融合技术主题互补性和合作网络相似性的企业技术合作伙伴预测研究*

关鹏[1]　王曰芬[2]　傅柱[3]
（1.巢湖学院经济与法学学院，合肥 238000；2.天津师范大学大数据科学研究院，天津 300387；
3.江苏科技大学经济管理学院，镇江 212003）

摘要：针对专利密集型产业技术合作决策中的信息不对称问题，构建企业技术合作伙伴预测模型，辅助企业技术合作决策。基于专利文献提取技术知识特征，本研究利用共同专利权人关系构建合作网络，通过Node2vec算法量化企业合作网络结构相似性，再应用作者-主题模型抽取技术主题向量，并结合JS散度计算技术主题互补性指标，构建“技术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型。光刻机领域的实证研究表明，相较于传统单一合作网络预测方法，融入技术主题互补性的预测方法性能更优，且可有效识别战略协同型、突破创新型、资源整合型和机会探索型4种类型潜在合作伙伴。研究方法可为专利密集型行业提供数据驱动的合作伙伴推荐工具，帮助企业快速识别潜在合作伙伴。

关键词：技术主题互补性；合作网络相似性；技术合作；预测；专利
中图分类号：G315 文献标识码：A　DOI：10.3772/j.issn.1673-2286.2026.02.004
引文格式：关鹏，王曰芬，傅柱. 融合技术主题互补性和合作网络相似性的企业技术合作伙伴预测研究[J]. 数字图书馆论坛，2026，22（2）：31-42.

随着西方技术封锁升级，我国面临严峻的“卡脖子”困境，亟需在一些关键核心技术领域实现创新突破。《中共中央关于进一步全面深化改革 推进中国式现代化的决定》强调要“强化企业科技创新主体地位，建立培育壮大科技领军企业机制，加强企业主导的产学研深度融合，建立企业研发准备金制度，支持企业主动牵头或参与国家科技攻关任务”[1]。而企业作为创新主体，“其技术创新能力还存在多重短板，与大学/科研院所的技术衔接及合作创新需求较为迫切，有必要促进产学研创新的集群式发展，产出更多有助于企业发展所需的科学技术成果”[2]。现有关于中小企业的研究认为，研究组织通过提供有效的人才、知识和技术支持能够促进中小企业技术创新绩效的提升[3]。

随着技术复杂度越来越高，产业链与创新链的深度融合成为突破关键核心技术的关键。一方面，“开放式”创新背景下，企业重点关注如何依据自身知识基础、技术能力和资源条件寻找合适的合作伙伴，以获取异质性知识和互补性资源，实现关键技术突破；另一方面，从知识服务角度，如何利用海量科技文献资源，帮助企业厘清技术领域协作分布的状态及趋势，支撑企业发现潜在合作伙伴，成为重要的研究课题。为此，本研究利用专利权人合作关系和专利文本信息，构建“技

收稿日期：2025-12-06
*本研究得到国家社会科学基金项目“面向AI4S的场景化智慧知识服务框架研究”（编号：24CTQ029）、安徽省哲学社会科学基金项目“‘科学-技术-产业’关联视角下新兴技术预测研究”（编号：AHSKY2025D58）、天津市哲学社会科学规划重点项目“面向创新链产业链深度融合的京津冀专利技术转移影响机制及提升路径研究”（编号：TJTQ24-001）资助。

术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型，为企业技术合作提供定性与定量相结合的决策工具。

1 相关研究

随着科学研究的复杂性不断提升和跨领域研究需求的持续增长，利用海量科技文献资源快速找到与自身兴趣契合的科研合作伙伴，日益成为提升学者科研效率和创新能力的关键，也成为图书情报领域一个越来越值得重点关注的课题[4]。

1.1 技术合作预测研究

当前开展技术合作预测主要利用专利文献，通过专利合作网络和专利文本挖掘方法识别潜在的合作伙伴。以创新主体专利合作关系（合作发明人关系、共同专利权人关系等）为基础，构建专利合作网络，在合作网络中利用网络拓扑结构或网络属性特征等挖掘创新主体之间有较高合作潜力的合作对象[5]。这类方法主要利用合作网络提取创新主体的网络结构属性，通过链路预测达到合作伙伴预测目的。但随着研究深入，学者发现仅通过合作网络结构属性无法准确地预测合作伙伴，可将创新主体自身的特有属性融入链路预测中，以此提升合作伙伴预测的准确性。如韩菁等[6]在基于专利数据集的合作网络和知识网络及其“知识-合作”从属网络基础上，构建多层网络链路预测模型；谢小东等[7]利用专利发明人之间的直接信任关系、间接信任关系、领域偏好关系构建不同类型的专利发明人合作网络，并将发明人专利摘要文本信息作为网络节点属性，利用异构图卷积网络识别专利发明人潜在合作伙伴。

亦有学者以创新主体的专利文本信息为基础，利用文本挖掘技术抽取创新主体的技术属性，通过相似度计算或聚类方法识别潜在合作伙伴的技术特征，以评估企业与潜在合作伙伴的技术兼容性，以此寻找合作伙伴[8]。例如，吕源等[9]基于专利文本相似度计算预测企业的产学研潜在合作伙伴。

1.2 技术互补性相关研究

国际专利分类（International Patent Classification，IPC）对专利涉及的技术进行分类，在海量专利文献检索、分析和挖掘中起到重要作用。在情报学领域，技术互补性指两个机构在共有的广义技术领域内关注不同狭义技术领域的程度，通常使用专利IPC分类号计算机构间技术互补的程度[10]。Jaffe[11]提出可利用专利IPC分类号构建创新主体技术结构向量，计算创新主体间的技术相似度。Frenken等[12]提出利用技术大类和技术小类的分布分别测算创新主体之间的不相关技术与相关技术的多样性。郑江淮等[13]提出两个地区在同一大类技术中的不同小类技术有很多部分会组合在某一类产品上，称之为地区间技术互补；反之，两个地区在同一细分技术中形成大量专利，称之为地区间技术竞争；技术互补可以视作从两个地区技术大类相似度中减去同一技术小类发生的相似关系。

除应用IPC分类号外，Chang等[14]使用专利引用网络发现了技术互补性在企业专利收购战略中的重要性；Wang等[15]利用专利文本内容挖掘技术主题并测算不同专利权人之间的技术互补性。部分学者使用技术互补性开展专利合作伙伴预测研究，如使用专利的IPC分类信息识别具有互补性技术的专利权人作为潜在合作伙伴[16]；利用专利文本信息抽取SAO语义结构构建技术分布图，识别技术热点和优势企业，以此寻找具有特定技术的潜在合作伙伴[17]。关键技术突破往往依赖于引入互补技术，具有技术互补性的企业或组织通过开展技术合作，有利于吸收对方优势技术弥补自身不足，提高技术创新效率、降低创新成本和风险[18]。很多学者从企业间技术互补[19]、区域间技术互补[20]的角度研究了技术互补对企业经济效益的影响，研究表明技术互补对战略性新兴产业全要素生产率提升具有直接效应，且对缩小企业间、地区间差距进而推动地区均衡发展具有促进作用。

综上，当前技术合作预测研究主要利用专利合作网络链路预测方法寻找网络结构相似或技术相似的合作伙伴，研究结果存在的偏向性可能导致预测系统更易推荐技术同质化伙伴，错过突破性创新所需的技术互补型合作机会。少量研究采用技术互补性识别具有互补性技术的潜在合作伙伴，但未考虑协同创新网络这一重要因素。基于此，本研究融合专利合作网络和专利文本信息，开展企业技术合作伙伴预测研究，构建可同时量化合作网络结构相似性与技术主题互补性的双维度潜在合作伙伴预测模型，以预测和识别支撑企业技术创新所需的不同类型的合作伙伴。基于专利文

本的主题挖掘可以从语义内容层面发现和分析技术主题，是技术预测的重要方法。借鉴现有技术互补性测算方法，本研究提出技术主题互补性计算方法。需要说明的是，本研究提出的技术主题互补性概念与技术互补性概念不同，侧重于技术主题在语义表达上的差异性，而非技术分类上的不同。

2　研究设计

专利作为技术信息的载体，是技术知识发现过程的累积形态，广泛应用于技术主题识别。因此，本研究使用专利文献表征技术知识，综合利用网络表示学习、主题建模等方法，预测企业技术合作伙伴，研究框架如图1所示。

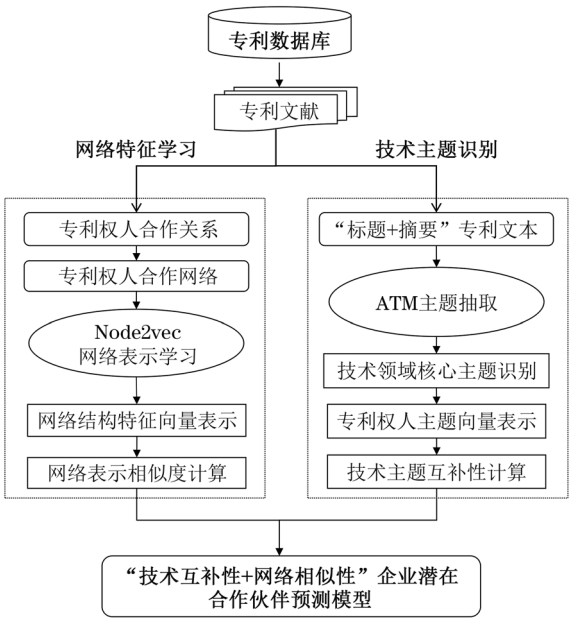

图1　研究框架

整体研究框架包括3个部分。首先，进行专利合作网络构建及网络表示学习。针对特定的关键技术领域从专利数据库中下载授权专利文献题录信息，对专利文献进行去重、“标题+摘要”分词、技术术语抽取等数据预处理，利用共同专利权人合作关系构建专利人合作网络，利用Node2vec网络表示学习方法学习专利权人的合作网络结构属性，将专利权人表示为网络结构属性向量，并计算专利权人之间的网络相似性。其次，构建技术领域作者-主题模型（Author-Topic Model，ATM），挖掘专利权人的专利技术主题，将专利权人表示为主题向量，并计算专利权人之间的技术主题互补性。最后，以技术主题互补性和合作网络相似性为决策指标，利用二维四象限分析法，构建“技术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型，识别战略协同型、突破创新型、资源整合型、机会探索型4种类型的潜在合作伙伴。

2.1　合作网络构建及网络表示学习

以专利权人为节点，以共同专利权人为边，构建专利权人合作网络$G=(V, E)$，其中V是节点集合，由企业、高校和研究所等机构构成，不包括个人作为专利权人，E为网络边集合，以机构之间的合作次数作为边的权重。

Node2Vec是对DeepWalk算法的改进，通过有偏随机游走提取网络信息，将图数据（通常情况下为高维稀疏矩阵）映射为低维稠密向量，对图中的每个节点产生一个嵌入表示向量作为输出。Node2Vec算法为了捕捉结构和功能相似的网络节点，在DeepWalk完全随机游走的基础上，设计了广度优先采样（Breadth-first Sampling，BFS）和深度优先采样（Depth-first Sampling，DFS）两种搜索策略，并通过调整p和q参数的数值使搜索策略倾向于BFS或者DFS，实现有偏随机游走。本研究设置p=1.0，q=0.5，偏向远离起始节点的深度探索（DFS特性），从而能更好地反映节点在网络中的结构性角色。即使这些节点不属于同一个社区，只要它们在网络中扮演相似的结构性角色（如枢纽节点或桥梁节点），其嵌入表示也会相近。这种策略能帮助模型识别出图中具有相似结构的节点，兼顾局部和全局信息，有助于提高节点嵌入的质量，使嵌入结果不仅反映节点的邻近关系，还能体现其在网络中的更广泛影响。

假设网络G中节点A和节点B经Node2Vec网络表示学习后分别表示为向量$V_A=(A_1,A_2,\cdots,A_M)$和$V_B=(B_1,B_2,\cdots,B_M)$，向量维度M由Node2Vec算法参数设置决定。则机构A和B的合作网络相似性S_{AB}可由向量的余弦相似度计算得到。

2.2　技术主题抽取和技术主题互补性计算

ATM是一种从大规模文本数据集中抽取文本主题的非监督机器学习算法。Rosen-Zvi等[21]在LDA（Latent Dirichlet Allocation）主题模型的基础上加

入作者层，构建ATM。本研究将机构（专利权人）视为专利文献的作者，利用ATM抽取机构-技术主题概率分布。

通过ATM获得机构-技术主题概率分布矩阵（见表1），假设技术领域共有机构m个，抽取的技术主题共n个，则机构i可以表示为主题向量$P_i=(p_{i1},p_{i2},\cdots,p_{in})$，其中$\sum_{j=1}^{n}p_{ij}=1$，$p_{ij}$是第$j$个主题在机构$i$的主题向量中的概率。

表1　机构-技术主题概率分布矩阵

机构编号	Topic 1	Topic 2	Topic 3	…	Topic n
机构1	p_{11}	p_{12}	p_{13}	…	p_{1n}
机构2	p_{21}	p_{22}	p_{23}	…	p_{2n}
机构3	p_{31}	p_{32}	p_{33}	…	p_{3n}
…	…	…	…	…	…
机构m	p_{m1}	p_{m2}	p_{m3}	…	p_{mn}

利用两个机构的主题向量计算机构之间技术主题互补性指标，互补性反映两个机构在技术主题分布上的差异程度，侧重于主题语义内容之间的差异，分布差异越大，互补性越高。使用JS散度（Jensen-Shannon Divergence）度量概率分布的差异，技术主题互补性指标（Technical Topic Complementarity，TTC）计算方法如公式（1）所示。

$$TTC(A,B)=\frac{1}{2}\sum_{k=1}^{n}\left[p_{Ak}\,\mathrm{lb}\left(\frac{2p_{Ak}}{p_{Ak}+p_{Bk}}\right)+p_{Bk}\,\mathrm{lb}\left(\frac{2p_{Bk}}{p_{Ak}+p_{Bk}}\right)\right] \quad (1)$$

式中：机构A和机构B的技术主题概率分布分别为$P_A=(p_{A1},p_{A2},\cdots,p_{An})$、$P_B=(p_{B1},p_{B2},\cdots,p_{Bn})$，其中，$p_{Ak}$为机构$A$的第$k$个主题在主题分布$P_A$中的概率值，$p_{Bk}$为机构$B$的第$k$个主题在主题分布$P_B$中的概率值，$n$为技术领域主题总数，$k$为主题序号。

当$P_A=P_B$时，机构A和机构B的技术主题分布完全相同，无互补性，互补性指标取值为0；当P_A与P_B互斥时，机构A和机构B的技术主题分布完全不重叠，互补性指标取值为1。

2.3　企业潜在合作伙伴预测模型

以技术主题互补性和合作网络相似性为决策指标，利用二维四象限分析法，构建“技术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型（见图2）。根据技术主题互补性指标和合作网络相似性指标值的大小，将潜在合作关系分为战略协同型、突破创新型、资源整合型、机会探索型4种类型，企业可根据实际情况进行合作策略选择。

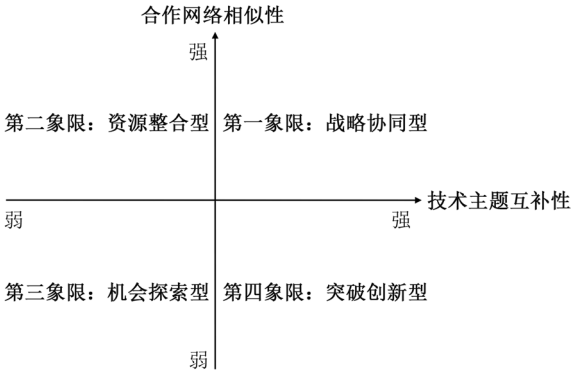

图2　“技术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型

（1）战略协同型。在技术层面呈现出“强互补+强合作”的双重特性，机构间既有技术缺口填补潜力，又处于同一网络社区或存在高度相似的合作网络结构特征。强网络相似性意味着双方已嵌入相似的合作生态（如共同合作伙伴或相似的网络结构角色），而强主题互补性（专利技术主题差异显著）能形成“技术拼图效应”。已有研究[22]表明，结构对等（网络相似）且技术互补的盟友更易实现联合专利产出最大化。机构通过合作获取互补性资源，相似的合作网络结构可降低协调成本。

（2）资源整合型。在技术层面呈现出“弱互补+强合作”的双重特性，机构间技术主题重叠度高且处于同一网络社区或存在高度相似的合作网络结构特征。根据知识吸收与创新绩效的相关研究，高网络相似性（如同产业联盟）可以强化组织间信任，具有相似技术背景可以增强组织知识吸收能力[23]，适用于渐进式创新，即使技术同质化，也可通过规模效应降低研发成本。但过度依赖现有网络可能导致技术路径固化，网络中的企业易形成“技术孤岛”。

（3）机会探索型。在技术层面呈现出“弱互补+弱合作”的双重特性，机构间既没有技术缺口填补潜力也缺乏合作基础（在网络中表现为弱连接），在市场上处于相互竞争的关系。从弱连接优势理论分析，弱连接（跨网络连接）带来非冗余信息流[24]，加之技术相似性较高，可能开辟边缘创新机会，存在潜在技术组合创新机会。

（4）突破创新型。在技术层面呈现出“强互补+弱

合作”的双重特性，机构间有技术缺口填补潜力但缺乏合作基础。根据创新网络结构对颠覆性创新的影响相关研究[25]，低网络相似性表明双方处于不同知识社群，而技术主题互补性提供了知识重组机会，若机构处于创新网络结构洞位置，则对颠覆性创新具有显著正向影响，但合作成本较高，需建立机构间跨网络信任机制。

3　实证分析

本研究选取光刻机领域作为实证分析案例，验证“技术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型的有效性。在集成电路制造过程中，光刻是核心、复杂的工艺，而光刻工艺所需的光刻机是集成电路制造中最为核心的高端装备[26]。光刻机通过光学曝光方式，将光掩膜版上的集成电路图形转移至硅衬底表面的感光光刻胶层，实现精密图形转印[27]。

3.1　数据采集与处理

本研究的专利数据来源于国家知识产权局专利检索及分析数据库，检索时间为2025年3月16日，以“光刻机”为关键词进行检索，获取相关发明专利和实用新型专利共5 602件，其中发明专利4 679件，专利申请时间为1985—2024年，其中合作专利（联合申请专利）364件。进一步统计该时间范围内授权专利的年度分布情况，其授权专利数量走势如图3所示。我国第一件光刻机领域的发明专利是中国科学院上海光学精密机械研究所于1985年申请的扫描式投影光刻机物镜（专利号：CN85104369A）。从图3可以看出，2018年以来，国内光刻机领域专利申请量呈现快速增长态势。原因在于部分海外经济体针对我国半导体产业实施技术与设备出口管制。为聚焦产业与科研主体，下文统一以机构专利权人为分析对象，不再纳入个人专利权人。剔除个人专利权人后，共有机构专利权人1 508家：企业1 179家、高校168家、科研院所（研究院、研究所和实验室）161家。

其中，1985—2009年专利数量为448件（占总量的8%），合作专利数量（联合申请专利数量）仅为12次（占合作专利总量的3.3%），大多数专利合作发生于2010年以后。同时考虑到专利合作规模不大，本研究对专利合作网络不进行分阶段动态分析，采用静态快照方式进行分析，以确保对专利合作网络结构属性度量的有效性。

3.2　专利合作网络表示学习

以共同专利权人为关系构建光刻机领域专利合作网络结构图（见图4），该网络包含290个节点、198条连边，网络密度为0.005，网络平均聚类系数为0.141，最大联通子集的平均最短路径长度2.626。从合作网络的结构属性看，光刻机领域专利合作网络稀疏且规模不大，机构之间合作不够紧密。

使用基于Python语言的Node2Vec工具包实现光刻机领域专利合作网络表示学习，参数设置为：嵌入维度64，随机游走长度30，节点游走次数为200次，

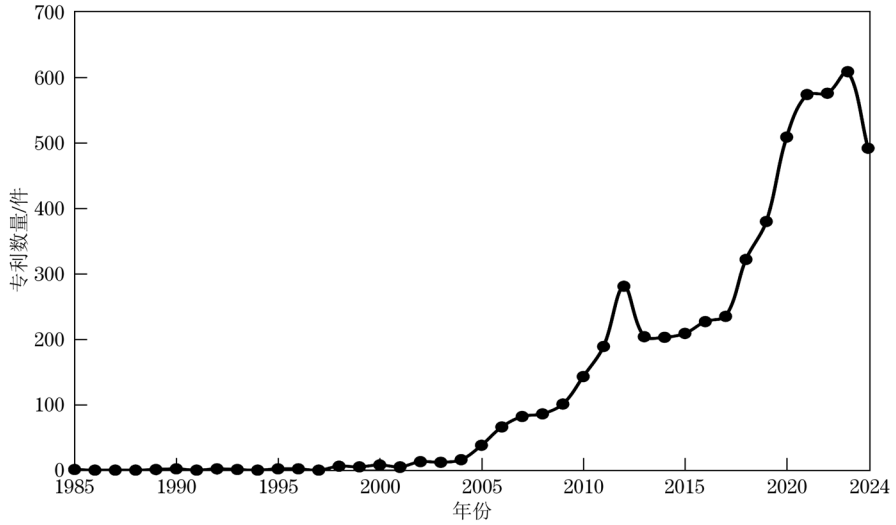

图3　1985—2024年授权专利数量走势图

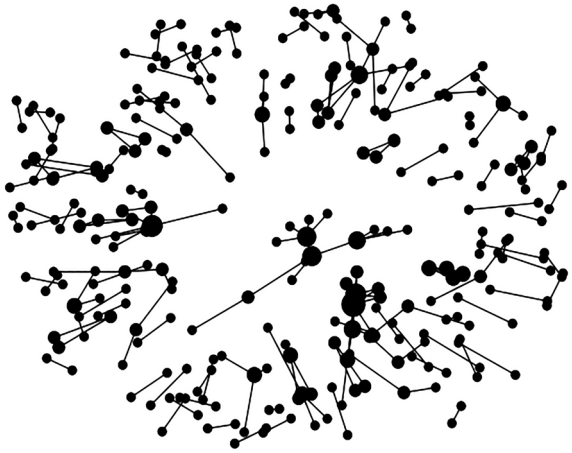

图4　光刻机领域专利合作网络结构图

p=1.0，q=0.5。Node2Vec生成的嵌入向量将被输入到t-SNE中，其中t-SNE将64维向量表达进行降维。经过Node2Vec生成的嵌入向量如表2所示，各机构均以64维向量形式表示。

通过计算各机构合作网络表示向量的余弦相似度来衡量机构间网络相似度。该指标取值范围为0~1，数值越接近1代表网络结构相似度越高，机构自身的网络相似度为1，构建机构间合作网络相似度矩阵部分结果如表3所示。例如，北京华卓精科科技股份有限公司（机构2）与清华大学（机构1）的网络相似度为0.99，与无锡华润上华半导体有限公司（机构3）的网络相似度为0.43。

表2　机构合作网络学习表示向量（部分）

机构编号	机构名称	维度1	维度2	维度3	…	维度64
机构1	清华大学	0.07	0.23	−0.10	…	0.06
机构2	北京华卓精科科技股份有限公司	0.09	0.26	0.07	…	0.11
机构3	无锡华润上华半导体有限公司	0.66	−0.16	0.55	…	0.10
机构4	上海微电子装备（集团）股份有限公司	0.62	−0.11	0.03	…	−0.50
机构5	上海微高精密机械工程有限公司	−0.29	−0.04	−0.13	…	−0.21
机构6	中国科学院微电子研究所	−0.34	−0.03	0.48	…	−0.79
机构7	真芯（北京）半导体有限责任公司	−0.21	−0.01	0.34	…	−0.91
机构8	北大方正集团有限公司	0.85	0.07	−0.54	…	0.38
机构9	中芯国际集成电路制造有限公司	−0.05	−0.56	0.13	…	−0.23
机构10	湖南大学	0.79	−0.03	−0.73	…	−0.80

表3　机构间合作网络相似度矩阵（部分）

机构编号	机构1	机构2	机构3	机构4	机构5	机构6	机构7	机构8	机构9	机构10
机构1	1.00	0.99	0.42	0.32	0.32	0.40	0.41	0.40	0.39	0.36
机构2	0.99	1.00	0.43	0.33	0.33	0.41	0.41	0.38	0.39	0.37
机构3	0.42	0.43	1.00	0.35	0.32	0.38	0.43	0.39	0.35	0.33
机构4	0.32	0.33	0.35	1.00	0.98	0.36	0.36	0.37	0.34	0.38
机构5	0.32	0.33	0.32	0.98	1.00	0.36	0.36	0.37	0.35	0.38
机构6	0.40	0.41	0.38	0.36	0.36	1.00	0.98	0.30	0.42	0.35
机构7	0.41	0.41	0.43	0.36	0.36	0.98	1.00	0.31	0.41	0.36
机构8	0.40	0.38	0.39	0.37	0.37	0.30	0.31	1.00	0.42	0.41
机构9	0.39	0.39	0.35	0.34	0.35	0.42	0.41	0.42	1.00	0.41
机构10	0.36	0.37	0.33	0.38	0.38	0.35	0.36	0.41	0.41	1.00

3.3　技术领域主题抽取与技术主题互补性计算

使用基于Python语言的gensim工具包构建ATM。使用主题一致性指标确定最佳主题数量。主题一致性指标用于衡量主题内部词语的语义关联强度，综合评估主题抽取质量，其核心思想是高质量的主题应包含在文档中频繁共现且语义相似的词语[28]。光刻机领域专利文献ATM主题一致性得分曲线如图5所示，当主题数量为5和11时主题一致性指标得分较高，经咨询领域专家，结合光刻机领域主要研究方向，确定主题数为11。通过分析各主题下相关度Top 30关键词并结

合领域专家建议，本研究将国内光刻机技术领域专利文献技术主题划分为：光罩设备（Topic 0）、电极制备（Topic 1）、传感器（Topic 2）、曝光系统（Topic 3）、光刻技术（Topic 4）、光刻胶工艺（Topic 5）、实用新型光刻机装置（Topic 6）、物镜系统（Topic 7）、工件台与掩模台系统（Topic 8）、光源系统（Topic 9）、检测设备（Topic 10），光刻机领域各技术主题核心关键词（Top 5）如表4所示。

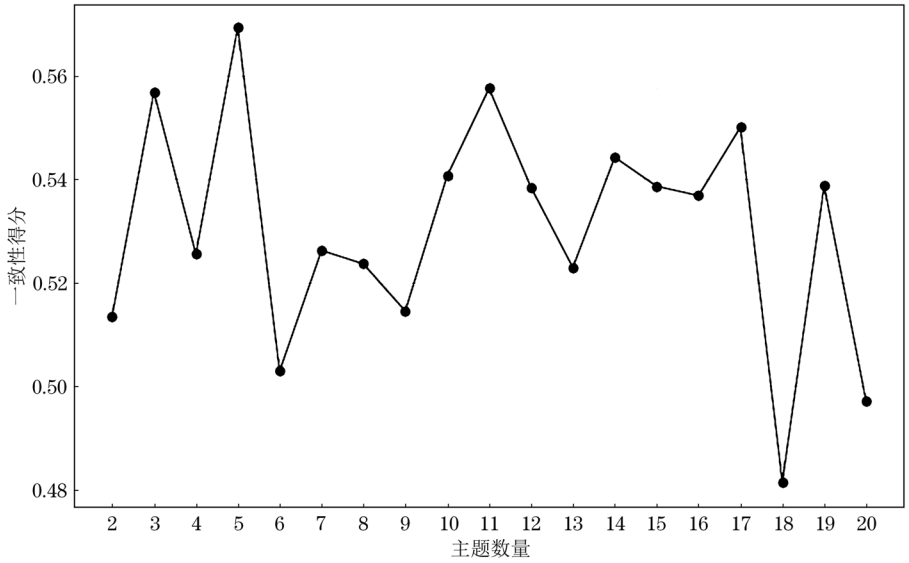

图5　光刻机领域专利文献ATM主题一致性得分曲线

表4　光刻机领域各技术主题核心关键词（Top 5）

排名	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
1	光罩	电极	传感器	曝光	结构	光刻胶	连接	透镜	硅片	单元	模块
2	结构	制备	波导	图形	制备	工艺	安装	物镜	运动	光源	检测
3	单元	薄膜	光纤	对准	纳米	光刻	机构	装置	工件	系统	控制
4	凹槽	器件	结构	标记	芯片	芯片	设置	投影	驱动	光栅	数据
5	表面	金属	输出	光刻	技术	形成	装置	浸没	装置	照明	系统

利用ATM抽取的文档-主题概率分布矩阵可计算主题强度。在抽取的11个主题中，电极制备、曝光系统、光刻胶工艺和实用新型光刻机装置主题强度值分别为0.11、0.14、0.15和0.15，其他主题强度值均小于0.10。因此，以上4个主题是国内光刻机领域的热点研究主题。

国内光刻机领域机构-主题概率分布矩阵（部分）如表5所示。其中，清华大学（机构1）在工件台与掩模台系统、光源系统重点布局；上海微电子装备（集团）股份有限公司（机构4）在曝光系统、工件台与掩模台系统、光源系统、实用新型光刻机装置和物镜系统这5个主题上的概率值较高，表明公司重视在这5个技术主题上的创新研发投入。

利用机构-主题概率分布结果，构建机构间技术主题互补性矩阵。技术主题互补性指标取值范围为0~1，其中机构自身的技术互补性数值为0，机构间技术主题互补性矩阵（部分）如表6所示。例如，北京华卓精科科技股份有限公司与清华大学之间的技术主题互补性为0.08，与无锡华润上华半导体有限公司之间的技术主题互补性为0.53。

3.4　企业技术合作伙伴预测方法评价

为衡量本研究所提融合技术主题互补性与合作网络相似性的企业技术合作伙伴预测模型的综合性能，选择正确率、精确度、召回率和F1值作为评价指标。本研究遵循预测算法通用实验规则，采用8:2比例划分训练集与测试集，根据专利申请年份，以2022年为时间节点将数据集划分为两部分，使用1985—2022年的数据作为训练集，2023—2024年的数据作为测试集，数据集描述见表7。

表5　光刻机领域机构-主题概率分布矩阵（部分）

机构编号	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
机构1	0.01	0.04	0.05	0.05	0.05	0.05	0.05	0.03	0.42	0.20	0.04
机构2	0.01	0.01	0.04	0.06	0.03	0.01	0.20	0.06	0.28	0.23	0.07
机构3	0.01	0.03	0.00	0.51	0.03	0.28	0.01	0.02	0.04	0.02	0.04
机构4	0.02	0.01	0.04	0.21	0.02	0.02	0.13	0.12	0.17	0.16	0.09
机构5	0.01	0.01	0.06	0.10	0.02	0.02	0.16	0.04	0.33	0.12	0.13
机构6	0.08	0.18	0.03	0.18	0.04	0.22	0.08	0.03	0.03	0.05	0.07
机构7	0.01	0.04	0.02	0.33	0.05	0.15	0.11	0.04	0.05	0.09	0.10
机构8	0.01	0.05	0.01	0.34	0.00	0.33	0.01	0.03	0.05	0.01	0.14
机构9	0.06	0.03	0.04	0.26	0.02	0.21	0.10	0.03	0.02	0.09	0.13
机构10	0.02	0.40	0.05	0.07	0.15	0.24	0.02	0.02	0.02	0.00	0.00

表6　机构间技术主题互补性矩阵（部分）

机构编号	机构1	机构2	机构3	机构4	机构5	机构6	机构7	机构8	机构9	机构10
机构1	0.00	0.08	0.46	0.14	0.08	0.31	0.26	0.43	0.31	0.47
机构2	0.08	0.00	0.53	0.07	0.03	0.35	0.24	0.50	0.29	0.61
机构3	0.46	0.53	0.00	0.33	0.45	0.17	0.11	0.05	0.13	0.35
机构4	0.14	0.07	0.33	0.00	0.06	0.24	0.11	0.31	0.16	0.55
机构5	0.08	0.03	0.45	0.06	0.00	0.30	0.19	0.39	0.24	0.59
机构6	0.31	0.35	0.17	0.24	0.30	0.00	0.09	0.13	0.06	0.15
机构7	0.26	0.24	0.11	0.11	0.19	0.09	0.00	0.12	0.03	0.34
机构8	0.43	0.50	0.05	0.31	0.39	0.13	0.12	0.00	0.10	0.34
机构9	0.31	0.29	0.13	0.16	0.24	0.06	0.03	0.10	0.00	0.34
机构10	0.47	0.61	0.35	0.55	0.59	0.15	0.34	0.34	0.34	0.00

表7　数据集描述

单位：个

项目	训练集	测试集
专利数	4 501	1 101
合作专利数	302	70
机构数	1 200	587
合作机构数	224	94

为测试融合技术主题互补性与合作网络相似性的企业技术合作伙伴预测方法（方法1）的有效性，本研究使用基于合作网络相似性的预测方法（方法2）作比较对象，两种方法性能对比结果如表8所示。

表8　两种方法性能对比结果

方法	正确率	精确度	召回率	F1值
方法1	0.94	0.54	0.86	0.66
方法2	0.88	0.33	0.71	0.45

结果表明，方法1的正确率显著高于方法2。这表明相较于使用合作网络相似性的单一模型，融入技术主题互补性的模型能更准确地区分整体样本中的正负样本，有效降低误判率，从而提升模型分类稳健性。从精确度与召回率两项指标来看，融合技术主题互补性与合作网络相似性的模型，其预测正例中的实际正例占比更高，可有效降低假正例干扰；同时能够覆盖更多真实正例，减少假负例遗漏问题。F1值作为精确度与召回率的调和平均，其优势凸显了融合技术主题互补性与合作网络相似性的预测方法的平衡性——既能精准筛选正例，又能覆盖多数相关样本。因此，本研究提出的融合技术主题互补性与合作网络相似性的预测方法在预测性能上优于传统基于合作网络相似性的预测方法，具有可行性和有效性。

3.5　企业潜在技术合作伙伴预测模型构建

综合机构之间的技术主题互补性与合作网络相似性计算结果，构建企业潜在技术合作伙伴二维四象限决策模型。在指标阈值的选择方面，中位数是一个简单有效的阈值选择，尤其适合非正态分布或含异常值的

数据。本研究计算所得合作网络相似性和技术主题互补性均不符合正态分布。选择合作网络相似性和技术主题互补性的中位数作为阈值，旨在捕捉技术主题互补性与合作网络相似性的协同效应临界点，避免因过度依赖统计分布而低估潜在合作价值。对于机构A，将其与其他机构之间的技术主题互补性和合作网络相似性结果，分别构成一维向量TTC_A和S_A，则选择TTC_A和S_A的中位数α、β分别作为技术主题互补性阈值和合作网络相似度阈值。据此，将企业的潜在技术合作对象划分为4种类型，分别为战略协同型（技术主题互补性$>\alpha$，合作网络相似度$>\beta$）、资源整合型（技术主题互补性$\leqslant\alpha$，合作网络相似度$>\beta$）、机会探索型（技术主题互补性$\leqslant\alpha$，合作网络相似度$\leqslant\beta$）、突破创新型（专利主题互补性$>\alpha$，合作网络相似度$\leqslant\beta$）。

本研究以北京华卓精科科技股份有限公司为例，绘制企业潜在技术合作伙伴二维四象限决策模型，如图6所示。经计算，北京华卓精科科技股份有限公司与其他机构的专利主题互补性和合作网络相似性的中位数分别为0.55和0.41，故专利主题互补性和合作网络相似性的阈值分别设置为α=0.55和β=0.41。结合表5可以得出，北京华卓精科科技股份有限公司的专利主要集中于实用新型光刻机装置、工件台与掩模台系统和光源系统3个主题，与清华大学和北京优微精密测控技术研究有限公司具有合作关系。综合技术主题互补性和合作网络相似性双重维度，从图6可知，北京华卓精科科技股份有限公司的潜在技术合作伙伴可划分为4个类型。

（1）战略协同型合作伙伴。中国电子科技集团公司第十三研究所（机构1491）、乌镇实验室（机构1190）、同方威视技术股份有限公司（机构1164）、吉林大学（机构469）、广东中科半导体微纳制造技术研究院（机构288）、北京科华微电子材料有限公司（机构98）、东旭光电科技股份有限公司（机构218）与北京华卓精科科技股份有限公司（机构2）的网络相似度分别为0.99、0.99、0.94、0.64、0.63、0.62、0.62，技术主题互补性分别为0.82、0.49、0.57、0.69、0.79、0.87、0.45。从专利合作网络角度看，中国电子科技集团公司第十三研究所、同方威视技术股份有限公司、乌镇实验室与北京华卓精科科技股份有限公司具有共同的专利合作伙伴清华大学，处于局部紧密的合作网络社区内。同时，吉林大学、广东中科半导体微纳制造技术研究院、北京科华微电子材料有限公司与北京华卓精科科技股份有限公司具有相似的网络连接模式，因此具有较高的网络相似度。从技术主题角度看，中国电子科技集团公司第十三研究所的技术主题包括电极制备、光刻胶工艺，同方威视技术股份有限公司的技术主题分布包括

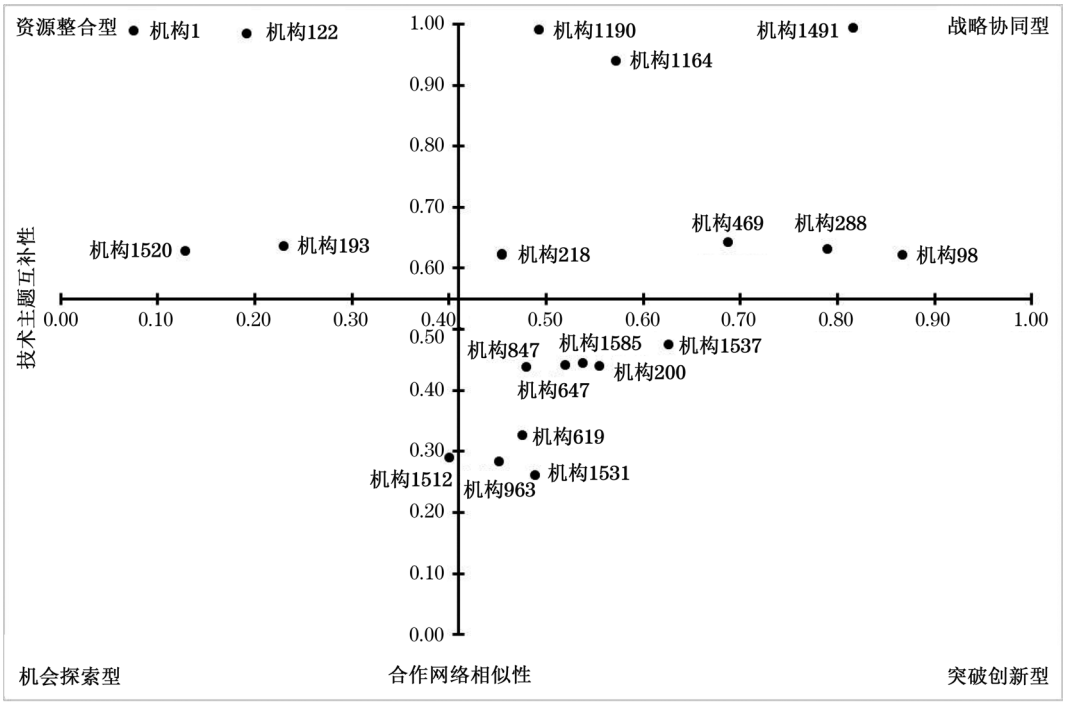

图6　北京华卓精科科技股份有限公司潜在技术合作伙伴预测模型

光罩设备、曝光系统、光刻胶工艺、实用新型光刻机装置、物镜系统，乌镇实验室的技术主题包括光罩设备、传感器、光刻技术，3家机构与北京华卓精科科技股份有限公司的技术主题重叠度较低，因此具有较高的技术主题互补性。综上分析，北京华卓精科科技股份有限公司的潜在战略协同型合作伙伴，是以高技术主题互补性和高合作网络相似性为双轮驱动，围绕光刻机核心系统攻关构建的创新联合体。这类合作不仅具备技术上的强互补特征，更依托清华大学这一核心枢纽，形成了高度稳定、信任基础深厚的研发网络。

（2）资源整合型合作伙伴。北京华卓精科科技股份有限公司（机构2）与清华大学（机构1）、北京优微精密测控技术研究有限公司（机构122）有密切的合作，网络相似度分别为0.99和0.98，但技术互补性为0.08和0.19。从技术主题角度看，清华大学的技术主题包括工件台与掩模台系统和光源系统，北京优微精密测控技术研究有限公司的技术主题包括实用新型光刻机装置、工件台与掩模台系统和检测设备，这3家机构在实用新型光刻机装置和工件台与掩模台系统主题上具有较高的重叠性，导致其技术互补性较低。除已具备合作基础的清华大学、北京优微精密测控技术研究有限公司，电子科技大学（机构193）、中国科学院光电技术研究所（机构1520）与北京华卓精科科技股份有限公司的网络相似度和技术互补性分别为0.64、0.63，0.23、0.13，表现出高合作网络相似性和低技术主题互补特性。综上分析，以上机构可归类为资源整合型合作伙伴。这类合作的典型特征是机构间技术路线高度同频，知识基础相近，能快速整合研发、生产、市场资源，极大提升工程化效率与系统稳定性。

（3）突破创新型合作伙伴。中国科学院苏州纳米技术与纳米仿生研究所（机构1537）、东方晶源微电子科技（北京）有限公司（机构200）、中冶赛迪技术研究中心有限公司（机构1585）、闽都创新实验室（机构647）、上海华虹宏力半导体制造有限公司（机构847）、联华电子股份有限公司（机构619）、中国科学院上海技术物理研究所（机构1531）、深圳市大族半导体装备科技有限公司（机构963）与北京华卓精科科技股份有限公司（机构2）的网络相似度分别为0.48、0.44、0.45、0.44、0.44、0.33、0.26、0.28，技术互补性分别为0.63、0.55、0.54、0.52、0.48、0.48、0.49、0.45。从合作网络角度看，以上机构与北京华卓精科科技股份有限公司无直接或间接合作关系，网络结构角色也不同，因此网络相似度较低。从技术主题角度看，中国科学院苏州纳米技术与纳米仿生研究所的技术主题包括曝光系统、光刻胶工艺，东方晶源微电子科技（北京）有限公司的技术主题包括曝光系统、检测设备，中冶赛迪技术研究中心有限公司的技术主题包括电极制备、传感器、物镜系统，闽都创新实验室的技术主题包括电极制备、曝光系统、光刻技术、光刻胶工艺、光源系统，上海华虹宏力半导体制造有限公司的技术主题包括光刻胶工艺，联华电子股份有限公司的技术主题包括光刻胶工艺、新型光刻机装置，中国科学院上海技术物理研究所的技术主题包括光罩设备、电极制备、光刻胶工艺，深圳市大族半导体装备科技有限公司的技术主题包括曝光系统、光源系统、检测设备，以上机构除了闽都创新实验室和深圳市大族半导体装备科技有限公司在光源系统上有重叠外，其他技术主题均不同，与北京华卓精科科技股份有限公司的技术主题具有较强的互补性。综上分析，以上机构可归类为突破创新型合作伙伴。这类合作以强技术主题互补性驱动创新跃迁，以弱合作网络相似性引入外部异质性知识，打破路径依赖、填补技术盲区，具备催生突破性成果的潜力。这类合作的挑战是网络连接弱，沟通成本高，合作深度可能不稳定，需投入资源建立信任。

（4）机会探索型合作伙伴。中国科学技术大学（机构1512）与北京华卓精科科技股份有限公司（机构2）的网络相似度为0.29，技术主题互补性为0.40，均低于网络相似度阈值0.41和技术主题互补性低于阈值0.55。从合作网络角度看，该公司与北京华卓精科科技股份有限公司无直接或间接合作关系，网络结构角色也不同，因此网络相似度较低。从技术主题角度看，中国科学技术大学的技术主题主要涉及电极制备、光刻胶工艺、光源系统，与北京华卓精科科技股份有限公司在光源系统上重叠。综上分析，北京华卓精科科技股份有限公司与中国科学技术大学的潜在合作关系属于典型的机会探索型，既无深厚的合作网络基础，也未形成显著的技术互补，但局部技术交集为其未来潜在合作保留了可能性。如北京华卓精科科技股份有限公司在精密运动控制、激光设备等领域的技术优势，与中国科学技术大学在EUV光源预研、激光等离子体光源方面的专长结合，有机会在光源系统研发方向催生新的合作项目。这类合作的挑战在于需要克服网络连接弱、沟通成本高等问题，合作稳定性可能不如战略协同型高。

4　结语

本研究将专利合作网络表示学习与专利文献技术主题挖掘相结合，构建了“技术主题互补性+合作网络相似性”企业潜在技术合作伙伴预测模型，突破了传统单一维度（如合作关系或技术相似性）的局限性，在挖掘专利合作网络中结构和功能上相似机构的基础上，进一步考虑机构之间的技术主题互补性，为企业技术合作伙伴预测提供了新方法。

本研究在光刻机领域开展实证研究，研究结果表明，相较于传统单一网络预测方法，融入技术主题互补性的合作伙伴预测模型，为光刻机等专利密集型行业提供数据驱动的合作伙伴推荐工具，帮助企业快速识别具有不同合作特征的合作伙伴（战略协同型、资源整合型、机会探索型、突破创新型），降低技术合作试错成本，提升研发效率。另外，预测模型可迁移至新能源电池、生物医药等高技术领域，支持跨行业技术趋势预测、竞争对手技术布局分析等场景，服务于企业战略规划与技术创新管理。

本研究存在两方面不足。①数据局限性。仅利用专利数据表征技术知识，忽视了技术标准、学术论文等非专利技术载体，可能导致技术主题覆盖不全。②动态性缺失。由于实证研究领域专利合作规模不大，专利合作网络与技术主题均采用静态分析，未能考虑和验证技术演进与企业战略调整带来的动态互补需求变化。下一步，将从多源数据融合以增强技术主题表征的完整性、动态网络建模捕捉技术合作关系的演化规律以提升预测方法的时效性等方面进一步深化研究。

参考文献

[1] 中共中央关于进一步全面深化改革推进中国式现代化的决定[EB/OL]. [2025-11-21]. https://www.gov.cn/zhengce/202407/content_6963770.htm?1sRedirectHit=11454859.

[2] 李姝影，张鑫，许轶，等. 产学研关联创新技术主题分析及实证研究[J]. 科技管理研究，2020，40（14）：165-173.

[3] 解学梅. 中小企业协同创新网络与创新绩效的实证研究[J]. 管理科学学报，2010，13（8）：51-64.

[4] 袁永旭，成韬. 基于node2vec模型的多维度科研合作学者推荐研究[J]. 数字图书馆论坛，2024，20（12）：77-86.

[5] 于凯，郭煜婕. 基于链路预测的潜在专利合作关系研究[J]. 云南民族大学学报（自然科学版），2024，33（3）：377-385.

[6] 韩菁，唐箫，余乐安. 基于多层网络链路预测的潜在合作关系识别研究[J]. 系统工程理论与实践，2021，41（4）：1049-1060.

[7] 谢小东，吴洁，盛永祥，等. 基于异构图卷积网络的专利发明人潜在合作伙伴识别方法研究[J]. 现代情报，2025，45（6）：14-25.

[8] 琚春华，诸惠，曹倩雯. 基于多元融合的企业技术创新合作预测方法研究[J]. 中国科技论坛，2024（3）：108-119.

[9] 吕源，刘敏榕. 基于专利文本的产学合作预测研究[J]. 现代情报，2022，42（10）：71-78.

[10] MAKRI M，HITT M A，LANE P J. Complementary technologies，knowledge relatedness，and invention outcomes in high technology mergers and acquisitions[J]. Strategic Management Journal，2010，31（6）：602-628.

[11] JAFFE A. Technological opportunity and spillovers of R&D: evidence from firms' patents，profits and market value[J]. American Economic Review，1986，76（5）：984-1001.

[12] FRENKEN K，VAN OORT F，VERBURG T. Related variety，unrelated variety and regional economic growth[J]. Regional Studies，2007，41（5）：685-697.

[13] 郑江淮，陈喆，孙志燕，等. 从竞争到互补：区域技术结构变迁的测度与理论假说[J]. 经济评论，2022（1）：13-29.

[14] CHANG P C，CHANG Y H，SU F P，et al. The study on patent acquisition from complementarity and supplementarity: evidence from smartphones of Apple and Samsung[C]// Proceedings of the Portland International Conference on Management of Engineering & Technology. Piscataway：IEEE，2014：2996-3003.

[15] WANG X F，QIAO Y L，HOU Y J，et al. Measuring technology complementarity between enterprises with an hLDA topic model[J]. IEEE Transactions on Engineering Management，2021，68（5）：1309-1320.

[16] WANG M Y，CHIU T F，CHEN W Y. Exploring potential R&D collaborators based on patent portfolio analysis: the case of biosensors [C]// Proceedings of the Portland International Conference on Management of Engineering & Technology. Piscataway：IEEE，2009：322-330.

[17] 温亮，邱鹏君，马萍萍，等. 基于SAO语义分析的潜在技术合作伙伴识别[J]. 北京理工大学学报（社会科学版），2017，19（4）：91-96.

[18] 徐晟，袁德玉，赵惠芳，等. 互补性专利下创业企业协同创新投资策略研究[J]. 科技进步与对策，2013，30（6）：75-79.

[19] 章激扬，邵记友，冯金梅. 技术互补如何促进战略性新兴产业效率提升：以长三角为例[J]. 经济问题探索，2025（1）：67-91.

[20] 郑江淮，金晟男. 地区间技术互补、均衡发展与企业全球价值链攀升[J]. 中国工业经济，2024（7）：85-104.

[21] ROSEN-ZVI M，GRIFFITHS T，STEYVERS M，et al. The author-topic model for authors and documents[C]//Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence. AUAI Press，2004：487-494.

[22] AHUJA G. Collaboration networks, structural holes, and innovation: a longitudinal study[J]. Administrative Science Quarterly，2000，45（3）：425-455.

[23] COHEN W M，LEVINTHAL D A. Absorptive capacity: a new perspective on learning and innovation[J]. Administrative Science Quarterly，1990，35（1）：128-152.

[24] GRANOVETTER M S. The strength of weak ties[J]. American Journal of Sociology，1973，78（6）：1360-1380.

[25] 张金福，黄雪晴. 创新网络结构对颠覆性创新的影响机制：双元性学习的中介作用[J]. 科技管理研究，2020，40（8）：7-16.

[26] 胡楚雄，周冉，付宏，等. 集成电路装备光刻机发展前沿与未来挑战[J]. 中国科学：信息科学，2024，54（1）：130-143.

[27] 张霞，刘宏波，顾文，等. 全球光刻机发展概况以及光刻机装备国产化[J]. 无线互联科技，2018，15（19）：110-111，118.

[28] RÖDER M，BOTH A，HINNEBURG A. Exploring the Space of Topic Coherence Measures[C]//Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. New York：ACM，2015：399-408.

作者简介

关鹏，男，博士，教授，硕士生导师，研究方向：数据挖掘与知识发现、技术创新管理、知识工程与知识挖掘，E-mail：guanpeng1983@163.com。
王曰芬，女，博士，教授，博士生导师，研究方向：数据挖掘与数智服务、竞争情报与知识管理、科学计量与创新评价、数字政府数据资源开发。
傅柱，男，博士，副教授，硕士生导师，研究方向：知识挖掘与知识管理。

Research on the Prediction of Enterprise Technology Partners Based on the Complementarity of Technology Topics and the Similarity of Cooperation Networks

GUAN Peng[1] WANG YueFen[2] FU Zhu[3]
(1.School of Economics and Law, Chaohu University, Hefei 238000, P.R. China;2.Institute for Big Data Science, Tianjin Normal University, Tianjin, 300382, P.R. China;3.School of Economics and Management, Jiangsu University of Science & Technology, Zhenjiang 212003, P.R. China)

Abstract: To address the information asymmetry in technology collaboration decision-making within patent-intensive industries, this study constructs a prediction model for enterprise technology partners, aiming to assist enterprises in making technology collaboration decisions. Based on patent literature, technological knowledge features are extracted, and collaboration networks are built using co-patentee relationships. The Node2vec algorithm quantifies enterprise network structural similarity, while the AT model extracts technology topic vectors combined with JS divergence to calculate the complementarity of technology topics, establishing a dual-dimensional potential partner identification model integrating both technology and network perspectives. Empirical research in the field of lithography machines shows that this model outperforms traditional single cooperative network prediction methods. It integrates the complementarity of technical themes and can effectively identify four types of potential partners: strategic synergy type, breakthrough innovation type, resource integration type, and opportunity exploration type. The research method can provide data-driven partner recommendation tools for patent-intensive industries, helping enterprises quickly identify potential partners.

Keywords: Complementarity of Technology Topics; Similarity of Cooperative Networks; Technology Collaboration; Prediction; Patent

（责任编辑：管清漾）