

# 基于复杂网络的产业核心技术识别预测研究

## ——以能源产业为例

王燕鹏<sup>1</sup>, 赵宇<sup>2</sup>, 刘盼盼<sup>1,3</sup>, 韩涛<sup>\*1,3</sup>

(1. 中国科学院文献情报中心, 北京 100190; 2. 中国科协创新战略研究院, 北京 100038;  
3. 中国科学院大学经济与管理学院图书情报与档案管理系, 北京 100190)

**摘要:** 立足情报研究视角, 探索建立科学完整的科技文献信息挖掘及核心技术识别预测方法体系, 揭示战略前沿领域的核心技术, 明确核心技术的重点研发方向和潜在研发趋势。在界定核心技术概念及类型的基础上, 利用文献知识聚类识别热点技术, 通过RAKE算法抽取技术主要研究内容, 以各项热点技术为节点构建复杂网络, 通过节点二次聚类和可视化方法展现技术结构网络, 采用结构洞理论分析网络和节点特性, 从中遴选出关键技术; 利用复杂网络链路预测算法, 预测技术结构网络中的缺失边产生连接的可能性, 判断现有热点技术未来的交叉融合, 以此识别潜在新兴技术。以能源产业为例开展实践, 对该领域的科技文献开展深入挖掘, 揭示领域核心技术, 初步验证了方法的可操作性和有效性。

**关键词:** 核心技术, 关键词抽取, 复杂网络, 结构洞, 链路预测

### 1. 引言

习近平总书记在2018年两院院士大会上提出要以关键共性技术、前沿引领技术、现代工程技术、颠覆性技术创新为突破口, 努力实现关键核心技术自主可控。我国《“十三五”国家战略性新兴产业发展规划》明确指出要把战略性新兴产业摆在经济社会发展更加突出的位置, 大

力构建现代产业新体系, 推动经济社会持续健康发展。

然而, 在中美贸易摩擦加剧的背景下, 我国由于技术发展不平衡而面临的一系列关键核心技术威胁日渐凸显, 如高端光刻机、智能核心算法、锂电池隔膜、光伏逆变器等, 已对我国经济发展、社会进步和人民福祉等各方面产生了前所

---

**作者简介:** 王燕鹏, 男, 硕士, 助理研究员, 中国科学院文献中心, 主要从事特色情报数据资源建设与分析、量化技术预测预见、自然语言处理等研究。

赵宇, 博士, 副研究员, 中国科协创新战略研究院创新评估研究所副所长, 研究方向为创新创业评估、技术预见、数据挖掘与算法。

刘盼盼, 女, 硕士, 在读研究生, 中国科学院文献情报中心, 研究方向为新兴技术发现识别研究。

韩涛, 男, 博士, 研究员, 中国科学院文献中心, 业务管理处处长, 主要从事智能情报方法和技术研究。\* 通讯作者。

**基金项目:** 中国科协创新战略研究院科研项目“典型产业关键技术领域信息挖掘与核心(前沿)技术识别分析研究”(项目编号: 2019ys1-1-2-9)。

未有的影响<sup>[1]</sup>。在新一轮科技革命和产业变革的机遇下，清楚认识战略前沿领域的核心技术，明确核心技术的重点研发方向和潜在研发趋势，对我国前瞻性布局核心技术研发具有重要意义。鉴于此，本文旨在探索建立一套较为科学完整的科技文献信息挖掘及核心技术识别预测方法，结合机器学习、自然语言处理、复杂网络和文献计量等理论方法，深度挖掘科技文献大数据，识别典型产业领域核心技术并预测潜在研发趋势，揭示核心技术下的重点研发方向。

## 2. 国内外研究现状

实现技术识别和预见的目标，遴选可能产生最大化经济效益和社会效益的关键技术，很大程度上取决于科学合理的理论和方法体系。随着政府、学术界和产业界的深度研究实践，核心技术识别和预见方法也在不断丰富和完善。目前，常见的技术识别和预见方法包括德尔菲法、多指标分析法、技术路线图法、模型分析法、文献计量法和机器学习法等，各类方法内涵、特性以及优劣势等如表1所示。

总结来看，德尔菲法、多指标分析法和模型分析法在开展技术识别预见时，都需要该领域专业人员的深度参与，侧重对技术的定性分析。文献计量法和机器学习法侧重定量分析，通过多源文本数据和显性分析过程实现技术识别预见的目的，具有识别操作过程成本低、可重复性强、识别结果可解释性强的特点。同时，也应注意到，文献计量法和机器学习法仍处于不断研究和尝试的阶段，尚未形成统一、认可度较高的方法体系和框架。因此，本文从文本内容和关系出发，融合机器学习方法和复杂网络理论，构建一套较为科学合理且可复用推广的方法体系，为科技文献信息挖掘及核心技术识别提供一种新的问题解决思路。

## 3. 研究方法和技术路线

本文在界定核心技术概念及类型的基础上，利用文献知识聚类识别热点技术，通过RAKE算法抽取技术主要研究内容，以各项热点技术为节点构建复杂网络，通过节点二次聚类和可视化方法展现技术结构网络，采用结构洞理论分析网络和节点特性，从中遴选出关键技术；利用复杂网络链路预测算法，预测技术结构网络中的缺失边产生连接的可能性，判断现有热点技术未来的交叉融合，以此识别潜在新兴技术。研究总体技术路线及方法如图1所示。

### 3.1 核心技术类型及概念界定

在文献调研基础上，本文将核心技术定义为三类：① 热点技术：某学科领域内，近年来受到科研人员广泛关注并已产出相应研究成果的主要研究方向和技术主题，反映了学科领域的研发现状和技术结构全貌；② 关键技术：对学科领域内其他研究方向和技术主题产生广泛影响的，其研究成果可供参考借鉴和分享使用的一类技术，反映了学科领域的重要研发基础或技术交叉前沿，是学科领域内的“思想源泉”或“集大成者”；③ 潜在新兴技术：某学科领域内，未来由于学科交叉融汇而产生的新兴研究方向和技术主题，反映了学科领域内具有重要研发前景、值得深入探索的技术交叉点。

### 3.2 基于多维遴选策略的领域核心科技文献集构建

确定领域核心科技文献数据集是开展领域信息挖掘和核心技术识别的重要环节，直接影响了分析结论的科学性和可信度。因此，本文从数据新颖性、影响力、价值度等方面出发，构建了领域核心科技文献的多维遴选策略。具体方法如下：

(1) 针对科学论文数据：从论文发表时

表1 各类技术识别预见方法的总结与对比

方法类别	德尔菲法	多指标分析法	技术路线图法	模型分析法	文献计量法	机器学习法
内涵	借助问卷调查或访谈的方式向领域专家进行多轮征询, 以获得对技术发展前景较为科学和权威的判断	构建多指标的评估框架, 通过计算排序或专业分析, 识别核心技术	运用文字、图形和表格等形式描绘技术发展和变化趋势, 从路线图中识别核心技术	按照一定的理论框架和准则, 通过统计等方式构建相关模型, 对核心技术进行评估和判断	基于科技文献数据, 利用文献计量、文本挖掘和网络分析等手段从中挖掘潜在的核心技术	利用机器学习算法及模型从海量数据中挖掘、预测潜在的核心技术。其常用算法包括深度学习、神经网络、LDA主题模型等
类型	定性分析为主	定性和定量分析相结合	定性分析为主	定性分析为主	定量分析为主	定量分析为主
实施机构	政府及相关机构	政府及相关机构、学术界	政府及相关机构、产业界、学术界	学术界	学术界、产业界	学术界
优势	凝聚专家智慧, 识别结果具有较强的科学性和权威性	多维度指标体系可以对技术进行全面评估, 科学性和客观性较强	明确技术发展方向和实现目标所需条件, 理清产品和技术之间的关系	借鉴经济学、运筹学和数学等学科的基础理论知识, 形成相对完整的分析框架和准则, 更加系统客观地识别核心技术	科技文献反映了技术的诞生阶段, 适用于萌芽早期技术的预测和发现, 数据易于获取, 分析工具相对成熟	可处理全领域海量数据, 发现数据中潜藏的特征和模式, 预测技术未来的发展潜力与演变方向, 分析结果更具广度和深度
不足	受专家主观认识影响较大, 新技术、新想法出现的初期, 难以形成专家共识, 遗漏部分潜在的新兴技术	部分指标依然需要专家进行打分, 对方法客观性造成一定影响	不利于识别和预见具有非连续性发展轨迹的技术	对研究者的专业素养和实践技能有着较高的要求, 可操作性较差	分析方法的科学性有待进一步加强, 且主要关注技术本身的研究内容和影响力, 缺少对产品和市场的考量	对核心技术反映在客观数据上的特征定义和提取不明确, 分析方法的科学性有待进一步加强
典型案例	日本科学技术预见 <sup>[2]</sup> 英国《技术与创新未来: 英国2020年的增长机会》 <sup>[3]</sup>	中国工程院“引发产业变革的颠覆性技术内涵与遴选研究” <sup>[4]</sup>	Uchihira N“Future direction and roadmap of concurrent system technology” <sup>[5]</sup>	Sood A“Demystifying Disruption: A New Model for Understanding and Predicting Disruptive Technologies” <sup>[6]</sup>	Joung J“Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data” <sup>[7]</sup>	Lee C“Early identification of emerging technologies: A machine learning approach using multiple patent indicators” <sup>[8]</sup>

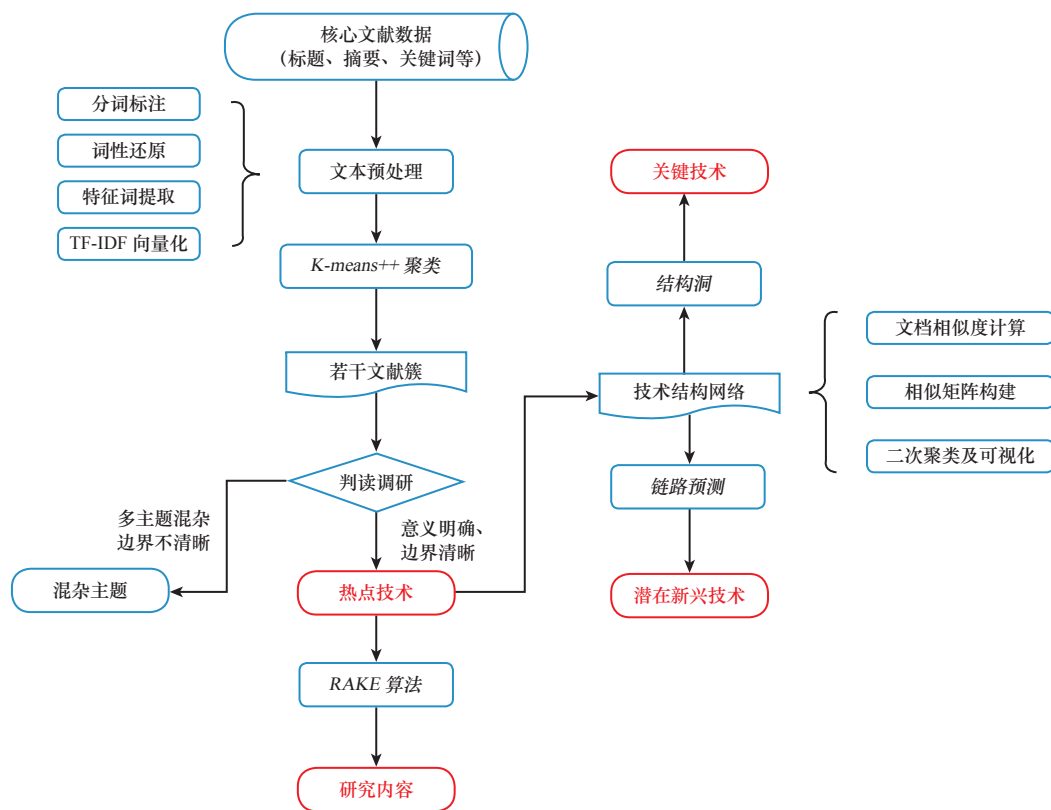


图1 研究技术路线

间、论文被引情况（高被引论文（Highly Cited Papers）<sup>1</sup>、热点论文（Hot Papers）<sup>2</sup>等）、论文来源（即出版期刊或会议的影响力<sup>3</sup>）等角度对科学论文数据质量进行评估，确定待分析的数据集；

（2）针对技术专利数据：从专利优先权时间、专利布局情况（三方专利、五方专利等）、专利价值度（数据库商根据自身的评价体系对专利价值的打分）等角度对技术专利数据质量进行评估，确定待分析的数据集。

### 3.3 基于多次聚类的领域技术结构及热点技术识别方法

热点技术代表了当前学科领域的主流研究方向，基本涵盖了该学科科研人员的主要研究关

注点和着力点。相较K-means算法，K-means++算法在选取聚类质心方面做出了改进，使得聚类质心相互之间的距离最大化，使得各个聚类簇划分更加明显，优化聚类效果<sup>[9]</sup>。针对遴选出的核心科技文献，基于文献语义关系，采用K-means++文本聚类方法进行文献聚合，计算得到论文/专利文献类簇，进而通过二次聚类的方式形成由多个文献类簇组成的综合类簇。然后，结合领域技术人员的学科知识以及关键词自动抽取对文献类簇和综合类簇进行专业判读，识别领域内的热点技术，归纳梳理领域的整体技术结构。由此，建立由单篇文献、文献类簇、综合类簇到单一研究、热点技术、技术结构的分析映射关系。

<sup>1</sup> 高被引论文指过去10年所发表的论文中，被引次数在同学科、同发表年度论文中排名前1%的论文。

<sup>2</sup> 热点论文指近2年内所发表的论文中，近2个月的被引频次在同学科、同发表年度论文中排名前0.1%的论文。

<sup>3</sup> 部分国内学会或高校会定期发布国内外高质量学术期刊和会议推荐目录，该类目录通常由相关领域专家审议评定，具有较高的权威性。例如，中国计算机学会发布的《中国计算机学会推荐国际学术会议和期刊目录》。

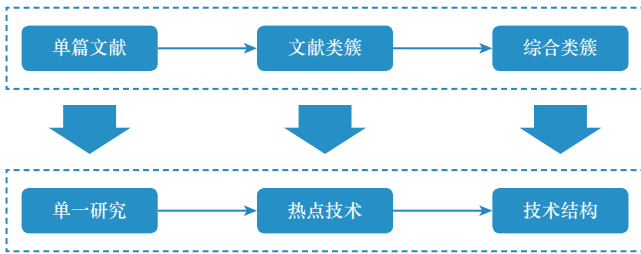


图2 热点技术、技术结构与科技文献的分析映射关系

### 3.4 基于RAKE算法的热点技术研究内容抽取

热点技术的研究内容反映了当前学科领域的主要研究问题及方向，能够更全面、更详细的揭示热点技术的研发细节。RAKE算法的全称是Rapid Automatic Keyword Extraction，可以实现快速的文档关键词抽取，且准确性较高。RAKE算法首先使用标点符号（如半角的句号、问号、感叹号、逗号等）将一篇文档分成若干分句，然后对于每一个分句，使用停用词作为分隔符将分句分为若干短语，这些短语作为最终提取出的关键词的候选词。此后，以空格分割候选词形成单词，利用候选词作为单元计算单词与其他所有单词的共现频次，除以单词本身的词频，作为该单词的得分。最后，对于每个候选词，将其中每个单词的得分累加，并进行降序排序，RAKE将候选词总数的前三分之一认为是抽取出的关键词。本文利用RAKE算法抽取每项热点技术对应文档的重要关键词，以此作为每项热点技术的主要研究内容。

### 3.5 基于结构洞理论的关键技术识别

关键技术反映了当前学科领域的研发基础或技术交叉前沿，对该学科领域的发展起到了重要推进作用。结构洞理论于1992年提出，表示非冗余的联系<sup>[10]</sup>，指在社会网络中的某个或某些个体和有些个体发生直接联系，但与其他个体不发生直接联系的现象，从网络整体看好像网络结构中出现了洞穴。Raider H J<sup>[11]</sup>的研究表明，结构洞占

据者的位置对信息控制、识别以及交易起着重要的作用。由此可知，结构洞反映了节点对网络资源的控制能力，占据结构洞的节点能够获得来自多方面的非重复性信息，占据结构洞的研究方向更有可能是“思想源泉”或“集大成者”，即领域关键技术，可对学科领域内其他研究方向产生广泛影响。基于此，本文以热点技术文献簇质心向量为节点，以向量间相似性为连边依据，以此构建复杂网络，通过可视化和质心向量二次聚类的方法展现技术结构网络。基于结构洞理论，计算网络中各节点的限制度指标，以此反映节点在网络中运用结构洞的能力，节点的限制度指标值越小，代表节点拥有的结构洞越多，运用结构洞的能力越强，以此识别领域的关键技术。限制度计算公式为：

$$C_{ij} = (P_{ij} + \sum_q P_{iq} P_{qj})^2 \quad \text{式 (1)}$$

其中，节点 $q$ 是节点 $i$ 和节点 $j$ 的共同邻接点； $P_{ij}$ 表示在节点 $i$ 的所有邻接点中节点 $j$ 所占的权重比例。则节点 $i$ 的限制度为：

$$C_i = \sum_j C_{ij} \quad \text{式 (2)}$$

### 3.6 基于复杂网络链路预测的潜在新兴技术识别

跨学科领域的交叉融合和集成创新已成为当代科学技术创新的鲜明特征，通过预测多个研究方向和技术主题产生交叉融合的可能性，将有助于识别潜在新兴技术。复杂网络中的链路预测<sup>[12]</sup>是指如何通过已知的网络结构等信息预测网络中尚未产生连边的两个节点之间产生连接的可能性。基于此，本研究以上述构建的技术结构网络为基础，利用复杂网络链路预测方法，探测技术结构网络中尚未产生连接但未来可能产生较强连接的节点对，分析热点技术交叉融合促进科技创新形成的现象（即潜在的科学发现），以此识别潜在新兴技术。

具体的，本文选用链路预测中基于网络结构相似性的方法，该方法计算复杂度低、准确性好，并且网络的拓扑结构性质能够帮助选择合适的相似度指标。由于上述技术结构网络为加权无向网络，选择使用基于网络结构相似性的含权的Adamic-Adar指标（简称AA指标）计算产生连接的可能性，以综合考虑节点共同邻居数量及节点度对连接的影响，使链路预测结果更加准确。

#### 4. 能源产业核心技术识别预测实践

基于上述的研究方法和技术路线，本文以能源产业为例开展实践，对该领域的科技文献（仅包括论文数据）开展深入挖掘，揭示领域核心技术。

##### 4.1 热点技术分析

以Web of Science学科分类中“能源与燃料”（ENERGY&FUEL）领域为研究范围，选取该学科2015年~2020年5月间的高被引论文（Highly Cited Papers）作为分析对象，共计获得4223篇论文，以论文的标题、摘要、关键词字段作为分析数据，进行文本预处理、向量化表示处理和K-means++聚类，通过计算轮廓系数，确定聚类簇数量K=40时，具有较好的聚类效果。本步骤共获取得到40个聚类簇，其中含39个有效类和1个混杂类，经解读，形成能源产业的热点技术。通过对上述39个聚类簇进行可视化和模块划分，又可以形成7个技术大类，如下图所示，每种颜色代表1个研究大类。

##### 4.2 关键技术分析

以能源产业热点技术分布图（图3）为对象，计算网络中各节点的限制度指标并升序排列，取Top10的热点技术作为能源产业的关键技术（表2）。

##### 4.3 潜在新兴技术分析

针对能源产业热点技术分布网络（图3），利用链路预测方法，预测尚未产生连接或连接较弱，但未来可能产生连接的节点对。按照节点对产生连接可能性降序进行排列，取Top10的热点技术节点对作为未来可能产生交叉融汇的潜在新兴技术（表3）。

#### 5. 结语

近年来，随着“数据驱动”观念的深化，机器学习、大数据技术和复杂网络理论等的逐步普及为领域信息挖掘及核心技术识别预测带来了新的问题解决思路。本文立足情报研究视角开展方法探索和实践，以领域信息挖掘及核心技术识别为应用目的，以科技创新基本规律为理论基础，以文献知识聚类和技术结构复杂网络构建为分析主线，以文本聚类、结构洞理论和复杂网络链路预测为分析方法，进行热点技术、关键技术和潜在新兴技术的挖掘和识别，探索构建了一套较为科学合理且可复用推广的方法体系。在此基础上，对能源产业开展信息挖掘，识别该领域的热点技术、关键技术和潜在新兴技术，识别预测结果得到了领域课题组的认可，其中部分识别结果纳入到该领域课题组的技术备选清单中，并进入德尔菲调查环节，供专家遴选评估，对其技术预见工作起到了良好的支撑作用。

综上，基于文献知识，结合机器学习、自然语言处理、复杂网络等理论方法和分析技术，可实现对领域信息的深度挖掘，实现对领域热点技术、关键技术和潜在新兴技术的识别和发现，研究方法体系化，研究过程量化，研究结果具有较强科学性、合理性和可解释性。

需要指出的是，本文所提出的方法体系仍有

表2 能源产业关键技术及研究内容

序号	关键技术	限制度 (对数值-ln)	技术研究内容
1	电池阴极材料	-2.155	高能锂离子电池阴极制造、石墨烯及其衍生材料、石墨烯/氧化镧纳米复合材料、富镍层状氧化物阴极、层状氧化物阴极材料、碳基二维储能材料、二维金属有机骨架、富锂锰正极材料
2	新型燃料电池技术	-2.153	阴离子交换膜电池、质子交换膜燃料电池、乙醇燃料电池、固体氧化物燃料电池、氢氧化物交换膜燃料电池、质子陶瓷燃料电池、直接液体燃料电池、微生物燃料电池
3	二氧化碳捕获和储存	-2.146	低温二氧化碳捕集技术、非水胺基吸收剂、氨基酸盐溶液、离子液体、多孔吸附剂、钙环吸附剂、含盐含水层CO <sub>2</sub> 封存效率
4	氢的生产/储存/运输及其应用	-2.134	太阳能-地热联合制氢系统、浮式光伏发电制氢系统、化学链制氢、金属有机骨架型析氢催化剂、镁-石墨复合材料水解制氢、液态有机氢载体储氢、镁基复合材料、储氢用氨、金属氢化物固态储氢材料、氨硼烷、液态有机和无机化学氢化物氢运输燃料
5	电能存储系统与可再生能源	-2.126	可再生能源整合和转化、全球能源系统、配电系统储能调度、现代分布式发电系统中主机容量、可再生能源电池储氢混合系统优化、第四代分布式区域供热系统优化、电能存储技术、材料和系统:集中太阳能发电厂热能储存、填充床气液储能系统、太阳能热发电用固体气体热化学储能系统、抽水蓄能独立光伏发电系统、热电联产蓄能供热系统、混合电池超级电容器储能器、风力发电集成支持储能系统
6	热循环和回收技术	-2.122	有机朗肯循环、余热回收和储存技术、吸附换热储热新材料、太阳能冷热电联供系统、电热网联合分析、微冷热电联产系统、纳米流体换热系统
7	可再生能源的应用与发展	-2.119	家庭能源消费、城市能源消费、可再生能源投资决策、新能源系统转型、可再生能源系统管理与优化、欧盟可再生能源政策、英国能源系统模型、北欧低碳能源转型、美国能源转型、巴西光伏太阳能利用
8	石油采收和油水分离技术	-2.114	地热系统、稠油开采技术、油水分离技术、提高稠油采收率、纳米颗粒与低盐度水、聚合物表面活性剂、纳米颗粒稳定泡沫等
9	太阳能综合利用技术	-2.114	太阳能集热器的设计与效率提升、太阳能蒸汽发电薄膜技术、太阳能蒸馏器、纳米流体的光学性质和应用、太阳能驱动海水淡化技术、太阳能热电联产系统
10	生物质热解技术	-2.100	生物质热解反应动力学参数、生物质热解催化、木质纤维素生物质热解、生物质制氢热化学途径、生物质与废塑料共热解、微藻热解、水热液化制备生物原油、生物质水热碳化

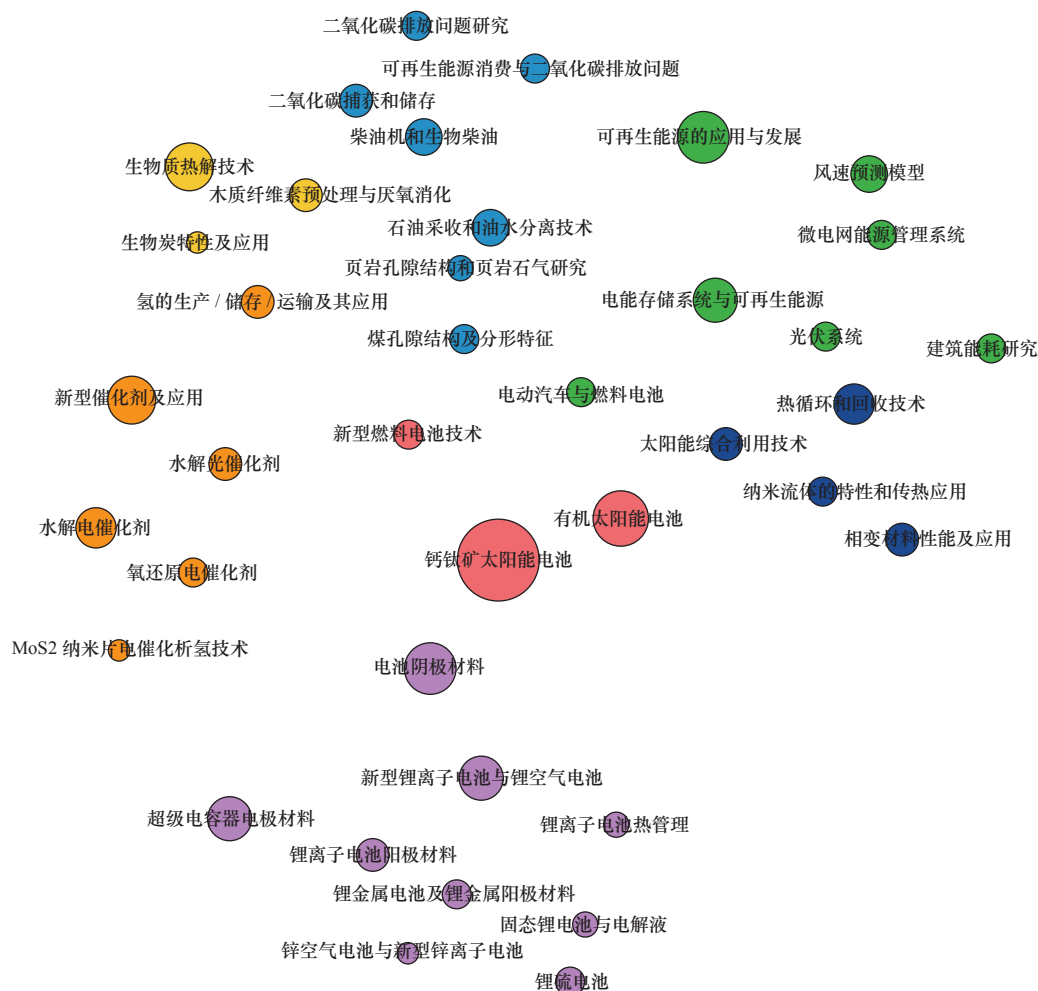


图3 能源产业热点技术分布图

注：图中每个节点代表一个聚类文献簇，即热点技术，节点大小代表聚类簇下文章数量多少，节点颜色代表该节点归属的技术大类，共7类。

表3 未来可能产生交叉融汇的潜在新兴技术

序号	潜在新兴技术	产生连接的可能性	
1	石油采收和油水分离技术	煤孔隙结构及分形特征	0.1231
2	可再生能源的应用与发展	二氧化碳捕获和储存	0.0795
3	电能存储系统与可再生能源	二氧化碳排放问题研究	0.0740
4	页岩孔隙结构和页岩气研究	生物质热解技术	0.0611
5	氧还原电催化剂	氢的生产/储存/运输及其应用	0.0573
6	页岩孔隙结构和页岩气研究	二氧化碳捕获和储存	0.0560
7	水解电催化剂	新型锂离子电池与锂空气电池	0.0544
8	新型燃料电池技术	太阳能综合利用技术	0.0537
9	页岩孔隙结构和页岩气研究	热循环和回收技术	0.0516
10	锂离子电池热管理	电能存储系统与可再生能源	0.0496



一定的改进空间。首先, 文本预处理过程中, 词形还原的准确性有待进一步提高。其次, 将结构洞理论与其它网络及客观度量指标相结合, 多维度共同揭示关键技术。

面向未来的技术预见工作, 上述方法体系可在如下方面进行持续深入的探索和完善。首先, 技术识别精细化。本文方法体系可落脚到单篇文献层面, 但科技文献中还有丰富的技术参数、技术实施细节、技术先进性评价等内容, 对此类数据进行抽取和挖掘, 将实现更加精细化的技术识别。其次, 技术评价全面化。本文中主要通过结构洞和链路预测方法开展关键技术和潜在新兴技术的评价, 未来可结合更多的科技发展、产业实践、财务金融、社交媒体等数据对技术进行全方位的评价, 保障评价的全面化和合理性。最后, 研究过程工具化。未来将逐步对研究方法、过程及源代码进行整合, 形成领域信息挖掘和核心技术识别的成熟工具, 降低定量化技术预见的操作门槛, 保障科学性和合理性的同时, 追求简单化和便捷化。

责任编辑: 李琦 校对: 贺茂斌 陈峰

### 参考文献

[1] 汪雪锋, 赖院根, 朱东华. 技术威胁理论研究[J]. 科学学研究, 2009, 27(02): 166-169.

[2] 文部科学省科学技术·学术政策研究所. 科学技术予测センター.NISTEP.REPORT No.183. 第11回科学技术予测调查S&T Foresight 2019综合报告书[EB/OL]. 2019.11.

[3] Foresight Horizon Scanning Centre. Technology and innovation futures: UK growth opportunities for the 2020s[R]. 2010.

[4] 孙永福, 王礼恒, 孙棕檀, 等. 引发产业变革的颠覆性技术内涵与遴选研究. 中国工程科

学, 2017, 19(5): 9-16.

[5] Uchihira N. Future direction and roadmap of concurrent system technology. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2007, 90(11): 2443-2448.

[6] Sood A, Tellis G J. Demystifying Disruption: A New Model for Understanding and Predicting Disruptive Technologies[J]. Marketing Science, 2010, 30(2): 339-354.

[7] Joung J, Kim K. Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data[J]. Technological Forecasting & Social Change, 2017(114): 281-292.

[8] Lee C, Kwon O, Kim M, et al. Early identification of emerging technologies: A machine learning approach using multiple patent indicators[J]. Technological Forecasting & Social Change, 2018, 127(2): 291-303.

[9] Arthur D, Vassilvitskii S. k-Means ++: The advantages of carefull seeding[J]. Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete algorithms, Society for Industrial and Applied Mathematics, 2007, 11(6): 1027-1035.

[10] 刘军. 整体网分析讲义: UCINET软件实用指南[M]. 上海: 格致出版社/上海人民出版社, 2009.

[11] Raider H J. Market structure and innovation[J]. Social science research, 1998, 27(1): 1-21.

[12] 吕琳媛. 复杂网络链路预测[J]. 电子科技大学学报, 2010, 39(05): 651-661.

# Research on the identification and prediction of industrial core technologies based on complex networks: Taking the energy industry as an example

Wang Yan-peng<sup>1</sup>, Zhao Yu<sup>2</sup>, Liu Pan-pan<sup>1,3</sup>, Han Tao<sup>\* 1,3</sup>

(1. National Science Library, Chinese Academy of Sciences, Beijing 100190, China;

2. National Academy of Innovation Strategy, China Association for Science and Technology, Beijing 100038, China;

3. Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academic of Sciences, Beijing 100190, China)

**Abstract:** Based on the perspective of information research, this paper tries to establish a scientific and complete S&T literature information mining and core technology identification method, in order to reveal the core technology in the strategic frontier field, and clarify the key R&D direction and potential R&D trend of core technology. On the basis of the definition, types and concepts of core technology, this paper uses literature knowledge clustering technology to identify heated technologies, extracts the main research content of the technology through RAKE algorithm, constructs a complex network with hot technologies as nodes, presents the technical structure network through nodes secondary clustering and visualization method, analyzes the network and node characteristics by using the structure hole theory, and selects the key technologies. Using the link prediction algorithm of complex network, the possibility of connection generated by missing edges in the technical structure network is predicted, and the future cross fusion of existing heated technologies is judged, so as to identify potential emerging technologies. Taking the energy industry as an example to carry out the practice, the in-depth excavation of S&T Literature in this field is carried out to reveal the core technology of the field, and the operability and effectiveness of the method are preliminarily verified.

**Key words:** core technology; keywords extraction; complex network; structure hole; link prediction

---

(上接第49页)

## Research on foresight of new-energy key technologies

Liu Jin-ping, Lu Shi-gang

(China Automotive Battery Research Institute Co. Ltd, Beijing 101407,China)

**Abstract:** in order to promote the technology innovation of china in the field of new-energy key technologies, the delphi method is adopted to carry out technology foresight research in the key field of new-energy electricity, new-energy vehicles and electrical power storage, etc., and the technology realization time, important technical topics, restrictive factors and leading countries or regions are clarified, especially 10 key generic technology topics for priority development are identified in this study. the analysis results show that the 10 priority technology topics will achieve large-scale application around 2030, supporting the accelerated and integrated development of new-energy industries and the construction of new-energy system with a high proportion of new-energy sources will be able to guarantee china's energy security, improve the quality of people's lives, and promote the sustainable development of society and economy.

**Key words:** new-energy technology; new-energy electricity; new-energy vehicle; electrical power storage; technology foresight; Delphi method