

情报理论与实践
Information Studies: Theory & Application
ISSN 1000-7490, CN 11-1762/G3

《情报理论与实践》网络首发论文

题目： 基于技术主题演化的颠覆性技术识别研究
作者： 单晓红, 韩晟熙, 刘晓燕
网络首发日期： 2023-03-15
引用格式： 单晓红, 韩晟熙, 刘晓燕. 基于技术主题演化的颠覆性技术识别研究[J/OL]. 情报理论与实践.
<https://kns.cnki.net/kcms/detail/11.1762.g3.20230313.1515.004.html>



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。

●单晓红, 韩晟熙, 刘晓燕(北京工业大学经济与管理学院, 北京 100124)

基于技术主题演化的颠覆性技术识别研究*

摘要: [目的/意义]颠覆性技术对技术和市场具有双重影响, 无论是对科技发展还是市场结构调整都起到十分重要的作用, 准确识别颠覆性技术能够解决可能出现的技术壁垒, 对国家和企业具有重要意义。[方法/过程]从技术主题的演化趋势和演化形式两个维度提出颠覆性技术识别框架。首先, 利用 LDA 模型和基于全局和局部影响力的社区探测方法识别领域内的技术主题。其次, 基于颠覆性技术的“突变性”和“融合性”特征, 引入技术主题时序共现网络和主题引用网络对主题间的演化趋势和演化形式进行判别, 实现颠覆性技术的识别。最后, 以人工智能领域为例验证本文提出的颠覆性技术识别方法的可行性和有效性。[结果/结论]该方法能够准确实现颠覆性技术的识别, 研究识别出人工智能领域的七项颠覆性技术, 包括“机器人模块化设计方法及应用”“相关领域机器人技术及应用主题”“仿生机器人技术及应用”“自动无人智能系统”“自动驾驶技术”“服务机器人技术及应用”和“机器人部件优化技术”这有利于我国抓住科技创新主动权, 加快颠覆性技术布局。

关键词: 颠覆性技术; 技术主题演化; 突变性; 融合性; 人工智能

Research on Disruptive Technology Identification Based on the Evolution of Technology Themes

Abstract: [Purpose/significance] Disruptive technologies have a dual impact on technology and the market, and they are crucial to both technological advancement and market restructuring. Accurately identifying disruptive technologies can help nations and businesses overcome potential technical barriers, which is very important. [Method/process] This study offers a methodology for identifying disruptive technologies in two dimensions: the evolutionary trend and the evolution of the technology topic. To begin, technical themes in the domain are found using LDA models and community identification approaches based on global and local influences. Second, based on the "mutate" and "convergent" qualities of disruptive technologies, we introduce a technology topic co-occurrence network and a topic citation network to discriminate the evolutionary trends and evolutionary forms among the subjects, helping us to identify disruptive technologies. Finally, by using the artificial intelligence industry as an example, the viability and efficacy of the disruptive technology detection approach described in this research are confirmed. [Result/conclusion] According to the study's findings, the seven disruptive technologies in the field of artificial intelligence include robot component optimization techniques, related field robotics and application themes, bionic robotics and applications, automated unmanned intelligent systems, autonomous driving technologies, service robotics and applications and robot component optimization technologies. The methodology provides precise identification of disruptive technologies, allowing China to seize the lead in science and technology innovation and accelerate the development of disruptive technologies.

Keywords: disruptive technologies; evolution of technology themes; transformative; convergent; artificial intelligence

0 引言

世界各国发展历程表明, 尽早识别和培育颠覆性技术, 将其作为科技和产业变革的重要抓手, 有助于加强颠覆性技术的主动布局, 提高国际竞争优势。在我国, “颠覆性技术”被写入《国家创新驱动发展战略纲要》《“十三五”国家科技创新规划》和党的十九大报告中。2018年, 两院院士大会上指出“以颠覆性技术创新为突破口”, 将颠覆性技术创新提到了战略高度。因此, 在未来的国家发展道路上, 颠覆性技术将成为引领技术和产业变革的风向标, 准确识别颠覆性技术, 能够迭代选择出未来技术需求方向, 及时甄别和培育未来产业, 抓住科技发展良机, 推动我国未来产业发展。

颠覆性技术识别研究是颠覆性技术研究中的重要内容之一, 早期的识别研究往往从静态视角出发, 在单一的静态时间切片下通过颠覆性技术的特征进行识别, 存在主观性较强、识别准确性不高的不足。为弥补静态研究的缺陷, 近年来学者尝试从专利引用变化和主题变化的动态视角对颠覆性技术进行识别, 但是由于只局限于单个技术的演化, 没有充分考虑技术

* 本文为国家自然科学基金后期资助项目“社交媒体数据驱动的决策支持理论、方法与实践”(项目编号: 21FGLB042)和国家自然科学基金后期资助项目“技术创新网络关系治理机制研究”(项目编号: 20FGLB004)的成果。

间的作用关系及演化关系对技术发展的影响,难以清晰阐释技术演化的趋势和形式。技术主题是相同技术属性的一个以上的技术方案集合,反映了领域内的技术特征,技术主题的演化趋势和演化形式能够体现领域内技术的发展过程以及技术间的相互关系。因此,本文从技术主题演化的角度出发,基于技术主题的准确划分,对技术主题所对应技术的演化趋势和形式进行分析,通过技术主题的突变性和融合性特征,实现颠覆性技术的识别,从而帮助有关部门提早制定产业发展战略,优先布局和规划。

1 相关研究

1.1 颠覆性技术识别方法研究

颠覆性技术识别方法研究通常从静态和动态两个视角展开。

静态视角下的颠覆性技术识别方法研究,通过分析技术或技术主题的特征建立指标体系,并对技术或技术主题进行评价,从而实现颠覆性技术的识别。颠覆性技术往往具备一般技术所不具备的特性,孙永福等^[1]、冯倩倩等^[2]、曹悦等^[3]、刘志辉等^[4]、S. Hardman 等^[5]、Jia Weifeng 等^[6]、Chen Xiaoli 等^[7]通过归纳总结颠覆性技术的特征,建立了多层级的评价指标体系,通常包括技术和市场两个方面的特征,利用指标体系进行颠覆性技术识别的研究。由于颠覆性技术通常具有较长的潜伏期,从已经显现的技术中识别可能会有遗漏,因此张秀妮等^[8]、马永红等^[9]、韩芳等^[10]利用文本挖掘、主题模型等方法与工具,从技术专利信息中提取出技术主题,对技术主题进行评价,深层次地挖掘潜在的颠覆性技术。

为更好地理解颠覆性技术在某个应用领域引进和发展的动态过程,动态视角下的研究,往往增加了时间维度信息,从专利引用变化和主题变化两个方面进行颠覆性技术识别。专利引用变化方面,具有时序的专利引用信息可以从多角度体现出技术的价值,于光辉等^[11]、曹艺文等^[12]、王康等^[13-14]、苏敬勤等^[15]、陈育新等^[16]、A. Momeni 等^[17]、王海军等^[18]通过使用社会网络分析、引文曲线、JAL 型曲线、评价模型等方法完成对颠覆性技术的识别。主题变化方面使用突变检测的方法,通过对技术主题的“突变性”进行颠覆性技术的判别,陈虹枢等^[19]借助社会网络分析方法对具有突变性的主题进行识别,刘忠宝等^[20]、赵玉桐等^[21]、Jia Weifeng 等^[22]则借助 Citespace 软件中提供的功能进行突变检测,张金柱等^[23]、马铭等^[24]进一步创新性地提出了突变程度的测度方法,以识别潜在的颠覆性技术。

1.2 技术主题识别方法

技术主题识别是颠覆性技术识别的基础。目前多数采用改进的 LDA 算法,或是将 LDA 算法与 K-Means 等聚类方法相结合的方法^[25-27]。也有学者通过文本挖掘的方法,包括关键词聚类分析^[23,28]、SAO 结构识别^[29]等方法,从专利数据中提取相关信息,并进行聚类分析,以完成技术主题的划分。随着研究的深入,张金柱等^[30]将相互包容系数应用于关键词共现网络的构建,并采用简单中心算法分析关键词间共现关系的强度,并进行聚类,以提取技术主题。李慧等^[31]构建了专利价值评价指标并结合 LDA 主题模型生成的概率分布进行技术创新主题的挖掘。刘小玲等^[32]利用专利的文本内容、引用关系和分类信息来构建专利的多属性向量,并利用聚类方法获得技术主题。

综上,颠覆性技术识别研究无论从定义和特征上,还是从方法上都尚未形成完整的研究体系,现有研究中还有很多值得深入研究的问题。首先,颠覆性技术是在技术主题演化过程中形成的,静态视角的研究往往忽略了技术的演化过程对颠覆性技术产生的关键作用。其次,技术主题演化过程中,技术主题会表现出扩展、收缩,或者产生新的主题的趋势,技术主题之间也会呈现出衍生、融合和分裂的不同形式,这些演化的趋势和形式是在技术演化过程中识别颠覆性技术的关键,但是现有动态视角的研究则只考虑了单个技术在发展中的演化趋势,没有综合考虑技术主题的演化形式,对颠覆性技术的识别具有一定的片面性。第三,常见的技术主题提取方法,如 LDA,通常只考虑了专利内容,没有考虑专利内容之间的共现关系,忽略了关键词之间关系的紧密程度,即同一专利的关键词要比不同专利的关键词联系更加紧密。此外, LDA 忽略了主题与主题之间的相似度,没有进行相似主题的融合,可能会导致主题区分度不高。

基于上述问题,本文提出了基于技术主题演化的颠覆性技术识别方法。首先利用自然语言处理方法处理专利数据,构建关键词共现网络,并使用基于全局和局部影响力的社区探测方法,完成社区划分。将专利与社区进行匹配,能更为准确地识别技术主题。进而从技术主题演化的趋势和形式两个角度,构建主题时序共现网络,对网络中的技术主题演化趋势进行分类,完成突变性主题的识别。然后,将专利之间的引用关系转化为主题之间的引用关系

构建主题引用网络，并在之前识别的突变性主题基础上进行演化形式的判别，最终实现颠覆性技术主题的识别。最后，为证实该模型的有效性，选取人工智能领域开展实证研究，通过该领域的技术主题演化分析识别该领域的颠覆性技术。

1.3 基于技术主题演化识别颠覆性技术的可行性

在技术不断地发展过程中，技术主题也会发生变化，技术主题的演化包括技术主题的演化趋势和演化形式两个方面，如图 1 所示。每个时间段都会出现一个或多个主题，主题用圆圈表示，圆圈的大小表示当前时间段内该主题内关键词的数量，圆圈间的有向连线代表两个主题之间的引用关系^[33]。

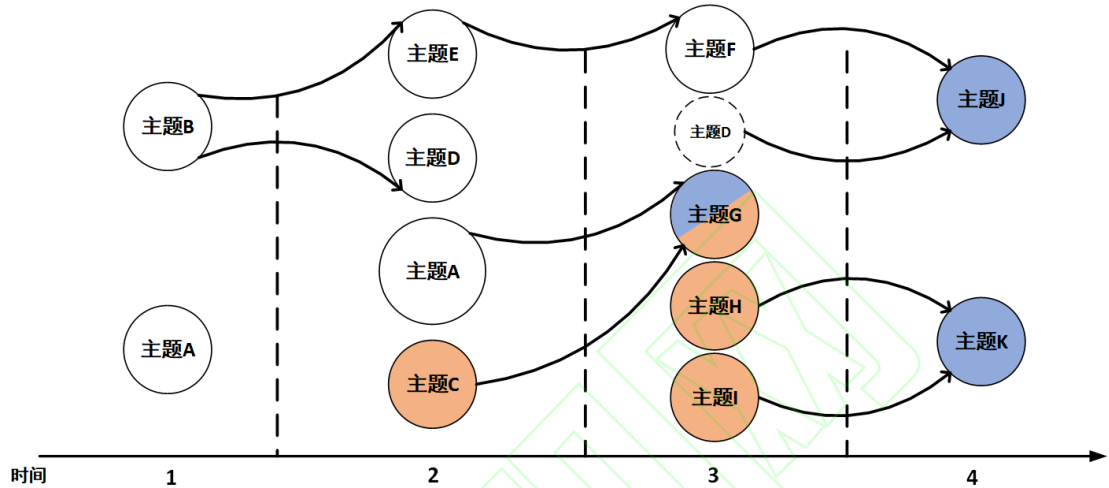


图1 技术主题的演化趋势和演化形式示意图

Fig.1 The schematic of evolutionary trends and evolutionary forms of technical themes

单个技术主题的演化趋势包括新生主题、扩张主题和收缩主题^[33]。新生主题是指在时间段 i 不存在，但在 $i+1$ 时间段中出现的主题，该类主题与已经存在的主题仅有较弱的共现关系或完全没有共现关系，如图 1 中时段 2 中主题 C 是新生主题。扩张主题是时间段 i 存在的主题，在 $i+1$ 时间段内继续存在，且包含关键词的数量增多，即规模扩张，如图 1 中时段 2 中主题 A。收缩主题是时间段 i 存在的主题，在 $i+1$ 时间段内继续存在，且包含关键词的数量减少，即规模缩小，如图 1 中时段 3 中的主题 D。王海军等^[18]在对智能语音领域的颠覆性技术进行识别时，将该领域的发展阶段分为了萌芽阶段、成长阶段和成熟阶段，通过比较萌芽和成长阶段的技术主题可以发现该领域的语音增强主题从萌芽阶段到成长阶段规模明显扩张，属于扩张主题。

技术主题的演化形式包括融合、衍生和分裂三种。融合是指由多个现有主题融合或现有主题与新生主题任意融合或新生主题之间进行融合生成的主题，如图 1 中主题 G 同时引用主题 A 和主题 C；主题 K 同时引用主题 H 和主题 I；主题 J 同时引用主题 F 和主题 D，主题 G、K 和 J 都属于融合主题。衍生是指两个主题间关系联系紧密但不同属于一个主题，如图 1 中主题 F 与主题 E 之间的关系。分裂是指在时间段 i 存在的主题，在时间段 $i+1$ 消失（或规模减小），并且同时被两个主题引用，如图 1 中主题 B 在时间 1 存在，但在时间 2 消失，且同时被主题 E 和主题 D 引用。

根据技术主题的演化趋势和演化形式，技术主题在演化过程中形成多种路径，包括新生—融合、新生—衍生、扩张—融合、扩张—衍生、收缩—融合。新生主题是在过去研究中没有出现的主题，体现了技术的突变性，往往能够突破现有技术的桎梏，在创新活动中产生的重大技术范式变革，成为颠覆性技术。融合的技术能够体现基础科学的跨学科、跨领域交叉，更有可能带来突破性创新。同时具有新生的演化趋势和融合的演化形式的主题往往能够形成颠覆性技术。如智能制造装备领域中，智能数字控制技术在 2006 年明显形成规模，且融合了机械技术、计算机技术及控制技术^[9]，形成了具有新生—融合演化特征的技术主题，并成为智能制造装备领域的颠覆性技术。技术主题如果只属于新生主题，而不进行融合形式的演化，则仍然是在原先的领域中进行创新；技术主题如果只是进行融合形式的演化，则没有突破旧技术的束缚，仍然是在原有技术的基础之上进行改进。例如，在显示屏制造技术的发展过程中，实力克从只能提供单一计算器显示屏逐步改进技术，到可以对多种设备的液晶

显示屏进行供应^[34]，形成了具有扩张—融合特征的技术主题；在油气技术的发展过程中，超级压裂技术的出现属于新生主题，其在传统压裂技术的基础上进行了改进，压裂强度大幅度提升^[35]，形成了具有新生—衍生演化特征的技术主题，液晶显示屏的制造技术和油气技术没有同时包含新生和融合，所以在各自的领域只是实现了渐进性创新。

本文将技术主题演化过程中的新生主题识别为突变性主题^[33]，重点关注具有融合性的突变性主题，这些主题能够体现出技术随时间发展过程中显现出的“新颖性”和“交叉性”，从而实现从主题的演化趋势和演化形式的角度对颠覆性技术的识别。

2 基于技术主题演化的颠覆性技术识别研究框架

在技术主题的演化过程中，可能呈现出多种演化趋势和形式的组合。颠覆性技术通常具有“突变性”和“融合性”双重特征，对应着新生的演化趋势和融合的演化形式。因此，通过探究技术主题是否同时属于新生主题和融合性主题，能够完成颠覆性技术的识别。本文通过基于全局和局部影响力的社区探测方法进行技术主题的提取，并采用技术主题时序共现网络和和技术主题引用网络分析方法分析技术主题的演化过程，对颠覆性技术进行识别。整体研究框架如图 2 所示。

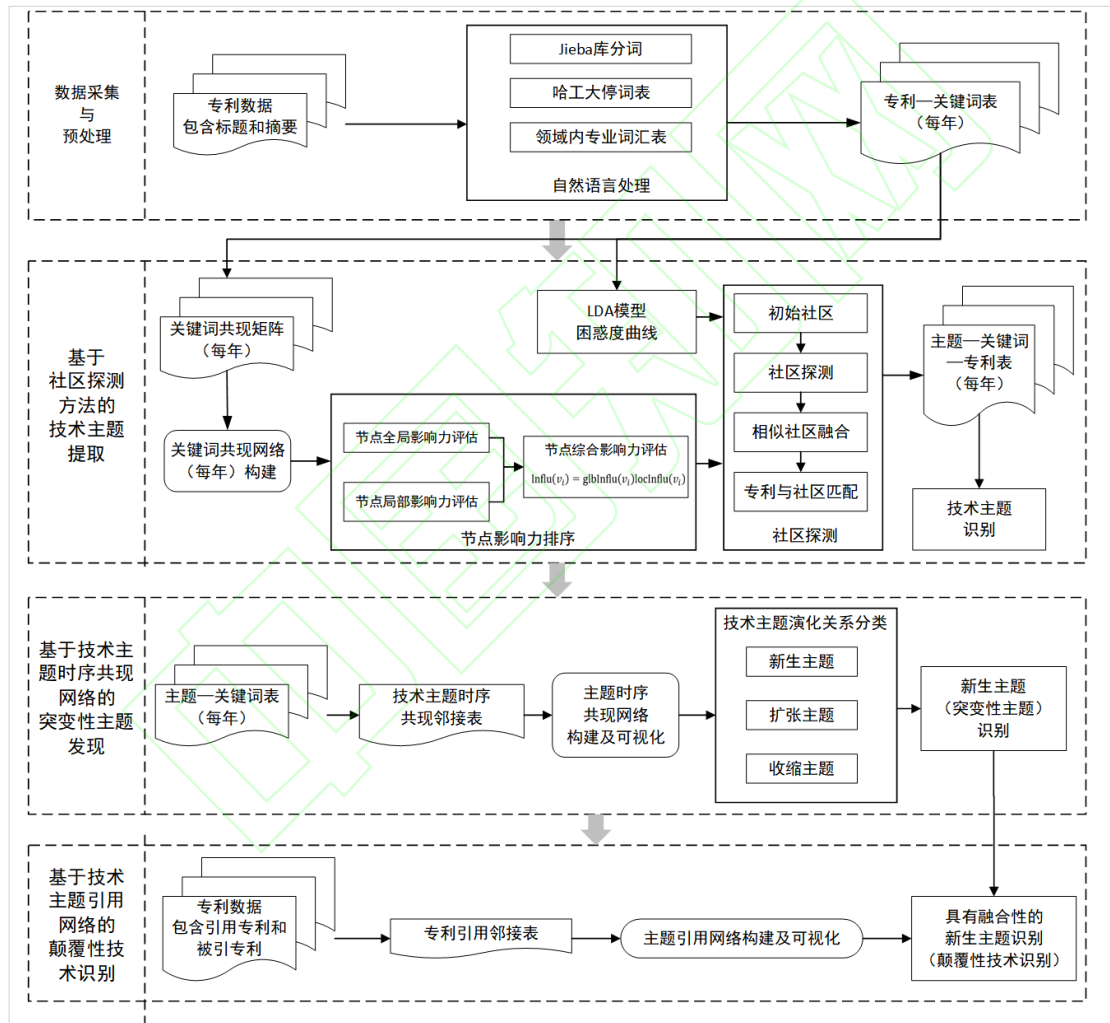


图2 研究框架

Fig.2 The research framework

首先，从专利数据库下载相关领域中具有时序性的专利数据，并通过 python 中的 jieba 库对专利数据的标题和摘要字段进行分词操作，构建专利—关键词表。其次，利用词表构建关键词共现网络，并使用基于全局和局部影响力的社区探测方法，完成技术主题的识别。再次，将专利与技术主题进行匹配，构建技术主题—关键词—专利表，以其中的主题和关键词字段构建技术主题时序共现网络，通过对网络中的技术主题演化关系进行分类，完成突变性主题的识别。最后，通过提取专利之间的引用关系，构建专利引用邻接表，并将专利之间的引用关系转化为主题之间的引用关系。最后，构建主题引用网络，并完成突变性主题中融合

性主题的识别，从而完成颠覆性技术的识别。

2.1 数据采集及预处理

从专利数据库中下载相关领域近 10 年的专利数据。利用 Python 软件对专利数据进行预处理，并使用 Jieba 库进行分词，使用哈工大停用词表和自定义的领域内专业词汇表优化分词结果，保留词频为 3 次以上的关键词。最后，得到每一年的专利—关键词表。

2.2 基于社区探测方法的技术主题提取

采用基于社区探测方法的技术主题提取方法，能够对节点的影响力进行测度与排序。依据影响力排序确定初始社区并进一步完成社区探测，最后评估社区间的社区重叠率，对重叠率较高的社区进行融合，确保技术主题提取的准确性。同时，还通过潜在狄利克雷分布 (LDA) 计算初始社区数量^[36]来减少迭代次数，提高效率。

2.2.1 基于 LDA 的技术主题初始数量确定 首先，利用 Python 软件中“sklearn”库的

“LatentDirichletAllocation”函数建立 LDA 模型；其次，对专利的标题和摘要进行文本分类，将已经进行分词后得到的专利—关键词表输入模型，通过 Python 绘制困惑度曲线判断最佳的技术主题数；最后，得到初始的技术主题数量 K 作为社区探测的基础。

2.2.2 节点综合影响力排序 通过分析每年的专利—关键词表中的关键词共现关系构建关键词共现网络 $G(V, E, W)$ 。其中 V 表示节点集合， E 表示边的集合， W 表示边的权重。每一个关键词可以抽象为一个节点，关键词间的共现关系则可以抽象为边。

为了准确评价关键词共现网络中节点的影响力，不仅要考虑节点在整个网络中的影响力，还应考虑其邻居节点的影响及其自身资源的多少，即融合全局结构和局部结构两种不同视角对节点的影响力进行评价。因此，节点 v_i 的影响力是将其局部影响力 $locInflu(v_i)$ 和全局影响力 $gblInflu(v_i)$ 的组合，如公式 (1) 所示。其中， $Influ(v_i)$ 即为节点 v_i 的影响力数值，数值越高，影响力越高，表明节点在共现网络中占有越重要的地位。

$$Influ(v_i) = gblInflu(v_i)locInflu(v_i) \quad (1)$$

节点 v_i 的全局影响力，用节点与其所有邻居节点的信息交互能力来综合评估，计算方法如公式 (2)， $d(v_i)$ 表示节点 v_i 的度； A 是一个常数，将其设置为 $A = 1.1^{[37]}$ ； $\#Com(v_i, v_j)$ 表示 v_i 和 v_j 之间的公共邻居节点数量，公共邻居节点可由公式 (3) 式得出。

$$gblInflu(v_i) = d(v_i) \sum_{v_j=Nei(v_i)} pow(A, \#Com(v_i, v_j)) \quad (2)$$

$$Com(v_i, v_j) = Nei(v_i) \cap Nei(v_j) \quad (3)$$

节点 v_i 的局部影响力，是节点在局部网络中的自身属性及周围节点的影响的组合，利用节点的度中心性和贡献概率进行计算，如公式 (4) 所示：

$$locInflu(v_i) = \sum_{v_j=Nei(v_i)} DC(v_j)p(v_j) \quad (4)$$

式中， $DC(v_j)$ 表示节点 v_j 的度中心性； $p(v_j)$ 表示节点 v_i 的邻居节点对节点 v_i 的影响力做出的贡献概率。如公式 (5) 所示：

$$p(v_j) = \frac{1}{\sum_{v_j=Nei(v_i)} d(v_j) / \#Nei(v_i)} \quad (5)$$

分母表示节点 v_i 所有邻居节点的平均度； $\#Nei(v_i)$ 表示节点的邻居节点数量。

通过上述公式可以对引用关系网络中的专利数据的局部影响力进行测度。最后，采用最大—最小规范化的方法对网络中节点影响力进行归一化。通过对网络中所有节点的影响力进行排序，即可得到节点的重要性排序。

2.2.3 社区探测 根据 LDA 主题模型得到的初始社区数量 K ，以及关键词共现网络中节点

影响力的重要性排序，确定初始社区的核心节点。将这些节点及其最近邻居节点分配到同一社区中，完成初始社区 $\{C_1, C_2, \dots, C_k\}$ 的构建。用 C_u 表示未分配节点的集合； node_i 表示社区中的核心节点； $\text{neighbors of node}_i$ 表示 node_i 的邻居节点。整个网络及初始社区可表示为公式（6）：

$$\begin{aligned} C_1 \cup C_2 \cup \dots \cup C_k \cup C_u &= G \\ C_i &= \{\text{node}_i, \text{neighbors of node}_i\}, i \in 1, 2, \dots, k \end{aligned} \quad (6)$$

初始社区构建后，将网络中待分配的节点先分配到每个社区的候选集 $\text{NS}(C_i)$ 中，对于候选集中未分配的节点按照影响力进行排序，并用 v_j 表示。定义 $l(C_i, v_j)$ 用来度量未分配节点 v_j 和社区 C_i 之间的接近程度。它由两部分组成，第一项是节点与社区之间的相似程度，用来衡量节点与同社区其他节点的相似程度。当两个节点同属于一个社区的时候，二者应该有许多公共邻居节点，即社区内节点具有更高的相似度。第二项是节点与社区之间的距离，用来衡量节点与同社区其他节点的距离之和，距离越短，说明社区内节点联系越紧密^[38]。如公式（7）所示：

$$l(C_i, v_j) = \sum_{i \in C} \frac{|\text{Nei}(v_i) \cap \text{Nei}(v_j)|}{|\text{Nei}(v_i)|} + \frac{1}{\sum_{i \in C} \frac{|\text{Nei}(v_i) \cap \text{Nei}(v_j)|}{\sqrt{|\text{Nei}(v_i)| * |\text{Nei}(v_j)|}}} \quad (7)$$

式中， $\text{Nei}(v_i)$ 和 $\text{Nei}(v_j)$ 表示节点 v_i 和节点 v_j 的邻居节点； $|\text{Nei}(v_i) \cap \text{Nei}(v_j)|$ 表示它们的公共邻居节点数量； $|\text{Nei}(v_i)|$ 和 $|\text{Nei}(v_j)|$ 表示邻居节点数量。

通过公式（7）测度每个社区候选集中的节点，按照接近程度降序排列，将关联更紧密的节点优先分配到对应社区当中，实现初始社区的扩展。如果节点 v_j 对多个社区具有相同的接近程度，则将节点 v_j 同时分配给这些社区，节点 v_j 称为重叠节点。对候选集中的节点进行一一分配，最终将所有社区候选集中的节点都分配到一个或多个社区当中，完成社区的探测。

2.2.4 相似社区合并 社区探测后，社区中会存在许多重叠节点，即一个节点同属于很多社区。如果社区之间的重叠节点过多，则证明两个社区具有较高的相似度，这样探测出的社区并不具有“低内聚”的特点，探测结果并不理想。因此，将相似社区进行合并，优化社区探测结果。

以两个社区之间的重叠率为依据，来进行相似社区之间的合并。社区重叠率用 δ 表示，定义如公式（8）所示：

$$\delta = \frac{|C_i \cap C_j|}{\min\{C_i, C_j\}} \quad (8)$$

式中， $|C_i \cap C_j|$ 表示两个社区中相同节点的数量； $\min\{C_i, C_j\}$ 表示两个社区中规模较小的节点数量。如果两个社区之间的重叠率超过设置的阈值，说明两个社区相似度很高，则这两个社区将会被合并。对现有社区进行两两评估，将重叠率高的社区进行合并，得到最终探测出的社区用 $C = \{C_1, C_2, \dots, C_t\}$ 表示。

2.2.5 技术主题提取 在识别出最终社区之后，为了更加准确地识别技术主题，在总结各个社区所对应的主题内容前，先将每个主题与该时间段下的专利进行匹配。

利用 Python 将专利—关键词表与技术主题进行对照，并通过多次实验，计算合适的专利与主题的关键词重叠率 θ ，如公式（9）所示。当专利包含的关键词与主题包含的关键词重叠率大于阈值时，则将该项专利分配到该主题下。

$$\theta = \frac{|W_i \cap W_j|}{\min\{W_i, W_j\}} \quad (9)$$

式中， $|W_i \cap W_j|$ 表示专利与主题之间相同的关键词的数量； $\min\{W_i, W_j\}$ 表示专利和主题二者中规模较小的关键词数量。

通过以上步骤，可以得到技术主题—关键词—专利表，表中包含了技术主题包含的关键词以及与该主题所匹配的专利数据。

2.3 基于技术主题演化趋势的突变性主题识别

2.3.1 构建技术主题时序共现网络

第一步,根据时间切片对已有专利数据集进行划分。采用固定时间段,即设定时间长度 i 为一个时间切片,将专利数据集按照 i 划分为 n 份。然后,利用 2.2 节中基于社区探测方法的技术主题提取方法对每一份专利数据集进行技术主题识别。

第二步,若词汇 W_a 和 W_b 分别处于时间段 i 和时间段 $i+1$ 的主题当中,则将词汇 W_a 和 W_b 之间的共现关系,转化为词汇 W_a 和 W_b 所属于的主题之间的共现关系,并且共现关系强度为 n 。

第三步,利用 Python 完成主题共现网络的构建。对于主题和主题之间的共现关系可以抽象为无向有权图 $G(V, E, W)$, 其中, V 表示节点集合, E 表示边的集合, W 表示边的权重。每一个主题可以抽象为一个节点,主题间的共现关系则可以抽象为边,并按照时序和主题规模对各个时间段的主题进行排列,为后续研究做准备。

2.3.2 突变性主题识别 颠覆性技术具有的本质特征之一是“突变性”,即一个技术主题是从无到有,通过技术创新而来的。在技术主题时序共现网络中,体现为在时间段 i 不存在、在时间段 $i+1$ 突然出现的技术主题,可以被识别为突变性主题。

2.4 基于技术主题演化形式的颠覆性技术识别

2.4.1 技术主题引用网络构建

首先,利用专利数据在引用专利和被引用专利这两个字段中包含的引用和被引用关系重新整理引用关系,形成专利引用邻接表。

其次,将专利间的引用关系转化为施引专利所在主题与被引证专利所在主题之间的引用关系,将所有关系进行整合,从而形成主题与主题之间的引用关系邻接表。若两个主题之间有 n 项专利对存在引用关系,则两个主题之间存在权值为 n 的引用关系。

第三,由于专利引用关系是内含有时序的,即后序专利引用前序专利。因此,在主题引用网络中也不再强调时序概念。对不同时段的相同主题的专利列表进行合并,最终得到领域内的主题—专利表。

第四,利用 Python 实现对主题引用网络的构建。对于主题和主题之间的引用关系可以抽象为有向有权图 $G(V, E, W)$, 其中 V 表示节点集合, E 表示边的集合, W 表示边的权重。每一个主题可以抽象为一个节点,主题间的引用关系则可以抽象为边,主题间的引用关系强度表示为权重。最后,完成主题引用网络的构建并进行可视化。

2.4.2 颠覆性技术识别 在技术主题时序共现网络中,具有“突变性”的技术主题是技术创新的产物。将每个时间段的突变性主题对应到主题引用网络当中,判别他们是否为融合性主题。融合性主题标志着这个技术主题是继承了现有技术发展而来的,即在技术主题引用网络中拥有较强的引用与被引用关系的主题。因此,颠覆性主题是具有融合形式的突变性主题。

3 实证研究

人工智能领域是学界普遍认为的具有很大发展前景的领域之一。它是计算机科学的一个分支,它通过计算机、自动化等技术,期望能够赋予机器人人类的智慧,来让这些机器能够辅助人们工作。事实上,人工智能已经深入人们的生活当中,许多智能设备都包含来人工智能的软件或硬件。该领域的研究不仅有可以完成各种工作的智能机器人还有语言识别、图像识别、自然语言处理等这些实用技术,是未来科技发展的重点研究对象。因此,以人工智能领域为例开展实证研究,期望能够验证本文对颠覆性技术识别的有效性。

3.1 人工智能领域的专利数据采集和预处理

以 Incopat 专利数据库为数据源,检索国内有效的人工智能领域的专利。借鉴陈挺等^[39]构建人工智能领域专利检索式的经验及其研究成果,在 Incopat 专利数据库中设置检索式为: $TI = (\text{“深度学习算法与模型” OR “经典机器学习算法” OR “自然语言处理” OR “数据挖掘信息分析算法” OR “生物特征抽取” OR “人脸识别” OR “图像处理” OR “动作捕捉” OR “智能算法” OR “智能电器” OR “网络通信” OR “生物医学” OR “机器人” OR “自动驾驶” OR “燃料电池”})$, 同时将专利申请年份设置为“2011—2021”,共采集 11 年的数据。专利有效性设置为“有效”,专利申请人国别设置为“中国”,专利公开类型设置为“发明申请”。

检索时间集中在 2022 年 3 月 30 日，共得到人工智能专利数据 39709 条，其中包括标题、摘要、公开（公告）号、引用专利、被引用专利等字段。为构建技术主题时序共现网络，需要将得到的原始数据集按时间段（每年）进行切片，将原始数据集分为 2011—2021 年共 11 个数据集，为后续研究做准备。

3.2 技术主题识别

首先，构建 2011—2021 年的主题词共现网络。以 2015 年为例，展示度数在 40 以上的节点，如图 3 所示。

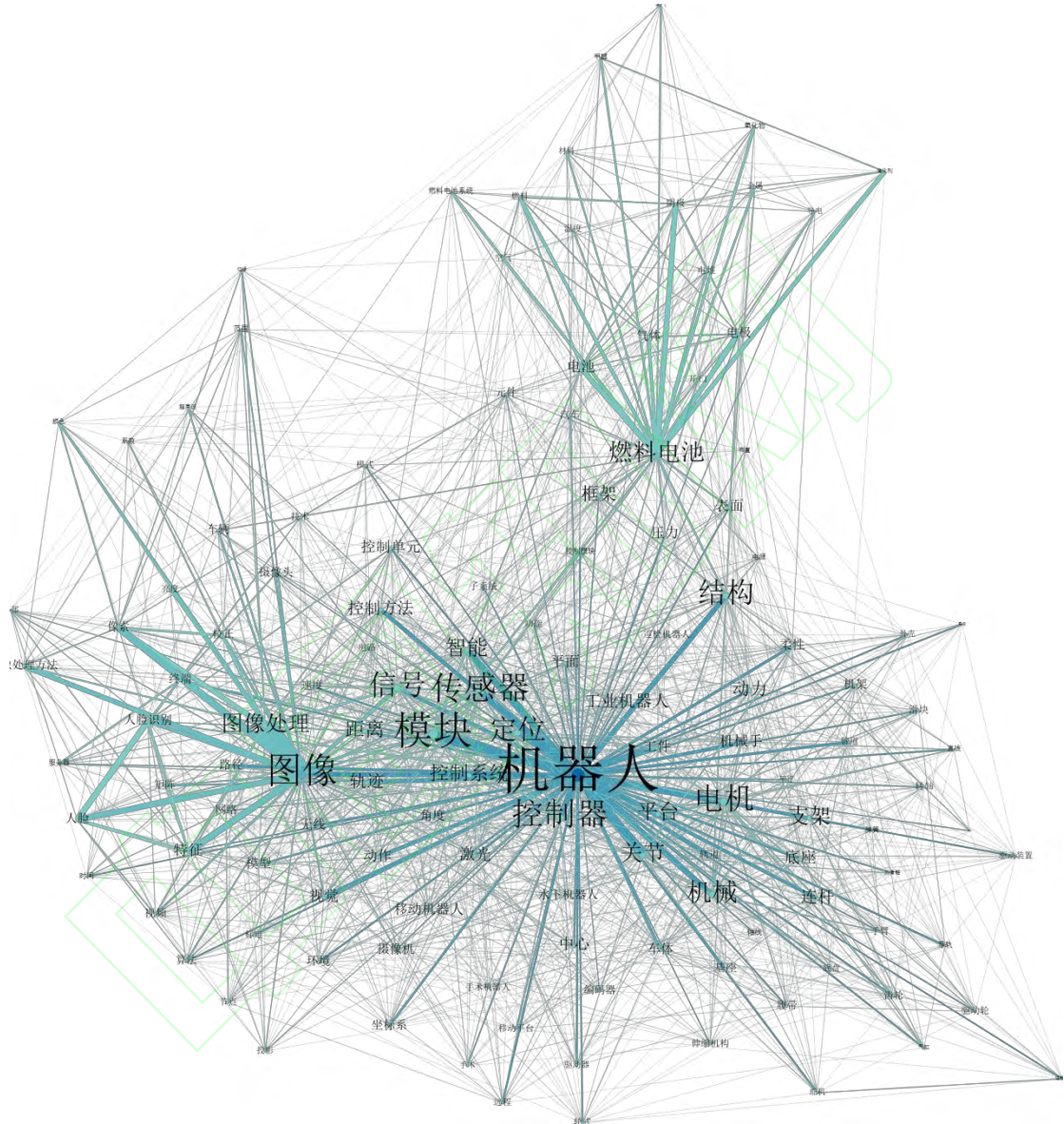


图3 2015年关键词共现网络图

Fig.3 The diagram of keyword co-occurrence network in 2015

然后，利用 sklearn 库中函数进行 LDA 模型的建立。通过绘制困惑度曲线，找到每一个时段的合适主题数量。利用 Networkx 库中函数对各时间段共现网络中的节点影响力进行测度，并对网络中节点的影响力进行归一化，最终得到各年度节点综合影响力排名。为方便展示，仅以 2015 年为例，见表 1。

表1 2015年节点综合影响力排名

Tab.1 The influence rankings for nodes in 2015

序号	关键词	节点影响力
1	机器人	1
2	图像	0.131973334

3	模块	0.038754682
4	电机	0.00453184
5	图像处理	0.001854914
6	传感器	0.000239476
7	关节	0.000142157
8	控制器	2.25861E - 05
9	机械	1.13915E - 05
10	定位	9.45614E - 06

综合每个时间段的困惑度曲线和节点综合影响力排名确定技术主题的初始个数,即社区探测方法中的 K 值。通过初始社区的扩展和相似社区的融合,最终得到各时间段的技术主题,以及各个主题中包含的关键词。在相似社区融合时,为提高各个主题的分度,进行反复实验,对比不同重叠率阈值下各时间段的主题数量和主题的分度,最终确定重叠率的阈值设置为 0.8。

提取出 2011—2021 年具有的技术主题个数分别为: 9、24、14、15、13、6、4、10、8、6、6。然后,对主题和专利数据进行匹配,得到技术主题—关键词—专利表。以 2015 年的主题—关键词表(部分)为例,见表 2。

表2 2015年主题—关键词表(部分)
Tab.2 The theme-keyword list in 2015 (partial)

主题	关键词
人脸识别技术及应用	人脸识别, 图像, 特征, 像素, 距离, 无线, 模块, 终端, 照片, 局部
燃料电池及其制备工艺	电极, 燃料电池, 氧化物, 导电, 机器人, 处理器, 表面, 距离, 传感器, 改性
图像处理技术及方法	终端, 图像, 亮度, 图像处理, 机器人, 处理器, 距离, 图像处理方法, 传感器, 无线,
机器人控制单元优化技术	控制单元, 燃料电池, 电器, 图像, 亮度, 特征, 图像处理, 机器人, 处理器, 传感器
自动定位与导航技术	算法, 图像, 特征, 图像处理, 机器人, 距离, 图像处理方法, 传感器, 模块, 移动平台

3.3 人工智能领域的突变性主题识别

首先,构建技术主题时序共现网络。利用 Gephi 软件对技术主题共现网络图进行可视化,可以更清晰地展示技术主题时序演化轨迹。根据时序人工进行横向时间轴的建立,完成技术主题时序网络的可视化,如图 4。

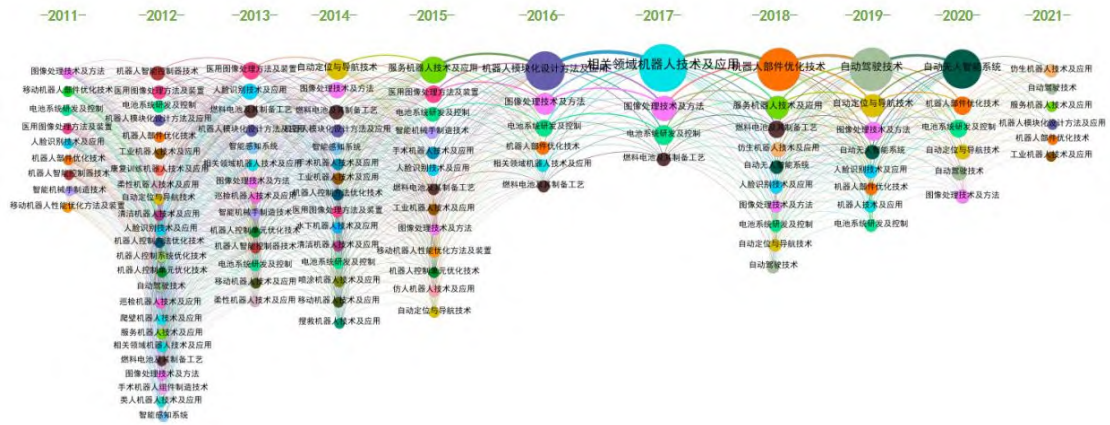


图4 2011—2021年人工智能领域技术主题时序共现网络

Fig.4 The time-series co-occurrence network of technology themes in artificial intelligence for 2011-2021

图4详细展示了人工智能领域技术主题在时间变化下的演化情况。王康等^[40]将人工智能领域的发展从1965—2020年分为了3段，分别为1965—1991年的萌芽期、1992—2011年的缓慢发展期、2012年以来的迅速发展期。从图4也可以看出2011年的主题数明显小于2012年，在2012年技术主题的数量出现了一次激增，说明进入了迅速发展期。这也说明本文的技术主题数量是恰当的。

根据图4对人工智能领域的技术主题演化进行特征分析，可以得到如下6点结论。

第一，2012年处于缓慢发展期和迅速发展期的过渡阶段，学者们的研究方向比较多元，从各个角度进行人工智能领域方面的研究，所以技术主题较为分散。2012年的主题规模在2011的基础上继续扩大。2011年出现的技术主题有：“机器人智能控制器技术”“医用图像处理方法及装置”“电池系统研发及控制”“机器人部件优化技术”“人脸识别技术及应用”，2012年出现的新生主题有：“机器人模块化设计方法及应用”“工业机器人技术及应用”“康复训练机器人及应用”“柔性机器人技术及应用”“清洁机器人技术及应用”等，大多为各种具有专业功能的机器人技术，即人工智能领域在其他领域的应用。由于2012年新出现的主题很多，按照规模判定主要的新生主题有：“机器人模块化设计方法及应用”“工业机器人技术及应用”“康复训练机器人技术及应用”“柔性机器人技术及应用”。

第二，在2013年持续扩张的主题有：“医用图像处理方法及装置”“人脸识别技术及应用”“燃料电池及其制备工艺”“机器人模块化设计方法及应用”。没有新生主题。但是，有两个主题在2011年出现的规模较小，而在2013年规模激增，即“智能机械手制造技术”主题和“移动机器人技术及应用”主题。

第三，2014年的技术主题有很明显的集中趋势。在2012年曾出现的技术主题，在2014年再次出现，且规模较2012年明显扩大，如“自动定位与导航技术”“清洁机器人技术及应用”“工业机器人技术及应用”“机器人控制方法优化技术”。此外，2014年的新生主题有：“手术机器人技术及应用”“水下机器人技术及应用”“喷涂机器人技术及应用”“搜救机器人技术及应用”。

第四，在2015年，2011年曾出现的主题服务机器人技术及应用再次出现，规模明显大于前几年且居于2015年之首。在2011年和2013年出现的智能机械手制造技术再次出现，可以初步认为该技术具有周期性，即当其他技术主题发展到一定程度，需要该技术主题的突破，才可以继续发展。2015年的新生主题有“仿人机器人技术及应用”。

第五，在2016年，曾分别在2012年、2013年和2014年出现过的“机器人模块化设计方法及应用”技术主题再次出现且规模最大。在2011年和2012出现过的机器人部件优化技术也再次出现，且规模扩大。2016年的新生主题有“相关领域机器人技术及应用”，当相关领域利用人工智能技术去大力发展自动化的时候，对于人工智能领域来说，所使用的技术相似度较高，所以会被分配到同一个技术主题下，因此将其概括为“相关领域机器人技术及应用”。

第六，在2017—2021年期间，邻域内技术主题更加趋于稳定，新生主题减少。在2018年仅有两个新生主题：“仿生机器人技术及应用”和“自动无人智能系统”。还有一个曾在

2012 年出现但规模很小、在 2018 年规模激增的主题，即“自动驾驶技术”。

通过分析人工智能领域的技术主题时序网络，得到各个时间段的新生主题即突变性主题，见表 3。

表 3 2012—2021 年各时间段突变性主题表
Tab.3 The Table of mutate themes by period 2012-2021

年份	突变性主题
2012	机器人模块化设计方法及应用、工业机器人技术及应用 康复训练机器人技术及应用、柔性机器人技术及应用
2014	手术机器人技术及应用、水下机器人技术及应用 喷涂机器人技术及应用、搜救机器人技术及应用
2015	仿人机器人技术及应用
2016	相关领域机器人技术及应用
2018	仿生机器人技术及应用、自动无人智能系统

3.4 人工智能领域的颠覆性技术识别

首先，利用 2.4.1 节中提出的基于专利引用构建主题引用网络的方法，利用刘玉琴等^[41]开发的软件 ItgInsight（正乙科技，<http://cn.Itginsight.com/>），对 2011—2021 年人工智能领域主题引用网络进行可视化，如图 5 所示。

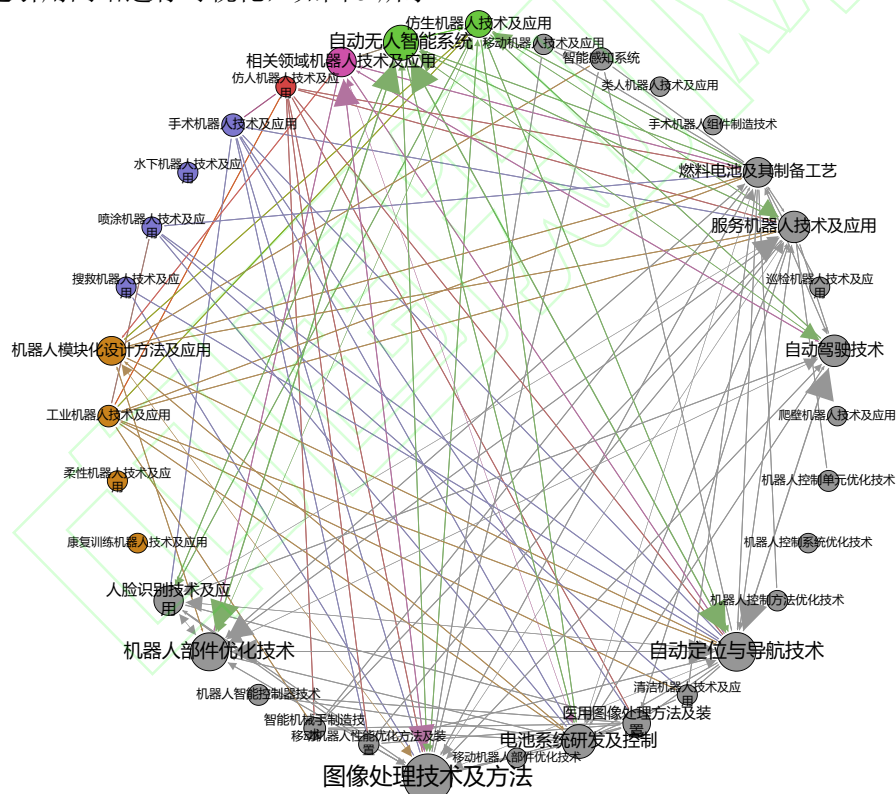


图5 2011—2021年人工智能领域主题引用网络图

Fig.5 The thematic citation network in artificial intelligence for 2011-2021

然后，将 3.3 节中的突变性主题对应到主题引用网络中，进一步判断突变性主题中演化形式为融合主题，如表 4 所示。

表4 突变性主题的融合主题表

Tab.4 The table of convergent themes for mutate themes

年份	突变性主题	融合主题
2012	机器人模块化设计方法及应用	图像处理技术及方法、电池系统研发及控制、 机器人部件优化技术

	工业机器人技术及应用	无
	康复训练机器人技术及应用	无
	柔性机器人技术及应用	无
2014	手术机器人技术及应用	无
	水下机器人技术及应用	无
	喷涂机器人技术及应用	无
	搜救机器人技术及应用	无
2015	仿人机器人技术及应用。	无
2016	相关领域机器人技术及应用	图像处理技术及方法、电池系统研发及控制、燃料电池及其制备工艺
2018	仿生机器人技术及应用	服务机器人技术及应用、机器人部件优化技术、人脸识别技术及应用
	自动无人智能系统	自动定位与导航技术、机器人部件优化技术、自动驾驶技术

因此，技术主题“机器人模块化设计方法及应用”“相关领域机器人技术及应用”“仿生机器人技术及应用”“自动无人智能系统”都同时具突变性和融合性，可以将它们识别为人工智能领域的颠覆性技术。

由于 2011 年之前的主题情况不可知，在 2011 年首次出现的“服务机器人技术及应用”和“机器人部件优化技术”这两个技术主题，根据定义没有办法判定为突变性主题。但是，由于它们在 2011 年的规模都很小，之后再次出现时规模激增，所以也可以认为是作为一个新主题出现的。并且，它们与多个技术主题有较强的引用关系，具有融合性特征，因此，也可以将它们识别为颠覆性技术。另外，还有一些主题在首次出现时规模极小，当其再次出现时规模激增，并且与多个主题具有较强的引用关系，也同理可将它们识别为颠覆性技术。例如，“服务机器人技术及应用”主题在 2011 年首次出现后很长时间未出现，而其在 2015 年再次出现时则与“机器人部件优化技术”“仿生机器人技术及应用”“人脸识别技术及应用”“图像处理技术及方法”等多个主题具有很强的引用关系；“自动驾驶技术”首次在 2012 年出现，但当时规模极小，2018 年再次出现时规模激增，引用了包括“自动定位与导航技术”“自动无人智能系统”“机器人部件优化技术”在内的多个技术主题，且与这些主题都联系紧密；在 2011、2012 年出现过的“机器人部件优化技术”，在 2016 年再次出现，并且规模有所扩张，其与“自动无人智能系统”“自动定位与导航技术”“服务机器人技术及应用”等多个主题具有较强的引用关系。

综上所述，人工智能领域的颠覆性技术共 7 项，分别为：“机器人模块化设计方法及应用”“相关领域机器人技术及应用主题”“仿生机器人技术及应用”“自动无人智能系统”“自动驾驶技术”“服务机器人技术及应用”和“机器人部件优化技术”。其中，“相关领域机器人技术及应用”主题包括很多专业性强的机器人技术及应用。

3.5 结果对比

如表 5 所示，对只通过突变性和融合性识别的主题与共同具有两种特性的主题进行对比发现，突变性主题中的“工业机器人技术及应用”“康复训练机器人技术及应用”“柔性机器人技术及应用”等专业性较强的技术主题，并不具备融合性，没有对专业技能以外的技术产生影响，因此并不能作为技术发展中的颠覆性创新。另一方面，在融合主题中的“人脸识别技术”“图像处理技术及方法”等技术主题，由于出现时间较早且技术的颠覆性会随时间不断减弱，不能被识别为颠覆性技术。由此可见，通过两种特征识别，可以将不完全具有颠覆性技术特征的识别结果进行更加准确地筛选。

表5 不同特征下识别的结果对比
Tab.5 Comparison of the results of recognition with different features

突变性主题	融合主题	具有融合性的突变性主题
机器人模块化设计方法及应用	自动无人智能系统	机器人模块化设计方法及应用
工业机器人技术及应用	自动驾驶技术	相关领域机器人技术及应用
康复训练机器人技术及应用	自动定位与导航技术	仿生机器人技术及应用
柔性机器人技术及应用	医用图像处理方法及装置	自动无人智能系统
手术机器人技术及应用	图像处理技术及方法	自动驾驶技术
水下机器人技术及应用	人脸识别技术	服务机器人技术及应用
喷涂机器人技术及应用	燃料电池及其制备工艺	机器人部件优化技术
搜救机器人技术及应用	机器人模块化设计方法及应用	
仿人机器人技术及应用。	相关领域机器人技术及应用	
相关领域机器人技术及应用	机器人部件优化技术	
仿生机器人技术及应用	服务机器人技术及应用	
自动无人智能系统	仿生机器人技术及应用	
	电池系统研发及控制	

欧盟委员会在 2019 年发布前瞻性研究报告《面向未来的 100 项颠覆性技术创新》，通过对研究前沿的科技文献进行大规模文本挖掘，并且结合各领域的专家评论，最终筛选出 100 项可能对全球经济产生重大影响的颠覆性技术。如果将本文识别出的人工智能领域的颠覆性技术与该报告中人工智能领域的颠覆性技术进行对比，“自动驾驶技术”“各类机器人技术及应用”均在报告中有提及，这表明本文提出的识别方法具有一定的可行性与合理性。

4 结论

从技术主题的演化出发，利用 LDA 模型和基于全局和局部影响力的社区探测方法识别领域内的技术主题，并依据颠覆性技术具备的“突变性”和“融合性”特征进行颠覆性技术识别。“突变性”特征主要通过构建技术主题时序共现网络，并分析主题演化趋势来确定；“融合性”则通过主题引用网络，判别突变性主题中具备“融合性”的主题，从而实现颠覆性技术的识别。最后，以人工智能领域为例，验证本文提出的颠覆性技术识别方法的可行性和有效性。

在理论上，从技术主题演化的动态视角分析了颠覆性技术的产生过程，为后续研究提供了理论基础。在方法上，首先，从技术主题演化的趋势和形式两个维度，提出了基于技术主题演化的颠覆性技术识别研究框架。通过测度技术主题的突变性和融合性，充分体现了颠覆性技术随时间变化的新颖性和交叉性特征。其次，提出了基于全局和局部影响力的技术主题提取方法。通过准确和全面评估每一个关键词在共现网络中的影响力，实现了对技术主题的高质量提取。

本文也存在一些不足之处。一是构建领域内专业词汇表并不充分，有些专业词汇没有添加进来，主题候选词的专业性有待提高。二是技术主题时序共现图的时序维度通过人工进行可视化，效率有待提高，后续可进行自动化绘制研究。三是实证研究的专利数据仅以中国人工智能领域的授权专利为代表，研究样本有限，未来可纳入不同国家、领域的专利数据，扩大分析样本，提高颠覆性技术识别的科学性。□

参考文献

[1] 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[J]. 中国工程

科学, 2017, 19(5): 9-16.

[2] 冯倩倩, 张光宇, 戴海闻, 等. 颠覆性技术遴选的指标体系与流程设计——基于扎根理论的多案例研究[J]. 科技管理研究, 2021, 41(24): 50-59.

[3] 曹悦, 白晨, 张英杰, 等. 颠覆性技术识别模型研究——以工业机器人领域为例[J]. 中国科技资源导刊, 2022, 54(2): 81-92.

[4] 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.

[5] HARDMAN S, STEINBERGER-WILCKENS R, VAN DER HORST D. Disruptive innovations: the case for hydrogen fuel cells and battery electric vehicles[J]. International Journal of Hydrogen Energy, 2013, 38(35): 15438-15451.

[6] JIA Weifeng, WANG Shuo, XIE Yongping, et al. Disruptive technology identification of intelligent logistics robots in AIoT industry: based on attributes and functions analysis[J]. Systems Research and Behavioral Science, 2022, 39(3): 557-568.

[7] CHEN Xiaoli, HAN Tao. Disruptive technology forecasting based on gartner hype cycle[C]//2019 IEEE Technology & Engineering Management Conference (TEMSCON). Atlanta, GA, USA: IEEE, 2019: 1-6.

[8] 张秀妮, 刘武英, 武茜. 基于动态赋值的存储芯片产业颠覆性技术识别研究[J]. 中国科技资源导刊, 2022, 54(4): 71-82.

[9] 马永红, 孔令凯, 林超然, 等. 基于异构数据的颠覆性技术识别研究——以智能制造装备领域为例[J]. 现代情报, 2022, 42(7): 92-104.

[10] 韩芳, 张生太, 冯凌子, 等. 基于专利文献技术融合测度的突破性创新主题识别——以太阳能光伏领域为例[J]. 数据分析与知识发现, 2021, 5(12): 137-147.

[11] 于光辉, 宁钟, 李昊夫. 基于专利和 Bass 模型的颠覆性技术识别方法研究[J]. 科学学研究, 2021, 39(8): 1467-1473,1536.

[12] 曹艺文, 许海云, 武华维, 等. 基于引文曲线拟合的新兴技术主题的突破性预测——以干细胞领域为例[J]. 图书情报工作, 2020, 64(5): 100-113.

[13] 王康, 陈悦, 王玉奇, 等. 基于专利引用变化的颠覆性技术识别研究[J]. 情报杂志, 2022, 41(1): 74-80,169.

[14] 王康, 陈悦. 技术融合视角下基于专利的颠覆性技术识别研究[J]. 情报杂志, 2022, 41(4): 29-36,134.

[15] 苏敬勤, 刘建华, 王智琦, 等. 颠覆性技术的演化轨迹及早期识别——以智能手机等技术为例[J]. 科研管理, 2016, 37(3): 13-20.

[16] 陈育新, 李健, 韩毅. 核心—边缘理论视角下的颠覆性技术识别研究[J]. 情报理论与实践, 2022, 45(8): 121-129.

[17] MOMENI A, ROST K. Identification and monitoring of possible disruptive technologies by patent-development paths and topic modeling[J]. Technological Forecasting and Social Change, 2016, 104: 16-29.

[18] 王海军, 于佳文. 基于专利发展路径的颠覆性技术识别: 以智能语音领域为例[J]. 科技管理研究, 2022, 42(6): 170-181.

[19] 陈虹枢, 宋亚慧, 金茜茜, 等. 动态主题网络视角下的突破性创新主题识别:以区块链领域为例[J]. 图书情报工作, 2022, 66(10): 1-14.

[20] 刘忠宝, 康嘉琦, 张静. 基于主题突变检测的颠覆性技术识别——以无人机技术领域为例[J]. 科技导报, 2020, 38(20): 97-105.

[21] 赵玉桐, 杨建林. 基于跨领域专利的颠覆性技术识别研究——以人工智能领域为例[J/OL]. 情报理论与实践: 1-13[2023-03-11]. <http://kns.cnki.net/kcms/detail/11.1762.G3.20220926.1903.004.html>

[22] JIA Weifeng, XIE Yongping, ZHAO Yanan, et al. Research on disruptive technology recognition of China's electronic information and communication industry based on patent influence[J]. Journal of Global Information Management, 2021, 29(2): 148-165.

[23] 张金柱, 张晓林. 基于被引科学知识主题突变的突破性创新识别[J]. 现代图书情报技术, 2016(Z1): 42-50.

[24] 马铭, 王超, 张伟然, 等. 突变视角下潜在颠覆性技术识别与分析方法研究[J]. 情报理论与实践, 2022, 45(3): 157-164,156.

-
- [25] 吴胜男, 田若楠, 蒲虹君, 等. 基于社交媒体的医药领域关联主题预测方法研究[J]. 数据分析与知识发现, 2022, 5(12): 98-109.
- [26] 宋凯, 朱彦君. 专利前沿技术主题识别及趋势预测方法——以人工智能领域为例[J]. 情报杂志, 2021, 40(1): 33-38.
- [27] 刘自强, 许海云, 岳丽欣, 等. 基于 Chunk-LDAvis 的核心技术主题识别方法研究[J]. 图书情报工作, 2019, 63(9): 73-84.
- [28] 王效岳, 赵冬晓, 白如江. 基于专利文本数据挖掘的技术预测方法与实证研究——以纳米技术在能源领域应用为例[J]. 情报理论与实践, 2017, 40(4): 106-110.
- [29] 周海炜, 吴成凤. 基于专利 SAO 结构和多指标评价的新兴技术识别研究——以手机芯片领域为例[J]. 情报杂志, 2022, 41(2): 86-94,48.
- [30] 张金柱, 吕品. 基于主题关联度改进的主题演变和突变分析[J]. 情报理论与实践, 2018, 41(3): 129-135.
- [31] 李慧, 玄洪升. 专利视角下融合多属性的技术创新主题挖掘方法——以芯片领域专利为例[J]. 图书情报工作, 2020, 64(11): 96-107.
- [32] 刘小玲, 谭宗颖. 基于专利多属性融合的技术主题划分方法研究[J]. 数据分析与知识发现, 2022, 6(Z1): 45-54.
- [33] 李乾瑞, 郭俊芳, 黄颖, 等. 基于突变-融合视角的颠覆性技术主题演化研究[J]. 科学学研究, 2021, 39(12): 2129-2139.
- [34] 尚增健. 渐进式技术创新:科技型中小企业的成长路径——成长型中小企业成长机理的个案分析[J]. 管理世界, 2002(6): 124-133.
- [35] 刘嘉, 张焕芝, 杨金华, 等. 全球油气勘探开发形势及技术发展趋势[J]. 世界石油工业, 2019, 26(6): 6-11.
- [36] BLEI D M. Probabilistic topic models[J]. Communications of the ACM, 2012, 55(4): 77-84.
- [37] SHENG Jinfang, DAI Jinying, WANG Bin, et al. Identifying influential nodes in complex networks based on global and local structure[J]. Physica A: Statistical Mechanics and its Applications, 2020, 541: 1-13.
- [38] MA Tinghuai, LIU Qin, CAO Jie, et al. LGIEM: global and local node influence based community detection[J]. Future Generation Computer Systems, 2020, 105: 533-546.
- [39] 陈挺, 邓启平, 李国鹏, 王小梅. 基于人工智能专利图谱的技术热点发现以及演化分析[J]. 中国发明与专利, 2021, 18(2): 13-21.
- [40] 王康, 陈悦. 技术融合视角下基于专利的颠覆性技术识别研究[J]. 情报杂志, 2022, 41(4): 29-36,134.
- [41] 刘玉琴, 汪雪锋, 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.

作者简介: 单晓红, 女, 1976 年生, 副教授, 博士生导师。研究方向: 信息管理, 商务智能。韩晟熙, 男, 2000 年生, 硕士生。研究方向: 信息管理。刘晓燕 (通信作者: jgliuxiaoyan@bjut.edu.cn), 女, 1974 年生, 副教授, 博士生导师。研究方向: 组织理论与战略管理。

作者贡献声明: 单晓红, 设计、改进研究方案和框架, 修订论文。韩晟熙, 确定论文选题, 撰写、修改论文。刘晓燕, 撰写、修改论文。

录用日期: 2023-02-27