# 面向科技大数据的颠覆性技术多维度识别算法

赵 姝1,2,3 郭双瑞2 褚晓泉4 仇 瑜5 刘姝伶2

(1.智能计算与信号处理重点实验室,安徽合肥 230601; 2.安徽大学计算机科学与技术学院,安徽合肥 230601; 3.安徽省信息材料与智能传感实验室,安徽合肥 230601; 4.清华大学计算机系,北京 100084; 5.北京智谱华章科技有限公司,北京 100084)

摘要:颠覆性技术具备"重塑"科技与市场未来格局的潜力。为支持科技评估中颠覆性技术的识别任务,提出一种基于科技大数据的颠覆性技术自动化识别算法 (MRA-DT)。首先,基于科技出版物引文网络结构、时序化特征、跨领域影响和技术时效性等信息,设计了 4 个维度的颠覆性技术识别指标,并采用熵权法进行结果融合。然后,在计算机科学和无线电通信领域进行实证研究。实证研究表明,MRA-DT算法是一种可靠的工具,可以为科技情报分析人员和相关管理者提供科技创新决策依据,对于企业抢占科技制高点、国家提升科技和国防实力具有重要意义。

关键词: 颠覆性技术; 科技评估; 文献计量; CDt指数; 熵权法

中图分类号: G301 文献标识码: A

# Multidimensional Recognition Algorithm for Disruptive Technology in Sci-tech Big Data

ZHAO Shu<sup>1,2,3</sup>, GUO Shuangrui<sup>2</sup>, CHU Xiaoquan<sup>4</sup>, QIU Yu<sup>5</sup>, LIU Shuling<sup>2</sup>

- (1. Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, Hefei 230601;
- 2. School of Computer Science and Technology, Anhui University, Hefei 230601; 3. Information Materials and Intelligent Sensing Laboratory of Anhui Province, Hefei 230601; 4. Department of Computer Science and Technology, Tsinghua University, Beijing 100084; 5. Beijing Knowledge Atlas Technology Co., Ltd, Beijing 100084)

**Abstract:** Disruptive technologies have the potential to reshape the future landscape of both technology and markets. This paper aims to propose an automated algorithm for identifying disruptive technologies based on big data analysis, which can support the identification of disruptive technologies in technology assessment. The algorithm utilizes information such as the citation network structure of technology publications, temporal features, cross-disciplinary impact, and technological timeliness to design four dimensions of indicators for identifying disruptive technologies, and adopts the entropy weighting method for result fusion. Empirical research was conducted in the fields of computer science and wireless communication. The results demonstrate

收稿时间: 2023 年 4 月 19 日。

作者简介: 赵姝(1979—), 女,博士,安徽大学教授,博士生导师,主要研究方向为粒计算、图神经网络、知识图谱(通信作者);郭双瑞(1997—),男,硕士生,主要研究方向为自然语言处理、上下位关系抽取;褚晓泉(1994—),女,博士,清华大学助理研究员,主要研究方向为科技情报智能挖掘算法;仇瑜(1988—),男,博士生,主要研究方向为知识表示、知识图谱和科技大数据挖掘;刘姝伶(2001—),女,本科生,主要研究方向为复杂网络。

基金项目: 国家自然科学基金面上项目 "异质信息网络的多粒度表示与知识获取方法研究"(61876001);安徽省高等学校自然科学研究项目"面向异构信息的多粒度知识推理模型及应用研究"(KJ2021A0039)。

that the MRA-DT algorithm is a reliable tool that can provide decision-making support for technology intelligence analysts and related managers in technology innovation, which is of significant importance for companies to seize technological high ground, and for countries to enhance their technological and defense capabilities.

**Keywords:** disruptive technologies, science and technology assessment, bibliometrics, CDt index, entropy weight method

# 0 引言

自 1995 年美国哈佛大学教授 Christensen 提出颠覆性技术 ( Disruptive Technology ) 概念[1]以来,引起了科研、产业等多方讨论。颠覆性技术以其对市场和科技领域的革命性影响而闻名,通常表现为对旧技术的淘汰和新技术的崭露头角,主要的表现方式包括对已有技术的突破和由学科交叉而诞生的具备创新性的技术。颠覆性技术的概念在不同学者和领域有不同解读。一种观点是强调颠覆性技术改变市场格局,另一种观点是着重于技术本身的革新。如智能手机的崛起淘汰了传统手机行业,传统汽车产业也面临混合动力的新能源汽车技术带来的挑战。

鉴于颠覆性技术对经济发展和国家竞争力产生的重大影响,各国相继把培育和支持颠覆性技术发展列为战略重点。美国在1958年设立国防预先研究计划局(DARPA),先后在互联网、隐身技术等领域开展颠覆性技术预测和发展管理<sup>[2]</sup>。日本内阁府与科学技术振兴机构联合推行"颠覆性创新计划"(ImPACT)和登月型研发制度来找寻颠覆性技术,期望以此对社会创新管理体系进行根本性变革<sup>[3]</sup>。在创新驱动的时代背景下,党的十九大报告明确提出,我国必须加强建设创新型国家的进程,突出颠覆性技术创新,将颠覆性技术研究与发展作为重点工作研究课题<sup>[4]</sup>。

由于颠覆性技术本身具有不确定性、突变性等特点,对其进行识别则要求研究人员对科学技术预测方法有透彻的了解,同时要有相对应领域的知识储备。综合来看,当代技术创新过程日益复杂,各国相关部门、各学科学者对颠覆性技术的识别面临诸多挑战。目前,主流的识别方法包括专家研判法和模型预测法,但它们都存在一

定的局限性。专家研判法容易受到主观因素的干扰,模型预测法虽然减少了主观性,但现有方法的分析维度相对有限。

专家研判法是一种以专家知识为基础的方法,常用方法包括德尔菲法、技术路线图法和情景分析法等。专家咨询所带来的系统性主观偏见,致使这个方法难以消除这种偏见[5]。模型预测法依靠论文、专利等科技大数据,通过分析数据特征进行颠覆性技术识别,能够有效避免主观因素带来的影响,提升识别过程的科学性和可解释性[6]。但是,现有方法分析维度有限,尚未形成系统的、通用的和可操作性强的面向科技大数据的颠覆性技术识别方法。因此,辨析颠覆性技术的在含义,把握其本质特征,利用科学的文献计量和智能数据分析方法,实现从科技大数据中挖掘颠覆性技术,是亟待解决的重点难题。

鉴于此,本文深入分析颠覆性技术的本质特征,构建一种多维度识别算法,以科技大数据为基础,实现对颠覆性技术的定量评估与挖掘。首先,本文总结现有的颠覆性技术的识别方法,厘清现有方法的不足和突破点;然后,基于科技文献的引文构建引文网络,针对科技引文网络的引文结构、时序化特征、跨领域影响力和技术时效性,设计技术突破性、创新性、领域影响力、科技文献影响力4个维度的识别指标,并采用熵权法进行指标计算结果融合,实现面向科技大数据的颠覆性技术多维度识别算法。本文在两个典型领域进行实验,验证了此方法的有效性,为科技评估中颠覆性技术识别任务提供可靠的工具。

## 1 相关工作

## 1.1 专家研判法

截至目前,专家研判法仍是颠覆性技术识

别研究中普遍使用的方法。在对具体问题的领域 现状进行调研的基础上,汇集领域权威专家的力量,以发放问卷、开展研讨会等方式进行颠覆性 技术的定位和预测,是面临宏观分析需求时重 要的应对方式。德尔菲法(Delphi Method)、技术路线图法(Technology Roadmap)和情景分析 法(Scenario Analysis)是具有代表性的专家研判 方法。

- (1)德尔菲法是一种典型的方法,以专家意见为基础,首先设计技术领域的调查问卷,然后邀请领域专家进行回答,意见在多轮征询中趋于一致,从而决策未来技术发展方向,目前已经广泛应用于不同领域的研究中。李晓龙等问、但智钢等图以不同领域颠覆性技术预测为研究目标,分别利用德尔菲法对颠覆性技术进行分析;Kostoff等<sup>[9]</sup>采用德尔菲法结合文本挖掘,将潜在的颠覆性技术从科技文献中识别出来。从研究结果可以看出,德尔菲法能充分利用领域权威专家意见,具有较强的说服力。但是,这类方法的预测结果完全依赖于专家的主观意见和知识经验,具有很强的主观性与局限性。
- (2)技术路线图是一种图形化、结构清晰的技术规划管理工具,以便实现技术规划的结构化管理。利用探索技术、产品和市场之间的动态变化关系,发掘关键技术。技术路线图作为有益的工具,能够帮助学者探测技术性能在时间演进中的变化和跃迁,进而识别颠覆性技术。Uchihira<sup>[10]</sup>从技术的基本特征功能出发,通过预测目标市场的延伸功能需求,寻找可以应用技术的基本功能的市场。Momeni等[11]利用技术路线图评估光伏产业中具有潜力的颠覆性技术,并制定合理的专利发展路径。但是,在构建技术路线图的过程中需要考虑在技术、市场、支撑条件等多方面因素,存在不确定性且耗时耗力,在对颠覆性技术进行识别时,结果的客观性可能会受到一些影响。
- (3)情景分析法是一种基于对某种现象或 趋势进行假设的预测方法,可以预测未来可能出 现的情况。2014年,《游戏规则改变者:颠覆性

技术与美国国防战略》报告由新美国安全中心发布,对未来可能出现的技术应用情况进行情景分析,并研究颠覆性技术的需求和相对障碍,以实现技术识别[12]。Schoemaker等[13]认为情景规划法可以把握技术与市场之间的动态关系,预测新兴技术趋势,引导企业技术投资。但是,运用情景分析法进行预测需要投入大量时间和精力获取外部环境信息,对专家的要求较高,实际应用上存在一定的困难。

在人工智能和大数据飞速发展的新时代,面 对国家和企业对颠覆性技术预测的迫切需求,从 科技大数据中挖掘潜藏信息,弥补专家研判在科 学性、客观性方面的不足,已经成为科技管理和 技术预测的新趋势。因此,基于数据和模型的预 测方法已成为这个领域关注的重点。

# 1.2 科技大数据分析法

科技数据,如专利和论文,包含了丰富的技术信息。科学创新思想往往最初通过科学论文呈现。将科技大数据作为基础,并采用深度学习算法挖掘信息,进而对技术领域进行预测,被认为是更客观的颠覆性技术识别方法。目前,常见的方法有主题分析法(Thematic Analysis)、数据分析法(Data Analysis)和指标评估法(Index Evaluation)等。

(1) 主题分析法是一种用于文本形式的数据分析方法,首先对科技文献主题概念进行提取,然后采用定量方法对不同的主题进行提炼与取舍,从而实现技术识别。石慧等[14]借助多主题融合的思想,对论文数据中主题词的变更进行分析,同时结合专利手工代码的突变监测,实现对颠覆性技术的识别。Momeni等[15]利用主题模型和聚类分析相结合的方法,对专利技术路径进行分析,进而实现对颠覆性技术的识别。Dotsika等[16]利用词频网络共现分析方法,探究科学文献的时间演化,包括主题和结构的变化。但是,这类方法对技术的识别以主题为单位,实现了基于数据的分析和识别,适合宏观的评估需求,容易在分析过程中忽略具有发展潜力的细粒度技术点。

- (2)数据分析法是数据分析模型的应用,也是颠覆性技术识别重要方法[17]。黄鲁成等[18]在专利分析中尝试将物种入侵模型应用于颠覆性技术的识别。Cheng等[19]运用SIRS传染病模型,探究颠覆性技术的爆发规律,并从不同角度分析这个现象。Adner<sup>[20]</sup>利用用户需求构建的量化分析模型,识别颠覆性技术发生作用的条件。Sood等<sup>[21]</sup>提出技术颠覆性风险模型,利用风险系数预测颠覆性技术。但是,由于不同领域之间的知识内涵存在差异,很多研究无法直接应用于科技领域,同时对于知识扩散的测度研究较少,所得结论在指导实践方面还有待深入。
- (3) 指标评估法是对颠覆性技术的产生机 理与变革特征进行分析,聚焦于技术的原理性突 破、技术集成与跨领域应用。这个方法以论文、 专利等作为数据基础,在传统情报研究方法的基 础上,建立多种定量指标的评价机制来识别颠覆 性技术,并结合一定数量的技术实例对其进行验 证。Kaplan等[22]、马荣康等[23]将利用专利被引 次数作为衡量颠覆性技术的依据。根据 Martino [24] 等的观点,以及褚晓泉等[25]的研究,利用不同的 文献类型可以反映不同的技术发展阶段,如SCI 论文反映技术研究的基础阶段, EI论文反映技术 研究的应用阶段, 专利数据则反映技术研究的试 验发展阶段,从而对颠覆性技术继续识别。Park 等[26]基于引文网络提出的CD指标对颠覆性技术 的突破性进行衡量,实现了从引文网络结构层面 的颠覆性技术识别,被广泛应用于颠覆性技术分 析任务中, 但是这个指标仅能识别出具有特定引 文结构的技术点, 缺乏综合性的评估, 无法避免 识别结果的可信度。

虽然利用科技文献的外部特征可以在一定程度上识别颠覆性技术,但所识别出的技术主题范畴较为广泛。此外,不同的评估指标和识别模型从不同的角度和侧重点出发,致使指标模型的主观性较强且缺乏通用性,同时存在分析维度相对单一、识别标准缺少针对性等不足,现有研究结论存在局限性。因此,针对现有方法存在的不足,本文基于颠覆性技术发展特征逐一进行

突破,提出名为MAR-DT的颠覆性技术识别框架,结合大规模引文网络,提出技术突破性、创新性、领域影响力与科技文献影响力 4 个计量指标,采用熵权法进行指标融合,并通过实际数据进行验证。

## 2 颠覆性技术多维度识别算法

在科技领域,颠覆性技术的特征涵盖多个层面,对学术界和产业界都产生深远影响,且在学术文献中得以显著反映。为了更详细地分析这一复杂主题,本文引入了一个四维模型,涵盖技术的突破性、创新性、领域影响力以及科技文献影响力。

首先,技术的突破性特征反映了颠覆性技术 在特定领域取得的显著进展。这不仅包括技术在 该领域的技术性突破,更关注其对传统技术或思 维范式的激进性改变。如技术的突破性可能表现 为对传统产业链的颠覆性重构。

其次,技术的创新性特征强调颠覆性技术相对于现有技术的新颖性和独创性。这一维度不仅关注技术本身的创新,还考察其在应用方面可能呈现的全新思路和独特解决方案。如技术的创新性可能体现在其引入了全新的商业模式或解决了长期未解决的问题。

再次,技术的领域影响力特征表示颠覆性技术对特定领域的深刻影响程度,包括其应用范围和对该领域整体的重要性。这一方面可以体现为技术的广泛应用,另一方面可以体现技术对这个领域未来发展方向的引导作用。

最后,科技文献影响力特征反映了颠覆性技术在学术研究中的知名度和引用情况。高科技文献影响力通常意味着技术在学术界具有广泛认可和应用,为未来研究提供了坚实的理论基础。这可以通过技术在学术期刊中的发表次数、被引次数以及与其他重要文献的关联性等指标来衡量。

本文提出的颠覆性技术多维度识别算法(MAR-DT)可以划分为4个关键步骤:①获取与预处理科技大数据;②构建领域引文网络;③构建与评估多维指标;④识别颠覆性技术和量化

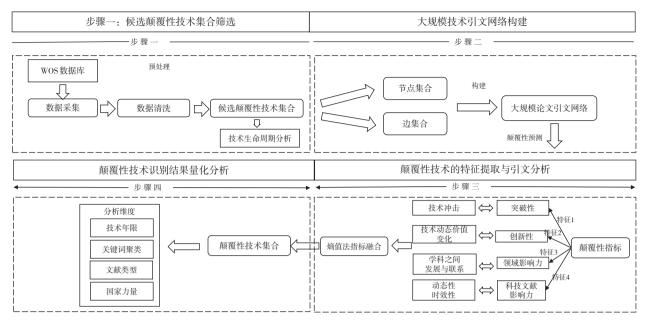


图 1 多维度颠覆性技术识别框架

分析结果。首先,根据分析目标设计文献检索策略,在Web of Science(WoS)科技论文数据库中收集颠覆性技术识别候选论文集,进行人工筛选和预处理,构建数据集;其次,借助数据集中引文字段提取节点集合与边集合,构建领域引文网络;再次,根据颠覆性技术特征构建面向技术突破性、创新性、领域影响力和科技文献影响力的评估指标,基于数据和引文网络进行指标计算并使用熵权法对指标矩阵进行融合计算,排序得出颠覆性技术候选清单;最后,从候选技术清单中分析、筛选得出颠覆性技术清单。算法总体流程如图 1 所示。

## 2.1 科技大数据收集与预处理

为了实现科技大数据分析,本文拟采用WoS 为数据获取平台。面对特定的颠覆性技术分析任 务,针对待分析领域设定的科学、全面的检索策 略,利用专业术语和专业的检索逻辑来确保原始 材料的准确性。完成检索后下载原始数据关键字 段信息。以论文数据为例,具体下载字段及其在 MAR-DT算法中的作用,具体见表 1。

#### 2.2 领域引文网络构建

科技文献中存在的引证关系体现着技术之间的流动,对于引文网络可以从结构的角度探究技术影响力,将有助于从网络的整体结构中动态地评价技术的重要性,也有助于探究技术在知识流动网络中的作用。本文根据科技论文数据创建节点集合  $Vertices = [doi_1, doi_2, doi_3, \cdots, doi_k]$  与边集  $Edges = [e_1, e_2, \cdots, e_t]$ ,构建大规模科技文献引文网络G = (V, E)。

#### 2.3 颠覆性技术识别指标获取

### 2.3.1 技术突破性指标

颠覆性技术出现后,对于科技文献的老化速 度具有催化作用,这种现象经常在科技文献的引 用关系中表现出来,当具有颠覆性的科技文献出

WoS字段	内容	作用
DI	数字对象标识符	论文数据唯一身份标识,通过该标识查询引文关系等信息
CR	引文信息	论文的引文信息,用于计算论文所代表技术的突破性
PY	发表年份	用于计算论文影响力
TI	论文标题	用于结果展示和主题验证
DE	论文关键词	用于结果展示和主题验证
SC	所属领域	用于计算论文所代表技术的领域影响力

表 1 论文数据子段说明

现后,科研人员对这个领域之前传统方法的关注 度急剧下降,转而开始频繁地引用具有颠覆性意 义的科技文献。

按照技术创新的方式,颠覆性技术可以划分为对现有技术的巩固和对现有技术产生冲击两种不同的类型。在科技文献数据中的体现是当一项技术以巩固性技术的形式出现时,引用文献通常会同时引用相关领域技术前辈的研究成果,以进行对比和分析;相反地,当一项颠覆性技术以技术冲击的形式出现时,引用文献对于相关领域先前研究的引用将减少。为了度量这个属性,本文引入Funk等[27]提出的CD指数进行颠覆性技术突破性(Breakthrough)的分析,其公式为:

Breakthrough = 
$$\frac{1}{n} \sum_{n=1}^{n} \left( -2f_i b_i + f_i \right)$$
 (1)

式中,n表示候选文献的施引文献数目;  $f_i$ 、 $b_i$ 用于度量文献引用情况。技术突破性指标范围为 [-1,1],指标值越接近-1,表明这项技术越具有

巩固性; 反之, 若突破性指标值越接近1, 表明此候选文献的技术在突破性方面具有越高的水平。引文结构如图2所示。

#### 2.3.2 技术创新性指标

颠覆性技术具有较高的创新性,发布后短时间内大概率将受到相关学者共同关注,呈现出快速扩张的趋势,在一定的时间单元内被引数量会显著增加。本文发现这种短期受到大量关注的现象可以根据引文网络中在一定时间内文献被引数量的变化进行衡量,因此针对颠覆性技术在问世后会存在短暂的快速关注期这一现象。本文通过对引文数量的短时间内的变化趋势进行创新性指标设计,以此解释颠覆性技术知识创新与传播的动态过程。创新性(Innovation)的计算方法为:

$$Innovation = \max \left( c_{v} - c_{v-1} \right) \tag{2}$$

式中, $c_y$ 表示目标论文在y年份的被引用次数; $c_y$ - $c_{y-1}$  表征候选科技文献在公开发表后每两年被引用数量的差距。技术创新性指标计算结果

同时引用目标参考文献和技术前辈的科技文献

引用技术前辈的科技文献

引用当前技术的科技文献

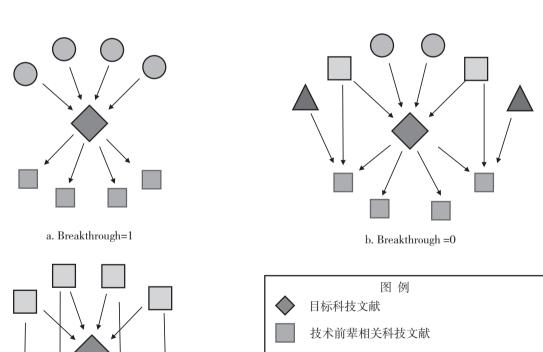


图 2 技术突破性指标示意

c. Breakthrough =-1

越高,代表技术在单位时间内的学术影响力越大,越具创新性。

#### 2.3.3 领域影响力指标

颠覆性技术存在发展与完善的过程,在过程中会随着知识跨领域的扩散,技术本身的影响逐渐扩散到其他领域。本文研究认为,颠覆性技术具有渗透或促进其他技术领域的潜力,对其他领域技术产生影响是其重要特征。技术融合可以使整个技术体系及其产品构成、性能等取得根本性的突破,技术融合还可以打破原有的技术格局,其影响深远。因此,本文设计了技术领域影响力这项指标,定量地评估颠覆性技术的跨领域影响力。这项指标是利用被引用的科技文献所属的技术领域信息进行度量的。技术影响力(Influence)指标表示一篇科技文献所引用的非原始科技文献在其技术领域内的影响程度。其计算公式为:

Influence = 
$$\frac{\sum_{n=i}^{n} dt_{i}}{\sum_{n=i}^{n} d_{i}}$$
 (3)

式中,n表示科技文献所有施引文献的数量; $dt_i$ 表示非原始科技文献自身领域的数量; $d_i$ 表示科技文献所引用的文献所涉及的技术领域。指标得分越高,这项技术的技术影响力越大。

#### 2.3.4 科技文献影响力指标

科技文献被认可的程度能够在被引用量的 变化上体现,但一些较新的科技文献为体现出文 献调研的全面性,不可避免地引用经典的科技文 献,对颠覆性技术的识别造成干扰。受到期刊影 响因子的启发,避免将被引用量作为单一的衡 量指标,引入时间维度的考量,提出技术即时影 响力的指标,聚焦于文献本身的动态性与短时变 化性,对颠覆性技术进行更细粒度的评估。技术 即时影响力是指一篇科技文献在其公开发表后当 年、发表后第二年以及发表后第三年的被引用量 之和与目标科技文献总被引量之间的比值,以此 突出科技文献的时效特性。

Instant Influence=
$$\frac{\left(c_{y}+c_{y+1}+c_{y+2}\right)}{\sum_{t=y}^{i}c_{t}} \tag{4}$$

式中, $c_y$ , $c_{y+1}$ 和 $c_{y+2}$ 分别代表目标科技文献在发表后当年、发表后第二年以及发表后第三年的被引用量。技术即时影响力指标得分越高,说明此目标科技文献的短时间技术影响力越大。

### 2.4 颠覆性技术识别指标融合

本文为保证所提各项指标融合时权值的客观性,运用客观赋值熵权法对本文 4 项评价指标进行了加权计算。熵权法是一种客观赋权法,计算指标的信息熵,并根据各指标相对于整体的动态变化程度对整个系统的影响,动态地确定各指标的权重。熵权法以数据的分布为基础,拥有较高的可靠性和准确性。

颠覆性技术指标矩阵满足
$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}$$
。

其中, $x_{ij}$ 代表i篇候选科技文献在第j项评价指标上的得分值。首先,利用式(5)对指标矩阵进行规格化处理。

$$\overline{x}_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$$
 (5)

式中, $\max(x_j)$ 表示在第j项指标上候选文献的最高得分值; $\min(x_j)$ 表示在第j项指标上候选文献的最低得分值; $\bar{x}_{ij}$ 代表归一化后的第i篇候选文献在第i项指标上的得分值。

进行归一化后得到的颠覆性技术指标矩阵

对归一化后的指标矩阵进行指标比重计算, 如式(6)所示。

$$P_{ij} = \frac{\overline{x}_{ij}}{\sum_{1}^{n} \overline{x}_{ij}} \quad i = 1, 2, ..., n; j = 1, 2, ..., m \quad (6)$$

然后,使用式(7)计算第j项指标的熵值:

$$e_{j} = -\varepsilon \times \sum_{i=1}^{n} P_{ij} \times \log(P_{ij}) \quad \varepsilon = \frac{1}{\ln(n)}$$
 (7)

利用式 (8) 对第j项指标的差异系数进行计算。

$$g_j = 1 - e_j \tag{8}$$

式中指标的差异系数越大,对于最后的颠覆 性技术评估的影响就越显著。

最后,利用式(9)对第j项指标的权重进行 计算。

$$W_j = \frac{g_j}{\sum_{h=1}^{m} g_h} \tag{9}$$

根据式(9)进行权重计算,分别得到各指标的权重,并综合突破性、创新性、技术影响力以及技术即时性指标,使用式(10)得到最终指标值Comprehensive。

$$Comprehensive_i = W_{ij} \times x_{ij}$$
 (10)

式中, $x_{ij}$ 表示第i篇文献的第j项指标的权重; $w_{ij}$ 表示其对应的权重。对最终指标值 *Comprehensive*进行逆序排序,选取排序结果的前 100 篇作为颠覆性技术识别结果。

# 3 颠覆性技术识别实证分析

本文采用计算机科学和无线电通信两个领域

的科技文献数据进行实证研究,深入分析并识别 其中的潜在的颠覆性技术,探索和揭示这些颠覆 性技术的潜力和影响。同时,本文还将提出相关 领域技术与产业发展方面的战略思考与建议,为 未来的创新和发展提供参考与借鉴,以促进技术 和产业的可持续发展。

### 3.1 数据预处理

本文采用的计算机科学与无线电通信领域的科技文献样本源于WoS(Web of Science)科技文献数据库,根据领域文献检索策略,下载WoS类别为计算机科学(Computer Science)与无线电通信(Telecommunications)领域的科技文献数据(其类型包括Article、Proceedings Paper、Review等),发表日期范围限定在1990—2020年,使用DOI号作为数字对象唯一标识符对数据进行清洗并去重,最终得到计算机科学领域科技文献数据1898805条,无线电通信领域科技文献数据423891条。计算机科学与无线电通信领域论文数据年份分布如图3所示。

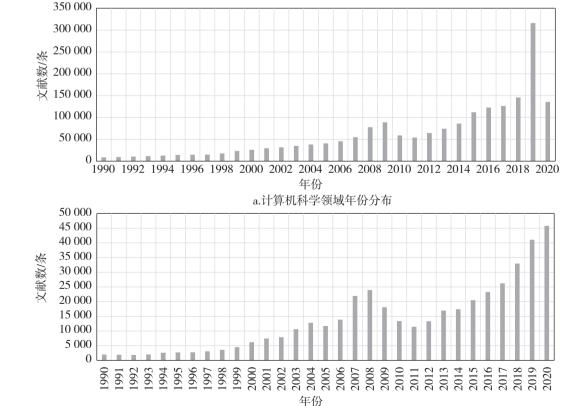


图 3 计算机科学与无线电通信领域论文数据年份分布

b.无线电通信领域年份分布

## 3.2 颠覆性技术识别结果分析

本文利用提出的方法,将计算机科学领域和 无线电通信领域的科技文献数据的DOI号作为网 络节点标识符,利用引文信息构建边集合,以此 搭建大规模科技文献引文网络。基于引文网络分 别计算突破性、创新性、技术影响力以及科技文 献影响因子指标,在对指标进行标准化处理后, 利用熵权法计算各文献的综合性指标。计算机科 学领域和无线电通信领域文献综合指标值前 20 位的排序分别见表 2、表 3。

从表 2 可以看出, 计算机科学领域中根据综合指标值排序的前 20 篇科技文献内容涉及CNN经典框架的AlexNet、Faster-RCNN以及对Faster-RCNN进一步改进的Mask R-CNN, 同时包含LSTM、SVM、Random Forest等机器学习经典算法, 其中来自美国的成果高达 12 篇, 中国

表 2 计算机科学领域文献综合指标值排序(Top 20)

表 2 月异机件子视域文献综合指标直排序(10月20)								
序号	DOI号	发表 年份	被引 频次	综合 指标值	国家	标题		
1	10.1145/3065386	2017	5 734	0.651 018	USA	ImageNet Classification with Deep Convolutional Neural Networks		
2	10.1109/TPAMI.2016.2577031	2017	3 484	0.625 254	China	Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks		
3	3 10.1109/ICCV.2017.322		1 836	0.595 496	USA	Mask R-CNN		
4	10.1109/CVPR.2017.243	2017	2 536	0.570 03	USA	Densely Connected Convolutional Networks		
5	10.1109/CVPR.2016.90	2016	12 994	0.541 894	China	Deep Residual Learning for Image Recognition		
6	10.1023/ B:VISI.0000029664.99615.94	2004	11 281	0.331 005	Canada	Distinctive Image Features from Scale-Invariant Keypoints		
7	10.1109/ICCV.2015.169	2015	2 505	0.322	USA	Fast R-CNN		
8	10.1109/CVPR.2016.308	2016	2 209	0.305 918	USA	Rethinking the Inception Architecture for Computer Vision		
9	10.1007/s11263-015-0816-y	2015	4 756	0.297 608	USA	ImageNet Large Scale Visual Recognition Challenge		
10	10.1145/1961189.1961199	2011	6 925	0.291 745	China	LIBSVNM: A Library for Support Vector Machines		
11	10.1023/A:1010933404324	2001	9 499	0.288 057	USA	Random Forests		
12	10.1109/CVPR.2016.91	2016	2 078	0.271 847	USA	You Only Look Once: Unified, Real-Time Object Detection		
13	10.1007/978-3-319-24574- 4_28	2015	4 195	0.263 184	Germany	U-Net: Convolutional Networks for Biomedical Image Segmentation		
14	10.1145/2647868.2654889	2014	3 407	0.260 485	USA	Caffe: Convolutional Architecture for Fast Feature Embedding		
15	10.1007/978-3-319-46448-0_2	2016	1 980	0.254 694	USA	SSD: Single Shot MultiBox Detector		
16	10.1162/neco.1997.9.8.1735	1997	7 852	0.240 581	Germany	Long Short-Term Memory		
17	10.1109/TIP.2003.819861	2004	6 739	0.224 71	USA	Image Quality Assessment: From Error Visibility to Structural Similarity		
18	10.1007/BF00994018	1995	3 283	0.222 45	USA	Support-Vector Networks		
19	10.1016/j.neunet.2014.09.003	2015	2 007	0.219 311	Switzerland	Deep Learning in Neural Networks: An Overview		
20	10.1109/ICCV.2015.123	2015	1 816	0.216 467	None	Delving Deep into Rectifiers: Surpassing Human– Level Performance on ImageNet Classification		

表 3 无线电通信领域文献综合指标值排序(Top 20)

序号	DOI 号	发表 年份	被引频次	综合 指标值	国家	标题
1	10.1109/ COMST.2016.2532458	2016	502	0.507 134	Korea	Next Generation 5G Wireless Networks: A Comprehensive Survey
2	10.1002/ett.4460100604	1999	2 393	0.485 503	USA	Capacity of Multi-Antenna Gaussian Channels
3	10.1109/49.730453	1998	3 111	0.476 108	USA	A Simple Transmit Diversity Technique for Wireless Communications
4	10.1109/JSAC.2014.2328098	2014	2 008	0.475 779	USA	What Will 5G Be?
5	10.1109/ COMST.2014.2368999	2015	652	0.440 898	Singapore	Wireless Networks With RF Energy Harvesting: A Contemporary Survey
6	10.1109/JSAC.2004.839380	2005	2 784	0.434 572	Canada	Cognitive Radio: Brain-Empowered Wireless Communications
7	10.1109/ MCOM.2014.6815890	2014	524	0.424 399	Germany	Scenarios for 5G Mobile and Wireless Communications: the Vision of the Metis Project
8	10.1109/ COMST.2015.2444095	2015	848	0.414 867	USA	Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications
9	10.1109/49.840210	2000	2 223	0.390 997	Italy	Performance Analysis, of the IEEE 802.11 Distributed Coordination Function
10	10.1109/ TWC.2013.031813.120224	2013	903	0.381 774	Singapore	MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer
11	10.1109/ MCOM.2014.6736746	2014	917	0.381 14	France	Five Disruptive Technology Directions for 5G
12	10.1109/ TWC.2010.092810.091092	2010	1 483	0.374 168	USA	Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas
13	10.1109/ ACCESS.2013.2260813	2013	1 446	0.370 216	USA	Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!
14	10.1109/ MCOM.2014.6736752	2014	533	0.364 356	Scotland	Cellular Architecture and Key Technologies for 5G Wireless Communication Networks
15	10.1109/ MCOM.2014.6736761	2014	902	0.363 007	Sweden	Massive MIMO for Next Generation Wireless Systems
16	10.1109/ MCOM.2002.1024422	2002	1 856	0.358 546	USA	A Survey on Sensor Networks
17	10.1109/98.788210	1999	1 965	0.354 408	Sweden	Cognitive Radio: Making Software Radios More Personal
18	10.1109/TWC.2002.804190	2002	1 603	0.329 361	USA	An Application-Specific Protocol Architecture for Wireless Microsensor Networks
19	10.1109/ TWC.2013.062413.122042	2013	640	0.328 827	Australia	Relaying Protocols for Wireless Energy Harvesting and Information Processing
20	10.1109/JSAC.2014.2330193	2014	1 816	0.327 843	USA	In-Band Full-Duplex Wireless: Challenges and Opportunities

位居第 2 名。从表 3 可以看出,无线电通信领域的 5G无线通信技术、Internet of Things 以及经典802.11 协议相关技术均在识别结果 Top 20 中。

为了进一步分析计算机科学和无线电通信 领域颠覆性技术的发展趋势,本文针对各领域 识别结果的Top 100进行时序分析。从图 4a可

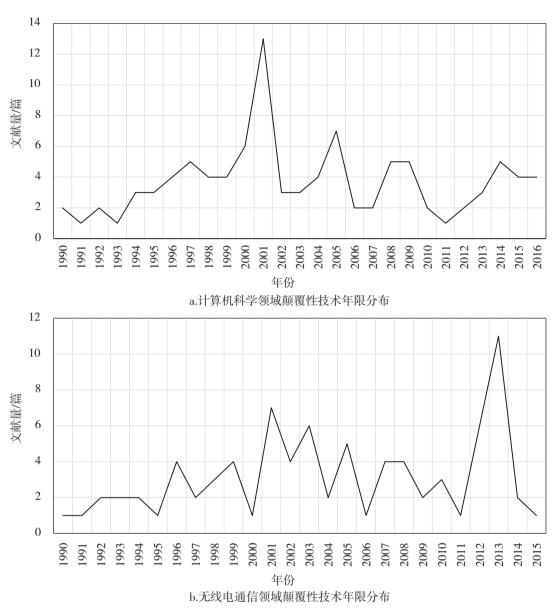
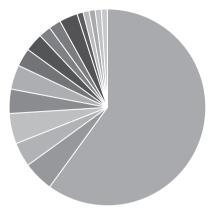


图 4 计算机科学领域与无线电通信领域颠覆性技术年限分布

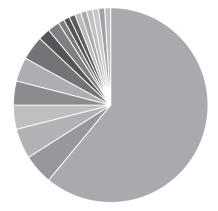
以看出, 计算机科学领域在 2002 年前后出现了多项颠覆性技术, 其中包含 NSGA-II 遗传算法、BLEU指标、k-anonymity模型、LBP特征 提取算法以及 SMOTE不平衡数据集处理方法等, 其中 NSGA-II 遗传算法是多目标进化领域中具有里程碑意义的技术, SMOTE是一种应用广泛的人工合成数据的综合采样算法。从图 4b可以看出, 无线电通信领域在 2014 年前后出现多篇颠覆性技术, 其中包含对于 5G无线通信技术的关键技术介绍(如多人多出技术 MIMO等)、应用的愿景与挑战,符合无线通信技术的发展规律。由此

可以看出,本文提出的框架能够对计算机科学领域与无线电通信领域具有颠覆性的方法或指标进行识别。

通过对颠覆性技术所属国家进行统计分析,可以在国家的层面了解这项技术在国际上的竞争力。因此,本文使用E-charts绘制了国家力量分布饼状图,如图 5 所示。美国作为世界上最具科技创新实力的国家,发布了数量最多的颠覆性技术成果。在无线电通信领域与计算机科学领域所拥有的颠覆性技术领域处于领先地位。中国在颠覆性技术突破方面仍具备较大的提升空间,加强







b.无线电通信领域国家力量分布

```
注: ■USA-61% ■ GERMAN-5% ■ SINGAPORE-5% ■ SWEDEN-4%
■ ITALY-4% ■ CANADA-4% ■ SOUTH KOREA-4% ■ SWITZERLAND-2%
6 ■ FRANCE-2% ■ JORDAN-1% ■ JAPAN-1% ■ ENGLAND-1%
■ FINLAND-1% ■ AUSTRALIA-1% ■ SCOLAND-1% ■ CHINA-1%
■ SPANISH-1% ■ UNKNOW-1%
```

图 5 计算机科学领域与无线电通信领域颠覆性力量分布

颠覆性技术的研发是实现技术领跑和维护领域技术安全的重要途径。

# 4 总结与展望

本文根据颠覆性技术发展特征、关联科技 大数据自身时序信息以及引文结构变化等,提 出一种定量化的颠覆性技术的多维度识别框架 (MRA-DT),综合考虑突破性、创新性、领域影 响力和科技文献影响力,设计了4种不同的颠覆 性技术识别指标,利用熵权法对颠覆性技术各特 征指标进行量化测度,基本实现了对论文数据的 颠覆性技术细粒度挖掘,解决了传统方法在科学 性、可解释性等方面存在的不足。

为验证此算法的可行性与有效性,本文利用 WoS数据库中在1990—2020年间计算机科学领域和无线电通信领域的论文数据对此方法开展实证研究。实证结果表明,本文提出的颠覆性技术多维度框架能够甄别出不同时期研究领域的代表性成果,达到预期的研究目标。我国在计算机技术和无线电通信方面颠覆性技术突破仍具较大的提升空间,需要加强关键领域科技研发,在稳步发展过程中确保国家科技安全。这个算法能够对科学文献领域的颠覆性技术进行识别,为科研人 员和相关情报从业人员以及国家管理者提供参考 与决策依据。

目前,面向颠覆性技术的方法主要针对科技论文开展分析,未来工作将进一步优化指标设计,考虑科技成果的时滞性,结合科技动态资讯、专利数据、产业信息等进行辅助分析与验证。

#### 参考文献

- [1] [1] BOWER J L, CHRISTENSEN C M. Disruptive technologies: catching the wave[J]. Harvard business review, 1995, 73(1):43–54.
- [2] Defense Advanced Research Projects Agency. Innovation at DAPRA[EB/OL]. [2023-04-10]. http://www.darpa.mil/attachments/DAPRA\_Innovation\_2016. pdf.
- [3] 中国科协创新战略研究院.日本颠覆性技术创新计划研究要点[J].中国高新科技,2017(1):93-95.
- [4] 习近平. 决胜全面建成小康社会夺取新时代中国特色社会主义伟大胜利[EB/OL]. [2023-04-10]. https://www.guancha.cn/politics/2017\_10\_27\_432557. html.
- [5] 石慧,潘云涛,苏成. 颠覆性技术及其识别预测方法 研究综述[J]. 情报工程, 2019, 5(3): 33-48.
- [6] 刘洪民,张世昌,李林遥,等.基于文献计量和颠覆性 潜力测度的颠覆性技术识别研究:以5G技术为例[J]. 科技和产业,2023,23(6):16-25.

- [7] 李晓龙,鲁平,李存斌.基于Delphi和DEMATEL法 影响国网的颠覆性创新技术影响因素综合排序分析 [J]. 科技管理研究, 2017, 37(6): 127-133.
- [8] 但智钢, 史菲菲, 王志增, 等. 中国环境工程科技 2035 技术预见研究 [J]. 中国工程科学, 2017, 19(1): 80-86.
- [9] KOSTOFF R N, BOYLAN R, SIMONS G R. Disruptive technology roadmaps[J]. Technological forecasting & social change, 2004, 71(1/2): 141–159.
- [10] UCHIHIRA N. Future direction and roadmap of concurrent system technology[J]. Ieice transactions on fundamentals of electronics, communications and computer, 2007, 90(11): 2443–2448.
- [11] 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(1): 16–29.
- [12] FITZGERALD B, SAYLER K, BRIMLEY S. Game changers: disruptive technology and U.S. defense strategy[EB/OL]. [2023-04-11]. https://www.files.ethz. ch/isn/170630/CNAS\_Gamechangers\_BrimleyFitzGeraldSayler\_0.pdf.
- [13] SCHOEMAKER P J, MAVADDAT V M. Scenario planning for disruptive technologies[J]. Wharton on managing emerging technologies, 2000(1): 206–241.
- [14] 石慧,潘云涛,赵筱媛,等.基于文献挖掘的颠覆性技术早期识别研究[J].中国科技资源导刊,2019,51(4):102-110.
- [15] 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(3): 16 –29.
- [16] DOTSIKA F, WATKINS A. Identifying potentially disruptive trends by means of keyword network analysis [J]. Technological forecasting and social change, 2017,

- 119(3): 114 -127.
- [17] 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.
- [18] 黄鲁成,成雨,吴菲菲,等.关于颠覆性技术识别框架的探索[J]. 科学学研究, 2015, 33(5): 654-664.
- [19] CHENG Y, HUANG L, RAMLOGAN R. Forecasting of potential impacts of disruptive technology in promising technological areas: elaborating the SIRS epidemic model in RFID technology[J]. Technological forecasting and social change, 2017, 117(4): 170–183.
- [20] ADNER R. When are technologies disruptive? a demand: based view of the emergence of competition [J]. Strategic management journal, 2002, 23(8): 667–688.
- [21] SOOD A, TELLIS G J. Demystifying disruption: a new model for understanding and predicting disruptive technologies [J]. Marketing science, 2011, 30(2): 339– 354.
- [22] KAPLAN S, VAKILI K. The double-edged sword of recombination in breakthrough innovation [J]. Strategic management journal, 2015, 36(10): 1435–1457.
- [23] 马荣康, 王艺棠. 知识组合多样性、新颖性与突破性 发明形成[J]. 科学学研究, 2020, 38(2): 313-322.
- [24] MARTION J. A review of selected recent advances in technological forecasting[J]. Technological forecasting and social change, 2003, 70(8): 719–733.
- [25] 褚晓泉,刘德兵,仇瑜,等.基于智能情报挖掘的技术 甄别发展现状及趋势预测研究[J].工程管理科技前 沿,2022,41(3):8.
- [26] PARK M, LEAHEY E, FUNK R. The decline of disruptive science and technology[J]. 2021. arXiv. 2106. 11184.
- [27] FUNK R J, OWEN-SMITH J. A dynamic network measure of technological change[J]. Management science, 2017, 63(3): 791–817.