

# 基于主题挖掘与专利评估的技术机会识别研究\*

## ——以智慧农业为例

■ 宋凯<sup>1</sup> 冉从敬<sup>2</sup>

<sup>1</sup> 山东师范大学图书馆 济南 250399 <sup>2</sup> 武汉大学信息管理学院 武汉 430072

**摘要:** [目的/意义] 技术机会识别是企业技术创新活动得以顺利开展的重要前提,及时发现和把握有价值的技术机会对技术创新突破意义重大。[方法/过程] 从主题挖掘与专利评估的视角提出一种技术机会识别方法,首先应用主题模型识别技术领域涵盖的技术主题并进行专利聚类;其次在技术主题层面展开细粒度分析,综合考虑技术机会应具备前沿性、价值性和时效性的重要特性,采用突变级数法和离群因子算法评估技术主题中的高价值专利和离群专利形成核心专利集,并计算每个技术主题的专利平均年龄;最后,将技术主题中的核心专利占比和专利平均年龄作为核心指标绘制技术机会识别地图,用于识别技术机会。[结果/结论] 以智慧农业领域为例,对所提方法进行实证,识别智慧农业领域的 5 个技术机会,为创新主体的技术研发提供决策支持。但识别结果的定量验证以及融合多源数据进行技术机会识别有待进一步探索。

**关键词:** 主题模型 价值评估 技术机会 智慧农业

**分类号:** G250

**DOI:** 10.13266/j.issn.0252-3116.2023.03.006

科技产业变革快速演进,技术创新成为驱动社会经济发展、提升企业技术竞争力的关键所在。而随着技术研发生命周期的不断加快,企业面临的研发竞争环境也愈加复杂,如何快速准确地帮助企业识别技术机会,在技术创新全流程的起点环节占据主动权,对企业抢占市场先机、获取竞争优势至关重要。但技术机会并不是以显性方式呈现,需要运用科学的方法和工具从论文、专利、研究报告等多源数据中进行挖掘。因此,帮助企业从海量数据中发掘潜在技术机会逐渐得到重视与发展。

技术机会是指“技术进步可能性的集合”,通过挖掘发现具有价值性和前沿性的技术机会,是企业技术创新活动中保持竞争优势的重要环节<sup>[1]</sup>。技术机会识别,也被称作技术机会分析或技术机会发现,尽管名称有所差异,但其内涵基本一致。在技术机会识别研究中,关键内容是依托一定的技术方法,通过检索、收集、处理和分析的过程,从情报数据中挖掘潜在技术机会的过程和方法,而技术方法的有效性和科学性决定了技术机会识别的合理性。因此,如何完善技术机会识别方法,成为技术机会分析研究的热点和难点。

鉴于此,笔者提出一种技术机会识别方法,从前沿性、价值性和时效性 3 个角度出发,首先,面向技术领域进行主题识别与专利聚类;其次,对各个主题内的专利进行价值评估和离群专利识别,将二者结合形成核心专利集;最后,基于各个主题的核心专利占比与专利平均年龄,绘制技术机会识别地图,探析技术领域涵盖的技术主题当前发展态势,从而识别技术机会,为企业进行研发投资提供决策支持。

## 1 相关研究

技术机会识别的概念是在 20 世纪 90 年代初,由佐治亚理工学院科技政策与评估中心的 A. L. Porter 教授团队提出。之后,技术机会识别方法取得了不断延伸与长足发展。

### 1.1 基于定性方法的技术机会识别研究

技术机会识别方法初始以定性方法为主,如采用德尔菲法<sup>[2]</sup>、层次分析法<sup>[3]</sup>等,此种方式主要依赖专家的经验学识进行判断。为了能够对复杂问题进行建模,形态分析法应运而生,该方法通过问题构建和参数分解的非定量建模过程,能够在语义层面可视化地展示领域知识结构。郭俊芳等<sup>[4]</sup>引入主语-谓语-

\* 本文系国家自然科学基金重大项目“大数据主权安全保障体系建设研究”(项目编号:21&ZD169)研究成果之一。

作者简介:宋凯,馆员,博士;冉从敬,教授,博士生导师,通信作者,E-mail: rancongjing@163.com。

收稿日期:2022-09-05 修回日期:2022-10-28 本文起止页码:61-71 本文责任编辑:王传清

宾语 (SAO) 语义挖掘方法, 提出一种新型的技术形态识别方法, 提高了技术形态识别的效率; K. Kim 等<sup>[5]</sup>采用 SAO 方法对人机交互领域的技术机会进行识别, 从而为企业研发投资提供支持; M. C. Chen 等<sup>[6]</sup>应用技术-服务演化分析探索物联网物流中的技术机会; L. J. Feng 等<sup>[7]</sup>提出一种基于形态分析和统一结构创新性思维的技术机会识别方法。基于定性方法的技术机会识别, 其优势在于依靠专家对技术领域的前沿性预测, 以获得具有创造性和洞察力的技术机会。但随着文献数量呈爆炸性增长, 专家难以凭借专长经验快速准确地分析海量数据, 无法应对和满足企业的实时需求。因此, 探索定量技术应用于客观数据分析以识别技术机会的重要性愈发凸显。

## 1.2 基于定量方法的技术机会识别研究

技术机会识别的定量方法由 A. L. Porter 等提出, 他们在分析过程中结合专利计量与专家意见以分析技术发展态势, 为后来的定量研究提供了思路和启发。随着数据挖掘、自然语言处理技术的快速发展, 这些技术也被逐渐应用于挖掘技术机会。其中, 基于专利地图的方法是技术机会识别的重要方法之一, 其核心思想是对专利文本进行自然语言处理后, 采用  $n$  维关键词向量表示专利, 进一步通过降维技术在二维空间对专利进行可视化展示, 将图中出现的空白点视作技术机会。常用的方法有自组织映射法 (self-organizing map, SOM)、生成式拓扑映射法 (generative topographic mapping, GTM) 等。S. Lee 等<sup>[8]</sup>依托主成分分析法, 提出一种基于关键词的专利地图绘制方法, 并以个人数字助理技术为例进行实证; E. Chun 等<sup>[9]</sup>采用 SOM 算法绘制技术地图, 识别了智慧农场的空白技术; J. Yu 等<sup>[10]</sup>应用 GTM 算法绘制了物流技术领域的技术地图, 并对空白技术和新兴技术进行挖掘; X. T. Han 等<sup>[11]</sup>基于专利申请和无效状态, 绘制技术发展地图用于识别技术机会; 桂美增等<sup>[12]</sup>采用 GTM 算法对新能源汽车的技术机会进行预测; 吴一平等<sup>[13]</sup>提出一种融合评论主题识别与技术属性多维度分析的技术机会发现方法。但采用基于技术地图的方法, 识别出的技术机会由关键词的组合表示, 准确解读技术含义具有一定的困难; 此外, 在进行降维过程中会造成部分关键信息丢失的情况, 所以应用该方法存在一定局限性。

为了弥补基于专利地图方法的不足, 研究人员将基于异常值检测的方法应用于技术机会分析中, 该方法的核心思想是将与主流技术存在明显差异的专利

视为“离群专利”, 这部分专利隐含着创新技术和技术进步的内容, 可为企业新技术研发提供决策参考。该方法的优势在于能够系统、定量地筛选出可能的前沿性专利, 所表达的技术机会含义相对明确。常用的方法有基于距离的 K-means 算法、基于密度的局部异常因子算法 (local outlier factor, LOF) 等, 其中 LOF 算法能克服 K-means 算法的不足, 应用范围更为广泛。B. Yoon 等<sup>[14]</sup>采用 LOF 算法提出了企业技术研发空白的识别方法; B. Kim 等<sup>[15]</sup>应用异常值检测算法, 提出一种专利引用网络潜在技术机会识别方法; X. Shi 等<sup>[16]</sup>应用主题模型、多维尺度和 LOF 算法, 识别了新能源汽车领域的技术机会; 杨辰等<sup>[17]</sup>采用文本挖掘与异常值检测方法识别了深度学习领域的技术机会。

此外, 基于链路预测的方法也是技术机会识别的新方向, 该方法的优势是能揭示网络未出现的链接, 实现追踪技术的动态变化与新的知识组合。翟东升等<sup>[18]</sup>利用专利之间的引用关系, 提出一种基于链路预测的技术机会挖掘方法。I. Park 等<sup>[19]</sup>通过构建引文网络, 采用链路预测法识别生物技术与信息技术领域的技术机会; X. T. Han 等<sup>[20]</sup>结合形态分析中的 SAO 语义发掘方法与链路预测法, 对医学领域皮肤恶性肿瘤的技术机会进行研究; B. Yoon 等<sup>[21]</sup>结合技术机会地图与链路预测方法, 提出一种划分新兴、稳定和衰退等级的技术机会识别方法; J. H. Lee 等<sup>[22]</sup>依托多维专利索引系统——F-term, 采用链路预测识别企业技术组合中的技术机会。但应用链路预测识别的技术机会, 其输出结果形式为“IPC 对”或“关键词对”, 不利于技术细节的揭示。

经过上述分析发现, 当前关于技术机会识别的方法主要有定性和定量两种方法, 而基于定性的方法已不能适用于不断递增的文献量级。随着新兴技术的发展, 基于定量方法具备的无监督、高效率等特征, 在技术机会识别中得到广泛应用。结合已有研究, 技术机会的本质是领域内技术进步的可能性, 应当具备 3 个重要特性: 首先, 应具备前沿性, C. Lee 等<sup>[23]</sup>认为前沿性是技术机会的关键特性, 具备前沿性的技术机会才能为企业提前开展新技术研发布局提供决策参考。其次, 要具备价值性, X. X. Zhao 等认为技术机会识别的前提是要获取技术领域内的高价值专利, 以此为基础挖掘的技术机会才值得企业研发投入<sup>[24]</sup>。最后, 要具备时效性, J. Choi 等<sup>[25]</sup>和黄鲁成等<sup>[26]</sup>认为技术机会是在特定领域内发现的潜在技术趋势

和方向,是近期隐现、在未来具备发展潜力的新技术。而目前基于定量方法进行技术机会识别,较少从前沿性、价值性和时效性3个方面进行综合考量。因此,笔者综合考量技术机会的3个重要特性,应用主题模型、突变级数法、LOF算法等,构建一种基于主题挖掘与专利评估的技术机会识别方法,为企业明晰研发方向,发掘技术创新点提供决策参考。

## 2 基于主题挖掘与专利评估的技术机会识别模型构建

为了实现技术机会识别的目标,笔者以专利数据为基础,结合专利计量相关技术方法,构建基于主题挖掘与专利评估的技术机会识别模型,如图1所示:

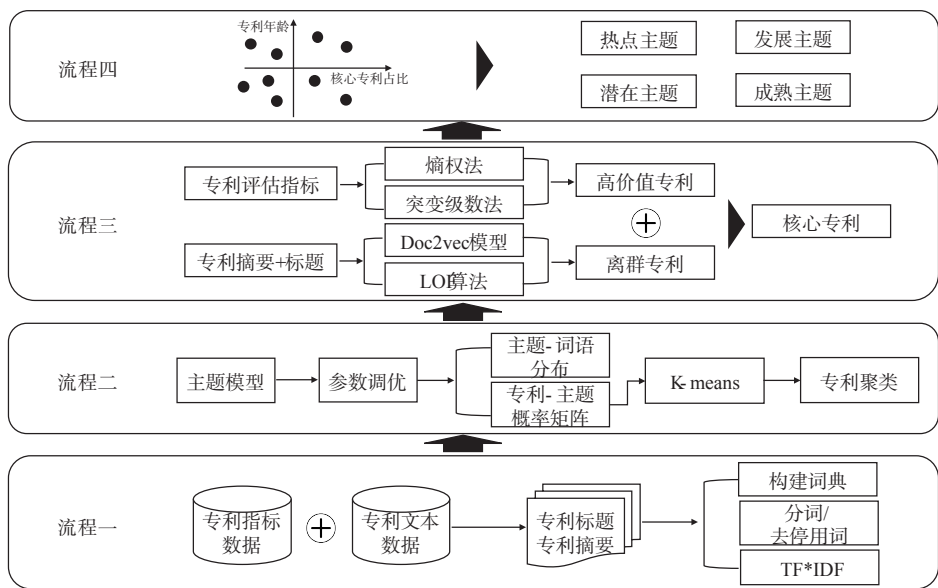


图1 技术机会识别模型

### 2.1 流程一：数据检索与预处理

首先,选取特定技术领域,围绕技术领域形成组配词,在数据库中进行专利检索,获取技术领域下的权利有效专利;其次,对检索到的专利数据进行预处理,根据分析的需要,分别获取专利指标数据用于核心专利评估,以及专利文本数据用于技术领域主题挖掘。考虑到专利标题和摘要是对专利技术内容的高度凝练,因此抽取专利标题和摘要,应用python的jieba库和gensim库进行分词、去停用词、TF\*IDF的自然语言处理过程,在分词过程中,通过获取同领域学术论文中的关键词以及专利技术功效词形成领域专业词典。通过上述过程,最终获得实验语料集。

### 2.2 流程二：主题识别与专利聚类

应用python的gensim库对实验语料集进行主题抽取,借助pyLDAvis库的可视化呈现确定最优主题数目;通过设置主题数目,调整alpha和eta两个隐含参数,形成最优主题模型,分别得到每个技术主题下的相关主题词和“专利—主题”概率矩阵。其中,每个技术主题下的相关主题词用于对技术主题名称进行总结,“专利—主题”概率矩阵将与K-means算法结合进行专利聚类<sup>[27]</sup>。

### 2.3 流程三：技术主题核心专利评估

在实现专利聚类后,将每一件专利划分到一个技术主题中;进而对每个技术主题中的专利进行细粒度评估,以获取核心专利。核心专利是具备较高应用价值和经济价值的专利,应用价值体现在核心专利刻画的技术内容具有前沿性,C. Lee等<sup>[28]</sup>提出一种基于LOF算法的专利前沿度判别方法,将通过LOF算法识别的离群专利视为具备前沿性并应用于技术机会识别;经济价值体现在核心专利具备价值性,冉从敬等<sup>[29]</sup>选取专利价值评估指标,通过采用熵值法与神经网络算法识别高价值专利以辅助技术交易决策。因此,对技术主题核心专利的评估主要从两个方面开展:一方面是计算专利价值性;另一方面是计算专利前沿性。最后将结果取并集以保证核心专利既具备价值性,又具备前沿性。

#### 2.3.1 评估专利价值,识别高价值专利,凸显核心专利的价值性

在专利价值评估过程中,参考前期研究成果<sup>[30-31]</sup>,通过查阅国家知识产权局的《专利价值分析指标体系操作手册》,将保护范围、研发水平、市场前景3个维度作为一级指标,并进一步选取10个二级指标,具体解释如表1所示:



表 1 专利价值评估指标

一级指标	二级指标	解释
保护范围	权利要求数量	反映专利权的法律稳定性，权利要求数量越多，专利保护范围越大
	IPC 分类号数量	反映专利应用前景，IPC 数量越多，技术应用领域越广泛
	文献页数	反映专利技术复杂程度，专利文献页数越多，技术复杂程度越高
研发水平	家族引证次数	反映专利技术背景，家族引证次数越多，表示相对已有技术做了显著改进，专利权更加稳定
	家族被引证次数	反映专利的原创水平和影响力，同族专利在不同国家和地区被引证次数总和越高，表示专利创新性和影响力越大
	发明人数量	反映专利的人才基础和资源投入状况，发明人数量越多，表明技术研发投入规模大、知识基础扎实
市场前景	同族个数	反映专利布局情况，同一项技术在多个国家申请专利，表示专利布局全面，应用前景较好
	许可 / 转让次数	反映专利的市场应用前景，如果专利实现转让或许可，表示专利具备较好的应用前景
	技术先进性	incoPat 数据库给出的专利指标，分值在 1-10 之间，分数越高，技术先进性越好
	技术稳定性	incoPat 数据库给出的专利指标，分值在 1-10 之间，分数越高，技术稳定性越好

选取专利评估指标后，需要基于指标计算对专利价值进行评估。已有的专利商业数据库（如 incoPat）提供了专利价值度的评估指标，但相关研究发现，incoPat 提供的专利价值度范围为 1-10，分值范围较窄，价值区分度相对较低<sup>[32]</sup>。所以，为了扩大指标区分度，在已有研究中，较多采用多元统计方法评估专利价值，如层次分析法<sup>[33]</sup>、模糊评价综合法<sup>[34]</sup>等。但这类方法较难解决指标量化问题，导致研究人员的主观评价对结果合理性影响较大。而突变级数法能有效弥补过往研究方法的不足，在评估过程中不需要对指标赋予权重，减少了主观性，使得评估结果客观科学，在生态环境<sup>[35]</sup>、城市竞争力<sup>[36]</sup>、产业发展水平<sup>[37]</sup>等评估场景中得到广泛的应用。突变级数法的理论基础是突变理论中的各类初等突变模型，并结合模糊数学隶属度函数，通过归一化的过程得到评估结果，用于对评估对象进行排序<sup>[38]</sup>。虽然突变级数法相对已有方法，具备评估结果科学合理的优势，但在具体应用过程中需要先确定评估指标的重要程度。因熵权法可依据原始数据确定指标权重，计算结果客观合理，可应用于确定评估指标权重，并进一步结合突变级数法进行专利价值评估。熵权法与突变级数法的结合可最大限度避免评估过程中的主观性因素影响，具体计算过程如下：

（1）确定专利评估指标各层级突变类型。常见的突变系统类型有 4 种，其中，只有 1 个子指标为折迭突变系统，有 2 个子指标为尖点突变系统，有 3 个子指标为燕尾突变系统，有 4 个子指标为蝴蝶突变系统<sup>[39]</sup>。在每个突变系统中都有一个状态变量，而每个状态变量依据突变系统类型的不同，由控制变量计算得到。在专利价值评估场景中，控制变量表示为专利评估指标，而控制变量的重要性则由熵权法计算的权重确定。

（2）利用归一公式进行评估。根据突变级数法的取值原则，控制变量的取值区间为 0-1。所以，二级评估指标如果为正向指标，则直接应用 Max-Min 标准化方法进行处理。如果一级评估指标的二级指标之间存在明显关联关系，则将二级指标的关系定为“互补型”，那么一级指标的评估值为通过二级指标计算得到状态变量的平均值；如果二级指标之间无明显关联关系，则将二级指标的关系定为“非互补型”，那么一级指标的评估值按照大中取小原则对状态变量进行取值。

### 2.3.2 挖掘离群专利，识别前沿技术，凸显核心专利的前沿性

已有研究将离群点检测技术应用到技术机会识别中，通过该技术可识别得到一定时期内某技术领域的离群专利，离群专利反映的技术内容与主流技术存在明显差异，具有较高的前沿性，可视为前沿技术，能够为企业研发创新提供重要的技术导向<sup>[40]</sup>。鉴于离群点检测方法在技术判别应用中取得的良好效果，笔者采用该方法对技术主题中的离群专利进行识别。具体计算过程如下：

（1）专利文本数据预处理。在专利聚类的基础上，获取每个技术主题涵盖的专利。获取每件专利的标题和摘要，应用 python 中的 jieba 库对专利摘要进行分词、去停用词的自然语言处理过程，形成“专利—关键词”实验语料集。

（2）专利文本向量化表示。对文本进行向量化的过程，通常应用较多的是隐含狄利克雷分布（Latent Dirichlet Allocation, LDA）模型，从主题层面对专利进行向量化表示。但该方法没有考虑文本中的词项顺序、语法句法等上下文信息，导致向量化后的文本缺乏语义信息。2014 年，T. Mikolov 等提出了段落向量化的深度学习模型 Doc2vec，该

模型不仅加入文本语义信息,同时具有更好的泛化能力,在学术研究中进行文本向量化表示得到广泛的验证和应用<sup>[41-42]</sup>。因此,笔者选用 Doc2vec 模型对各个技术主题下的专利文本进行向量化表示。

(3) 离群专利识别。通过对技术主题中的专利文本进行向量化表示,进一步采用离群点检测方法识别技术主题内的离群专利。在离群点检测方法中,基于密度的离群检测 LOF 算法在专利分析领域的可用性得到有效检验<sup>[43]</sup>,因此,笔者选择 LOF 算法对技术主题中的离群专利进行识别。其核心逻辑为:对技术主题中每一个专利向量赋予 LOF 值,表示该专利向量所处空间的密集程度。当专利向量的 LOF 值大于 1 时,表示所处空间较为稀疏,专利前沿性潜力越大;反之,当 LOF 值小于 1 时,表示所处空间较为密集,专利前沿性潜力越小。

### 2.3.3 将高价值专利和离群专利合并,形成核心专利集

基于上述步骤,将技术主题中识别出的高价值专利和离群专利取并集,并删除重复专利,从而形成核心专利集。通过此种方式,保证核心专利集中的专利既具备前沿性又具有价值性。

## 2.4 流程四:技术机会识别地图绘制

通过上述流程,分别识别出技术领域涵盖的技术主题,并对每个技术主题中的核心专利进行评估。核心专利是具备较高前沿性和价值性的专利集合,如果一个技术主题涵盖的核心专利比例较高,就能反映该技术主题相对其他技术主题具有更高的前沿性和价值性潜力。因此,在对技术主题的前沿性和价值性计算中,将应用技术主题核心专利占比来表征,并计算每个技术主题的专利平均年龄,表征技术主题的时效性。G. Zanella 等<sup>[44]</sup>提出了一种非监督式专利分析框架,在专利聚类的基础上挖掘新颖性专利并结合专利平均年龄绘制技术地图,识别区块链领域的新兴技术。该研究为本文的技术机会识别地图绘制提供了理论参考和借鉴,笔者将结合技术主题的核心专利占比与专利平均年龄两个指标绘制技术机会识别地图。通过设定阈值,将技术主题类型划分为热点技术主题、发展技术主题、潜在技术主题和成熟技术主题,从而帮助企业识别不同技术主题的发展态势,明晰技术机会。

## 3 结果分析

### 3.1 数据来源

本文数据来源于 incoPat 数据库<sup>[45]</sup>,选择分析的技术领域为智慧农业。选择 incoPat 数据库是因为其

完整收录全球 100 余个国家/组织/地区 1 亿余件专利信息,具有数据覆盖全面、更新速度快、多维深入加工等优势,在学术研究中得到广泛的使用<sup>[46-47]</sup>。而选择智慧农业领域开展分析的原因是:在《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》中指出,要把建设智慧农业作为“十四五”时期以及面向 2035 年提高农业质量效益与竞争力的重要内容,对赋能乡村振兴具有重要意义。因此,对智慧农业领域进行分析,通过技术机会识别帮助企业寻求研发创新,从而激发产业振兴活力。

通过文献调研选取组配词,最终确定检索式为:标题/摘要=(智慧 or 智能) and (农场 or 农田 or 农业 or 种植 or 养殖) or 标题/摘要=((云计算 or 物联网 or 大数据 or 人工智能 or 区块链 or 传感器) and (农业 or 农场 or 农田 or 农作物 or 种植 or 病虫害 or 施肥 or 灌溉 or 养殖)),筛选得到权利有效的专利,共获得数据 17 614 条,分别获取专利指标数据和专利文本数据。

### 3.2 主题挖掘与文本聚类

根据技术主题挖掘流程,当主题数目设置为 10 个,  $\alpha = 0.1$ ,  $\eta = 0.01$  时,基于 pyLDAvis 模型呈现的主题可视化结果中,各主题之间交叉度较小,主题之间相对独立,表明主题数目最优。依据最优模型的输出结果,首先依据主题词对技术主题名称进行凝练,再应用“专利—主题”概率矩阵进行专利聚类,结果如表 2 所示:

表 2 智慧农业领域技术主题汇总

序号	主题名称	专利数量/件
1	智能养殖技术	1 465
2	智能灌溉	2 792
3	农业智能控制	1 446
4	农业无人机	2 085
5	农业 RFID 技术	1 345
6	智能农业机械	1 423
7	智能农业照明	1 795
8	农业智能监测	1 171
9	农业太阳能	2 276
10	农业智能机器人	1 816

### 3.3 技术主题核心专利评估

在完成专利聚类后,能够获取各个技术主题下涵盖的专利,进一步按照核心专利评估的流程展开分析。首先识别各个技术主题中的高价值专利,应用熵

权法计算专利价值评估的二级指标权重，再将二级指标权重相加得到一级指标权重，并根据指标数量确定突变系统类型，结果如表 3 所示：

表 3 专利价值评估指标权重及突变系统类型

总指标	突变类型	一级指标	突变类型	二级指标
专利综合价值 分值 A	燕尾突变系统	保护范围 B <sub>1</sub> (0.098)	燕尾突变系统	IPC 分类号数量 C <sub>1</sub> (0.056 3)
				文献页数 C <sub>2</sub> (0.026 7)
				权利要求数量 C <sub>3</sub> (0.015 4)
		研发水平 B <sub>2</sub> (0.427)	燕尾突变系统	家族被引证次数 C <sub>4</sub> (0.194)
				家族引证次数 C <sub>5</sub> (0.162)
				发明人数量 C <sub>6</sub> (0.071)
				许可 / 转让次数 C <sub>7</sub> (0.249)
		市场前景 B <sub>3</sub> (0.474)	蝴蝶突变系统	同族个数 C <sub>8</sub> (0.171)
				技术先进性 C <sub>9</sub> (0.047)
				技术稳定性 C <sub>10</sub> (0.007)

在确定指标权重和突变类型后，对评估数据进行标准化处理，以“主题 1→智能养殖技术”中的“一种用于蚕丝生产过程的智能温控装置”专利为例，对专利价值评估过程进行说明。

一级指标“保护范围”包含 3 个二级指标，属于燕尾突变，指标重要度排序为  $C_1 > C_2 > C_3$ ，且指标之间有一定相关关系，属于互补型突变系统，则：

$$X_{B1} = \frac{1}{3}(\sqrt{C_1} + \sqrt[3]{C_2} + \sqrt[4]{C_3}) = \frac{1}{3}(\sqrt{0} + \sqrt[3]{0.694} + \sqrt[4]{0}) = 0.137 \quad \text{公式 (1)}$$

一级指标“研发水平”包含 3 个二级指标，属于燕尾突变，指标重要度排序为  $C_4 > C_5 > C_6$ ，且指标之间有一定相关关系，属于互补型突变系统，则：

$$X_{B2} = \frac{1}{3}(\sqrt{C_4} + \sqrt[3]{C_5} + \sqrt[4]{C_6}) = \frac{1}{3}(\sqrt{0} + \sqrt[3]{0} + \sqrt[4]{0}) = 0 \quad \text{公式 (2)}$$

一级指标“市场前景”包含 4 个二级指标，属于蝴蝶突变，指标重要度排序为  $C_7 > C_8 > C_9 > C_{10}$ ，且指标之间有一定相关关系，属于互补型突变系统，则：

$$X_{B3} = \frac{1}{4}(\sqrt{C_7} + \sqrt[3]{C_8} + \sqrt[4]{C_9} + \sqrt[5]{C_{10}}) = \frac{1}{4}(\sqrt{0.2} + \sqrt[3]{0.029} + \sqrt[4]{0.556} + \sqrt[5]{0.75}) = 0.615 \quad \text{公式 (3)}$$

总指标系统共包含 3 个一级指标，属于燕尾突变，指标重要度排序为  $B_3 > B_2 > B_1$ ，且指标之间有一定相关关系，属于互补型突变系统，则：

$$X_A = \frac{1}{3}(\sqrt{B_3} + \sqrt[3]{B_2} + \sqrt[4]{B_1}) = \frac{1}{3}(\sqrt{0.615} + \sqrt[3]{0} + \sqrt[4]{0.137}) = 0.464 \quad \text{公式 (4)}$$

通过上面的步骤，评估出“主题 1→智能养殖技术”中的“一种用于蚕丝生产过程的智能温控装置”专利的综合价值分值为 0.464，据此可进一步对技术主题内的其他专利进行评估。依据已有研究得出的技术领域高价值专利分布比例为 2%-15%<sup>[48-49]</sup>，因此，将专利综合价值的阈值设置为 0.75，控制每个技术主题中的高价值专利比例不超过 15%，将高于 0.75 的专利视为高价值专利。

在评估完成各个技术主题中的高价值专利后，进一步对技术主题中的离群专利进行识别。

首先，应用 python 的 jieba 库对技术主题中的专利标题和摘要进行分词和去停用词处理，生成实验语料集。其次，应用 python 的 gensim 库，选用 Doc2vec 模型进行摘要向量化表示。综合考量实验数据规模，以及过往研究中 Doc2vec 的参数设定经验<sup>[50-51]</sup>，采用 100 维向量描述每一篇专利文本。通过 Doc2vec 训练后，形成 17 614\*100 的向量矩阵。最后，应用 python 的 sklearn 库，选用 LOF 算法进行离群专利识别。在应用 LOF 算法过程中，需要设定专利的临近点个数 k，用于测算专利在高维空间中的离群程度。为了从更微观层面对智慧农业领域的专利进行细粒度评估，此处 k 值设置为 3。

通过上述过程，对技术主题中每件专利的 LOF 值进行计算，依据判定标准，获得技术主题中的离群专利。对每个主题下的高价值专利和离群专利取并集，删除重复值后的汇总结果如表 4 所示：

表 4 技术主题核心专利汇总

序号	主题名称	高价值专利 / 件	离群专利 / 件	核心专利 / 件	核心专利占比 / %
1	智能养殖技术	29	32	61	4.16
2	智能灌溉	44	42	86	3.08
3	农业智能控制	28	25	53	3.67
4	农业无人机	18	72	90	4.32
5	农业 RFID 技术	67	26	93	6.91
6	智能农业机械	9	28	37	2.60
7	智能农业照明	52	26	78	4.35
8	农业智能监测	59	41	100	8.54
9	农业低碳减排	46	33	79	3.47
10	农业智能机器人	121	75	196	10.79

### 3.4 技术机会识别

对各个技术主题中的核心专利进行评估后，计算核心专利的占比情况，进一步对各个技术主题的专利平均年龄进行统计，为了与技术快速迭代更新态



势相适应, 在计算过程中以月为单位, 计算时间截至 2022 年 7 月。将核心专利占比与专利平均年龄作

为 X 轴和 Y 轴坐标映射到二维空间, 绘制智慧农业领域的技术机会识别地图, 如图 2 所示:

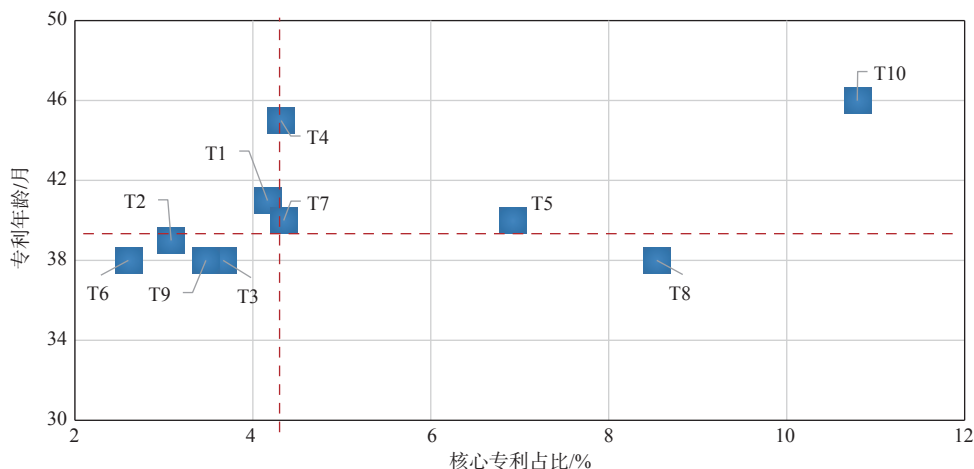


图 2 智慧农业领域技术机会识别地图

分析图 2 可以发现, 为了清晰界定技术主题之间的界限, 选定各个技术主题的核心专利占比和专利平均年龄的中位数 (4.24%, 39.5) 作为阈值, 可将技术主题划分为 4 种类型:

(1) 第一维度→发展技术主题, 核心专利占比 >4.24%, 专利平均年龄 >39.5 个月。处于该维度的技术主题, 具有较高比例的核心专利, 存在一定的研发创新机会, 但专利平均年龄较大, 表明该技术主题研发持续时间较长, 已建立起一定的技术基础, 企业如果选择该技术主题进行研发, 需要在已有技术基础上尝试寻求创新突破。

(2) 第二维度→成熟技术主题, 核心专利占比 <4.24%, 专利平均年龄 >39.5 个月。处于该维度的技术主题, 核心专利占比较小, 但专利平均年龄较大, 表明该技术主题技术研发机会较少, 相关技术已处于成熟阶段, 企业可减少对该技术主题的关注。

(3) 第三维度→潜在技术主题, 核心专利占比 <4.24%, 专利平均年龄 <39.5 个月。处于该维度的技术主题, 核心专利占比较少, 且专利平均年龄较少, 表明相关技术研发处于萌芽期, 具备较大的技术研发机会, 但也存在一定的创新阻力, 企业可以通过寻求校企合作, 降低研发投入风险。

(4) 第四维度→热点技术主题, 核心专利占比 >4.24%, 专利平均年龄 <39.5 个月。处于该维度的技术主题, 核心专利占比较大, 且专利平均年龄较少, 表明该技术主题正处于快速上升阶段, 市场关注强烈, 技术研发机会较多, 企业应重点关注, 加大研发投入力度, 及时建立竞争优势。

基于上述分析发现, 潜在技术主题和热点技术主题, 因其处于技术萌芽期与快速成长期, 蕴含的技术研发效益大, 是技术领域需要重点关注的技术机会。在智慧农业领域, “主题 8→农业智能监测” 属于热点技术主题。由于城市迁移和人口老龄化, 农业劳动力变得越来越稀缺, 日益加剧的气候变化正在以难以预测的方式改变生长条件, 因此, 农业智能监测的重要性不断提升。目前, 农业智能监测的核心专利主要集中在病虫害监测、农场气象实时监测、土壤墒情精准监测等, 通过实时监测田间虫情、气象数据、土壤数据等, 帮助农场制定防治措施、精准控制环境指标。而随着农业精细化、高效化的不断发展, 需要全面检测农场, 精准把控生产问题, 为作物播种、施肥、收获活动提供决策支持。此外, 极端气象的频发, 需要对极端气象及时预报, 提醒农户及时做好防灾防险准备, 减少农业损失。因此, 面向上述方向的研究可视为智慧农业领域的技术机会, 相关企业可将其视为重点方向并加大研发投入力度。

潜在技术主题包括“主题 2→智能灌溉”“主题 3→农业智能控制”“主题 6→智能农业机械”和“主题 9→农业低碳减排”。其中, “主题 2→智能灌溉”在水资源日趋紧张的形势下愈发重要, 尤其是在我国淡水资源比较紧缺的环境下, 通过智能灌溉技术创新, 不仅可以提高水资源利用率, 还可以增加农作物的产量, 降低农产品的成本。“主题 3→农业智能控制”是依托农业智能网关和传感器, 通过种养殖业数据实时采集与分析, 实现可控化生产和管理, 相关技术的研发突破, 对实现提高种养殖业生产和管理效率、提

升现代农业建设水平意义重大。此外,我国是一个农业大国,目前也是世界第一的农机生产和使用大国,农业耕地面积广,但农业生产效率低,生产成本较高。因此,“主题6→智能农业机械”要向农业机械全寿命周期延伸和发展,提升用户对农机的管理能力,从农业机械的“管、用、维”3个方面的深度支撑,实现农机智能化与信息化为核心目标的综合化解决方案,这既对企业创新研发提出了挑战,同时也指明了技术攻关的重要方向。最后,随着我国双碳目标的提出,新能源与新农业是绿色发展之未来与民生之根基的完美结合,“光伏+农业”正在成为一种新兴的农业形式,而“主题9→农业低碳减排”中针对光伏种植、光伏养殖、光伏水利、光伏村舍的持续研发,将具备广阔的市场前景。

为了验证本文识别的技术机会具有科学性和合理性,通过广泛查阅我国、美国、欧盟、日本等国家和地区在支持智慧农业发展的相关政策、战略和报告,考察强调的重点方向,与本文识别的技术机会进行比对。

针对“主题2→智能灌溉”,2022年,我国中央网信办等十部门出台的《数字乡村发展行动计划(2022-2025年)》提出,加快农村水利工程智慧化、水网智能化<sup>[52]</sup>;国务院发布的《中共中央 国务院关于做好2022年全面推进乡村振兴重点工作的意见》提出,推进黄河流域农业深度节水控水,通过提升用水效率、发展旱作农业,稳定粮食播种面积<sup>[53]</sup>。

针对“主题3→农业智能控制”,2019年美国国家科学院发布的《至2030年推动食品和农业研究的科学突破》提出,强化农业传感器的研发与应用,实现农业生产控制向高端化发展<sup>[54]</sup>;2021年日本经济省发布的《生物技术将培育第五次工业革命》提出,使用人形实验机器人和针对每步操作的模块化设备组合,将自动化引入研发和产品开发阶段<sup>[55]</sup>;2021年全国人大颁布的《中华人民共和国乡村振兴促进法》提出,发展智慧农业,建立农业农村大数据体系,推动新一代信息技术与农业生产经营深度融合<sup>[56]</sup>。

针对“主题6→智能农业机械”,2019年欧洲农机协会发布的《农业技术2030技术》提出,加强自动化、机器人、数字连接和人工智能等具有重要潜力技术领域的研究资助,以推动农业数字技术的采用<sup>[57]</sup>;2020年日本颁布的《食品、农业、农村发展五年计划纲要》提出,大力发展智慧农业,扩大无人农机的应用范围、加强农业生产自动化系统的开发

及运用<sup>[58]</sup>;2021年国务院发布的《中共中央 国务院关于全面推进乡村振兴加快农业农村现代化的意见》提出,提高农机装备自主研制能力,支持高端智能、丘陵山区农机装备研发制造,加大购置补贴力度,开展农机作业补贴<sup>[59]</sup>。

针对“主题8→农业智能监测”,2019年日本科学技术与学术政策研究所发布的《日本第11次科技预见(至2050)》提出,使用信息技术和空间技术对全球环境和资源进行监测、评估和预测<sup>[60]</sup>;2020年美国农业部的《科学蓝图—2020至2025年科研方向》提出,重点研发作物病虫害监测、早期发现和快速反应类传感器<sup>[61]</sup>;2021年我国农业农村部等6个部门发布的《“十四五”全国农业绿色发展规划》提出,加强外来入侵物种防控,开展外来入侵物种普查和监测预警<sup>[62]</sup>。

针对“主题9→农业低碳减排”,2020年丹麦政府公布国家投资绿色研究、技术和创新的首项战略,提出开发多种技术和解决方案,显著减少传统食品和有机食品生产与农业等对气候和环境的影响<sup>[63]</sup>;2021年我国国家能源局发布的《国家能源局关于2021年风电、光伏发电开发建设有关事项的通知》提出,积极推进分布式光伏发电的建设,结合乡村振兴战略启动“千乡万村沐光”行动<sup>[64]</sup>;2022年我国农业农村部、国家发展改革委联合印发的《农业农村减排固碳实施方案》提出,以农业农村绿色低碳发展是关键,要求大力发展农村可再生能源,建立完善监测评价体系,强化科技创新支撑<sup>[65]</sup>。

综上,通过结合近3年国内外关于智慧农业的重要政策、战略和报告,本文识别的关于智慧农业领域的技术机会均在相关文件中得到了重点强调,进一步验证了笔者提出方法的可行性和有效性。

## 4 结语

技术机会识别方法的研究对于创新资源的优化配置、助力相关创新主体的研发决策具有重要意义。当前关于技术领域的技术机会识别,在技术机会前沿性、价值性和时效性的综合考量方面需要进一步深化。据此,笔者依据技术机会应具备的3个特性,从研究主题细粒度层面进行分析,对每个技术主题涵盖的专利,依托突变级数法和LOF算法识别高价值专利和离群专利,将二者合并形成核心专利集,彰显技术主题的价值性和前沿性;计算每个技术主题中专利的平均年龄,体现技术主题的时效性;将核心专



利占比与专利平均年龄作为两个核心指标绘制技术机会识别地图, 通过阈值设定, 将技术主题划分为 4 种类型, 依据不同技术主题的发展态势识别技术机会。从方法上看, 从前沿性、价值性、时效性进行研究是对现有基于定量分析的技术机会识别方法的重要补充。

从结果上看, 以智慧农业为例, 对所提方法进行验证, 挖掘了智慧农业领域的 5 个技术机会, 并结合国内外智慧农业相关政策报告验证了方法的有效性。但本文还存在进一步深化的空间: 首先, 没有构建起定量的技术机会识别结果验证方法; 其次, 在技术机会识别过程中仅考虑了专利数据, 需要进一步扩展数据多源性; 最后, 仅以智慧农业为例对所提方法进行实证分析, 未来需要在其他技术领域展开分析, 验证方法的科学性和可靠性。

#### 参考文献:

- [1] 韩晓彤, 朱东华, 汪雪锋. 科学推动下技术机会发现方法研究 [J]. 图书情报工作, 2022, 66(10): 19-32.
- [2] YOON J, PARK H, SEO W, et al. Technology opportunity discovery (TOD) from existing technologies and products: a function-based TOD framework[J]. Technological forecasting and social change, 2015, 100: 153-167.
- [3] CHO J, LEE J. Development of a new technology product evaluation model for assessing commercialization opportunities using Delphi method and fuzzy AHP approach[J]. Expert systems with applications, 2013, 40(13): 5314-5330.
- [4] 郭俊芳, 汪雪锋, 李乾瑞, 等. 一种新型的技术形态识别方法——基于 SAO 语义挖掘方法 [J]. 科学学研究, 2016, 34(1): 13-21.
- [5] KIM K, PARK K, LEE S. Investigating technology opportunities: the use of SAOx analysis[J]. Scientometrics, 2019, 118(1): 45-70.
- [6] CHEN M C, HO P H. Exploring technology opportunities and evolution of IoT-related logistics services with text mining[J]. Complex & intelligent systems, 2021, 7(5): 2577-2595.
- [7] FENG L J, NIU Y X, LIU Z F, et al. Discovering technology opportunity by keyword-based patent analysis: a hybrid approach of morphology analysis and USIT[J]. Sustainability, 2020, 12(1): 136.
- [8] LEE S, YOON B, PARK Y. An approach to discovering new technology opportunities: keyword-based patent map approach[J]. Technovation, 2009, 29(6): 481-497.
- [9] CHUN E, JUN S, LEE C. Identification of promising smart farm technologies and development of technology roadmap using patent map analysis[J]. Sustainability, 2021, 13(19): 10709.
- [10] YU J, HWANG J G, HWANG J, et al. Identification of vacant and emerging technologies in smart mobility through the GTM-based patent map development[J]. Sustainability, 2020, 12(22): 9310.
- [11] HAN X T, ZHU D H, LEI M, et al. R&D trend analysis based on patent mining: an integrated use of patent applications and invalidation data[J]. Technological forecasting and social change, 2021, 167: 120691.
- [12] 桂美增, 许学国. 基于深度学习的技术机会预测研究——以新能源汽车为例 [J]. 图书情报工作, 2021, 65(19): 130-141.
- [13] 吴一平, 白如江, 刘明月, 等. 融合评论主题识别与技术属性多维度分析的技术机会发现研究 [J]. 图书情报工作, 2021, 65(10): 56-67.
- [14] YOON B, PARK I, YUN D, et al. Exploring promising vacant technology areas in a technology-oriented company based on bibliometric analysis and visualization[J]. Technology analysis & strategic management, 2019, 31(4): 388-405.
- [15] KIM B, GAZZOLA G, YANG J, et al. Two-phase edge outlier detection method for technology opportunity discovery[J]. Scientometrics, 2017, 113(1): 1-16.
- [16] SHI X, CAI L F, SONG H F. Discovering potential technology opportunities for fuel cell vehicle firms: a multi-level patent portfolio-based approach[J]. Sustainability, 2019, 11(22): 1-22.
- [17] 杨辰, 王楚涵, 陶琬莹, 等. 基于专利的技术机会识别: 深度学习领域的案例分析 [J]. 科技管理研究, 2021, 41(12): 172-176.
- [18] 翟东升, 刘鹤, 张杰, 等. 一种基于链路预测的技术机会挖掘方法 [J]. 情报学报, 2016, 35(10): 1090-1100.
- [19] PARK I, YOON B. Technological opportunity discovery for technological convergence based on the prediction of technology knowledge flow in a citation network[J]. Journal of informetrics, 2018, 12(4): 1199-1222.
- [20] HAN X T, ZHU D H, WANG X F, et al. Technology opportunity analysis: combining SAO networks and link prediction[J]. IEEE transactions on engineering management, 2021, 68(5): 1288-1298.
- [21] YOON B, MAGEE C L. Exploring technology opportunities by visualizing patent information based on generative topographic mapping and link prediction[J]. Technological forecasting and social change, 2018, 132: 105-117.
- [22] LEE J H, KO N, YOON J, et al. An approach for discovering firm-specific technology opportunities: application of link prediction to F-term networks[J]. Technological forecasting and social change, 2021, 168: 120746.
- [23] LEE C, LEE G. Technology opportunity analysis based on recombinant search: patent landscape analysis for idea generation[J]. Scientometrics, 2019, 121(2): 603-632.
- [24] ZHAO X X, ZHANG X M, EMMANUEL A. Research and demonstration of technology opportunity identification model

- based on text classification and core patents[J]. Computers & industrial engineering, 2022, 171: 108403.
- [25] CHOI J, JEONG B, YOON J. Technology opportunity discovery under the dynamic change of focus technology fields: application of sequential pattern mining to patent classifications[J]. Technological forecasting and social change, 2019, 148: 119737.
- [26] 黄鲁成, 李晓宇, 李晋. 基于专利的 ABOD-RFM 技术机会识别方法研究 [J]. 情报理论与实践, 2020, 43(9): 144-149.
- [27] 宋凯, 李秀霞, 赵思喆. 基于 CTM 模型与 K-means 算法融合的文本聚类研究 [J]. 情报理论与实践, 2017, 40(11): 135-138.
- [28] LEE C, KANG B, SHIN J. Novelty-focused patent mapping for technology opportunity analysis[J]. Technological forecasting and social change, 2015, 90: 355-365.
- [29] 冉从敬, 李旺, 宋凯, 等. 混合智能下的高校专利价值评估方法 [J]. 图书馆论坛, 2021, 41(7): 78-86.
- [30] 冉从敬, 宋凯. 高校专利价值评估模型构建——以云计算领域为例 [J]. 图书馆论坛, 2021, 41(1): 91-98.
- [31] 冉从敬, 宋凯. 高校可转化专利识别模型构建——以人工智能领域为例 [J]. 情报理论与实践, 2020, 43(11): 79-85.
- [32] 谢智敏, 范晓波, 郭倩玲. 专利价值评估工具的有效性比较研究 [J]. 现代情报, 2018, 38(4): 124-129.
- [33] HUANG Z L, LI J L, YUE H G. Study on comprehensive evaluation based on AHP-MADM model for patent value of balanced vehicle[J]. Axioms, 2022, 11(9): 481.
- [34] 赵辉, 杨瑞琦, 林芳芳. 基于模糊综合评价的科技成果转化分析 [J]. 科技管理研究, 2016, 36(10): 30-34.
- [35] 徐丹丹, 刘凯元, 曾章备, 等. 我国区域农村金融生态环境评价研究——基于突变级数法的分析 [J]. 农业经济问题, 2016, 37(4): 70-80.
- [36] 鲍珂宇, 何刚, 金兰, 等. 基于突变级数法的安徽各市相对资源承载力研究 [J]. 地域研究与开发, 2021, 40(4): 39-44.
- [37] 吴敬茹. 创新驱动视角下我国制造业技术创新能力评价 [J]. 工业技术经济, 2020, 39(10): 74-80.
- [38] 李柏洲, 苏屹. 基于改进突变级数的区域科技创新能力评价研究 [J]. 中国软科学, 2012(6): 90-101.
- [39] 田红娜, 刘思琦, 尹洪艳. 基于熵值—突变级数法的制造业绿色工艺创新能力评价 [J]. 科技管理研究, 2019, 39(24): 153-161.
- [40] 翟东升, 郭程, 张杰, 等. 采用异常检测的技术机会识别方法研究 [J]. 现代图书情报技术, 2016(10): 81-90.
- [41] 曹祺, 赵伟, 张英杰, 等. 基于 Doc2Vec 的专利文件相似度检测方法的对比研究 [J]. 图书情报工作, 2018, 62(13): 74-81.
- [42] LIU T L, HSIEH L H, HUANG K C. Applying natural language processing and TRIZ evolutionary trends to patent recommendations for product design [J]. Applied sciences-basel, 2022, 12(19): 10105.
- [43] 李佳佳, 马铁驹. 利用局部离群因子算法探测核心技术发展趋势——以中国风能专利数据为例 [J]. 情报杂志, 2017, 36(3): 119-124, 195.
- [44] ZANELLA G, LIU C Z, CHOO K K R. Understanding the trends in blockchain domain through an unsupervised systematic patent analysis [J/OL]. IEEE transactions on engineering management: 1-15[2021-12-28]. <https://ieeexplore.ieee.org/document/9428346>.
- [45] 北京合享智慧科技有限公司. incoPat [EB/OL]. [2022-10-27]. <https://www.incopat.com>.
- [46] 郑荣, 魏明珠, 高志豪, 等. 基于 SCAN-CPM 的产业新兴技术识别与演化路径分析: 以新能源汽车产业为例 [J]. 图书情报工作, 2022, 66(11): 100-109.
- [47] 刘明信, 李丹丹, 李荣, 等. 专利视阈下产业技术创新合作网络演化分析——以肿瘤疫苗为例 [J]. 图书情报工作, 2022, 66(9): 117-127.
- [48] 郑思远, 王学昭. 专利转让视角下技术转移特征指标体系研究 [J]. 图书情报工作, 2020, 64(7): 94-102.
- [49] 宋凯. 高校专利技术转移价值评估研究——基于熵权 TOPSIS 模型和梯度提升树算法 [J]. 情报杂志, 2021, 40(7): 52-57.
- [50] 阮光册, 夏磊. 基于 Doc2Vec 的期刊论文热点选题识别 [J]. 情报理论与实践, 2019, 42(4): 107-111, 106.
- [51] 张卫卫, 胡亚琦, 翟广宇, 等. 基于 LDA 模型和 Doc2vec 的学术摘要聚类方法 [J]. 计算机工程与应用, 2020, 56(6): 180-185.
- [52] 中华人民共和国中央人民政府. 中央网信办等十部门印发《数字乡村发展行动计划(2022-2025 年)》[EB/OL]. [2022-09-01]. [http://www.gov.cn/xinwen/2022-01/26/content\\_5670637.htm](http://www.gov.cn/xinwen/2022-01/26/content_5670637.htm).
- [53] 中华人民共和国中央人民政府. 中共中央 国务院关于做好 2022 年全面推进乡村振兴重点工作的意见 [EB/OL]. [2022-09-01]. [http://www.gov.cn/zhengce/2022-02/22/content\\_5675035.htm](http://www.gov.cn/zhengce/2022-02/22/content_5675035.htm).
- [54] National Academies of Sciences, Engineering, and Medicine. Science breakthroughs to advance food and agricultural research by 2030[EB/OL]. [2022-09-01]. <https://nap.nationalacademies.org/catalog/25059/science-breakthroughs-to-advance-food-and-agricultural-research-by-2030>.
- [55] Ministry of Economy, Trade and Industry. Bio-industry subcommittee's report, 'Fifth industrial revolution' cultivated with biotechnology, compiled[EB/OL]. [2022-09-01]. [https://www.meti.go.jp/english/press/2021/0202\\_001.html](https://www.meti.go.jp/english/press/2021/0202_001.html).
- [56] 中华人民共和国乡村振兴促进法 [EB/OL]. [2022-09-01]. <http://www.npc.gov.cn/npc/c30834/202104/8777a961929c4757935ed2826ba967fd.shtml>.
- [57] European Agricultural Machinery Association. AgriTech 2030: CEMA's plan for Europe's agricultural machinery industry[EB/OL]. [2022-09-01]. [https://www.cema-agri.org/images/publications/position-papers/2019\\_CEMA\\_Agritech\\_2030\\_web.pdf](https://www.cema-agri.org/images/publications/position-papers/2019_CEMA_Agritech_2030_web.pdf).
- [58] 冯献, 李瑾, 崔凯. 中外智慧农业的历史演进与政策动向比较

- 分析 [J]. 科技管理研究, 2022, 42(5): 28-36.
- [59] 中华人民共和国中央人民政府. 中共中央 国务院关于全面推进乡村振兴加快农业农村现代化的意见 [EB/OL]. [2022-09-01]. [http://www.gov.cn/zhengce/2021-02/21/content\\_5588098.htm](http://www.gov.cn/zhengce/2021-02/21/content_5588098.htm).
- [60] National Institute of Science and Technology Policy Library. 第11回科学技術予測調査 S&T Foresight 2019 総合報告書 [EB/OL]. [2022-09-01]. [https://nistep.repo.nii.ac.jp/?action=pages\\_view\\_main&active\\_action=repository\\_view\\_main\\_item\\_detail&item\\_id=6657&item\\_no=1&page\\_id=13&block\\_id=21](https://nistep.repo.nii.ac.jp/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=6657&item_no=1&page_id=13&block_id=21).
- [61] U.S. Department of Agriculture. USDA casts vision for scientific initiatives through 2025 [EB/OL]. [2022-09-01]. <https://www.usda.gov/media/press-releases/2020/02/06/usda-casts-vision-scientific-initiatives-through-2025>.
- [62] 中华人民共和国农业农村部.“十四五”全国农业绿色发展规划 [EB/OL]. [2022-09-01]. <http://data.mofcom.gov.cn/upload/file/07.pdf>.
- [63] Ministry of Higher Education and Science. Green solutions of the future - Strategy for investments in green research, technology, and innovation [EB/OL]. [2022-09-01]. <https://ufm.dk/en/publications/2020/green-solutions-of-the-future-strategy-for-investments-in-green-research-technology-and-innovation>.
- [64] 中华人民共和国中央人民政府. 国家能源局关于2021年风电、光伏发电开发建设有关事项的通知 [EB/OL]. [2022-09-01]. [http://www.gov.cn/zhengce/zhengceku/2021-05/27/content\\_5612874.htm](http://www.gov.cn/zhengce/zhengceku/2021-05/27/content_5612874.htm).
- [65] 中华人民共和国中央人民政府. 农业农村部 国家发展改革委关于印发《农业农村减排固碳实施方案》的通知 [EB/OL]. [2022-09-01]. [http://www.moa.gov.cn/govpublic/KJJYS/202206/t20220630\\_6403715.htm](http://www.moa.gov.cn/govpublic/KJJYS/202206/t20220630_6403715.htm).

#### 作者贡献说明:

宋凯: 负责模型构建和实验设计、论文撰写和修改;

冉从敬: 对论文的选题、思路进行指导。

## Research on Technology Opportunity Identification Based on Topic Mining and Patent Evaluation: A Case Study of Smart Agriculture

Song Kai<sup>1</sup> Ran Congjing<sup>2</sup>

<sup>1</sup>Shandong Normal University Library, Jinan 250358

<sup>2</sup>School of Information Management, Wuhan University, Wuhan 430072

**Abstract: [Purpose/Significance]** The identification of technological opportunities is an important prerequisite for the smooth development of technological innovation activities of enterprises. It is of great significance to discover and grasp valuable technological opportunities in time for technological innovation breakthrough. **[Method/Process]** In this paper, a technology opportunity identification method was proposed from the perspective of topic mining and patent evaluation. Firstly, the technology topics covered by the technology field were identified by the topic model and the patent clustering was carried out. Then, the fine-grained analysis was carried out at the level of technology topics. This paper comprehensively considers the three important characteristics of technological opportunities: frontier, value and timeliness. The mutation progression method and outlier factor algorithm were used to evaluate the high-value patents and outlier patents in the technology topics to form the core patent set, and the average age of patents for each technology topic was calculated. Finally, the proportion of core patents in the technology topic and the average age of patents were taken as the core indicators to draw the technology opportunity identification map to identify the opportunities in the technology field. **[Result/Conclusion]** Taking the field of smart agriculture as an example, the proposed method is empirically demonstrated, and five technological opportunities in the field of smart agriculture are identified, which provides decision support for the technology R&D of innovation subjects. However, the quantitative verification of identification results and the fusion of multi-source data for technical opportunity identification need to be further explored.

**Keywords:** topic model value evaluation technology opportunity smart agriculture