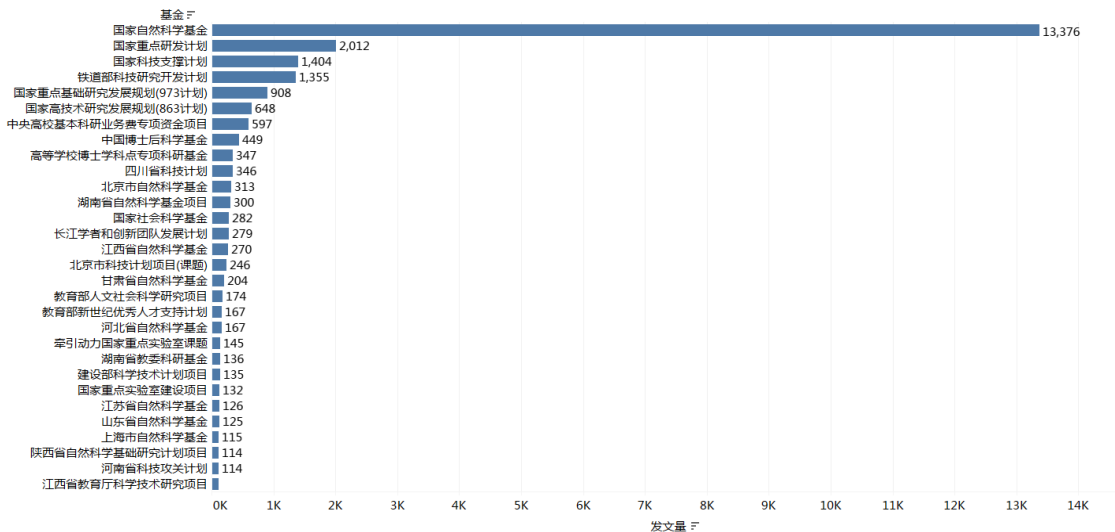


图6 支撑发文章量最多前40类基金项目

## 2 研究热点分析

某领域在一定时期内开展的有内在联系的科学问题或专题被视为研究热点<sup>[7]</sup>。样本数据集中,论文的关键词维度代表相应论文的研究领域、主题、专题、问题、方向和内容<sup>[8-9]</sup>,因此,样本数据集中出现的频次较高的关键词,可被认为是铁路运输领域的学术研究热点<sup>[10-11]</sup>。

### 2.1 高频关键词分析

将样本数据集导入至CiteSpace中,得到按词频数值降序排列的关键词列表,选取前10个关键词,如图7所示,可以看到词频数值最高的关键词为“高速铁路”,其词频数值也就是样本数据集中关键词的最高词频值为

2015。

Visible	Count	Central...	Year	Keywords
<input checked="" type="checkbox"/>	2015	0.00	2008	高速铁路
<input checked="" type="checkbox"/>	1388	0.00	2008	城市轨道交通
<input checked="" type="checkbox"/>	1056	0.00	2008	高速列车
<input checked="" type="checkbox"/>	906	0.00	2008	数值模拟
<input checked="" type="checkbox"/>	854	0.00	2008	地铁
<input checked="" type="checkbox"/>	616	0.00	2008	轨道交通
<input checked="" type="checkbox"/>	537	0.00	2008	无砟轨道
<input checked="" type="checkbox"/>	425	0.00	2008	地铁车站
<input checked="" type="checkbox"/>	409	0.00	2009	动车组
<input checked="" type="checkbox"/>	395	0.00	2008	铁路运输

图7 词频最高的前10个关键词

根据普莱斯定律<sup>[12-13]</sup>,高频关键词的最低词频数值 $T = 0.749 \times \sqrt{P_{\max}}$ ,其中, $P_{\max}$ 表示样本数据集中的关键词最高词频值,由图7可知 $P_{\max}=2015$ ,计算得出 $T=33.62$ ,则词频大于34的关键词为高频关键词,共有213个,生成高频关键词词云图如图8所示。



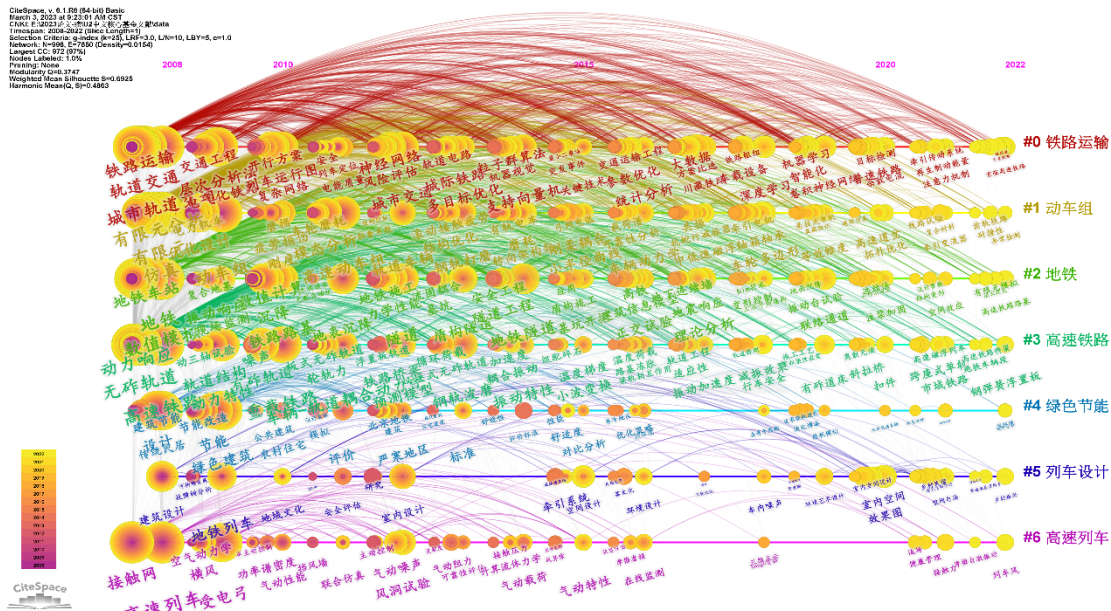


图 10 关键词聚类时间线图

对于 #1 动车组大类, 2008—2012 年, 研究热点包括有限元分析法、疲劳损伤、刚度、车轮磨耗、车体模态分析等, 主要集中于车体、车轮、底架等车体关键承载部件。2013—2017 年, 研究热点包括有轨电车、结构优化、刚柔耦合、可靠性分析、小半径曲线、共振等, 主要集中于与车辆动力学相关的内容, 研究热点较之前五年更为细化。2018—2022 年, 研究热点包括牵引电机、悬挂式单轨、高速道岔、线路试验、齿轨铁路、耐撞性等, 可看出研究热点聚焦于适应更加复杂多变的沿线地形与环境, 为动车组奔驰在更加广袤的祖国大地上提供技术保障。

对于 #2 地铁大类, 2008—2012 年, 研究热点主要聚焦于修建地铁相关的内容, 包括地铁车站、复合地基、振动响应、沉降、地铁施工等。2013—2017 年, 研究热点包括盾构隧道、安全工程、地铁隧道、隧道工程、盾构施工、建筑信息模型等, 这些研究热点主要集中在隧

道与盾构技术方面, 这个阶段也是中国地铁盾构技术飞速发展的阶段, 到 2017 年, 中国的盾构机销量排世界第一<sup>[14]</sup>。2018—2022 年, 研究热点有 BIM 技术、变形控制、注浆加固、结构变形、空间效应等。在这一阶段, 地铁建设规模不断扩大, 数字化技术的应用与确保基坑安全的注浆加固技术越来越受到人们关注<sup>[15]</sup>。

对于 #3 高速铁路大类, 2008—2013 年, 研究热点聚焦于轮轨方面, 主要包括无砟轨道、轨道结构、有砟轨道、轮轨力、车辆-轨道耦合动力学、板式无砟轨道、双块式无砟轨道等。2014—2022 年, 关于高速铁路的很多方面陆续成为研究热点, 包括 2014—2016 年的路基冻胀、温度荷载、温度梯度等与高寒铁路相关的内容; 2017 年、2018 年的振动加速度、减振效果、行车安全等与高铁安全相关的内容; 2019 年以后的高速磁悬浮列车、市域铁路等最近几年在城市发展需求驱动下不断出现的研究热点。







## 4 研究机构分析

通过 CiteSpace 对样本数据集进行 Burst 突发检测,不仅能够识别不同时间段的研究前沿,

还可以识别不同时间段发文量增长速率激增的研究机构,从而分析出铁路运输领域不同类型的研究机构在不同时段发挥的作用,对研究机构的突发检测结果如图 12 所示。



图 12 近 15 年铁路运输领域研究机构的突发检测结果

从分析结果中可看出,从 2008—2011 年,产出核心论文的研究机构主要包括西南交通大学牵引动力国家重点实验室、中南大学土木建筑学院、西南交通大学交通运输学院、西南交通大学土木工程学院等,主要以高校为主。同时,2008—2016 年前后,北京交通大学土木建筑工程学院、中科院寒区旱区环境与工程研究所冻土工程国家重点实验室、中国铁道科学研究院金属及化学研究所、中国铁道科学研究院

铁道建筑研究所等研究机构作为基金核心论文产出的主要研究机构,持续近十年。也就是说从 2008 年开始,近 15 年里的前期、前中期,主要研究机构主要以高校、科研院所为主。

再看最近 15 年里中期、中后期、后期的情况,主要研究机构包括南车青岛四方机车车辆股份有限公司、中国铁道科学研究院运输及经济研究所、中国铁道科学研究院通信信号研究所、中国铁路设计集团有限公司、北京纵横机

电科技有限公司、广州地铁设计研究院股份有限公司、中车青岛四方机车车辆股份有限公司等,可看出在这几个阶段,有基金核心论文产出的主要研究机构以企业、科研院所为主。

也就是说,在最近 15 年的时间里,基金核心论文成果的主要产出机构在由高校、科研院所向企业、科研院所转变,与科研成果从研究到应用转化的过程相吻合,体现出高校、科研院所、企业在铁路运输领域不同时期的科研工作中发挥的重要作用,也体现出科研院所作为兼顾科研与应用的研究机构,在多个时期中发挥的持续性作用,为铁路运输领域科技创新提供重要支撑。

## 5 结束语

在中国高铁诞生至今的 15 年里铁路运输领域有基金支撑的核心期刊论文数据进行的多维度分析中,运用了创新性的分析角度与分析方法,提出相关建议,具体如下:

(1) 创新性地关注基金支持对核心论文发表产生的积极作用,并基于历史数据做出预测。

统计并预测了发文总量,综合分析了期刊来源、学科与主题、支撑基金分布数据,发现铁路运输领域整体具有发文逐年增加、学科间交叉渗透、国家自然科学基金成为重要支撑等特点。建议更加充分地发挥其他各类基金的支撑作用,持续增加高质量论文成果的产出。

(2) 首次结合高频词定义,全面综合地对铁路运输领域学术论文产出数据进行多维度关联分析,并结合铁路行业不同阶段的发展特点,

对分析结果进行解读。

铁路运输领域研究热点较多,共现网络中各研究热点关联程度较高,研究热点与实际应用紧密结合,建议依据分析结果对各个聚类进行持续深入的研究与完善。同时,对研究热点进行阶段性识别、分析,可基于分析结果对科研发展趋势进行跟踪分析,为科研工作提供指导。

(3) 创新性地核心期刊论文的产出机构进行分类分析,识别出高校、科研院所、企业在科技发展不同阶段的机构优势。

可参考分析结果,充分发挥各类研究机构在各个研究阶段的作用,进一步促进产业创新与成果转化,让学术成果更加高效地得到应用,惠及民众,为高质量发展实现科技赋能。

通过此项分析与研究,从多维视角揭示研究主题热点的发展与关键技术的衍生过程,为铁路运输领域研究发展提供全面解读与决策支持。

## 参考文献

- [1] 中华人民共和国中央人民政府. 截至 2022 年底综合交通运输网络总里程超 600 万公里 [EB/OL]. (2023-02-27) [2024-01-09]. [http://www.gov.cn/xinwen/2023-02/27/content\\_5743414.htm](http://www.gov.cn/xinwen/2023-02/27/content_5743414.htm).
- [2] 罗春晓. 世界高速列车图鉴 [M]. 北京: 中国铁道出版社有限公司, 2020 :50-59.
- [3] 王德, 张航, 邵鸣, 等. 基于文献统计分析的我国铁路运输管理领域研究进展及热点 [J]. 中国铁路, 2022(11): 102-114.
- [4] 高俊宽. 文献计量学方法在科学评价中的应用探讨 [J]. 图书情报知识, 2005(2): 14-17.
- [5] 陈伟, 白福臣. 我国海洋产业增加值的组合预测分析 [J]. 生态经济, 2010(2): 44-46.
- [6] 国家自然科学基金委员会. 机构概况 [EB/OL].

- (2023-03-02) [2024-01-09]. <https://www.nsf.gov.cn/publish/portal0/jgsz/01/>.
- [7] 李雪山. 基于知识图谱的铁路运输管理工程领域研究热点及趋势分析 [J]. 铁道运输与经济, 2017, 39(1): 81-87.
- [8] 刘则渊, 尹丽春. 国际科学学主题共词网络的可视化研究 [J]. 情报学报, 2006, 25(5): 634-640.
- [9] 刘秀琴. 国外科学数据管理研究量化分析 [J]. 图书馆工作与研究, 2016(10): 65-71.
- [10] 马费成, 张勤. 国内外知识管理研究热点 -- 基于词频的统计分析 [J]. 情报学报, 2006, 25(2): 163-171.
- [11] 郝强, 唐锡晋. 基于社会网络分析的“系统科学大会”可视化探析 [J]. 系统科学与数学, 2018, 38(5): 521-536.
- [12] 羊晚成, 胡孙婕, 张若愚. 我国编辑出版学合作网络与合作研究热点分析 [J]. 中国科技期刊研究, 2017, 28(10): 959-964.
- [13] 陈晓钰, 马海群. 开放政府数据领域文献计量学相关定律实证分析 [J]. 图书馆研究与工作, 2022(4): 26-36.
- [14] 中国产业经济信息网. 中铁装备: 十年磨剑, 让中国盾构机迈上世界之巅 [EB/OL]. (2023-03-05) [2024-01-09]. <http://www.cinic.org.cn/zgzz/pp/640800.html>.
- [15] 百度百科. 地铁注浆加固 [EB/OL]. (2023-03-05) [2024-01-09]. <https://baike.baidu.com/item/%E5%9C%B0%E9%93%81%E6%B3%A8%E6%B5%86%E5%8A%A0%E5%9B%BA/17679541?fr=aladdin>.
- [16] 陈悦, 陈超美, 刘则渊, 等. CiteSpace 知识图谱的方法论功能 [J]. 科学学研究, 2015, 33(2): 242-253.
- [17] 魏晓俊. 基于科技文献中词语的科技发展监测方法研究 [J]. 情报杂志, 2007(3): 34-36, 39.



开放科学  
(资源服务)  
标识码  
(OSID)

# 国内外顾客互动研究进展与热点分析 ——基于科学知识图谱视角

乐承毅 吴嘉炜 周佳辰

华东交通大学 经济管理学院 南昌 330013

**摘要:** [目的/意义] 随着信息技术的不断发展,顾客互动成为企业实践关注的焦点,然而既有的综述文献不仅时间偏早且多为主观定性分析,随着研究的深入,国内外相关文献不断涌现,对现有研究进行梳理十分必要。[方法/过程] 以 SSCI 和 CSSCI 文献数据为样本,通过 Citespace 软件进行文献计量分析,从前沿文献特征、科研合作、基础知识结构和研究热点等方面,对中外顾客互动领域的研究进展和前沿进行对比研究与全景展示,最后提出了未来的研究趋势。另外,针对目前中文期刊 Citespace 研究范式存在的问题,利用 DEAN 数据清洗流程与系统化的检索策略优化数据质量,增强了文献计量分析结论的准确性。[结果/结论] 国内外顾客互动领域的研究合作均较为紧密,且国外研究在近年更为活跃;国内外基础文献的演进脉络主线都是从顾客参与、互动到顾客契合的转变,但二者的支线脉络与显示出的研究前沿有所不同;国内外前沿文献的研究内容本同末异,但国外前沿文献的研究情境与方法要更加丰富。

**关键词:** 顾客互动; 科学知识图谱; Citespace; 研究进展; 热点分析

**中图分类号:** G35; F274

## Domestic and International Research Progress and Hotspot Analysis of Customer Interaction: Based on the Perspective of Scientific Knowledge Graph

LE Chengyi WU Jiawei ZHOU Jiachen

School of Economics and Management, East China Jiaotong University, Nanchang 330013, China

**Abstract:** [Objective/Significance] With the continuous development of information technology, customer interaction has become the focus of corporate practice. However, the existing literature reviews are not only early, but also mostly subjective and qualitative analysis. As the research work went further, the related domestic and international literature has increased rapidly in recent years, and it is necessary to sort out the existing research. [Methods/Processes] With literature data of SSCI and CSSCI as sample, this article uses Citespace software for bibliometric analysis to compare the domestic and international research

**基金项目** 国家自然科学基金项目“开放式创新下企业-用户知识在线互动与知识共创研究”(72161013);国家自然科学基金项目“移动社交网络环境下企业虚拟社区的用户知识贡献行为研究”(71761012);江西省研究生创新基金项目“企业虚拟社区用户持续反馈行为研究”(YC2021-S476)。

**作者简介** 乐承毅(1982-),教授,博士生导师,主要研究方向为知识管理与用户行为;吴嘉炜(1997-),硕士研究生,主要研究方向为知识管理与数据挖掘, E-mail: wujiaweiwy@163.com;周佳辰(1999-),硕士研究生,主要研究方向为用户行为与数据分析。

**引用格式** 乐承毅,吴嘉炜,周佳辰.国内外顾客互动研究进展与热点分析——基于科学知识图谱视角[J].情报工程,2024,10(1):107-125.



progress and cutting-edge in the field of customer interaction from four aspects including cutting-edge literature features, research cooperation, basic knowledge structure and research hotspots. Finally, the future research trend in the field of customer interaction is proposed. In addition, in view of the problems existing in the current research paradigm of Citespace in Chinese journals, this paper uses DEAN data cleaning process and systematic retrieval strategy to optimize the data quality and enhance the accuracy of bibliometric analysis conclusions. [Results/Conclusions] We come to the conclusion that research cooperation in the field of customer interaction at home and abroad is relatively close, and international research have become more active in recent years; the main line of evolution of basic literature at domestic and international is the transformation from customer participation and interaction to customer engagement, but the evolution branch lines are different and the focus of research front are different; the research content of domestic and international frontier literature is the same, but the research context and methods of international frontier literature are more abundant.

**Keywords:** Customer Interaction; Scientific Knowledge Graph; Citespace; Research Progress; Hotspot Analysis

## 引言

互动概念广泛存在于社会学、心理学等多个领域，互动理论也是服务营销和关系营销理论的重要基石。随着营销理论的发展，尤其是顾客主导逻辑取代产品主导逻辑后，价值共创理论逐步成熟，顾客互动受到了前所未有的关注<sup>[1]</sup>。互联网时代下，体验经济等新经济形态的产生促使企业对于运营逻辑和创新思维进行重新思考，许多企业开始重新审视和挖掘更适应于时代发展的实践活动<sup>[2]</sup>，顾客互动的实际应用情境不断延展。因此，多元化情景下顾客互动的研究焦点各有侧重，如大数据背景下顾客互动是用户参与产品创新的研究重点<sup>[3]</sup>，而社交商务环境下顾客间互动受到了众多学者的关注<sup>[4]</sup>。

与此同时，梳理顾客互动领域研究脉络的文献较少，仅有的几篇综述文献时间偏早<sup>[1,5-6]</sup>。并且参与（participation）、契合（engagement）和互动（interaction）作为最常见的理论概念，在国内外研究中常有不同程度的混淆。国内文献多用参与一词，而 involvement 在新产品开发

和虚拟社区领域的英文文献中又常与 participation 同义。Engagement 的概念内涵和测量标准尚不统一<sup>[7]</sup>，包含了多个维度，其中认知与情感维度与 involvement 类似，而行为维度又与 participation 近似，因此国内部分学者将其译作契合加以区分，但更多地统称为参与<sup>[8]</sup>。故随着研究进程的发展，厘清相关理论概念，梳理分析近年来顾客互动领域的国内外研究成果，是十分必要的。另一方面，与大多数领域相类似，顾客互动研究热点与趋势分析的相关文献仍以定性研究为主，鲜有定量分析，也缺乏对研究成果与结论的可视化展示。

基于此，本文从文献计量、信息可视化的视角出发，利用 Citespace 科学知识图谱对国内外顾客互动相关研究文献进行定量统计分析和定性描述探索，通过梳理分析近年来该领域的研究成果，厘清了相关的理论概念与演进脉络。另外，针对中文期刊 CiteSpace 研究范式中数据质量难以保证，从而导致分析结论缺乏信度的问题，本文通过 DEAN 数据清洗流程与系统化的检索策略对数据质量进行优化，增强了本文结论的可靠性。最后，顾客互动领域的文献综

述往往仅从国内或国外的研究出发，缺乏两者的对比分析与全景展示。因此本文从前沿文献特征、科研合作、基础知识结构和研究热点等多个方面对国内外顾客互动的中英文研究进展进行对比研究，并提出了未来的发展趋势与方向，为其他学者提供了有益参考。

## 1 数据收集

### 1.1 数据来源

为保证检索到的文献数据质量，本文以中文社会科学引文索引（CSSCI）和 Web of Science（WOS）核心合集下的社科引文索引（SSCI）与论文集引文索引（CPCI-SSH）作为数据源。相关索引作为中国人文社科评价领域的标志性工程以及国际公认的反映科学研究水准的数据库，能够较全面地反映顾客互动领域的研究情况，其检索到的相关文献可被视为该领域的前沿文献。

中文期刊的 CiteSpace 研究范式在数据收集阶段往往缺少系统的检索策略以及关键的数据清洗步骤<sup>[9]</sup>。因此本文的检索策略依据 Citespace 作者陈超美<sup>[10]</sup>教授的意见，优先保证查全率，再兼顾查准率。为保证检索样本的全面性与准确性，本文通过以下步骤构建检索策略：

首先，将顾客和互动各自对应的中英文近义词拼接为检索词组合，中英文检索词之间分别使用“与”和“NEAR/0”连接，以保证检索词组合的整体性，使用的相关词汇如表 1 所示。其次，英文检索式以“OR”连接检索词组合，中文检索针对 CSSCI 数据库的限制，将结果通过脚本自动加以整合。最后余下的检索条件设置如下，时间跨度均设置为 2012—2021 年，对 WOS 和 CSSCI 数据库分别选择“Article”和“论文”的文献类型，通过“Topic”和“所有字段”进行检索，得到 2871 条英文记录和 622 条中文记录。

表 1 顾客互动领域中英文文献检索词汇

中文检索词	中文近义词	英文检索词	英文近义词
顾客	客户、消费者、用户	Customer	Client、Consumer、User
互动	交互、参与、契合、融入	Interaction	Participation、Involvement、Engagement

### 1.2 数据清洗

对检索结果进行数据清洗可进一步提高其有效性，潘玮等<sup>[11]</sup>采用的 DEAN 数据清洗流程可以有效地提升研究热点识别的准确性，本文参考该流程对以下 4 类对象进行数据清洗：

（1）Duplicates（重复数据）：本文以标题作为待查重字段，使用 Note Express 的数据去重功能进行数据处理，共去除英文重复文献

记录 3 条，中文重复文献记录 31 条。

（2）Errors（不合检索需求的数据）：本文按以下标准进行筛选，一是将文献类型为综述、会议通知、编者按及广告等记录予以排除；二是研究对象上仅考虑互动对象是企业的文献；三是在研究内容上主要选择经管类文献，不涉及具体算法；四是 involvement 相关文献仅选取与 participation 同义的文献。最终筛选出符合要求

的 392 条英文文献记录以及 347 条中文文献记录。

(3) Alias(相似数据): 主要包括两类情况, 一是语法异构, 即英文大小写与单复数造成的关键词重复, 以及 Citespace 自动将关键词小写并转化为单数而造成的识别错误。二是语义异构, 包括缩写和同(近)义词。本文通过编写 python 脚本获取中英文文献所有关键词, 并利用 Citespace 的节点合并与排除功能, 将识别错误的英文关键词进行修复, 同时合并相同概念的中英文关键词, 共创建英文节点规则 495 条, 中文节点规则 222 条。

(4) Noises(干扰数据): 主要指聚类时包含节点较少的子网络以及部分干扰关键词。本文在进行可视化与解读时会忽略包含节点较少的子网络, 对于由 WOS 数据库以及 IEEE 期刊自动分配的部分关键词(中文期刊不提供)

予以剔除, 仅使用作者提供的关键词。

## 2 前沿文献基本特征

### 2.1 发文量变化趋势

学科领域内每年发文量的变化能够反映其研究热度的变化趋势, 通过对顾客互动领域的中英文文献刊发情况进行初步统计分析(见图 1), 可以发现该领域每年发文量在 2012 年最低, 中文文献发文量于 2017 年达到峰值后开始回落, 至 2021 年又出现激增, 而英文文献发文量在 2016 年有所下降, 但 2019 年、2021 年均大幅增长。从整体来看, 顾客互动领域总发文量在 2016 年有所下降, 2017 年和 2021 年发文量剧增, 保持长期增长趋势, 因此该领域正处于高速发展期。

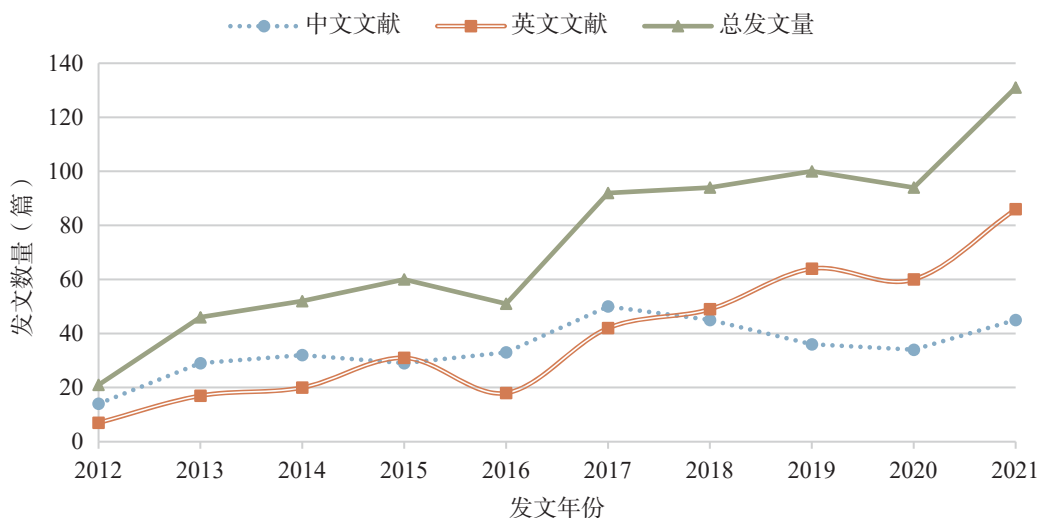


图 1 顾客互动文献逐年发文量

### 2.2 机构合作网络

学术机构发表论文数可以反映其在该领域的影响力, 机构的合作网络也能间接反映该领域学术研究的成熟度。本文选取每年发

文量位列前百的机构, 通过 pathfinder 算法绘制中、英文文献机构合作图谱。可得包含 167 所机构与 137 条合作关系的中文合作网络(图 2 左), 以及涵盖 560 所机构和 643 条合作关

系的英文合作网络(图2右)。图谱中节点大小代表该机构的发文量,节点存在紫色外圈则其中介中心性(网络中的重要度)显著(不

小于0.1),节点间连线粗细代表机构间的合作强度,连线颜色从深绿到浅绿代表合作时间的递增。

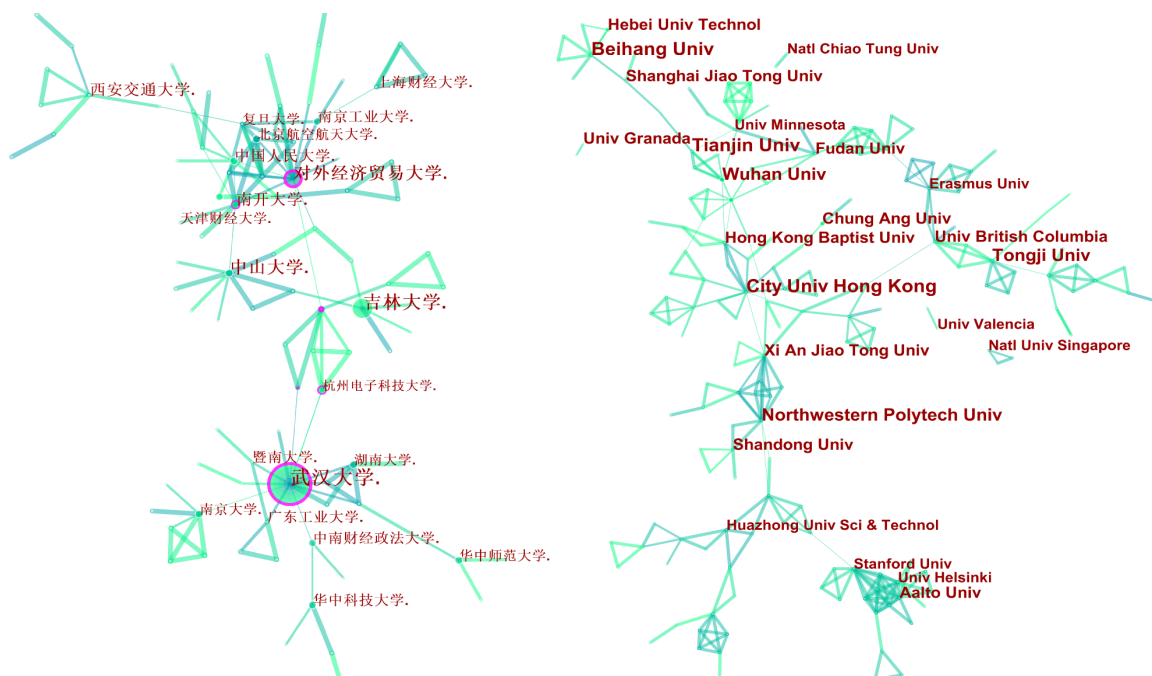


图2 中英文文献研究机构合作图谱

根据图2,顾客互动领域的相关中、英文文献都形成了连接性较高的合作机构网络,合作关系较为完备,研究领域成熟度较高。中文机构合作网络形成了一个由武汉大学与对外经济贸易大学为首的高校合作群,整体呈现“小而美”的特点,合作关系更为紧密,其高产机构发文量更大,影响力相对更高。而英文文献的合作网络中同样存在由香港城市大学、天津大学等组成的大型合作网络,但合作更为松散,整体规模较大,且一些国内高校是其重要组成部分。从发文量和中心性指标来看,英文文献中高产机构的影响力相对较低。由此可看出,关于顾客互动领域的研究,中、英文文献的研究主力都是国内高校和机构,英文文献中存在一

些国内和国际研究机构的紧密合作,但这种合作组织还比较少。

### 2.3 作者合作网络

核心作者能够推动学术研究的创新发展,其成果能够反映该领域的研究重点与发展动向,他们的学术合作保证了科研成果的知识共享与交流。本文选取每年发文量前100的作者,使用pathfinder算法绘制作者合作图谱(见图3,图中主要展示发文量为2篇以上的中文文献和英文文献作者)。另外,通过对发文量前10的高产作者进行统计分析(见表2),可以更好地展现顾客互动领域的核心作者合作情况。





图3 中英文文献作者合作图谱

中文文献发文量前十的高产作者中发文最多的姚山季、王永贵与刘德文、马双等学者组成了一个较大的合作团队，另外还存在姚唐等学者组建的高产团队，剩余的高产作者的合作关系则较为零散。非前十的学者中，张辉等也共同形成了一个较大的中文学者合作群。而英文文献作者合作网络中，发文量较多的高产作

者也有许多都是国内学者，形成了以 LINGYUN GUO 为中心的学者合作群，显然国内学者也发表了大量的英文文献，例如：LINGYUN GUO、TAIWEN FENG 等，是这个领域的研究主力，国内作者在英文国际文献中的作用在不断提升。整体而言，中英文文献中均存在高产作者合作团队，但更多的是较为分散的合作关系。

表2 中英文文献发文量前十的高产作者

发文量排名	中文文献作者	发文量	英文文献作者	发文量
1	姚山季	8	TAIWEN FENG*	5
2	王永贵	8	LINGYUN GUO*	5
3	刘德文	6	HUA (JONATHAN) YE*	4
4	邓胜利	6	MINGLI ZHANG*	4
5	马双	6	JIEUN LEE	3
6	于洪彦	6	XIAOLIANG SHEN*	3
7	张德鹏	5	PETRI PARVINEN	3
8	周涛	5	ATREYI KANKANHALLI	3
9	姚唐	5	YONGQIANG SUN*	3
10	宁连举	4	DAHUI LI	3

\* 英文期刊中高产的国内作者（通过 citespace 查询作者机构整理得到）

### 3 文献基础知识结构

#### 3.1 中文文献共被引分析

一个学科领域的被引文献通常被视为施引文献的知识基础，若同时被多篇论文引证则称

之为文献共被引，它能展现施引文献间的关系，从而对基础知识结构进行系统化的分析。故本文将参考文献作为节点类型，以文献 g 指数为阈值 (k=25)，使用 pathfinder 算法得到文献共被引图谱(见图 4)，红色外圈代表突现强度(被突发集中引用的强度)较高。

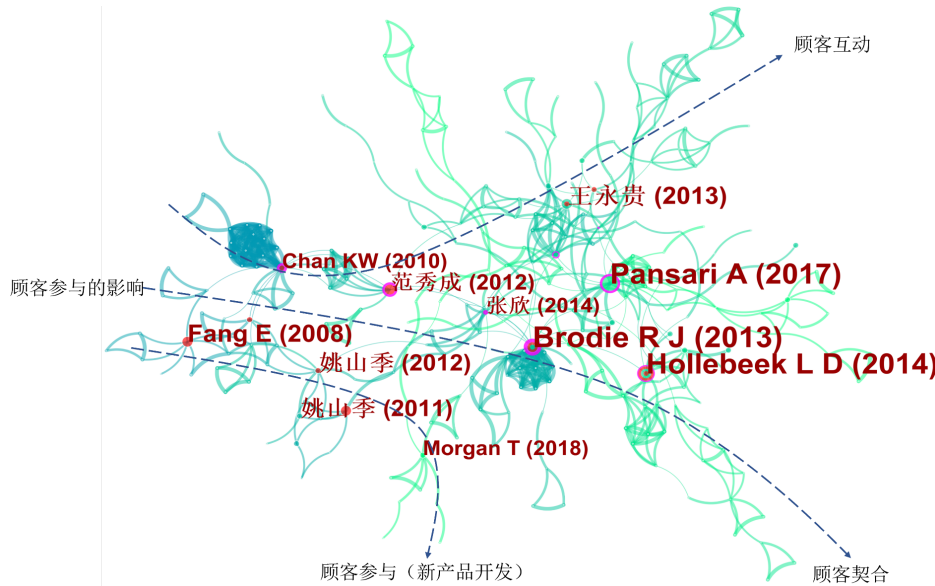


图 4 中文文献基础知识结构图谱

为揭示重要的被引文献信息，这里选取被引频次、中心性、突现强度前五，突现年份活

跃至 2022 年的文献作为关键被引文献(见表 3)，并以发文年份排序。

表 3 中文关键被引文献

第一作者	期刊	年份	被引频次	中心性	突现强度	突现年份
Fang	MARKETING	2008	9	0.04	5.2	2012-2013
Chan	MARKETING	2010	6	0.28	2.4	2012-2015
姚山季	科学学与科学技术管理	2011	8	0.09	3.59	2014-2016
姚山季	管理工程学报	2012	8	0.01	3.59	2014-2016
范秀成	管理评论	2012	7	0.33	2.59	2015-2017
Brodie	BUS RES	2013	15	0.5	3.27	2016-2018
王永贵	管理学报	2013	9	0.1	2.78	2016-2018
Hollebeek	INT MARKETING	2014	14	0.18	4.27	2017-2019
张欣	管理评论	2014	6	0.21	0	—
Pansari	ACAD MARKET SCI	2017	14	0.3	3.2	2020-2021
Chan	RES POLICY	2018	6	0.06	2.85	2020-2021

表3中, Brodie<sup>[12]</sup>于2013年发表的文献剖析了顾客契合的基本主题及其前因后果,为今后顾客契合的相关研究奠定了理论基础。该文献被引频次、中心性均为最高,突现强度位列第五,反映了其在中文被引文献网络中的重要地位。结合图4可以看出,该文献是基础知识结构的转折点,在2016年到2018年被前沿文献集中引用,而其他关键被引文献的突现年份均分布于2016年前后,故本文将国内基础知识结构的演变划分为2016年前后两个阶段。

(1) 2016年以前,前沿文献主要关注顾客参与新产品开发的基础研究,另一重要研究成果是对顾客参与“双刃剑”效果的检验与分析。

在顾客参与新产品开发领域,首先引起前沿文献关注的是Fang<sup>[13]</sup>首次探讨了客户参与对新产品开发绩效的影响。随后姚山季<sup>[14-15]</sup>发表的两篇论文成为焦点文献,前者探讨了顾客参与各维度对技术创新绩效的不同影响,后者引入了关系嵌入概念,对顾客参与对新产品开发绩效的影响机制进行了分析。此外,相关影响因素也得到了进一步的关注,如张欣<sup>[16]</sup>立足于关系视角对顾客参与新产品开发的驱动因素进行了实证分析。

另一方面,顾客参与存在利弊两方面的影响这一研究成果同样十分重要,有两篇中心性特别显著的相关文献(中心性排名第二和第三,见表3)。其中Chan<sup>[17]</sup>实证检验了顾客参与对顾客和员工的价值创造和满意度的影响有利有弊,是突现时间最长的被引文献。随后范秀成<sup>[18]</sup>也对顾客参与利弊得失的研究成果进行了综述。

(2) 2016年以后,由图4以及相关关键

被引文献信息可以看出,基础文献整体知识结构可以分为三个主要的研究方向,即顾客互动、顾客参与和顾客契合。

其中王永贵<sup>[19]</sup>在2013年发表的论文系统地剖析了顾客互动,运用实用—享乐理论探讨了顾客互动的驱动因素,该文献引用频次较高(第三),是2016至2018年前沿文献中相关研究的重要支撑。顾客参与方向上,Morgan<sup>[20]</sup>研究顾客参与新产品开发时,分析了创新性的中介效应以及企业吸收能力的影响,其研究成果受到了不少最新前沿文献的关注。顾客契合方向除Brodie外的研究成果还包括,Hollebeck<sup>[21]</sup>将顾客品牌契合概念化为认知、情感和行为三个维度;Pansari<sup>[22]</sup>对顾客契合的结构、前因与后果进行了理论分析,提供了新的概念框架,该文献被引频次并列第二,中心性十分显著,被2020至2021年前沿文献集中引用,其研究成果很可能成为今后的研究前沿的基础。

### 3.2 英文文献共被引分析

类似地,对英文文献重复上节的分析步骤,可得包含445篇英文文献、共1000条共现关系的共被引网络(见图5)和13篇英文关键被引文献(见表4)。

在所有英文关键被引文献中,Goh<sup>[23]</sup>在2013年发表的文献量化了用户生成内容(UGC)和营销人员生成内容(MGC)对消费者服装购买支出的影响,其中心性最为显著,自2017年起被多篇前沿文献引用,是整个英文共被引网络中最重要的节点。本文结合图5和表4,将国外基础知识结构的演变过程划分为2017年前后两个阶段。

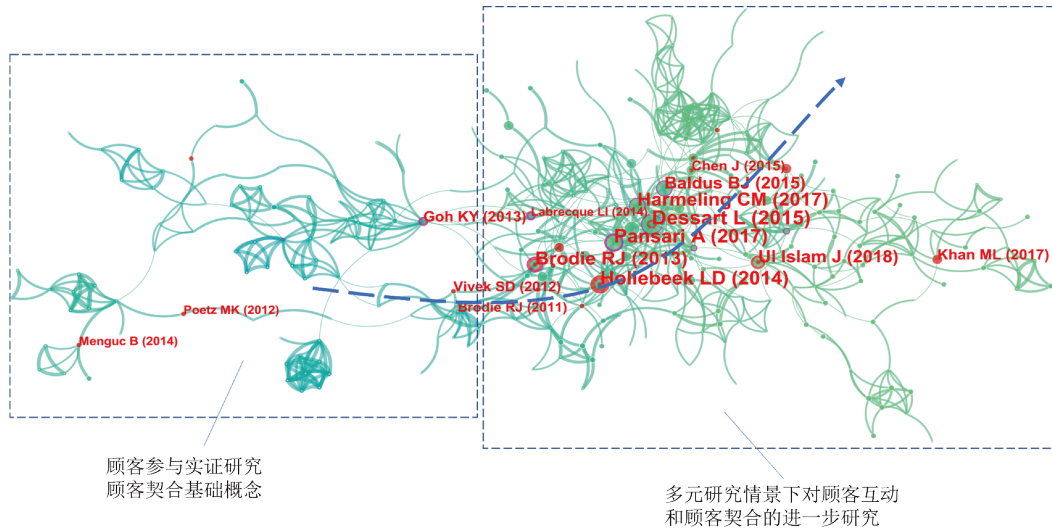


图 5 英文文献基础知识结构图谱

表 4 英文关键被引文献

第一作者	期刊	年份	被引频次	中心性	突现强度	突现年份
Brodie	SERV RES-US	2011	6	0.09	3.13	2014-2016
Vivek	MARKET THEORY PRAC	2012	7	0.01	3.53	2016-2017
Brodie	BUS RES	2013	16	0.17	4.54	2015-2018
Goh	INFORM SYST RES	2013	8	0.32	0	—
Hollebeek	INTERACT MARK	2014	18	0.07	4.48	2017-2019
Labrecque	INTERACT MARK	2014	5	0.21	0	—
Dessart	PROD BRAND MANAG	2015	19	0.14	3.1	2018-2021
Baldus	BUS RES	2015	12	0.2	0	—
Chen	DECIS SUPPORT SYST	2015	7	0.01	2.88	2019-2021
Pansari	ACAD MARKET SCI	2017	18	0.2	0	—
Harmeling	ACAD MARKET SCI	2017	15	0.13	0	—
Khan	COMPUT HUM BEHAV	2017	8	0.01	3.03	2020-2021
Islam	INTERNET RES	2018	13	0.06	3.28	2019-2021

(1) 2017 年以前，英文前沿文献关注的文献包括顾客参与和顾客契合两个主题。

除开 Goh<sup>[23]</sup> 的论文，从被引频次来看，顾客参与的研究成果还有 Menguc<sup>[24]</sup> 利用资源基础观研究客户和供应商参与对新产品性能的影响；Poetz<sup>[25]</sup> 验证了众包的有效性，首次将专业人员的想法与用户产生的想法进行了比较。另一主题顾客契合的主要研究成果有 Brodie<sup>[26]</sup> 基

于关系营销理论和服务主导 (S-D) 逻辑，探讨了顾客契合的理论基础。随后 Vivek<sup>[27]</sup> 探讨了顾客契合的性质和范围，并给出了包含契合前因后果的理论模型。

(2) 2017 年之后，英文基础文献更侧重于特定情境下互动和契合两大理论概念的研究。

2017—2019 年，顾客互动主题前沿文献关注的基础文献有 Labrecque<sup>[28]</sup> 探究了社交媒体



中准社会互动理论作为互动策略的影响。顾客契合主题下诸多契合定义及测量维度不断涌现,如 Brodie<sup>[12]</sup>、Hollebeek<sup>[21]</sup> 分别对顾客契合进行了理论分析与量表开发; Baldus<sup>[29]</sup> 对于顾客契合的定义与维度划分与前者迥然不同; Pansari<sup>[22]</sup>、Harmeling<sup>[30]</sup> 分别对顾客契合概念、营销理论进行了分析,也提出了新的契合定义。其中 Brodie、Hollebeek 和 Pansari 的研究在国内外均是关键被引文献,说明国内外顾客契合研究存在相近之处。

2020—2021 年,受到前沿文献集中引用的基础研究有 Chen<sup>[31]</sup> 对社交商务情境下消费者行为的研究,说明社交商务平台有潜力成为未来重要的研究情境。此外 Dessart<sup>[32]</sup> 在 Hollebeek 的研究基础上,确定了在线品牌社区消费者契

合的概念框架。Islam<sup>[33]</sup> 则同样于在线品牌社区情境下,研究自我品牌形象一致性和价值一致性对顾客契合的影响。上述两篇文献的被引频次均较高,前者中心性也较为显著,很可能成为今后的研究前沿。另外还有 Khan<sup>[34]</sup> 研究了 YouTube 用户契合,说明 Youtube 等视频社区可能成为今后热门的研究情境。

### 3.3 结果讨论

结合上述两节对国内外关键基础文献的分析,本文通过研读关键被引文献,对这三个理论概念的区别与联系进行梳理,同时根据国内外顾客互动领域知识结构的演变脉络进行对比分析,进一步揭示国内外基础知识结构的异同点。

表 5 基于关键被引文献的理论概念汇总

理论概念	第一作者及年份	定义	三个理论概念间的联系
顾客参与	Fang <sup>[13]</sup> , 2008 姚山季 <sup>[14]</sup> , 2011a 姚山季 <sup>[15]</sup> , 2011b	顾客参与新产品开发活动的程度	顾客参与新产品开发活动本质上就是一种顾客互动行为,且顾客参与需要通过顾客互动来实现最终的价值共创
	Morgan <sup>[20]</sup> , 2018 张欣 <sup>[16]</sup> , 2014		
	Chan <sup>[17]</sup> , 2010 范秀成 <sup>[18]</sup> , 2012	顾客涉入产品或服务的生产和传递的程度(投入的资源与行为)	
顾客契合	Brodie <sup>[26]</sup> , 2011 Brodie <sup>[12]</sup> , 2013 Hollebeek <sup>[21]</sup> , 2014 Dessart <sup>[32]</sup> , 2015 Islam <sup>[33]</sup> , 2018	消费者与品牌和/或社区其他成员之间的特定互动体验下的认知、情感和行为水平	顾客契合是由顾客的互动体验产生的,顾客参与是其前因
	Vivek <sup>[27]</sup> , 2012	个人参与产品、企业或顾客发起活动的强度	通过顾客参与活动的强度来定义,且企业和顾客之间的双向互动是顾客契合的核心
	Baldus <sup>[29]</sup> , 2015	与在线品牌社区互动的内在动机	直接以顾客互动动机或行为类型作为定义和测量维度
	Khan <sup>[34]</sup> , 2017	个人与媒体的互动行为	
	Pansari <sup>[22]</sup> , 2017 Harmeling <sup>[30]</sup> , 2017	顾客通过直接或/和间接贡献为公司增加价值的行为	类似于顾客参与的定义,且顾客贡献是通过与企业或其他顾客的互动产生的
顾客互动	王永贵 <sup>[19]</sup> , 2013	顾客为了获取或者贡献价值,在社区中进行的各种互动	特定的媒体环境下(如虚拟品牌社区)的顾客互动与顾客参与同样也是本质相同的现象
	Labrecque <sup>[28]</sup> , 2014	顾客通过互动感受到的,与品牌联系在一起的感觉	

(1) 从涉及的理论概念来看,都包含对参与、契合和互动这三个理论概念的研究。

参与、契合和互动三个理论概念都是关系营销、顾客主导逻辑和价值共创理论不断发展的产物。因此,互动作为上述理论的核心,自然也成为了三者间联系的重要桥梁。在顾客参与中最基础的理论框架就是通过顾客互动实现价值共创<sup>[17]</sup>。顾客契合的定义与测量维度尚无定论,但顾客互动也是其理论核心。例如受到广泛认可的契合定义是由 Brodie 等<sup>[26]</sup>提出的将契合视作由互动和体验产生的心理状态;Pansari 等<sup>[30]</sup>则参考顾客参与的定义,将契合与顾客的投入与贡献联系起来,而顾客贡献又是通过互动来产生的;Vivek<sup>[27]</sup>的观点则介于二者之间,以顾客参与活动的强度定义契合,维度与 Brodie 类似;Baldus<sup>[29]</sup>和 Khan<sup>[34]</sup>更是直接把顾客互动的动机或行为定义为顾客契合。顾客互动含义更加宽泛,但其本质与特定情境下的参与、契合是类似的,有学者指出参与和契合是两个本质相同的现象<sup>[32]</sup>,因此三者的相关文献也常作为彼此的参考文献出现。

(2) 从演进脉络来看,国内外基础知识结构的演变主线较为相似,但侧重点有所不同。

国内外基础文献的演进主线都是从参与、互动转变为契合,但国外英文文献主要侧重于对契合的研究,而国内还有新产品开发以及互动的支线。相较于国外英文文献,一方面,国内前沿文献对于契合基础概念定义的关注点有差异,如以顾客互动的动机或类型来定义的契合是国内前沿文献未曾重点关注的。另一方面,对契合在不同情境和理论背景下的研究,没有国外那么丰富和深入,例如 Islam<sup>[33]</sup>引入一致性

理论对在线品牌社区中顾客契合的前因后果进行实证检验的文献是国内前沿文献较少关注到的。

(3) 从基础文献中凸显的研究前沿来看,今后国内外的研究前沿可能也各有侧重。

关键被引文献所凸显出的国内外顾客契合研究前沿也大不相同,Pansari 等<sup>[22]</sup>提出的以顾客贡献为核心定义的概念框架,今后可能会成为国内研究的关注焦点,而 Brodie<sup>[12]</sup>提出的顾客契合概念框架在国外仍是主流,且其研究前沿更加注重在不同研究情境,如虚拟品牌社区、社交商务、视频社区下的实证研究。

## 4 研究热点与前沿

### 4.1 关键词共现分析

关键词是文献重要组成部分之一,其高度凝练了文献的研究方向和内容,通过分析关键词可以界定出研究热点。本文以关键词为节点类型,选取每年出现频次 top100 关键词,使用 pathfinder 裁剪算法得到关键词共现网络,以此得到共现频次前十的关键词(见表 6)。

可以看出,中文文献高频关键词中,顾客参与出现的频次最多,顾客契合、价值共创紧随其后,且中心度相对较高,这表明在国内研究中,参与和契合是最重要的概念,价值共创是其理论基础。另外,频次较高且中心性显著的顾客信任、影响因素和顾客互动等关键词说明信任等影响因素是较重要的研究对象。而余下的社交媒体是其常见的研究情境,用户信息交互、顾客行为则是其研究重点。

相较而言,英文文献中出现频次最多的关键词是 customer engagement, 而 customer participation 位列其后,说明国外关注的理论概念仍是参与和契合,且相较于中文文献很少将参与和契合加以区分,英文文献更重视契合。其后的 social media、online brand community、

online community 和 social commerce 点明了最常出现的研究情境,较之国内要更为丰富。余下的 value co-creation、customer innovation 则说明英文文献的理论背景与国内研究类似。最后高频关键词 social identification 表明英文文献对影响因素的侧重点与国内存在一定差异。

表 6 Top10 高频关键词

序号	中文关键词	频次	中心度	英文关键词	频次	中心度
1	顾客参与	103	0.68	customer engagement	59	0.41
2	价值共创	30	0.25	social media	50	0.44
3	顾客契合	28	0.16	customer participation	37	0.39
4	影响因素	17	0.09	online brand community	19	0.09
5	顾客信任	17	0.13	online community	18	0.2
6	顾客互动	15	0.11	value co-creation	18	0.12
7	顾客	11	0.11	social commerce	11	0.06
8	参与意愿	10	0.08	customer innovation	11	0.1
9	用户信息交互	9	0.04	customer behavior	9	0.1
10	知识共享	9	0.03	social identification	9	0.09

## 4.2 中文关键词聚类分析

关键词的共现展现了关键词之间的联系,因此根据关键词共现数据进行聚类分析,可以得到前沿文献的研究主题与研究热点。聚类效果主要由模块值(Q值)和平均轮廓值(S值)来衡量,一般Q值大于0.3则聚类效果显著,S值在0.7以上则聚类结果是令人信服的。本文得到的中文聚类图谱Q值为0.7333,S值为0.9269,表明聚类网络结构显著且聚类效果合理。

本文结合 citespace 自带的聚类算法得到的类名以及相应的关键词,微调后得到聚类标签,以时间线聚类图谱进行可视化分析(见图6)。该图谱显示,中文文献关键词主要可划分为五类,且所有聚类均持续至2021年,能够反映国内的研究热点。

#0 顾客参与和 #2 参与意愿这两个聚类涵盖了对顾客参与前因后果的研究成果。前者是包含中文关键词最多的聚类,主要关注对顾客参与的结果变量,特别是顾客满意与顾客参与创新的影响,且吸收能力在顾客参与创新中的影响是研究重点之一<sup>[35]</sup>。后者的研究重点是顾客参与的前因包括参与意愿和动机,其中顾客信任<sup>[36]</sup>和顾客间社会互动<sup>[37]</sup>分别在2015年和2019年左右成为该聚类的研究热点。

#1 用户聚类展现了国内研究以用户为中心的研究视角,社交媒体<sup>[38]</sup>一直是其常见的研究场景,这与高频关键词分析的结论类似。该视角下的研究内容从14年起集中于对相关影响因素的分析<sup>[39]</sup>,17年时研究影响因素的文献数量开始衰退,用户信息交互转而成为新一轮的研

研究热点<sup>[40]</sup>，结构方程模型是该聚类中常见的研究方法<sup>[41]</sup>。#3 在线互动聚类则是关于用户在线互动过程的相关研究，研究热点包括互动产生的心流体验<sup>[42]</sup>以及进一步形成的顾客契合<sup>[43]</sup>。

#4 价值共创是中文文献在顾客互动领域中常见的理论背景与研究方向，研究内容主要包括顾客价值<sup>[44]</sup>和感知价值<sup>[45]</sup>等，其常见的研究情境是在线品牌社区<sup>[46-47]</sup>。

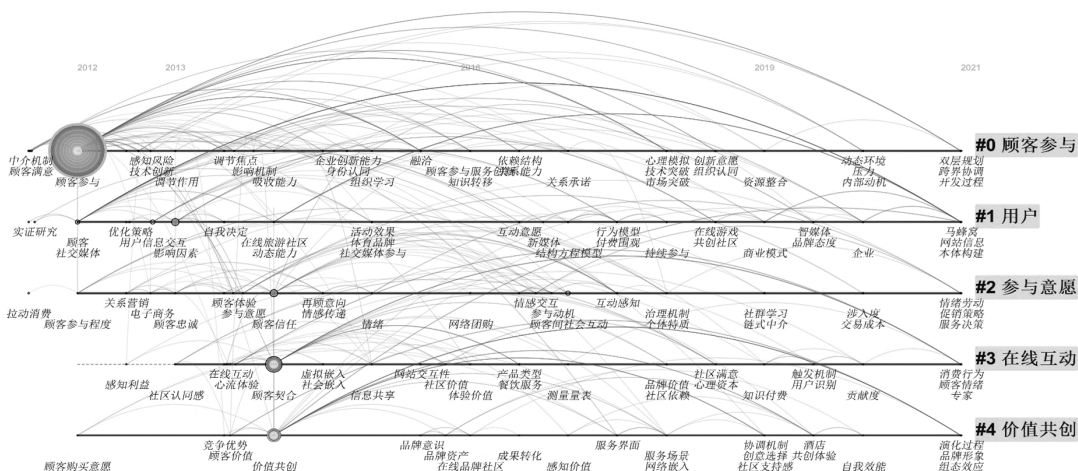


图6 中文关键词聚类图谱

综上所述，国内的研究热点可以分为三个主要维度，一是研究内容维度，对顾客参与前因后果的分析，特别是顾客参与与创新中吸收能力的影响是最新的研究热点，另外在线互动中形成的顾客契合也是目前的研究前沿。二是研究范式维度，从用户视角出发，以社交媒体为情境，使用结构方程模型进行的实证研究是主要的研究方式。三是理论背景维度，虚拟品牌社区下的价值共创是近来热门的理论背景。

### 4.3 英文关键词聚类分析

类似地，重复上节步骤得到英文聚类图谱(见图7)，其Q值为0.7562，S值为0.9074，说明聚类网络结构显著且聚类效果合理。英文文献关键词可划分为十类，且大部分聚类活跃至2021年，能够反映国外顾客领域的研究热点。图谱显示，#4 新产品开发是唯一在2021年中断的聚类，

研究内容包括以用户为中心的设计等<sup>[48]</sup>，案例研究是该聚类中常用的研究方法<sup>[49]</sup>。#0 社交媒体作为包含关键词最多的聚类，聚焦于社交媒体下的顾客互动，最新的研究热点是品牌帖子互动<sup>[50]</sup>。#3 在线品牌社区关注在线品牌社区互动，社交商务的兴起是其研究背景<sup>[51]</sup>，社会认同是其研究热点<sup>[52]</sup>。#9 在线社区是#3的拓展，成员间知识共享是其2021年的研究热点<sup>[53]</sup>。

上面三组聚类代表着重点关注的研究情境，而#1 顾客参与、#2 顾客契合和#5 顾客互动则是最重要的三个理论概念，企业社会责任<sup>[54]</sup>、顾客忠诚<sup>[55]</sup>和顾客关系管理<sup>[56]</sup>分别是三者近年的研究热点。#8 顾客契合行为是#2的衍生聚类，社交媒体营销<sup>[57]</sup>是其常见的研究情境，最近的研究热点是顾客信任<sup>[59]</sup>。#5 顾客互动集中展示了较热门的研究方法，如结构方程模型<sup>[60]</sup>、社会网络分析<sup>[61]</sup>、数据



挖掘<sup>[56]</sup>、内容分析<sup>[62]</sup>、潜在类别分析<sup>[63]</sup>等。  
#6 价值共创中社会互动<sup>[64]</sup>是其研究热点，#7

顾客创新中开放式创新<sup>[65]</sup>与开源<sup>[52]</sup>是较热门的研究背景。

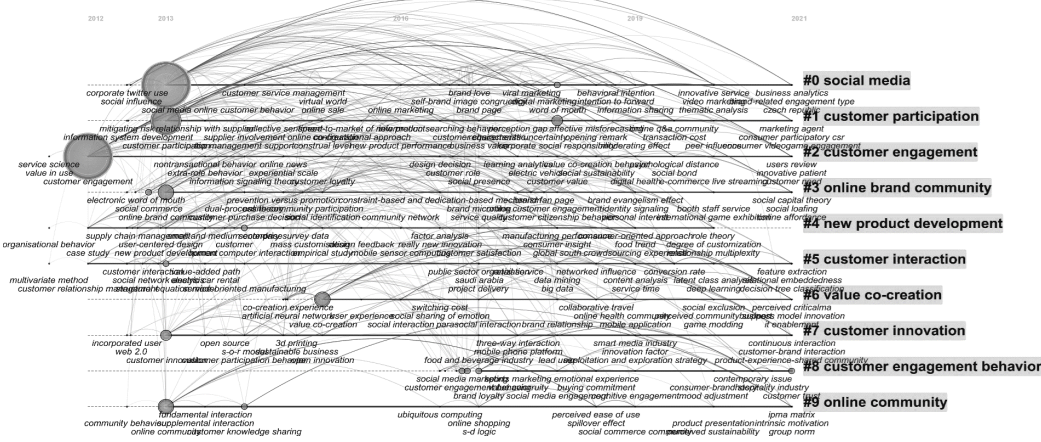


图7 英文关键词聚类图谱

总体而言，国外研究热点同样可以分为三个方面，首先研究情境上，随着社交商务的兴起，多以社交媒体和在线社区特别是在线品牌社区为研究情境，其中社交媒体品牌帖子互动以及在线社区成员间知识共享是近年来的研究热点。其次理论概念的研究上，主要关注参与、契合与互动三个概念，但更注重对契合的研究，采用的研究方法也更加丰富。最后理论背景方面与国内类似，都以价值共创与顾客创新为研究热点。另外，从各

聚类间节点的跨层连线来看，国内外对于顾客互动的研究并不局限于单一维度，集合多个热点的研究较为常见。

#### 4.4 关键词突现分析

在关键词共现分析的基础上，保持其他参数不变，进行突现词检测，关键词突现强度可以反映某一研究领域未来的发展前沿，从而揭示其发展趋势。如表7所示，从中英文文献中各选取10个突现关键词。

表7 Top10 核心突现关键词

序号	中文突现关键词	突现强度	突现时间	英文突现关键词	突现强度	突现时间
1	顾客互动	1.62	2012-2013	social commerce	2.19	2013-2014
2	新产品开发	2.65	2012-2014	user innovation	2.68	2014-2015
3	客户参与	3.78	2013-2014	customer involvement	1.27	2014-2019
4	电子商务	1.79	2013-2015	community participation	1.35	2015-2017
5	调节效应	1.74	2014-2016	community engagement	1.49	2016-2018
6	网络购物	1.58	2014-2015	social media	2.72	2017-2018
7	顾客参与创新	2.61	2015-2016	purchase intention	1.73	2017-2018
8	顾客间互动	2.34	2015-2017	online brand community	2.07	2018-2021
9	社会化媒体	1.65	2017-2018	brand loyalty	1.37	2018-2019
10	在线健康社区	2.12	2019-2021	customer loyalty	1.93	2019-2021

中文文献中顾客互动、新产品开发和客户参与作为最早一批突现的关键词，其突现强度也极高，与前文中文文献早期关注顾客参与新产品开发领域不谋而合。2014年左右中文文献中对电子商务、网络购物相关背景下的顾客互动研究逐步突现，对于影响机制的相关研究也逐步深入至调节效应。2015年后顾客参与创新和顾客间互动成为突现关键词，在前面的文献共被引分析与关键词聚类分析中，确实也体现了这两者已成为目前的研究热点。2017年以来社会化媒体和在线健康社区成为新的突现关键词，特别是后者的突现时间持续到了2021年，很有可能成为今后的研究热点。

相较而言，英文文献中社交商务、用户创新突现时间较早、突现强度较高，是较为重要的研究背景，且根据前文分析可知二者随后重新受到了学者的关注。顾客参与、社区参与和社区契合三个关键词突现强度一般，但均突现了较长时间，是该领域主要的研究对象。社交媒体是顾客互动领域重要的研究背景，它和购物意愿均在2017年后突现，且突现强度较高，结合前文分析可知是顾客互动领域的研究热点之一。在线品牌社区、品牌忠诚和顾客忠诚三个关键词有两个突现至2021年，且突现强度较高，说明品牌社区及其对顾客忠诚的影响很有可能成为今后的研究热点，这和前文的结论能够相互印证。

## 5 结论与展望

### 5.1 研究结论

本文运用 Citespace 文献计量可视化软件，

从前沿文献特征、基础知识结构和研究热点等多个方面，梳理顾客互动领域的相关研究，通过前文的分析和讨论，得出如下结论：

(1) 从前沿文献特征上来看，国内外顾客互动领域的研究合作均较为紧密，且国内高校在中外合作网络中占据重要地位。发文量趋势显示，近年来国内外期刊的发文量均处于高速发展期，预测今后顾客互动领域的研究将会迎来新的研究高潮。从机构与作者合作情况来看，国内外机构与作者之间均存在一定规模的合作团队，特别是研究机构间的合作是较为紧密的，并且国内高校在国内外的机构合作网络中的影响力均较高。

(2) 从研究基础上来看，国内外基础文献知识结构的演进脉络都包含从参与、互动转变到契合的主线，但国内还存在部分不同的演进支线，且显示出的研究前沿也各有侧重。从演进脉络来看，国内基础文献的研究范围要更加广泛，但对顾客契合相关研究的关注有所缺失，且对于顾客契合在不同应用情境和理论背景下的理论和实证研究，不如国外基础研究那么丰富和深入。从研究前沿来看，契合的关注点上国内与国外有所不同，且今后仍会关注顾客参与新产品开发等领域，而国外前沿文献会更加注重不同情境下对顾客契合的研究。

(3) 从研究热点上来看，国内外前沿文献的研究内容本同末异，但国外前沿文献的研究情境与方法相较国内要更加丰富。根据对关键词和聚类图谱的分析，可以发现国内外研究的理论背景都包含价值共创和顾客创新理论，研究对象中顾客行为均是其研究热点及研究前沿，且研究的主题都是参与、互动与契合这三个主

要的理论概念,但国内研究更关注参与,国外研究则偏向于契合。另外与文献共被引分析的结论类似,国外前沿文献对不同情境下的理论与实证研究要更为丰富,且国外前沿文献的研究方法更多样化。

## 5.2 研究展望

通过对基础文献演变过程和关键词聚类前沿关键词的分析,可以在一定程度上把握今后顾客互动领域的发展趋势,再根据国内外相关研究的差异提出以下研究展望:

(1) 研究视角上可以补充企业或环境视角下的相关变量。国内大部分文献都从顾客的视角出发研究相关影响变量,而考察企业或环境视角下能力、供给或者其他特质的影响,能够更加契合顾客和企业双向互动的特性。近几年来前沿文献引用了不少涉及相关变量的基础文献,说明分析企业或环境视角下的影响因素可能成为今后热门的研究趋势。

(2) 研究内容上探究多元化情境下的顾客互动,形成顾客互动和研究情境的良好匹配。国外前沿研究中对于不同情境下的理论应用研究相较国内要更加深入,能够更好地契合当下各类新兴的媒体或平台对顾客互动内容与形式的改变。根据对前沿热点和最新引用的基础文献的分析可知,目前社交商务平台、短视频平台、在线品牌社区、开源社区等很可能成为今后的热门研究情境,值得国内学者加以关注。

(3) 研究方法上可采用多元化的分析方法,从而得到更丰富的研究成果。从研究方法上来看,国内研究往往采用结构方程模型作为实证研究方法,而国外文献常使用数据挖掘、深度

学习、fsqca等研究方法,利用不同的分析框架,得到更为丰富的研究成果。由此提出两个可能的研究方向,一是通过更多样化的数据获取手段,将数据来源多元化;二是利用不同的分析方法,得到更丰富的研究成果,如fsqca等组态视角下的分析方法或结合质性研究与实证研究的混合研究方法等。

## 参考文献

- [1] 张欣,杨志勇,王永贵.顾客互动前沿研究——内涵、维度、测量与理论演进脉络述评[J].国际商务(对外经济贸易大学学报),2014(4):86-94.
- [2] 朱晓红,赵金国,张金霞.企业-用户互动对创新企业创新绩效的影响——迭代创新的研究视角[J].山东社会科学,2019(9):136-141.
- [3] 肖静华,胡杨颂,吴瑶.成长品:数据驱动的企业与用户互动创新案例研究[J].管理世界,2020,36(3):183-205.
- [4] LIU P, LI M, DAI D, et al. The effects of social commerce environmental characteristics on customers' purchase intentions: The chain mediating effect of customer-to-customer interaction and customer-perceived value[J]. Electronic Commerce Research and Applications, 2021, 48: 101073.
- [5] 李志兰.顾客间互动研究综述与展望[J].外国经济与管理,2015,37(12):73-85.
- [6] 刘好强.顾客间互动研究综述与未来展望[J].技术与管理研究,2014(10):56-60.
- [7] 宁连举,孙中原,刘茜.基于知识图谱的顾客契合研究热点与趋势评述[J].科研管理,2019,40(12):213-224.
- [8] 荆宁宁,李德峰.顾客契合研究综述[J].外国经济与管理,2015,37(7):33-45.
- [9] 贾维辰,李文光,余明媚.中文期刊知识图谱研究范式的优化[J].中国远程教育,2020(11):1-10,76.
- [10] CHEN CM. Science Mapping: A Systematic Review of the Literature[J]. Journal of Data and Information Science, 2017, 2(2): 1-40.
- [11] 潘玮,郑鹏,黄锦泉,等.基于数据清洗“DEAN”流程的健康信息领域研究热点探测[J].现代情报,

- 2018, 38(10): 73-77.
- [12] BRODIE R J, ILIC A, JURIC B, et al. Consumer engagement in a virtual brand community: An exploratory analysis[J]. *Journal of Business Research*, 2013, 66(1): 105-114.
- [13] FANG E. Customer Participation and the Trade-Off between New Product Innovativeness and Speed to Market[J]. *Journal of Marketing*, 2008, 72(4): 90-104.
- [14] 姚山季, 王永贵. 顾客参与新产品开发对企业技术创新绩效的影响机制——基于 B-B 情境下的实证研究 [J]. *科学学与科学技术管理*, 2011, 32(5): 34-41.
- [15] 姚山季, 王永贵. 顾客参与新产品开发及其绩效影响: 关系嵌入的中介机制 [J]. *管理工程学报*, 2012, 26(4): 39-48, 83.
- [16] 张欣, 姚山季, 王永贵. 顾客参与新产品开发的驱动因素: 关系视角的影响机制 [J]. *管理评论*, 2014, 26(5): 99-110.
- [17] CHAN KW, YIM CK, LAM S SK, et al. Is Customer Participation in Value Creation a Double-Edged Sword? Evidence from Professional Financial Services Across Cultures[J]. *Journal of Marketing*, 2010, 74(3): 48-64.
- [18] 范秀成, 杜琰琰. 顾客参与是一把“双刃剑”——顾客参与影响价值创造的研究述评 [J]. *管理评论*, 2012, 24(12): 64-71.
- [19] 王永贵, 马双. 虚拟品牌社区顾客互动的驱动因素及对顾客满意影响的实证研究 [J]. *管理学报*, 2013, 10(9): 1375-1383.
- [20] MORGAN T, OBAL M, ANOKHIN S. Customer participation and new product performance: Towards the understanding of the mechanisms and key contingencies[J]. *Research Policy*, 2018, 47(2): 498-510.
- [21] HOLLEBEEK L D, GLYNN M S, BRODIE R J. Consumer Brand Engagement in Social Media: Conceptualization, Scale Development and Validation[J]. *Journal of Interactive Marketing*, 2014, 28(2): 149-165.
- [22] PANSARI A, KUMAR V. Customer engagement: the construct, antecedents, and consequences[J]. *Journal of the Academy of Marketing Science*, 2017, 45(3): 294-311.
- [23] GOH K Y, HENG C S, LIN Z J. Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content[J]. *Information Systems Research*, 2013, 24(1): 88-107.
- [24] MENGUC B, AUH S, YANNOPOULOS P. Customer and Supplier Involvement in Design: The Moderating Role of Incremental and Radical Innovation Capability[J]. *Journal of Product Innovation Management*, 2014, 31(2): 313-328.
- [25] POETZ M K, SCHREIER M. The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas?[J]. *Journal of Product Innovation Management*, 2012, 29(2): 245-256.
- [26] BRODIE R J, HOLLEBEEK L D, JURIC B, et al. Customer Engagement: Conceptual Domain, Fundamental Propositions, and Implications for Research[J]. *Journal of Service Research*, 2011, 14(3): 252-271.
- [27] VIVEK S D, BEATTY S E, MORGAN R M. Customer Engagement: Exploring Customer Relationships Beyond Purchase[J]. *Journal of Marketing Theory and Practice*, 2012, 20(2): 127-145.
- [28] LABRECQUE L I. Fostering Consumer-Brand Relationships in Social Media Environments: The Role of Parasocial Interaction[J]. *Journal of Interactive Marketing*, 2014, 28(2): 134-148.
- [29] BALDUS B J, VOORHEES C, CALANTONE R. Online brand community engagement: Scale development and validation[J]. *Journal of Business Research*, 2015, 68(5): 978-985.
- [30] HARMELING C M, MOFFETT J W, ARNOLD M J, et al. Toward a theory of customer engagement marketing[J]. *Journal of the Academy of Marketing Science*, 2017, 45(3): 312-335.
- [31] CHEN J, SHEN X L. Consumers' decisions in social commerce context: An empirical investigation[J]. *Decision Support Systems*, 2015(79): 55-64.



- [32] DESSART L, VELOUTSOU C, MORGAN-THOMAS A. Consumer engagement in online brand communities: a social media perspective[J]. *Journal of Product & Brand Management*, 2015, 24(1): 28-42.
- [33] ISLAM J UI, RAHMAN Z, HOLLEBEEK L D. Consumer engagement in online brand communities: a solicitation of congruity theory[J]. *Internet Research*, 2018, 28(1): 23-45.
- [34] KHAN M L. Social media engagement: What motivates user participation and consumption on YouTube?[J]. *Computers in Human Behavior*, 2017(66): 236-247.
- [35] 许晖, 丁超, 王亚君, 等. 海外客户参与视角下的远程客户信任构建机制——基于外贸企业艾永特和海安格力的双案例研究[J]. *管理学报*, 2020, 17(9): 1277-1286.
- [36] 徐冬莉, 武文珍, 江若尘. 服务业顾客教育对顾客参与的影响——基于SOR理论和社会互动理论的视角[J]. *财贸研究*, 2015, 26(6): 97-103.
- [37] 周涛, 陈可鑫, 邓胜利. 社群学习用户持续参与行为机理研究[J]. *现代情报*, 2019, 39(1): 43-50.
- [38] 刘玉琦, 陈洁, 韦俊龙. 关系类型、感知相似性与用户互动意愿[J]. *软科学*, 2017, 31(9): 124-127.
- [39] 杨林, 张建宇, 孙纪寒. 体育休闲产业中基于体验经济的消费者参与合作创新研究[J]. *天津体育学院学报*, 2014, 29(5): 399-403.
- [40] 方文侃, 周涛. 社会交互对社会化商务用户行为作用机理研究[J]. *情报杂志*, 2017, 36(1): 167-172.
- [41] 张杨焱, 彭子健, 刘齐凯. 问答平台用户付费围观持续参与意愿的影响因素[J]. *图书馆论坛*, 2018, 38(6): 86-94.
- [42] 郭韧, 周飞. 顾客契合理论视角下的非交易型网络社区的成员心流体验研究[J]. *中央财经大学学报*, 2016(3): 122-128.
- [43] 姚山季, 王富家, 刘德文. 内容型虚拟社区中的用户互动和融入: 身份认同的中介效应[J]. *商业经济与管理*, 2018(2): 64-78.
- [44] 李强, 翁智刚, 高丁卉. 顾客参与能力识别: 内涵、方法与策略[J]. *外国经济与管理*, 2021, 43(6): 27-42.
- [45] 赵文军, 易明, 王学东. 社交问答平台用户持续参与意愿的实证研究——感知价值的视角[J]. *情报科学*, 2017, 35(2): 69-74, 91.
- [46] 杨楠. 顾客参与价值共创与品牌形象塑造的关系研究[J]. *科研管理*, 2021, 42(5): 87-93.
- [47] 乐承毅, 赵亚裴, 麻荣杰. 从浅层互动到深层互动: 虚拟品牌社区多层互动对知识共创的影响机制[J]. *信息资源管理学报*, 2022, 12(5): 77-88.
- [48] HUANG Z, BENYOUCEF M. From e-commerce to social commerce: A close look at design features[J]. *Electronic Commerce Research and Applications*, 2013, 12(4): 246-259.
- [49] LEWIS M A, BROWN A D. How different is professional service operations management? [J]. *Journal of Operations Management*, 2012, 30(1-2): 1-11.
- [50] DEMMERS J, WELTEVREDEN J WJ, DOLEN W M. Consumer Engagement with Brand Posts on Social Media in Consecutive Stages of the Customer Journey[J]. *International Journal of Electronic Commerce*, 2020, 24(1): 53-77.
- [51] HAMARI J. Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service[J]. *Electronic Commerce Research and Applications*, 2013, 12(4): 236-245.
- [52] DAHL D W, FUCHS C, SCHREIER M. Why and When Consumers Prefer Products of User-Driven Firms: A Social Identification Account[J]. *Management Science*, 2015, 61(8): 1978-1988.
- [53] OLAN F, SUKLAN J, ARAKPOGUN E O, et al. Advancing Consumer Behavior: The Role of Artificial Intelligence Technologies and Knowledge Sharing[J]. *IEEE Transactions on Engineering Management*, 2021(99): 1-13.
- [54] RAZA A, SAEED A, IQBAL M K, et al. Linking Corporate Social Responsibility to Customer Loyalty through Co-Creation and Customer Company Identification: Exploring Sequential Mediation Mechanism[J]. *Sustainability*, 2020, 12(6): 2525.
- [55] ABBAS M, GAO Y Q, SHAN S S H. CSR and Customer Outcomes: The Mediating Role of Customer Engagement[J]. *Sustainability*, 2018,

- 10(11): 4243.
- [56] LIU J W. Using big data database to construct new GFuzzy text mining and decision algorithm for targeting and classifying customers[J]. *Computers & Industrial Engineering*, 2019(128): 1088-1095.
- [57] KHOBZI H, LAU R Y K, CHEUNG T C H. The outcome of online social interactions on Facebook pages A study of user engagement behavior[J]. *Internet Research*, 2019, 29(1): 2-23.
- [58] YOSHIDA M, GORDON B S, NAKAZAWA M, et al. Bridging the gap between social media and behavioral brand loyalty[J]. *Electronic Commerce Research and Applications*, 2018(28): 208-218.
- [59] GUO L Y, HU X Y, LU J X, et al. Effects of customer trust on engagement in live streaming commerce: mediating role of swift guanxi[J]. *Internet Research*, 2021, 31(5): 1718-1744.
- [60] MARTINEZ-LOPEZ F J, ANAYA R, MOLINILLO S, et al. Consumer engagement in an online brand community[J]. *Electronic Commerce Research and Applications*, 2017(23): 24-37.
- [61] PHANG C W, ZHANG C H, SUTANTO J. The influence of user interaction and participation in social media on the consumption intention of niche products[J]. *Information & Management*, 2013, 50(8): 661-672.
- [62] LEE J YH, YANG CS, HSU C, et al. A longitudinal study of leader influence in sustaining an online community[J]. *Information & Management*, 2019, 56(2): 306-316.
- [63] WANG X Q, WONG Y D, TEO C C, et al. The four facets of self-collection service for e-commerce delivery: Conceptualisation and latent class analysis of user segments[J]. *Electronic Commerce Research and Applications*, 2020, 39: 100896.
- [64] GANGI P M D, WASKO M. Social Media Engagement Theory: Exploring the Influence of User Engagement on Social Media Usage[J]. *Journal of Organizational and End User Computing*, 2016, 28(2): 21.
- [65] RAYNA T, STRIUKOVA L. Open innovation 2. 0: is co-creation the ultimate challenge?[J]. *International Journal of Technology Management*, 2015, 69(1): 38-53.

# 《情报工程》征稿启事

《情报工程》期刊是由中国科学技术协会主管，中国科学技术情报学会与中国科学技术信息研究所主办的专业性学术刊物。该刊国际标准连续出版物号ISSN：2095-915X，国内统一连续出版物号CN:10-1263/G3。办刊目标是在情报分析领域中的理论研究、技术方法、工程应用、政府和企业实践等方面，提供一个专业的学术交流平台。

本刊为中国科技核心期刊（中国科技论文统计源期刊），已被数字化期刊全文数据库（万方）、中国学术期刊全文数据库（CNKI）、中国科技期刊数据库（维普）、《中国人文社会科学期刊评价报告（AMI）》引文数据库、国家哲学社会科学学术期刊数据库、超星学术期刊“域出版”等收录。

## 一、重点关注：

前沿图情理论与方法创新应用、人工智能与情报科学、情报工程与知识工程、数据科学与大数据分析、智库建设与决策支持、竞争情报服务模式与机制、创新管理与成果转化、政府和企业情报实践等。

## 二、约稿内容要求：

1. 稿件符合我刊办刊主题；
2. 稿件主题鲜明、数据可靠、文字通顺、符合学术规范；
3. 稿件要求包含学术论文的基本要件：标题（中英文）、作者姓名（中英文）、作者单位（中英文）、地址及邮编（中英文）、摘要（中英文）、关键词（中英文）、中图分类号、正文、参考文献、基金信息和作者简介与联系方式。稿件模版下载地址：[http://tie.istic.ac.cn/ch/first\\_menu.aspx?parent\\_id=20150331023349001](http://tie.istic.ac.cn/ch/first_menu.aspx?parent_id=20150331023349001)；
4. 来稿请用简化汉字，录稿用Word软件、五号宋体字、1.5倍行距、图表规范与文字相对应、图内文字清晰；
5. 来稿后编辑部将在一个月内反馈审稿意见，若未接收到稿件处理通知可向编辑部查询，修改稿逾一个月内未返回且不作说明，视为自动撤稿；
6. 请在来稿末尾附上作者详细通讯地址。包括作者所在省、市、区、街道名称、邮政编码、联系电话、电子信箱以及本人要求等，务必准确；
- 7 来稿一经采用，版权即归本刊编辑部，编辑部将发出《稿件录用通知》并在出刊后赠送当期刊物一本；
8. 请勿一稿多投。

## 三、投稿方式：

请直接通过<http://tie.istic.ac.cn>投稿。

本刊既厚名家、更重新人。欢迎国内外作者赐稿！

《情报工程》编辑部