

基于专利的颠覆性技术识别路径研究

卢文康

广州华立学院 广东广州 510000

摘要: 了解现有基于专利数据识别和预测颠覆性技术的研究现状和有效识别方法,为今后颠覆性技术识别的研究奠定基础。检索国内外相关文献,利用文献计量软件 HistCite 进行引用和关键词分析,在此基础上通过文献梳理总结颠覆性技术识别原理和方法,提出基于专利数据识别颠覆性技术的四条路径:基于时间序列的技术演化分析、挖掘专利文本实现聚类、技术特征的专利表述以及高颠覆性技术专利特征指标挖掘。

关键词: 专利; 颠覆性技术; 技术识别; 路径研究

引言

颠覆性技术关系到一个国家的竞争力和国际地位。通过科学方法准确识别和预测颠覆性技术的趋势,可以有效掌握技术发展的动态,调整国家科技战略布局,更好地在国际竞争中占据制高点。专利中蕴含着丰富可挖掘的技术信息,借助数据挖掘方法,能有效地表征技术的相关演化特征,可以从相对客观的角度较为准确地反映技术的颠覆性发展状况。本文通过文献调研并结合 HistCite 文献计量工具对基于专利的颠覆性技术识别研究论文进行文献计量和归纳总结,提出现有通过专利数据识别颠覆性技术的四条路径。

1. 检索式

为了较为全面地获取基于专利的颠覆性技术识别研究相关的国内外文献,本文以中国知网和 Web of Science 为数据源,中国知网的检索式为:(TKA = ‘颠覆性技术’ OR TKA = ‘颠覆性创新’ OR TKA = ‘颠覆式技术’ OR TKA = ‘颠覆式创新’) AND TKA = ‘专利’ 检索数据库为 CNKI,检索类型为 SCI 来源期刊、EI 来源期刊、核心期刊、CSSCI、CSCD,总共检索得到 81 篇文献,经过人工筛选去除无关文献后得到 68 篇相关文献。Web of science 数据库的检索式为:(TS=“disrupt* technolog*” or TS=“technolog* disrupt*” or TS=“disrupt* innovation*” or TS=“innovation disrupt*” or TS=“disrupt* potential” or TS=“potential disrupt*”) and TS=patent,检索数据库为“core collection”,

检索得到 45 篇文献,经过人工逐一判读,去除无关文献后最终得到 19 篇相关文献。

1.1. 关键词分析

运用 HistCite 软件对所有中英文文献进行关键词统计,人工合并相同或相近主题的关键词,形成表 1 关键词频次表如下:

表 1 关键词频次表

| 主题 | 关键词 | 合计频次 |
|---------|---|------|
| 技术识别 | 技术识别、技术预测、技术预见、早期识别、识别、预测研究、identification、technology forecasting、 | 29 |
| 专利数据可视化 | 专利地图、演化轨道、技术路线、visualization、technology roadmapping、development path、evolution pathways、main path | 9 |
| 引文网络 | 专利引用、引文分布、引文网络、citations、citation analysis、patent citations、patent citation network | 8 |
| 技术特征 | 新颖性、创新新、融合性、扩张性、扩散性、影响力 | 7 |
| 专利聚类 | 聚类分析、主题聚类、离群专利、突变检测、patent co-classification、outlier detection | 7 |
| 时序分析 | 时序分析、时间序列 | 2 |
| 回归分析 | quantile regression model、regression model | 2 |

频次最高的是“技术识别”主题,这也显示颠覆性技术相关研究中专利数据的主要作用就是在于如何运用专利数据开展识别和预测。其次是“专利数据可视化”主题,将专利数据转化为可视化图像能够更加直观的进行比较以及觉察趋势。出现频次第三位的是“引文网络”主题,引用和被引数据是研究专利各方面指标的重要数据。其他出现频次较多的还有“技术特征”、“专利聚类”、“时序分析”“回归分析”等主题。

1.2. 高被引参考文献分析

运用 HistCite 软件对检索到的文献的参考文献进行

统计, 被引频次最多的是 Funk R J 等在 2017 年的一项研究, 题为“A dynamic network measure of technological change”, 被引用次数达到 10 次。该研究提出一个颠覆性指数算法与引用了

例之差, 颠覆性指数的取值在 -1 到 1 之间。

2. 基于专利识别颠覆性技术的四条路径

根据以上引文分析和关键词分析的结果, 结合文献调

利识别颠覆性技术的路径如图 1

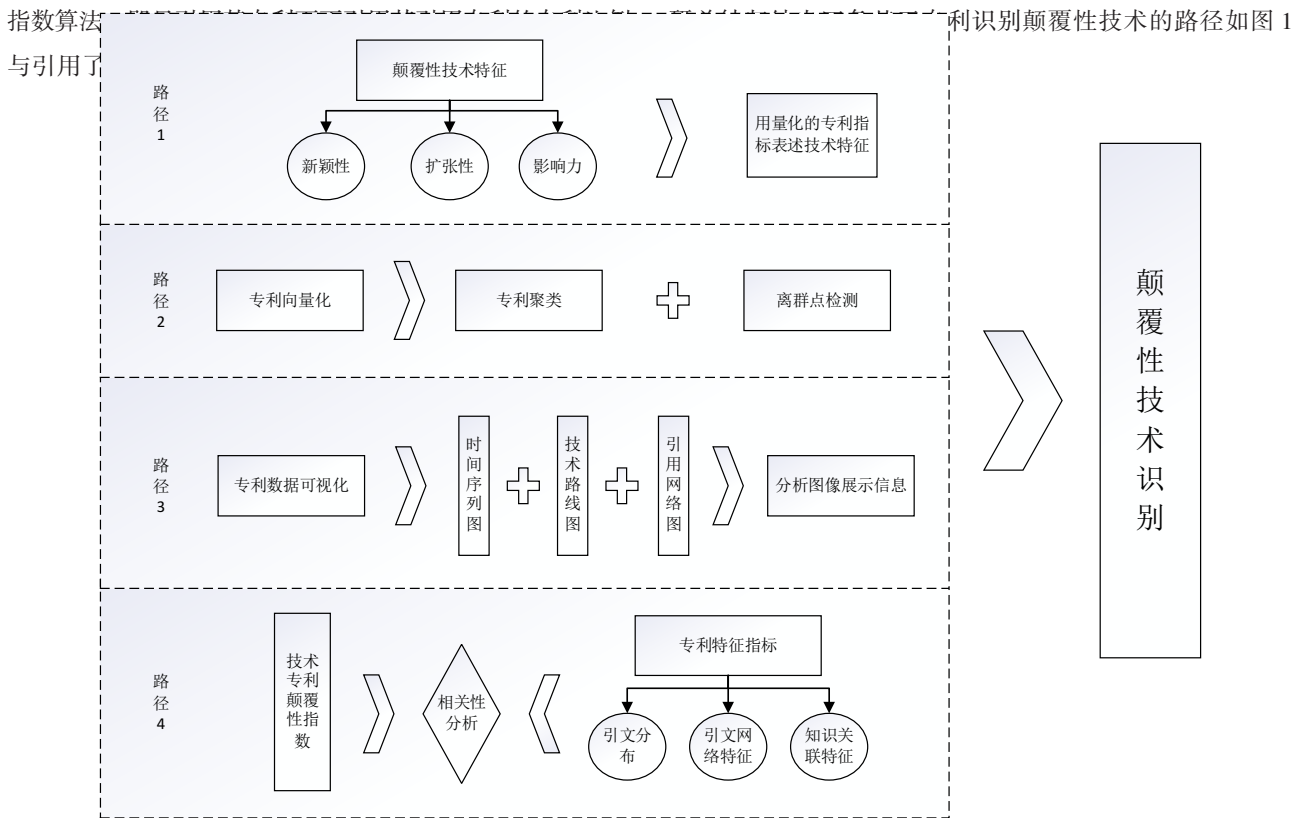


图 1 基于专利的颠覆性技术识别路径

路径 1: 技术特征的专利表述

相当部分学者是从颠覆性技术的技术特征为切入点, 基于专利数据, 围绕颠覆性技术的技术属性开展相关研究, 尝试量化专利各个维度的数据去表征技术属性。王海军等^[1]提出颠覆性技术的创新性、扩散性、转轨性、提升性及替代性等特征。综上, 颠覆性技术的技术特征可以归纳为三个方面: 新颖性、影响力、扩张性。

新颖性: 颠覆性技术往往是在创新活动中所产生的重大技术范式变革中形成的新技术, 或是现有技术不同组合, 颠覆了传统行业的运作方式。此外, 颠覆性技术还能够提供全新的功能、不连续的技术标准和新的所有制形式, 同时改变市场标准和消费者的期望, 为产品或服务带来全新的功能属性和技术特征。梁镇涛等^[2]将新颖性分解为三个维度: 技术知识量、技术知识影响力以及技术循环周期。分别用专利家族所参考的专利家族数量、专利家族所参考的专利家族的平均被引量以及专利家族所参考的专利家族的平均年龄

来表征。

影响力: Dosi G. 将突破性发明定义为对未来技术发展的影响, 能够成为未来许多发明基础的技术被认为是具有颠覆性的, 并通过对未来技术的影响实现技术的传递。于光辉等^[3]认为当颠覆性技术取得成功以后, 技术的影响力变高, 其后续的模仿者增多, 其与后续技术的专利引文结构就越相似。王丹等^[4]用主题词扩展度和专利被引次数两个指标进行刻画技术专利的影响力。主题词扩展度即某个主题词在经过一定时间之后的增长情况, 被引次数以 5 年的专利被引量作为评估标准。

扩张性: 技术领域内的知识扩张也称技术扩张、技术扩散。研究科学技术领域的知识扩散主要从引文下手, 专利可视为技术知识的代表, 潘一如等^[5]将专利知识扩散定义为专利申请人及审查人的引用行为驱动技术知识从源专利传递至施引专利的过程, 同时施引专利被下一代专利引用促使知识进一步传递, 如此往复, 源专利技术知识以辐射方式

向外围直接和间接施引的专利逐代扩散。王海军等^[1]用前向引用专利中 IPC 分类号的个数与研究对象本身 IPC 分类号个数的比值,用来表示研究对象的扩散应用程度。

路径 2: 挖掘专利文本实现聚类

这一路径通常用自然语言处理技术对专利进行文本挖掘,将专利向量化,根据向量化后的矩阵相似性对专利进行分类或识别出离群专利。在深度分类的技术发展网络中,颠覆性技术总是处于某一分类的根节点这一观点。由此使用 TF-IDF 技术和 SAO 结构提取技术特征,通过余弦相似性计算文本相似度,从而构建相似性矩阵,生成基于时间维度的技术发展网络,最后提出深度分类模型,根据技术发展网络的不同技术特征进行聚类,每个聚类的根节点即为颠覆性技术的起点。王丹等^[4]先利用 BERT 模型和 TF-IDF 算法将专利文本和专利分类号转化为可计算的向量,再通过离群点检测算法,以增量迭代的方式识别离群专利。

路径 3: 基于时间序列的技术演化分析

将专利数据转化成可视化图像,随着时间的变化显示出与传统技术不一样的技术轨迹或网络结构变化。国内最早的一篇通过此路径识别颠覆性技术的是苏敬勤等于 2016 年发表的论文,其根据专利量、专利引用量和专利引用率等作出随时间变化的智能手机与传统手机的 JAL 型曲线,将两者进行比较从而揭示颠覆性技术特殊的演化轨道和特征。李乾瑞等将专利各个维度的数据转化成了主题演化路径图和主题引用网络图,如果一个技术主题在主题演化路径图中有新主题的出现以及出现旧主题规模扩大、在主题网络图中引用关系随着时间不断增强,就认为该主题具有颠覆性。

路径 4: 高颠覆性技术专利特征指标挖掘

先用前人提出的较为有效的颠覆性指数计算方法计算专利颠覆性数值,再将专利的颠覆性数值与其他专利指标例如引文分布、引文网络特征等进行相关性分析,挖掘高颠覆性专利的特征指标。例如潘一如等^[5]等用 Wu 等提出的 DI 指标先确定了一批颠覆性技术专利,再从专利引文和专利耦合的角度选择两个对照组,对这三个组的引文分布指标和引文网络特征指标进行回归分析从而挖掘高颠覆性专利的特征。发现引文分布类型的首次引用年引用数、首次高峰间隔

年和首次高峰年引用数指标,引文网络特征类型的低引用代平均路径长度、平均聚类系数和连通性指标在识别高颠覆性专利方面具有重要意义。

结语

本研究通过文献计量的方法对基于专利的颠覆性技术识别领域文献进行了引用分析和关键词分析、运用文献调研的方法进行归纳总结,提出四条颠覆性技术识别路径:颠覆性技术特征专利表述、挖掘专利文本实现聚类、基于时间序列的技术演化分析以及高颠覆性技术专利特征指标挖掘。值得注意的是,专利数据只是在技术层面对颠覆性技术进行识别的依据,但颠覆性技术有时是在市场层面完成“颠覆”的,因此在今后的研究中还需要结合其他市场层面的数据源,例如微博、知乎等社交平台的数据进行多源数据分析。

参考文献:

- [1] 王海军. 基于专利和微博的颠覆性技术 ... 研究——以人工智能领域为例. 情报学报, 2023, 42(06): 649-662.
- [2] 梁镇涛, 毛进, 李纲. 融合“科学-技术”知识关联的高颠覆性专利预测方法. 情报学报, 2023, 42(06): 649-662.
- [3] 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. 科学学研究, 2021, 39(08): 1467-1473, 1536.
- [4] 王丹. 基于离群点视角的颠覆性专利预测研究. 情报学报, 2021, 40(08): 1467-1473, 1536.
- [5] 潘一如, 毛进, 李纲. 基于引文网络的高颠覆性专利知识扩散特征研究. 数据分析与知识发现, 2023, 7(10): 1-14.

作者简介:

卢文康(1994—),男,汉族,江苏扬州,助理馆员,硕士,专利分析、信息资源管理