



情报杂志
Journal of Intelligence
ISSN 1002-1965,CN 61-1167/G3

《情报杂志》网络首发论文

题目：基于应用场景的未来技术识别研究
作者：谢俊杰，孙希科，王智琦，韩盟，陈悦
网络首发日期：2024-01-04
引用格式：谢俊杰，孙希科，王智琦，韩盟，陈悦. 基于应用场景的未来技术识别研究[J/OL]. 情报杂志. <https://link.cnki.net/urlid/61.1167.G3.20240103.1725.012>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于应用场景的未来技术识别研究^{*}

谢俊杰¹ 孙希科² 王智琦^{1,3} 韩 盟¹ 陈 悦¹

((1. 大连理工大学 科学学与科技管理研究所暨 WISE 实验室 大连 116024;

2. 潍柴动力股份有限公司 潍坊 261061;

3. 大连理工大学 人文学院 大连 116024)

摘 要:[研究目的]未来技术作为传统产业的转型动力以及未来产业的形成基础,从技术的应用及其产业化视角出发,进行未来技术的识别与预测,对我国布局未来产业、增强发展优势具有重要意义。[研究方法]从技术单元与技术方案出发,构建基于应用场景的未来技术识别模型。首先,利用 LDA 主题模型对专利的用途进行主题聚类,识别技术的主要应用场景;其次,利用 KeyBERT 算法从专利的标题和新颖性文本中提取技术方案关键词,从技术方案的新颖性、关联性和重要性出发,筛选前沿技术方案;最后,定义产生前沿技术方案的技术单元为潜在未来技术,构建“技术影响力-技术生长力”坐标图,进行未来技术识别。[研究结论]以固体氧化物燃料电池领域为例进行实证研究,识别出 13 项未来技术,主要为 SOFC 批量化制造、可控性运行和实际化应用的技术,揭示了识别方法的有效性。

关键词:应用场景;未来技术;技术识别;技术方案;专利文本;固体氧化物燃料电池;KeyBERT 算法;LDA

中图分类号:G306.0

文献标识码:A

Research on Future Technology Identification Based on Application Scenarios

Xie Junjie¹ Sun Xike² Wang Zhiqi^{1,3} Han Meng¹ Chen Yue¹

(1. Institute of Science of Science and S&T Management and WISE Lab,

Dalian University of Technology, Dalian 116024;

2. Weichai Power Co., Ltd., Weifang 261061;

3. School of Humanities, Dalian University of Technology, Dalian 116024)

Abstract:[Research purpose]Future technologies serve as the catalysts for the transformation of traditional industries and the building blocks for the future industries. It is of great significance to identify and predict future technologies from the perspective of the application of technology and its industrialization, because they play a pivotal role in shaping our nation's strategy for the future industries and enhancing our competitive advantages in development. [Research method] This paper constructs an application scenario-based model for identifying future technologies in terms of technology units and technological solutions. Firstly, we identify the main application scenarios of the technology by employing Latent Dirichlet Allocation (LDA) to conduct topic clustering of the patents based on their intended uses. Secondly, we employ the KeyBERT algorithm to extract keywords from patent titles and novel text, enabling us to pinpoint frontier technology solutions based on their novelty, relevance and importance. Finally, we define the technology units that generate frontier technology solutions as potential future technologies, construct a "technology influence - technology growth" coordinate diagram, and identify future technologies. [Research conclusion] Taking the field of Solid Oxide Fuel Cell (SOFC) as the empirical field, the effectiveness of the identification method is validated, through which we successfully identified 13 future technologies, primarily centered around SOFC batch manufacturing, controllable operation and practical application.

基金项目:教育部哲学社会科学研究重大课题攻关项目“基础研究领域颠覆性科研成果识别与我国基础研究能力提升研究”(编号:22JZD021);国家重点研发计划项目“颠覆性技术识别理论、方法与专家预判系统”(编号:2019YFA0707201)研究成果之一。

作者简介:谢俊杰,男,1997 年生,硕士研究生,研究方向:科学学与科技管理;孙希科,男,1982 年生,博士,研究方向:信息管理、科技创新;王智琦,女,1990 年生,博士,助理研究员,研究方向:科学学与科学计量学、开放科学;韩 盟,男,1996 年生,博士研究生,研究方向:科学计量学、科技管理;陈 悦,女,1975 年生,博士,教授,博士生导师,研究方向:科学学理论、科学计量学、科技创新管理。

通信作者:陈 悦

Key words: application scenario; future technology; technology identification; technology solution; patent text; Solid Oxide Fuel Cell (SOFC); KeyBERT algorithm; LDA

0 引言

世界新一轮的科技革命和产业变革正在催生由未来技术所形成的新产品和新业态,由此而形成的未来产业是把握未来发展的关键,基于未来技术提前布局未来产业是具有战略性意义的举措。

未来技术的本质是影响未来的技术^[1],即面向未来发展方向和发展需求,当前尚未产业化但未来具有产业化潜力、有望赋能或替代传统产业或催生新的产业,进而重塑竞争格局、产生重大经济影响的技术。与未来技术相关的概念有颠覆性技术^[2]、新兴技术^[3]和前沿技术^[4],虽然它们都具有面向未来的属性,但对技术特征强调的视角不同,颠覆性技术强调技术的跃迁性,新兴技术强调技术新颖性和持续增长性,前沿技术强调技术的前瞻性和先导性。未来技术作为面向形成未来产业的技术,实质上是涵盖了不管技术具有任何特征,只要其在未来具有产生新的产业或对传统产业进行优化升级的潜力,均可以将其定义为未来技术,即颠覆性技术、新兴技术和前沿技术均可以看作未来技术的子集^[1]。

目前学界已有基于技术的固有特征及其生成方式对颠覆性技术、新兴技术和前沿技术的研究^[5-7],但少有从技术的应用及其产业化视角的考虑,即对未来技术的识别研究尚不充分。因此,本文从技术的应用场景出发,构建未来技术的识别模型,开展未来技术的识别研究。

1 研究设计

技术是人类为满足自身的需要,根据实践经验或科学原理所创造发明的各种手段和方式方法的总和^[8]。未来技术是以满足人民和经济社会发展的新需求为目标^[9],有望重塑人类社会、提升人类生产力和生活品质的新技术。可见,需求是技术创新活动的起点,也是技术创新得以实现的最终场所^[10],而技术的应用场景作为技术与需求的连接器^[11],可以有效地承载需求所提出的技术问题,以及技术提出的相应解决方案。本文设计了一种基于应用场景的未来技术识别模型(图1),新的技术应用场景的产生以及已有应用场景的优化升级会向技术单元提出新的技术问题,技术单元针对相应的技术问题会提出相应的技术方案,并将技术方案应用于相应的应用场景以解决技术问题,实现技术应用场景的适配与发展。随着技术应用场景所引发的需求深度不断增加,技术问题的复杂度也会随

之不断增加,已有技术单元将没有足够的能力产生解决新技术问题的技术方案,此时便需要有新技术单元的加入,产生前沿技术方案以解决技术问题。技术应用的需求深度越大,新的技术单元创造新的应用场景以及对已有应用场景进行转型升级的潜力也越高,该技术单元即为未来技术。

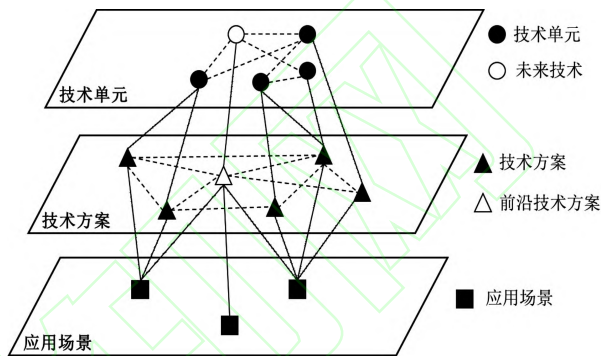


图1 未来技术识别模型

专利作为技术的有效载体,其摘要文本包含着面向技术需求的技术方案^[12],其IPC(国际专利分类号)可以作为专利所包含技术单元的有效划分^[13]。本文专利数据来源于最全面的深加工增强专利信息数据库德温特专利数据库(Derwent Innovations Index, DII),从技术方案的提出与应用出发,利用专利数据进行未来技术的逆向识别。首先,根据专利的主要用途识别出技术主要应用场景(图1中黑色方形);其次,从专利摘要中提取技术方案,并设计指标识别前沿技术方案(图1中白色三角形);最后,将产生前沿技术方案的技术单元定义为潜在未来技术,设计指标从中进行未来技术(图1中白色圆形)的识别。本文研究技术路径如图2所示。

1.1 技术应用场景识别

技术应用场景或领域是指技术发挥其功能的部门、产业、行业^[14],专利摘要中的Use部分描述了该专利的主要用途,对所有专利的Use文本进行主题聚类,即可以识别出技术的主要应用场景。本文对Use文本进行分词,去除停用词并对其余词进行词干化处理,将处理后的文本输入LDA主题模型^[15]中,计算困惑度与主题数之间的关系。根据“主题数-困惑度”曲线判断可能存在最优聚类的主题数,结合每个主题的关键词和实际情况获得最优的主题聚类,根据每个主题下的文本进行技术应用场景的识别。

1.2 前沿技术方案识别

专利的标题和摘要的Novelty部分包含了专利所涉及的主要技术,同时也是该专利面向应用场景所提

出的技术方案,对标题和 Novelty 文本进行关键词提取,即可提取该专利的技术方案。本文使用 KeyBERT 算法^[16]进行技术方案关键词的提取,将专利标题和摘要 Novelty 文本合并,去除停用词、标点与特殊符号,对高频词的单复数形式和时态进行合并,利用 KerBERT 算法对处理后的文本进行关键词提取。采用

关键词簇方法,对提取的关键词进行清洗,通过模糊匹配、关键词合并等方法建立关键词表,对具有相同含义的关键词进行合并和替换,最终在每个专利中提取的关键词即为技术单元针对技术应用场景所形成的技术方案。

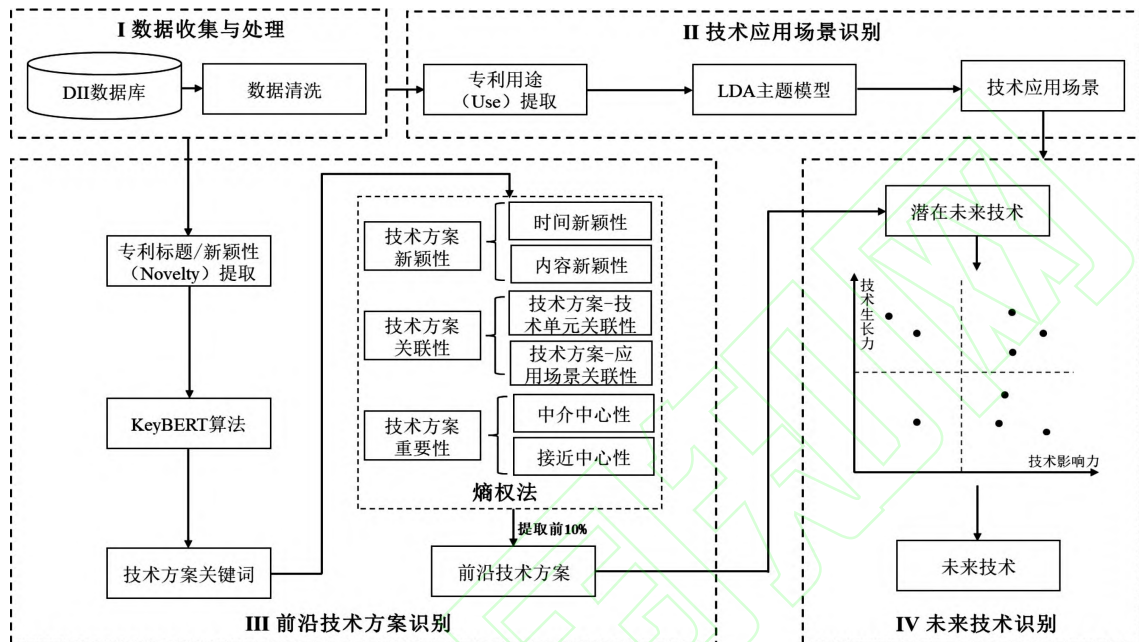


图2 面向应用场景的未来技术识别路径图

未来技术是面向未来,有望对科技发展、产业格局及人类生产生活方式带来重大影响的技术^[1],其面向技术应用场景生成的技术方案也应具有前沿性的特征。本文从新颖性、关联性和重要性三个维度出发,设计技术方案前沿性的测度指标。

(1) 技术方案新颖性

前沿技术方案应具有新颖性的特征,新颖性可以从时间和内容两个维度进行测度^[17]。时间新颖性(Time Novelty, TN)以专利的平均申请年表示,平均申请年越近,表明技术方案在时间上越具新颖性,其计算方式如公式(1)所示。

$$TN_i = \sum_{j=1}^N t_j^i / N \quad (1)$$

其中, t_j^i 表示技术方案 i 的第 j 件专利的申请年份; N 表示技术方案 i 的专利总量。

内容新颖性(Content Novelty, CN)以专利的年均增长率表示,研究的扩展速度越快,表明技术方案在内容上越具新颖性,其计算方式如公式(2)所示。

$$CN_i = (N_i^n / N_i^1)^{\frac{1}{n-1}} - 1 \quad (2)$$

其中, N_i^1 表示技术方案 i 在第一年申请的专利量; N_i^n 表示技术方案 i 在最后一年的专利量。

(2) 技术方案关联性

技术方案作为技术单元为解决应用场景的技术问

题所产生的技术,其与技术单元以及应用场景之间均存在关联性。技术单元产生了技术方案,但在应用场景并未实现有效的应用,或是在应用场景实现有效应用的技术方案并非由该技术领域的技术单元产生,这两种情况分别会导致技术方案与技术单元和应用场景的关联性偏低,即无法成为有效的前沿技术方案。

本文参考 Jaccard 相似系数引入关联性的计算方式^[18],技术方案-技术单元关联性(Technology Solution - Technology Unit, TT)由技术方案与每个技术单元关联性的平均值进行计算,TT 值越大表明技术方案与所有技术单元的联系越紧密,其计算方式如公式(3)所示。

$$TT_i = \sum_{j=1}^M [N_{ij} / (N_i + N_j - N_{ij})] / M \quad (3)$$

其中, N_i 表示技术方案 i 的专利量; N_j 表示技术单元 j 的专利量; N_{ij} 表示同时存在技术方案 i 和技术单元 j 的专利量; M 表示技术方案 i 所连接的技术单元总量。

同理,技术方案-应用场景关联性(Technology Solution - Application Scenario, TA)由技术方案与每个应用场景关联性的平均值进行计算,TA 值越大表明技术方案与所有应用场景联系越紧密,其计算方式如公式(4)所示。

$$TA_i = \sum_{j=1}^M [N_{ij} / (N_i + N_j - N_{ij})] / M \quad (4)$$

其中, N_i 表示技术方案 i 的专利量; N_j 表示应用场景 j 的专利量; N_{ij} 表示应用场景 j 中应用技术方案 i 的专利量; M 表示技术方案 i 所有的应用场景数量。

(3) 技术方案重要性

一项技术方案越重要,在技术方案网络中应处于越重要位置,表现出越高权力^[19]。本文以技术方案关键词作为节点,以同一项专利提取的关键词的共现关系作为边,构建技术方案关键词共现网络,利用中介中心性和接近中心性分析技术方案的重要性。中介中心性(Betweenness Centrality, BC)越高表明该技术方案在网络中作为中间中介者的能力越强,即对整个网络的资源控制程度越高,其计算方式如公式(5)所示。

$$BC_i = \frac{\sum_{i \neq j \neq k} b_{jk}(i) / b_{jk}}{(N-2)(N-1)/2} \quad (5)$$

其中, b_{jk} 表示节点 j 到节点 k 最短路径的个数; $b_{jk}(i)$ 表示连接节点 j 和节点 k 且经过节点 i 的最短路径数; N 表示节点总量。

接近中心性(Closeness Centrality, CC)越高表明该技术方案信息传递能力越强,居于网络中心的可能性越大,其计算方式如公式(6)所示。

$$CC_i = \frac{1}{\sum_{j=1}^N d_{ij} / (N-1)} \quad (6)$$

其中, d_{ij} 表示节点 i 到节点 j 的距离; N 表示节点总量。

由于上述指标的量纲不同,且其均为正向指标,使用 max-min 方法对其进行标准化处理, x'_i 为标准化之后的结果,其计算方式如公式(7)所示。

$$x'_i = \frac{x_i - \min\{x_1, \dots, x_n\}}{\max\{x_1, \dots, x_n\} - \min\{x_1, \dots, x_n\}} \quad (7)$$

采用熵权法确定各个指标的权重 w_1, w_2, \dots, w_6 , 并将归一化处理之后的各指标进行加权合并,形成技术方案前沿性(Frontier of Technology Solution, FTS)指标,其计算方式如公式(8)所示,提取指标评分前 10% 的技术方案为前沿技术方案。

$$FTS_i = w_1 TN'_i + w_2 CN'_i + w_3 TT'_i + w_4 TA'_i + w_5 BC'_i + w_6 CC'_i \quad (8)$$

1.3 未来技术识别

由于产出前沿技术方案的技术单元可能为该领域的基础技术,因此并非所有与前沿技术方案存在连接的 IPC 均为该领域的未来技术。例如, B60L(电动车辆动力装置)即为电动汽车领域的基础技术,所有应用于电动汽车相关应用场景的技术方案均与该技术单元有关,但其共性太强,难以将其判断为电动汽车领域的未来技术。提取与前沿技术方案存在连接的 IPC 作为潜在未来技术,基于陈捷等^[1]提出未来技术所具有

的方向前瞻性和影响深远性特征,设计技术生长力和技术影响力指标,构建潜在未来技术坐标图,并将同时满足高技术生长力和高技术影响力的技术单元识别为未来技术。

未来技术面向未来应用场景和未来需求,具有产业化潜力,当前处于正在萌芽和兴起的状况,具有较高的生长潜力。技术生长力(Technology Growth, TG)采用每年的专利申请量与过去五年的总专利申请量比值的平均值进行计算, TG 值越高表明技术的生长潜力越大,其计算方式如公式(9)所示,为更好地体现技术的方向前瞻性,本文以最近五年作为区间进行技术生长力的测度。

$$TG_i = \frac{\sum_{j=1}^5 (N_j^i / \sum_{k=j-4}^{k=j} N_k^i) / 5}{1} \quad (9)$$

其中, N_j^i 表示第 j 年申请的包含技术单元 i 的专利数量。

未来技术作为未来产业的底层技术,技术的突破与应用将对产业的发展与产业格局的重塑产生重要影响,本文以技术单元产生的技术方案作为技术对产业的影响方式,以技术方案对技术应用场景的应用能力作为技术影响力的衡量。如表 1, 类型 1 中技术影响力高的技术单元表现为单个(或少数)技术单元便可生成的一项前沿技术方案,并且该技术方案可以应用于多个应用场景;类型 2 中技术影响力低的技术单元表现为该技术单元必须与其他多个技术单元一起才能生成一项技术方案,且该技术方案只能用于单个(或少数)应用场景。基于此,技术影响力(Technology Influence, TI)采用公式(10)进行计算, TI 值越高表明技术的影响力越大。

表 1 技术影响力类型与特征

类型	特征	图示
类型 1	技术方案与技术单元连接少; 技术方案与应用场景连接多; 技术影响力高	
类型 2	技术方案与技术单元连接多; 技术方案与应用场景连接少; 技术影响力低	

$$TI_i = \sum_{j=1}^M (N_j^{apply} / N_j^{ipc}) / M \quad (10)$$

其中, j 表示技术单元 i 产生的前沿技术方案; N_j^{apply} 表示前沿技术方案 j 连接的应用场景数量; N_j^{ipc} 表示前沿技术方案 j 连接的 IPC 数量; M 表示技术单元 i 产生的所有前沿技术方案数量。

2 未来技术识别实证研究:以 SOFC 为例

2.1 数据收集与处理

固体氧化物燃料电池 (Solid Oxide Fuel Cell, SOFC) 是一种在高温下可以将存储在燃料和氧化剂中的化学能转换成电能的全固态能量转换装置,被誉为 21 世纪最具前景的绿色发电系统^[20]。作为一项重要的新能源技术,美国和日本早在上世纪六七十年代就已经开展了 SOFC 的相关研究,并且已经形成了一定的产业规模^[21]。我国对 SOFC 的研究起步较晚,且一度受到国外技术封锁的影响,在技术上与欧美等发达国家仍存在较大差距。对 SOFC 的未来技术进行识别研究,有助于我国把握 SOFC 未来的产业化方向,在此基础上对未来技术进行专项攻关,对我国突破国外技术封锁、把握发展的主动权具有重要意义。

结合相关技术资料 and 专家意见,从 SOFC 相关的关键词和 IPC 分类号两个方面设置检索策略,检索式为:((TS = "solid oxide") AND (TS = "fuel cell \$ " OR IP = B60K-006, 32 OR IP = B60L-003, 00 OR IP = B60L-053, 54 OR IP = B60L-050 * OR IP = B60L-058 * OR IP = B60L-009 * OR IP = B60W-010 * OR IP = H01M-010 * OR IP = H01M-012 * OR IP = H01M-004 * OR IP = H01M-008 * OR IP = H01M-002 * OR IP = C04B-035 * OR IP = C03C

-008 * OR IP = G01R-031 * OR IP = F02C-006 * OR IP = C22C-038 * OR IP = H02J-007 * OR IP = C01B-003 *)) OR TS = SOFC \$ OR TS = H-SOFC \$ OR TS = μ-SOFC \$ OR TS = TF-SOFC \$ OR TS = "Solid oxide electrolysis cell \$ " OR TS = SOEC \$, 检索时间为 2023 年 6 月 20 日。在 DII 数据库中共检索获取到在 2001—2020 年间申请的 1, 1791 条专利。由于数据库中的部分专利存在字段缺失,本文去除申请日期、标题、摘要、IPC 等关键字段缺失的专利,合并同族专利,去除重复专利,最终将得到的 1, 0487 条专利数据作为本研究的对象数据集。

2.2 SOFC 的技术应用场景

提取专利摘要中的 Use 文本,利用 NLTK 术语库对句子进行分词,结合 NLTK 内置的停用词库去除文本中的停用词,并进行词形还原和词干提取。同时,为减少高频主题词对划分效果的影响,将文本中的“sofc”、“solid”、“oxid”、“fuel”、“cell”等具有共性含义的高频主题词进行去除。在 Python 中调用 sklearn 库,利用 LDA 主题模型对预处理后的文本库进行主题聚类,并计算各个主题的困惑度值。筛选“主题数-困惑度”曲线为“肘形”的主题数进行主题划分与识别,结合实际情况确定最优的 14 个技术主题,即 SOFC 技术应用场景(见表 2)。

表 2 SOFC 技术应用场景

主题	应用场景	专利数量	主题词
Topic0	燃料电池结构	798	flat, plate, component, cylindrical, structure
Topic1	多类型燃料电池通用应用	1 058	carbon, acid, molten, polymer, phosphoric
Topic2	燃料重整	467	gas, hydrogen, production, reform, hydrocarbon
Topic3	高温处理与应用	1 298	electrochemical, high, temperature, module, electrolysis
Topic4	燃料电池汽车	1 133	vehicle, power, electricity, supply, energy
Topic5	分布式发电系统	1 345	generate, stack, distribute, integrate, plant
Topic6	电池制造	920	electrode, manufacture, material, air, anode
Topic7	电池组件互连	380	assemble, engine, exhaust, component, interconnect
Topic8	小型电池应用	528	electronic, device, battery, portable, mobile
Topic9	燃料生产	334	carbon, synthesis, turbine, dioxide, ammonia
Topic10	电池密封	1 126	ceramic, cathode, oxygen, prepare, seal
Topic11	电池材料	137	base, lanthanum, medium, solar, reversible
Topic12	电池温度控制	334	separator, coat, agent, surface, thermal
Topic13	电池组件设计	629	electrolyte, apparatus, membrane, single, film

由于 SOFC 技术处于由实验室向实际应用转化的阶段,部分技术尚不成熟,所涉及的商业化应用较少,技术应用的层次也存在较大差异,主要集中在 SOFC 的具体技术,如 SOFC 各组元材料性能优化、单体电池组装结构优化、催化剂制备、燃料制备与处理、温度控制、发电系统结构与优化等。技术的主要应用场景或领域如下:

(1) SOFC 内部构造的设计与优化方案,如燃料电池结构 (Topic0)、电池制造 (Topic6)、电池组件互连 (Topic7)、电池密封 (Topic10)、电池材料 (Topic11)、电池组件设计 (Topic13); (2) 电池燃料的来源与处理,如生产作为 SOFC 燃料的清洁能源 (Topic9) 和对燃料进行重整以提高燃料利用效率 (Topic2); (3) 已在实验室完成实验与设计,但尚未实现大规模推广的

具体应用,如可能会对传统燃油汽车产生替代的燃料电池汽车 (Topic4)、在部分发达国家已经实现推广的 SOFC 分布式发电系统 (Topic5) 以及应用于无人机和电脑等设备上的小型电池 (Topic8); (4) SOFC 反应时温度的控制以及余温利用,如电池工作时余温的处理与应用 (Topic3) 和让电池适应多种场景所进行的温度控制 (Topic12); (5) 除了 SOFC,还能在质子交换膜燃料电池 (PEMFC)、碱性燃料电池 (AFC)、磷酸燃料电池 (PAFC) 等多种燃料电池上通用的技术应用 (Topic1)。

2.3 SOFC 的前沿技术方案

首先提取专利摘要中 Novelty 文本,与专利标题进行合并;其次对文本进行预处理并去除“sofc”、“solid”、“oxid”、“fuel”、“cell”等具有共性含义的高频主题词;最后调用 KeyBERT 包并载入基于英文 BERT 的预训练模型 paraphrase-MiniLM-L6-v2 对处理后的文本进行关键词提取,其算法的伪代码如下:

算法:技术方案关键词提取算法
输入:专利数据 $\{P_1, P_2, \dots, P_n\}$;具有共性含义的高频主题词表 $\{W_1, W_2, \dots, W_m\}$

输出:技术方案关键词 $\{H_1, H_2, \dots, H_n\}$,其中 H_n 表示从专利 P_n 提取的技术方案关键词表 $\{h_{n1}, h_{n2}, \dots, h_{nk}\}$

```
1:for  $P_i$  in  $\{P_1, P_2, \dots, P_n\}$  do
2:  title  $\leftarrow$  extract title from  $P_i$ //提取专利的标题
3:  novelty  $\leftarrow$  extract novelty from  $P_i$ //提取专利摘要的新颖性
   字段
4:  text  $\leftarrow$  combine (title, novelty) //将标题与新颖性文本合并
5:  update text  $\leftarrow$  process text //文本预处理,去除停用词、标点与特殊符号,对高频词的单复数形式和时态进行合并
6:  for  $W_j$  in  $\{W_1, W_2, \dots, W_m\}$  do
7:    if  $W_j$  in text then
8:      update text  $\leftarrow$  text.remove( $W_j$ ) //去除文本中具有
      共性含义的高频主题词
9:    end if
10:  end for
11:   $H_i \leftarrow$  KeyBERT (text, ngram = (1, 2)) //运行 KeyBERT
   算法从文本中提取关键词,每个关键词设置为 1 或 2 个单词
12: end for
```

结合 SOFC 的技术主题词对提取的关键词构建技术方案词表,对具有相同含义的技术方案关键词进行合并和替换,提取的每个专利的技术方案关键词如表 3 所示。

表 3 技术方案关键词提取

专利编号	专利标题与摘要 (Novelty 部分)	技术方案关键词
P1	标题:Combined system of solid oxide fuel cell and industrial processing unit, heats power generating fuel of fuel cell, using exhaust gas of industrial process which uses exhaust gas from cell, as oxidant 摘要:A modifier heats and modifies power generating fuel of solid oxide fuel cell,	generate exhaust, process exhaust, exhaust gas, power generate, gas industrial
P2	标题:Solid electrolytic-type fuel cell e. g. solid oxide fuel cell has fuel and air electrode collectors arranged between electrode layers and separator, such that air holes of big and small size contact layers and separator 摘要:The fuel and air electrode collectors consists of air holes of different size,	electrode collector, electrolyte layer, electrode layer, electrolytic, electrolytic air
...
P10487	标题:Control apparatus of energy supply system used for server installed in data center, has main portion that is controlled to connect with grid of surrounding area and to supply electric power to grid of surrounding area 摘要:The apparatus has main portion that controls an energy supply system provided with	control energy, power supply, control apparatus, power grid, energy supply

结合技术方案关键词在关键词共现网络中的中介中心性,筛选出现频次大于 2 的技术方案关键词,共获取 2, 963 项技术方案。根据公式(1)—(6)对每项技术方案的新颖性、关联性和重要性进行计算,标准化处理之后,利用熵权法对各指标的权重进行计算,技术方案新颖性、技术方案关联性和技术方案重要性三个维度的权重计算结果分别为 0.017、0.407 和 0.576。

根据各指标的测度结果和权重,对各技术方案的前沿性进行评分,选择评分排名前 10% 的 296 项技术作为前沿技术方案 (表 4)。评分最高的电解质层 (electrolyte layer) 技术是 SOFC 的基础技术方案,与 46 个 IPC 存在关联,并且在所有场景都存在应用。部分技术方案具有较高的特定场景应用特征,如碳氢化合物处理 (hydrocarbon treatment) 的词频和评分均较高,但其是与燃料处理相关的技术方案,只能应用于

SOFC 燃料生产 (Topic9)。

表 4 前沿技术方案 (部分)

技术关键词	词频	关联 IPC 数	应用场景数	技术方案前沿性
electrolyte layer	405	46	14	0.628
power generation	165	37	10	0.427
heat exchanger	189	49	13	0.417
electrolyte	274	49	14	0.347
cathode	215	28	13	0.301
air electrode	162	22	14	0.288
electric power	141	27	9	0.282
electrode layer	160	26	11	0.262
hydrocarbon treatment	37	15	1	0.258
clamping device	3	3	1	0.243
electrode	198	27	13	0.237
stack	144	21	12	0.233
hydrocarbon containing	24	15	1	0.229
polymer electrolyte	117	27	8	0.227
coal formation	25	16	1	0.219

2.4 SOFC 的未来技术

为减少技术方案与技术单元和应用场景相关性较低的连接所造成的干扰,提高研究的准确性,本文筛选与前沿技术方案共现频次大于 2 的 IPC 作为技术单元,筛选前沿技术方案应用了 3 次及以上的场景为技术方案对应的应用场景。根据技术单元、前沿技术方案、应用场景间的共现关系构建矩阵,随后利用 Gephi0.10 绘制三者间的共现网络,并将相同类型的节点进行手动聚类,生成技术单元产生前沿技术方案并应用于相应应用场景的流程图(图 3),可以看出在 SOFC 领域,69 项技术单元产生的 296 项前沿技术方案应用在了 14 个应用场景。其中,H01M(用于直接转变化学能为电能的方法或装置)作为 SOFC 的基础技术单元,产生了 257 项前沿技术方案,C21B(铁或钢的冶炼)、D01F(专用于生产碳纤维的设备)、F24F(空气调节)等技术单元为在某一方面优化 SOFC 生产和应用的相关技术,各仅产生了一项前沿技术方案。

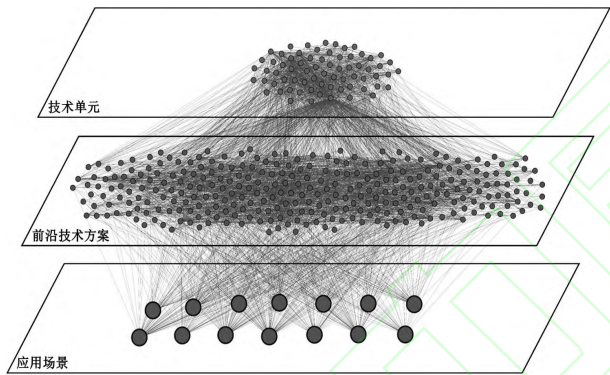


图 3 SOFC 技术单元应用流程

以 69 项技术单元作为潜在未来技术,测算每项技术的技术影响力及近五年的技术生长力,绘制 SOFC 领域潜在未来技术坐标图(图 4)。利用所有潜在未来技术生长力和技术影响力的均值^[22]将坐标轴划分为四个象限。其中,第一象限包括 F23D(燃烧气体燃料的燃烧器)、G01R(测量电变量;测量磁变量)、F24S

(太阳能热收集器;太阳能热系统)等,这部分技术具有较高的技术生长力和技术影响力,是 SOFC 领域的未来技术。第二象限包括 F16F(弹簧;减震器;减振装置)、F01B(一般的或变容式的机器或发动机)、B64D(用于与飞机配合或装到飞机上的设备)等,这部分技术的技术生长力较高,但产生的技术方案只适用于特定应用场景,如 B64D 为用于与飞机配合或装到飞机上的设备,飞行器用 SOFC 为热门研究方向,但其产生的技术很难应用于其他应用场景,因此技术影响力较低。第三象限包括 C23F(非机械方法去除表面上的金属材料)、H01G(电容器)、B23K(钎焊或脱焊)等,这部分技术的技术生长力和技术影响力均较低,能与其他技术单元共同产生前沿技术方案,但无法成为未来技术。第四象限包括 F28D(热交换设备)、F28F(通用热交换或传热设备的零部件)、B82Y(纳米结构的制造或处理)等,这部分技术具有较高的技术影响力,但由于在关键技术上存在限制等因素,导致近五年的技术生长力较低,如果关键技术实现突破,有望成为 SOFC 领域的未来技术。

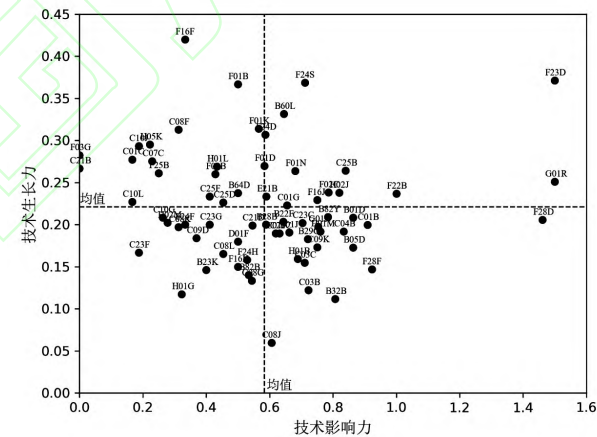


图 4 SOFC 领域潜在未来技术坐标图

SOFC 领域未来技术及其代表性技术方案如表 5 所示。

表 5 SOFC 领域未来技术

IPC	IPC 含义	代表性技术方案
F23D	燃烧气体燃料的燃烧器	burner, reaction chamber, gas outlet, gas turbine, compressor
G01R	测量电变量;测量磁变量	voltage measurement, distribution electrode, testing device, electrical property, galvanic
F22B	蒸汽的发生方法	steam generator, waste heat, power plant, water steam, boiler steam
C25B	生产化合物或非金属的电解工艺或电泳工艺	electrolyte layer, air electrode, cathode, electrode layer, anode
F24S	太阳能热收集器;太阳能热系统	heat storage, stainless steel, gas turbine, solar heat, electrolytic hydrogen
H02J	供电或配电的电路装置或系统;电能存储系统	electric power, power supply, control apparatus, hybrid power, hydrogen storage
F02C	燃气轮机装置	gas turbine, compressor, gas discharge, gas combustor, gas control
F16J	活塞;缸;一般压力容器;密封	sealing material, gas seal, assembly stack, structure sealing, gasket
B60L	电动车辆动力装置	power generation, control unit, power supply, heat exchanger, control apparatus
F01N	一般机器或发动机的气流消音器或排气装置	exhaust gas, combustion engine, catalytic converter, anode structure, gas guiding
F24D	住宅供热系统或区域供热系统	heat pump, unit heating, waste heat, heat storage, biomass gasification

续表 5 SOFC 领域未来技术

IPC	IPC 含义	代表性技术方案
C01G	含有不包含在 C01D 或 C01F 小类中之金属的化合物	electrode material, anode material, stabilized zirconia, perovskite, electrode electro-chemical
E21B	从井中开采油、气、水、可溶解或可熔化物质或矿物泥浆	synthesis gas, coal formation, pyrolysis hydrocarbon, carbon dioxide, process gas

可见,SOFC 领域未来技术大致可分为三个方向:

(1)提高 SOFC 批量化制造的未来技术。SOFC 要实现广泛的商业化应用,必须降低生产的成本,提高生产的效率和产品的合格率。早期 SOFC 的电解质是挤压成形的,而挤压工艺不可能形成足够薄的电解质层,进而限制了 SOFC 的小型化发展。C25B 所对应的电解工艺或电泳工艺与 SOFC 制备中的电泳沉积工艺和电化学气相沉积工艺相关,其中电泳沉积是将悬浮液中的带电胶体粒子在外加电场作用下移向沉积基体并形成材料的工艺,该工艺由于设备简单、制备薄膜密度高、厚度易控制等优势,是目前 SOFC 制备中的主要方法^[23]。SOFC 的电极、电解质等部件均是由陶瓷材料构成,这部分材料不仅要有较高的电导性和稳定性,还要有足够的机械强度,同时还应控制生产的成本,C01G 所对应的氧化钇稳定氧化锆(YSZ)、镧锶锰氧(LSMO)等化合物为 SOFC 的主要材料^[24]。

(2)增强 SOFC 可控性运行的未来技术。以风能、太阳能为代表的清洁能源无法完全取代传统化石能源的很大一部分因素在于能量输出的不稳定性,早期的 SOFC 由于无法有效控制化学反应的温度和效能也存在着同样的问题,F23D、G01R、F22B 分别是与 SOFC 反应控制、电力控制和温度控制相关的技术。合适的反应器可以使燃料的利用效率达到最大,并实现电力输出的稳定性和持续性;利用电力监测和控制系统,可以控制电池的电力输出功率,从而使其适应不同场景的应用;SOFC 也是使用水冷装置作为温度的控制方式,液态水在热交换系统中吸收热量并形成蒸汽排出,从而达到控制系统运行温度的目的,不仅保证了 SOFC 运行时的稳定性和可靠性,排出的蒸汽也可用于燃料重整和其他系统的供热。目前 SOFC 已经实现了大型固定发电设备的应用,移动供电应用是主要的研究方向和未来的发展方向,F16J 所对应密封技术可以让 SOFC 如蓄电池一样所有的燃料供应和反应流程在电池内部封闭进行,F01N 对应的排气和消音装置降低了电池工作时的噪音,进一步提升了移动端 SOFC 的可用性。

(3)扩展 SOFC 实际化应用的未来技术。作为一种可以对传统化石能源产生替代的清洁能源,SOFC 可以适用于绝大部分能源应用领域,F24D 对应的家庭热电联供系统是利用 SOFC 进行发电并利用余热进行家庭供热的系统,已经在美国、日本等国家实现了应

用;B60L 对应的 SOFC 动力汽车不仅具有比燃油汽车更清洁的特征,能源补充方式也比电动汽车更便捷,是各国新一代汽车的主要研究方向;SOFC 燃料的多样性也让其可以在 E21B 对应的油田、煤矿等矿物开采工作中发挥作用,利用采矿现场的废气为 SOFC 提供燃料,既可以减少燃气外泄造成的危害,产生的电力也可以用于采矿作业,促进安全生产和节能减排^[25]。SOFC 也可以与其他系统一起协作使用,提升系统的发电效率和资源的利用率,利用 F24S 对应的太阳能技术可以对燃料进行重整^[26],进一步提升燃料利用率,反应产生的废气还可以用于 F02C 对应的燃气轮机,组成联合发电系统^[27],进一步提升发电效率。

3 结论与启示

本文从技术单元产出技术方案以解决应用场景的技术问题出发,构建了基于应用场景的未来技术的识别模型。首先,利用 LDA 主题模型对专利摘要的用途(Use)文本进行主题聚类,识别技术的主要应用场景。然后,利用 KeyBERT 算法从专利标题和摘要的新颖性(Novelty)文本中提取技术方案关键词,并从技术方案的新颖性、关联性和重要性出发,进一步筛选前沿技术方案。最后,筛选产生前沿技术方案的技术单元作为潜在未来技术,构建潜在未来技术的坐标图,识别具有高技术影响力和高技术生长力的技术单元为未来技术。本文以固体氧化物燃料电池领域为例进行了实证研究,识别出 13 项未来技术,主要为 SOFC 批量化制造、可控性运行和实际化应用方面的技术。

本文的创新之处在于,首先,在理论上分析了未来技术的产生过程,阐明了技术单元面向应用场景的应用流程,并以此为基础提出了未来技术的识别方法。其次,在研究方法上采用了由应用场景到技术单元的逆向识别过程,从应用视角出发深入挖掘未来技术。考虑到未来技术是尚处萌芽期和生长期的技术,因此有些技术并未以专利形式进行记载,如何对此类技术进行追踪和识别是未来研究中值得考虑的问题。此外,本文的实证研究选用了固体氧化物燃料电池的专利数据,虽然研究领域具有较好的代表性,但缺乏与其他技术领域未来技术识别的比较,后续研究可将本文设计的未来技术识别模型和方法应用于更多领域进行进一步检验,并结合专家意见进行未来技术的识别。

参 考 文 献

- [1] 陈捷, 吴仲琦, 代涛. 未来技术风险识别框架研究——基于技术经济安全视角[J]. 中国科学院院刊, 2023, 38(4): 570-579.
- [2] Bower J L, Christensen M C. Disruptive technologies: Catching the wave[J]. Harvard Business Review, 1995, 73(1): 75-76.
- [3] Rotolo D, Hicks D, Martin B R. What is an emerging technology? [J]. Research policy, 2015, 44(10): 1827-1843.
- [4] 刘春文, 黄鲁成, 苗红, 等. 基于 ECT-Dim 的前沿技术识别方法——以养老环境辅助生活技术为例[J]. 情报杂志, 2023, 42(08): 77-82, 76.
- [5] 刘志辉, 张均胜, 林毅, 等. 基于隐性知识的潜在颠覆性技术评估方法研究[J]. 情报学报, 2021, 40(12): 1271-1278.
- [6] Jang W, Park Y, Seol H. Identifying emerging technologies using expert opinions on the future: A topic modeling and fuzzy clustering approach[J]. Scientometrics, 2021, 126(8): 6505-6532.
- [7] 赵新月, 赵筱媛. 基于高层次创新人才科研动向监测的前沿技术识别体系[J]. 情报学报, 2021, 40(12): 1279-1287.
- [8] 杨水飏, 石诚. 自然辩证法概论[M]. 北京: 国防工业出版社, 2017.
- [9] 发展未来产业要牢牢把握四个“未来”-新华网[EB/OL]. [2023-08-17]. <http://www.xinhuanet.com/tech/20230418/b60c3b519aa74c3f8dfa5249b9935331/c.html>.
- [10] 李平, 田朔. 市场需求对技术创新的门限特征分析[J]. 经济问题探索, 2014(10): 18-25.
- [11] 程鹏, 梁艳, 柳卸林, 等. 突破性产品创新群的实现机制——基于应用场景视角[J]. 科学学与科学技术管理, 2022, 43(12): 76-93.
- [12] 伊惠芳, 吴红. 多级需求分析视域下高校专利转移对象识别研究——以石墨烯为例[J]. 图书情报工作, 2020, 64(12): 118-126.
- [13] 俞荣建, 王雅萍, 赵一智, 等. 破解“卡脖子”技术难题: “情境—策略”非对称匹配视角[J]. 中国科学院院刊, 2023, 38(4): 580-592.
- [14] 吴菲菲, 李倩, 黄鲁成. 基于专利 SAO 结构的技术应用领域识别方法研究[J]. 科研管理, 2014, 35(6): 1-7.
- [15] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet Allocation [J]. Journal of Machine Learning Research, 2003, 3(1): 993-1022.
- [16] 宋培彦, 田文波, 冯超慧, 等. 融合选择性注意衰减模型的信息简报自动生成方法研究——以 UNESCO 科技报告为例[J]. 情报理论与实践, 2023, 46(4): 53-58+7.
- [17] 李昌, 杨中楷, 董坤. 基于多维属性动态变化特征的新兴技术识别研究[J]. 情报学报, 2022, 41(5): 463-474.
- [18] 江曼, 孙明汉, 余翔, 等. 协同创新视角下技术机会识别模型——以中国智能机器人为例[J]. 情报杂志, 2020, 39(7): 42-48.
- [19] 马永红, 孔令凯, 林超然, 等. 基于专利挖掘的关键共性技术识别研究[J]. 情报学报, 2020, 39(10): 1093-1103.
- [20] 王邵荣, 叶晓峰. 固体氧化物燃料电池技术[M]. 武汉大学出版社, 2015.
- [21] 许元武. 固体氧化物燃料电池性能演变评估与健康管控研究[D]. 华中科技大学, 2023.
- [22] 郑赛硕, 王学昭, 陈小莉. 共性技术识别方法构建与实证研究——以集成电路行业为例[J]. 图书情报工作, 2021, 65(15): 130-139.
- [23] 赵秉国, 刘亚迪, 胡浩然, 等. 制备固体氧化物燃料电池中电解质薄膜的电泳沉积法[J]. 化学进展, 2023, 35(5): 794-806.
- [24] 杜柯, 宋琛, 余敏, 等. 固体氧化物燃料电池氧化钪稳定氧化锆电解质的等离子喷涂制备及性能[J]. 硅酸盐学报, 2022, 50(7): 1929-1935.
- [25] 王鑫鑫. 基于固体氧化物燃料电池的煤矿低浓度瓦斯高效清洁利用研究[D]. 中国矿业大学, 2021.
- [26] 郑志美, 刘泰秀, 刘启斌. 太阳能热化学与燃料电池联合的发电系统[J]. 工程热物理学报, 2020, 41(11): 2627-2634.
- [27] 连琰珂, 明平文, 蔡黎明. 固体氧化物燃料电池/燃气轮机混合动力系统建模仿真研究进展[J]. 洁净煤技术, 2023, 29(3): 26-39.