● 郭剑明[1,2]，杨 杰[1,2]，方 静[1]，邓三鸿[1,2]
（1.南京大学信息管理学院，江苏 南京 210023；2.数据工程与知识服务省高校重点实验室（南京大学），江苏 南京 210023）

基于高阶网络结构的核心专利识别方法研究*

摘 要：［目的/意义］通过识别专利网络中频繁出现的局部交互模式，定义并计算专利在高阶网络结构中的重要性，丰富专利网络分析方法。［方法/过程］首先，定义并提取专利网络中的前馈循环模体结构；其次，将高阶结构特征与专利重要性联系起来，以此为依据提出两种基于前馈循环模体的专利评价模型PCDRank和PCSRank，识别核心专利；最后，利用信息传播动力学模型计算专利传播影响力，验证所提模型和其他中心性指标的识别效果。［结果/结论］专利引用网络的高阶结构包含了更为丰富的结构信息，有助于增加专利分析信息量；基于高阶网络结构的核心专利识别方法能获得更为准确的识别结果；利用信息传播动力学模型验证了所提识别模型具有良好的有效性和稳健性。［创新/价值］通过结合高阶结构特征与全局网络拓扑特性，构建了更加精准的核心专利识别框架，并为理解专利网络的结构与演化提供新视角。

关键词：高阶网络；核心专利；专利网络；节点重要性；传播动力学模型

DOI：10.16353/j.cnki.1000-7490.2026.01.013

引用格式：郭剑明，杨杰，方静，邓三鸿．基于高阶网络结构的核心专利识别方法研究［J］．情报理论与实践，2026，49（1）：106-114，149．

Research on Core Patent Identification Method Based on Higher-order Network Structure

Guo Jianming[1,2]，Yang Jie[1,2]，Fang Jing[1]，Deng Sanhong[1,2]
（1. School of Information Management，Nanjing University，Jiangsu Nanjing 210023；
2. Key Laboratory of Data Engineering and Knowledge Services in Provincial Universities （Nanjing University），Jiangsu Nanjing 210023）

Abstract：［Purpose/significance］By identifying frequent local interaction patterns in patent networks，defining and calculating the importance of patents in higher-order network structures，enriching patent network analysis methods，and providing a new perspective for understanding the structure and evolution of patent networks.［Method/process］Firstly，define and extract the feedforward loop phantom structure in the patent network. Secondly，the higher-order structural features are linked with the importance of patents，and based on this，two patent evaluation models PCDRank and PCSRank based on feedforward loop phantom are proposed to identify the core patents. Finally，the information dissemination dynamics model is used to calculate the influence of patent dissemination，and the recognition effect of the proposed model and other centrality indexes is verified.［Result/conclusion］The higher-order structure of patent citation network contains richer structural information，which is helpful to increase the amount of patent analysis information. The core patent recognition method based on high-order network structure can obtain more accurate recognition results. The dynamics model of information propagation is used to verify the validity and robustness of the proposed identification model.［Innovation/value］By combining high-order structural characteristics with global network topology characteristics，a more accurate core patent identification framework is constructed，and a new perspective is provided for understanding the structure and evolution of patent networks.

Keywords：higher-order network；core patent；patent network；node importance；spreading dynamics model

0 引言

随着全球技术创新的加速和专利数量的快速增长，专利信息系统已成为揭示技术演化规律、优化创新战略的重要载体[1]。从宏观视角看，专利网络是技术知识传播与扩散的复杂系统，反映技术创新的生态结构与动力学演化过程，深刻影响着国家、产业及企业的技术竞争和创新能

* 本文为国家自然科学基金青年学生基础研究项目（博士研究生）“融合创新测度和生成模型的科学技术知识空间预测研究”（项目编号：724B2015）和国家社会科学基金重点项目“基于领域集体智慧挖掘的颠覆性变革预测研究”（项目编号：25ATQ008）的成果。

力[2]。从微观视角看，核心专利作为技术领域的关键节点，承载了技术创新的首创性与产业化的核心价值，对技术研发方向的引领、技术布局的优化以及创新资源的配置起着决定性的作用[3]。因此，基于专利网络的核心专利识别研究，有助于揭示技术知识的流动与演化规律，对推动技术创新战略的制定和实施具有重要的政策价值和实践意义。

核心专利是专利分析的重要对象之一，从海量专利中识别核心专利对创新主体制定研发战略、优化技术布局具有关键作用[4]。传统核心专利识别方法主要依赖专利属性指标（如被引频次、权利要求数量）[5]或基于二元引用关系的简单网络分析（如节点中心性）[6]。然而，这些方法在捕捉专利间的复杂交互模式和网络结构特征方面存在局限[7]。基于二元关系的网络分析难以全面刻画专利网络中多节点间的群体关联与动态演化机制，忽略了技术知识流动中超越成对关系的多元交互行为[8]。单一属性指标或局部网络特征无法充分揭示专利在技术生态系统中的深层作用，限制了核心专利识别的精度与全面性。随着复杂网络科学的快速发展，高阶网络（Higher-Order Network）作为一种新兴的建模范式，突破了传统二元关系的局限[9]，通过引入模体（Motif）、超图（Hypergraph）等高阶结构，提供了捕捉专利网络中复杂交互模式与功能子系统的新视角。高阶网络不仅能更精确地描述专利网络的结构特征，还能揭示技术知识传播的动力学规律，为核心专利识别提供新的理论框架与分析工具。

1 研究背景

1.1 核心专利识别方法相关研究

核心专利是在某一领域具有首创性的并以此为核心被后续科技引用及产业化集聚必不可少的专利，具有后续科技之核心和产业经济之核心两个显著特征[10]。核心专利识别方法可大致归纳为利用专利属性指标和技术关系特征两大类。识别方法的早期研究主要着眼于识别指标的发现和应用[11]。基于专利属性指标的核心专利识别研究经历了从利用单一指标、组合使用多种指标到构建指标体系的过程。单一指标法将频次较高的专利作为核心专利[12]。专利指标组合方法通过使用多种属性指标识别核心专利[11]。专利指标体系法则是构建指标体系，利用主观或客观方法赋予指标权重，进而识别核心专利[13]。

基于技术关系特征的专利网络分析通常借助专利在网络中扮演的角色或所处的位置量化其重要程度。例如，有学者尝试使用了许多经典中心性算法，包括点度中心性、中介中心性[14]、PageRank和HITs中心性[4]等识别核心专利。也有学者提出借助专利在整体网络中的位置，利用拓扑结构特征测度专利重要性。例如，通过抽象专利网络的结构单元[15]或综合考虑专利度数及其邻域专利对间的相似程度[6]，定义结构重要性指标并识别核心专利。可见，结合专利网络拓扑特性测度专利重要性成为这类方法识别核心专利的关键。

1.2 高阶网络相关研究

高阶网络是相对于传统网络的一种复杂网络建模方式[16]，它强调在建模关系时考虑更复杂的上下文、依赖关系或路径顺序，而不是仅仅用节点和连边的简单连接来表示实体之间的关系。高阶网络结构是网络科学中重要且新兴的概念，其是对传统网络模型的扩展和深化，旨在更准确地捕捉和建模现实世界中复杂系统中超越成对二元关系的相互作用[17]。例如，单纯形和模体结构的引入，使得对节点间的群体关系建模成为可能，为识别关键节点提供了理论支持[18]。高阶网络所展现出的动态建模能力，能够揭示出传统网络无法捕捉的传播、扩散和演化机制[9]。可见，高阶网络不仅能更加精确地描述复杂系统的结构，而且能揭示系统中的动力学现象[19]。

综上，当前专利网络分析研究主要基于专利间的二元关系度量专利的重要性，较少涉及专利网络三元关系和社区等“中观”层面的研究。虽有学者对专利网络进行社区发现，在技术领域细分条件下提高了识别结果的精准性[5]，但识别方法仍存在一定局限性。而三元关系等高阶结构能够洞察网络中更深层次的个体关系和网络结构动因，对于专利网络分析具有重要意义。考虑到网络中局部专利间的特定交互模式，本文将专利网络的高阶结构特征与专利重要性相关联，旨在构建基于高阶结构的专利网络节点重要性评价框架并用于核心专利识别研究。

2 理论依据与识别方法

2.1 理论依据

专利信息包含了大量的技术创新、技术转让等方面的信息，应用网络科学理论和方法有助于更好地理解专利信息系统。专利网络描述了技术的发展演化过程，记录了专利网络上关键节点的产生和社区结构的形成与演化等动力学过程信息[20]。研究专利网络的高阶结构的价值体现在以下两个方面：一是更加精确地描述专利信息系统的结构；二是揭示专利信息系统中的动力学现象。

从结构描述的角度来看，高阶网络结构揭示了传统网络难以捕捉的知识结构的深层次演变，能够更准确地描绘科学技术知识的演化路径与创新潜力[21]。传统专利网络分析多基于节点属性或邻接关系，如度数中心性、PageRank或聚类系数等，尽管能够在一定程度上揭示专利的重要性，但往往难以捕捉专利在复杂网络中所处的群体组织关系或功能角色[21]。相比之下，高阶结构（如模体、k-核结构、社区结构、结构洞等）能够揭示节点间在更大规

模子结构中的互动模式，进而更准确地描绘出专利信息系统的组织方式与功能分层[19]。因此，将高阶结构纳入分析，有助于拓展专利网络的观察尺度，增强其结构解释力，从而更精确地揭示知识系统的内在构造。

从动力学机制的角度来看，高阶结构分析能够揭示传统方法难以观测到的演化过程与潜在规律。高阶结构遵循信息论中“非冗余性”的原则，提供了区别于属性指标与基础网络关系的信息增量[4]。例如，在节点度和PageRank上相近的两个专利，可能由于参与的模体类型或所处社区位置不同，而在技术演化路径上表现出显著差异。高阶网络分析关注三个及以上节点间的复杂交互关系，能够更全面地刻画专利间的知识关联与技术融合过程。例如，作为专利对之间的传递闭合结构，模体对于专利引用关系形成具有显著影响[8]。这一视角对于理解专利网络中创新扩散、关键路径形成以及结构演化规律具有重要意义，有助于提升对潜在关键专利的识别准确性，进而为科技战略制定和技术预测提供有力的理论支撑。

2.2 专利网络中的高阶结构

专利网络是以专利为节点，专利间的引用关系为边构建的网络，一般用$G=(V, E)$表示。在网络科学中，高阶结构通常使用模体、团（Clique）、超图和单纯形（Simplex）等相关概念来定义[22]。这些结构作为高阶网络中的特定模式，超越了简单的节点对或边，能够捕捉节点间的复杂关系。

模体作为一种常见的高阶网络结构，表征了网络中少数节点间反复出现的特定连接模式，其在真实网络中的出现频率显著高于随机网络[23]，因此本文将模体作为主要研究对象。模体以各种形式存在于真实世界中。例如，社会网络中的3-模体代表了社会关系的底层结构[24]，专利网络中的三元组（Triads）反映了专利网络的基础局部结构[25]。

在有向网络中，3-模体有13种可能的网络表现形式，如图1所示。专利引用网络具有单向性的基本特征，因此并不具有互惠（双向交互）关系。因此，前馈循环（Feed Forward Loops，FFL）M_4是专利引用网络3-模体中一般的网络表现形式。在这个结构中也包含M_1、M_2和M_3。前馈循环在信息传播和网络稳定性中起着关键作用，为理解网络中节点的重要性提供了新的见解[26]。因此，本文旨在探索前馈循环中的模体结构对专利网络节点重要性的影响。

2.3 基于高阶网络结构的核心专利识别模型

随着专利网络分析方法的兴起，基于专利网络的中心性指标逐渐成为专利价值评估与核心专利识别研究的热点[7]。在众多经典的中心性指标中，PageRank算法及其推广形式具有较好的评估效果和改进条件[27]。PageRank

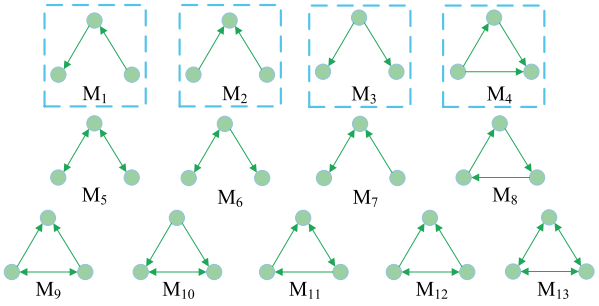


图1 有向3-模体及其网络表现形式

Fig. 1 Directed 3-motif and its network representation

算法的推广形式同时考虑了网络的整体性和每个节点的异质性，为识别模型的构建提供了思路。PageRank算法的计算过程为：

$$PR_i=\beta\sum_j A_{ij}\frac{PR_j}{k_j^{out}}+\frac{(1-\beta)}{n} \tag{1}$$

式中，β为阻尼因子，一般取值为0.85；A为邻接矩阵，$A_{ij}=1$表示存在专利j引用专利i；k_j^{out}表示专利j的出度。公式（1）表明专利i的PR值受到引证它的所有专利j的PR值的影响，PR值越高，专利越重要。

在PageRank的基础上，本文考虑在迭代过程中引入专利网络的高阶结构信息，提出基于前馈循环模体的专利网络节点重要性评估算法。本文假设存在于前馈循环模体中的专利更为重要。通过从专利网络G中提取前馈循环，得到有向网络G'。如果原专利网络G中两个专利间的有向边同时存在于不同的前馈循环结构中，则需在有向网络G'中赋权。在此情况下，可通过提取专利网络G中的全部前馈循环模体构造出有向加权网络G'。这个有向加权网络反映了原始网络的高阶结构信息，如图2所示。专利“2”和专利“3”之间的有向边存在于前馈循环{1，2，3}和{2，3，4}中。因此，在有向网络G'中，专利“2′”和“3′”之间的边权重为2。G中的一些专利可能不能与其他专利形成前馈循环结构，那么这个专利就是G'中的孤立节点，如图2中的专利“6”。

通过对专利网络G应用PageRank算法，可以得到每个专利的重要性得分。本文假设存在于前馈循环模体结构中的专利更为重要。因此，考虑放大存在于前馈循环模体中专利的重要性，即放大有向网络G'中专利的重要性，而不在有向网络G'中专利的重要性将被抑制。本文构建两种模型评估包含高阶结构的有向网络中专利的重要性。

1）根据有向网络G'中对应专利的入度（in-degree）更新专利的重要性。这种方法被称为耦合前馈循环模体入度的专利重要性排序模型（Patent Coupled In-Degree Rank，PCDRank），其公式为：

$$PCD_i=[\beta\sum_j A_{ij}\frac{PCD_j}{k_j^{out}}+\frac{(1-\beta)}{n}]\cdot d_i^{\theta} \tag{2}$$

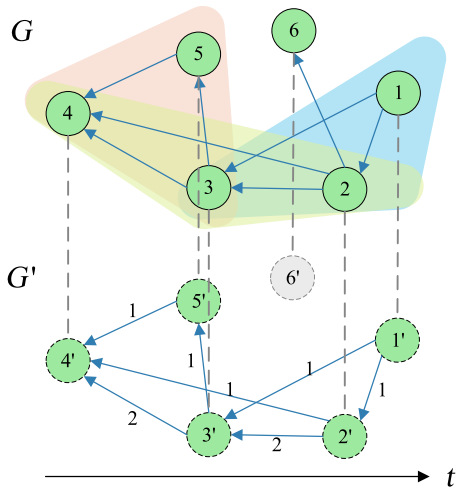


图2　基于前馈循环模体结构的专利网络节点重要性示意

Fig. 2　Importance of patent network nodes based on FFL motif structure

式中，θ是取值范围为0～1的调节参数。当$\theta=0$时，PCDRank退化为经典PageRank算法。为避免得分呈现指数增长，利用公式（3）对其进行归一化。

$$\mathrm{PCD}_i == \frac{\mathrm{PCD}_i}{\sum_j \mathrm{PCD}_j} \tag{3}$$

在初始时刻，将每个专利的重要性得分设为1，然后按照上述规则迭代计算，当专利重要性得分稳定时，得到专利重要性得分PCDRank。

2）本文假设频繁出现在不同前馈循环模体结构中的专利更重要。例如，在图2中，专利“3”存在于前馈循环｛1，2，3｝、｛2，3，4｝和｛3，4，5｝3个模体结构中，而专利“1”和专利“5”只存在于一个前馈循环中。在有向加权网络G'中，参与模体结构的专利的频率可以通过边的权重进行计算。将专利边权重之和记为强度（Strength，S）。继而可根据有向加权网络G'中对应专利的强度来迭代更新专利的重要性。这种方法被称为专利耦合前馈循环模体强度重要性排序模型（Patent Coupled Strength Rank，PCSRank），其公式为：

$$\mathrm{PCS}_i = [\beta \sum_j A_{ij} \frac{\mathrm{PCS}_j}{k_j^{\mathrm{out}}} + \frac{(1-\beta)}{n}] \cdot S_i^{\theta} \tag{4}$$

式中，θ是调节参数，与PCDRank模型相同。当$\theta=0$时，PCSRank退化为经典PageRank算法，其迭代过程与PCDRank模型一致。

2.4　识别模型评价方法

专利网络作为多知识组成的技术整体能够反映技术知识扩散过程，而技术知识扩散也能推动专利进化。随着初始专利被后续专利引用，其技术知识扩散范围不断扩大，并逐步进化为领域内占据重要位置的核心专利[15]。可见，利用技术知识扩散范围能够衡量专利在技术演化过程中的重要程度。因此，本文使用信息传播动力学中的SIR模型评价识别效果。

在SIR模型中，每个专利可以处于以下三种状态之一：易感状态（Susceptible，S）、感染状态（Infective，I）和恢复状态（Removed，R）[6]。在模型运行的每个时间步中，感染状态的专利以传播概率β将其相邻的易感染专利转变为感染状态，且它们会在下个时间步转变为恢复状态。一旦进入恢复状态，专利将不再被感染。

专利传播影响力通过其平均传播范围来衡量。信息传播概率β的计算方法为：

$$\beta = \frac{\langle k \rangle}{\langle k^2 \rangle} \tag{5}$$

式中，$\langle k \rangle$为网络平均度；$\langle k^2 \rangle$为网络二阶平均度。为减少独立实验随机误差的影响，定义专利i的实际传播影响力$\Phi(i)$，其计算方法为：

$$\Phi(i) = \frac{1}{L} \sum_{l=1}^{L} \Phi_l'(i) \tag{6}$$

式中，L表示实验次数；$\Phi_l'(i)$表示当专利i作为传播源时，在第l次实验中最终处于恢复状态的专利数量。最后，将实际传播影响力$\Phi(i)$排名前10%的专利视为“基准核心专利”，并将其与通过其他指标识别出的核心专利进行对比分析。

3　实证分析与讨论

钙钛矿太阳能电池是一种新兴的光伏技术，其核心材料为具有钙钛矿晶体结构的有机—无机杂化材料。近年来，钙钛矿电池以其快速提高的光电转换效率、低材料成本和低温加工特性而受到广泛关注，识别这个领域的核心专利对跟踪发展我国相关技术具有现实意义。

3.1　数据获取与处理

使用美国专利商标局（United States Patent and Trademark Office，USPTO）提供的PatentsView专利数据库作为数据来源。PatentsView专利数据库是由USPTO首席经济学家办公室（Office of the Chief Economist）和美国科学与政策研究中心（American Institutes for Research）等机构共同开发的，涵盖1976—2025年美国授权专利及专利申请公开数据。从PatentsView下载得到7966916条专利数据，包含了110014388条引用关系。Y02E 10/549是联合专利分类（Cooperative Patent Classification，CPC）系统中专门用于标识与钙钛矿太阳能电池（Perovskite Solar Cells，PSCs）相关技术的分类号，归属于绿色能源技术（Y02E）的大类。从数据集中抽取这个领域原始数据包含11333个节点和13356条边，发现构建的网络由一个巨连通分支和多个离散小分支组成。由于对仅有内部连接且与外部没有任何联系的小规模子群进行研究的意义不大[28]，同时，为保证识别模型算法能够收敛，抽取专利引用网络中最大连通图。经统计，抽取出的最大连通图包含4445个节点

和11688条连边，连边规模占全部边数的87.5%，较大程度保留了原始网络的结构信息。

3.2 核心专利识别结果

计算得到专利网络的传播概率为0.053。利用Python中的igraph包和自编函数计算专利的PCD、PCS与传播影响力。重复进行100次独立实验，得到“基准核心专利”452件。

3.2.1 识别结果的总体分析 为验证所提指标识别效果，选取常用的网络中心性：入度中心性（Indegree）、特征向量中心性（Eigenvector）、PageRank和专利传播影响力Φ作为比较指标。PCD和PCS指标均反映出专利价值的“幂律分布”特点。

表1统计了各指标得分的基本信息，入度中心性测度了专利在原始网络G中的局部结构重要性，特征向量中心性和PageRank反映了专利的全局影响力，Φ衡量了专利在技术知识扩散过程中的平均影响范围，而本文提出的PCD和PCS指标既保留了专利在高阶网络G'中的局部结构重要性，也迭代计算了专利在G中的全局影响力。偏度和峰度指标显示各指标偏度值均大于6，峰度值均大于55，不同指标得分基本呈现“右偏、尖峰、厚尾”的特点，说明专利重要性符合“幂律分布”特征，核心专利占比极少。

表1 各指标得分的描述性统计

Tab. 1 Descriptive statistics of scores of each index

指标	均值	中位数	最小值	最大值	标准差	峰度	偏度
Indegree	2.629	1	0	98	6.577	55.344	6.318
Eigenvector	0.03	0	0	1	0.066	93.572	9.155
PageRank	0	0	0	0.011	0	224.899	12.92
PCDRank	0.001	0	0	0.109	0.003	235.078	13.725
PCSRank	0.001	0	0	0.147	0.003	193.292	12.679
Φ	0.163	0.02	0	7.6	0.481	79.759	7.676

为评估模型改进效果，本文比较了PageRank、PCDRank和PCSRank与专利传播影响力间的斯皮尔曼（Spearman）相关系数，如图3所示。相关性系数越高，表明排序方法与模型仿真结果越接近，识别效果越好。根据模型设计，当$\theta=0$时，两个模型等价于经典的PageRank算法。此时，专利重要性得分与传播影响力间的ρ最小。随着θ的增大，PCDRank与传播影响力间的ρ不断增大，PCSRank则呈递减趋势。在相同参数θ下，PCDRank、PCSRank的相关系数始终高于PageRank，反映出改进的PCDRank和PCSRank模型优于PageRank。

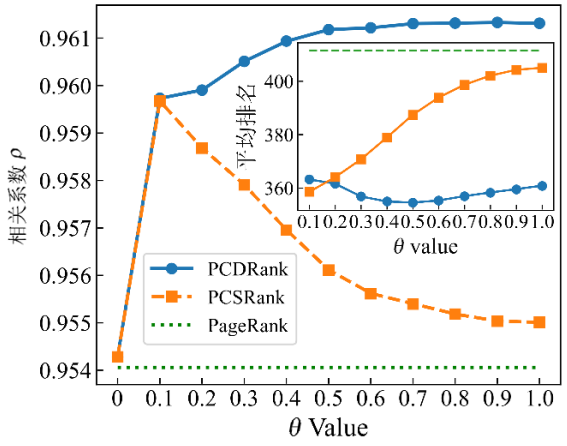


图3 Spearman相关系数及基准核心专利平均排名变化

Fig. 3 Spearman coefficient and average ranking changes of benchmark core patents

接着，选择专利传播影响力前10%的基准核心专利，计算不同参数θ下的PCDRank和PCSRank得分，并比较基准核心专利在上述指标中的平均排名，如图3所示。平均排名越靠前的核心专利越重要，因此对基准核心专利排名越高的指标越有效。当参数$\theta=0$时，两种模型都退化为PageRank，随着参数θ的增大，PCDRank对基准核心专利的平均排名逐渐降低，而PCSRank的平均排名呈上升趋势，但两模型对基准核心专利的平均排名始终小于PageRank，说明PCDRank和PCSRank模型对于核心专利的识别效果优于PageRank。

3.2.2 识别结果的具体分析 表2列举了PCDRank排名前15位的专利，并选取上述其他指标进行对比。从表2可见，各指标识别结果不存在显著差异，PCDRank排名前15位的结果中，入度中心性、PageRank和传播影响力结果中有15个核心专利，特征向量中心性结果中有14个核心专利，高PCD和PCS专利的其他指标值一般也较高，说明所提模型能够科学评估专利在网络中的重要性。其次，PCDRank和PCSRank得分排名前10位的专利相同，这说明两个模型具有一定的同质性。根据模型设计，PCDRank和PCSRank分别受到高阶网络结构中专利i的入度d_i和强度S_i控制，而两者具有较高的相关性。表3展示了专利的具体信息，其中“专利价值度”为壹专利价值度评估数值为1~100，从专利质量、技术性、经济性、发展前景、科研能力五大方面对专利进行量化评估，而得到的专利综合评分。

表2　PCD排名前15位专利指标对比

Tab. 2　Comparison of top 15 patent indicators of PCD

专利号	Indegree	Eigenvector	PageRank	PCDRank	PCSRank	Φ
US90736178A	80	67	1	1	1	34
US69410891A	63	51	2	2	3	23
US4125414A	75	68	12	3	2	17
US5757126A	9	309	3	4	4	3
US4127738A	199	266	9	5	5	139
US5171373A	377	235	4	6	8	147
US3935031A	249	323	14	7	7	316
US5686360A	13	224	5	8	9	6
US6781152B2	14	484	7	9	6	11
US4106951A	288	328	16	10	10	190
US5331183A	15	65	8	11	12	4
US5350459A	30	48	13	12	11	13
US5536588A	153	168	6	13	25	95
US5885368A	80	304	10	14	26	57
US5454880A	6	105	11	15	24	12

表3　PCD排名前15位专利具体信息

Tab. 3　Specific information of the top 15 PCD patents

专利号	专利权人	专利名称	IPC主分类号	专利价值度
US90736178A	伊士曼柯达公司	多层有机光伏元件	H01L31/06	57
US69410891A	南加利福尼亚大学	有机光电子器件及其制造方法	H01L29/28	50
US4125414A	伊士曼柯达公司	有机光伏元件	H01L31/04	58
US5757126A	通用显示器公司	具有交替聚合物和绝缘层的钝化有机器件	H05B33/04	94
US4127738A	阿莫科—安然太阳能	一种含有有机层的光伏装置	H01L31/04	59
US5171373A	美国国际电话电报公司	涉及富勒烯光行为的设备	H01L51/05	55
US3935031A	新英格兰INST公司	具有增强功率输出的光伏电池	C07D487/22	77
US5686360A	通用显示器公司	有机器件的钝化	H05B33/04	95
US6781152B2	半导体能源研究所	具有发光元件和保持彩色滤光片的粘合层的半导体器件	G02F1/136	92
US4106951A	UCE股份有限公司	使用有机材料作为活性层的光伏半导体器件	H01L31/06	38
US5331183A	加州大学	共轭聚合物—受体异质结；二极管、光电二极管和光伏电池	H01L31/10	77
US5350459A	理光株式会社	有机光伏元件	H01L31/06	52
US5536588A	东芝有限公司	非晶有机薄膜元件和非晶有机聚合物组合物	B41M5/26	62
US5885368A	默克专利有限公司	光伏电池	H01L31/04	98
US5454880A	加州大学	共轭聚合物—受体异质结；二极管、光电二极管和光伏电池	H01L31/10	77

技术领域一般可以细分为若干子领域，核心专利在各子领域当中应该具有一定的分布[5]。为此，使用Louvain算法对从专利网络G和专利高阶网络G'的社区结构进行了分析，如图4所示。

社区发现的思想是同一子领域内的专利技术相似性较高，不同子领域间的专利技术相似性较低。基于专利网络G与其对应的专利高阶网络G'，分别识别出452件和400件核心专利。两种网络识别出的核心专利数量相近，但其在不同社区中的分布特征差异显著。具体而言，专利网络G由4445个专利构成，核心专利占比约为10.2%，而高阶网络G'仅包含1733个专利，核心专利占比高达23.1%。尽管高阶网络中专利总数较少，但核心专利在各子领域中的占比普遍高于原始网络，且分布更加集中。

造成这一现象的原因在于两种网络结构特征的差异。原始网络基于二元引用关系构建，侧重于揭示知识传递的线性路径。而高阶网络引入了前馈循环模体结构，能够捕捉三元及以上节点间复杂的交互模式。这种基于局部结构嵌入关系的建模方式，不仅提升了网络的结构表达能力，还有效识别了在微观连接中处于关键位置的专利。高阶网络在结构层面对原始网络进行了“压缩”，剔除了未参与

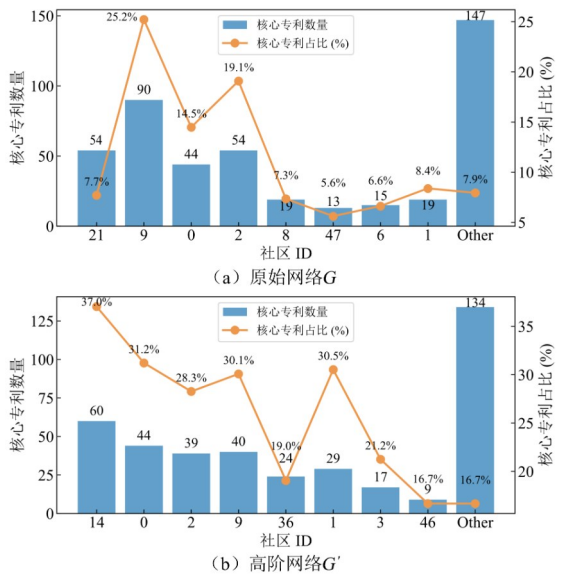


图4 核心专利在技术子领域中的分布

Fig. 4 Distribution of core patents in technical sub-fields

模体的边缘专利，使得保留下来的专利集中于互动密切的知识单元，从而在子社区内部形成了更高密度的核心专利聚集区。这种结构重组提升了核心专利识别的精度。上述结果反映出高阶结构包含了专利网络中更为丰富的结构信息，有助于增加专利分析信息量，优化核心专利识别效果。

3.3 识别模型评价

进一步地，使用基于网络的传播动力学模型，从专利网络的结构与功能视角尝试衡量不同指标的核心专利识别效果，分析识别模型的有效性和稳健性。

3.3.1 识别模型的有效性分析

首先，分析各指标的Top-L召回率，比较不同检测长度下识别到的基准核心专利的比例，并对每个指标计算当前L%对应的专利样本量。在此基础上，进行50次加权随机抽样并计算平均基准核心专利召回率，所有指标的召回曲线如图5所示。由图5可以发现，PCDRank和PCSRank在所有检测长度上都比特征向量中心性和PageRank具有更高的核心专利召回率。这说明在评估专利重要性时，考虑模体作为高阶结构，可以有效识别出有影响的专利，特别是频繁出现在网络模体中的专利可能更为重要。

其次，通过复杂网络上的信息传播过程进一步验证了PCDRank和PCSRank模型的有效性。专利网络中的传播过程可以看作是技术知识的传播。在专利引用网络中，每一件专利都代表一种新的技术，并将其技术知识扩散给其他施引专利。有一定概率接受这一技术知识的专利将进一步传播给后续专利。因此，更有影响力的专利应该能够更广泛地传播其技术知识。使用易感—感染（SI）传播模型来评估4种方法的有效性，并将平均传播范围视为目标专利的影响力。分别选取特征向量中心性、PageRank、PCDRank和PCSRank得到的前50件专利，依次将每个专利作为初始信息源进行传播。如果模型识别出的核心专利使得信息传播范围更广，则说明这个模型在专利重要性测度方面比其他模型更有效。图6显示了4种不同模型的平均传播范围。可以发现PCDRank和PCSRank的平均传播范围随着传播时间的增加而增长得更快。这表明，两种模型得到的初始传播源能够快速、广泛地影响专利网络中的其他专利，与特征向量中心性和PageRank相比，PCDRank和PCSRank在识别网络中有影响的专利方面更有效。

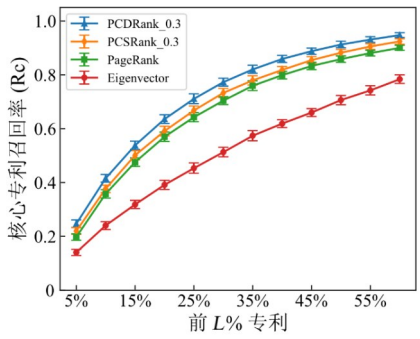


图5 各指标前L%随机样本中核心专利召回率比较

Fig. 5 Comparison of core patent recall rates among the top L% samples

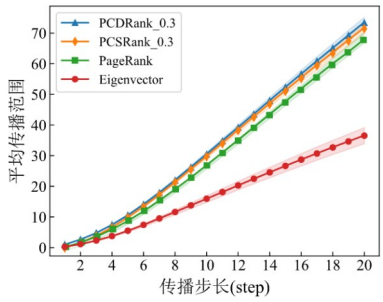


图6 各指标排名前100专利的传播范围

Fig. 6 Spread range of the top 100 patents in each index

3.3.2 识别模型的稳健性分析

与科学引文网络相似，专利引用网络中也存在引用操纵与引文噪声的现象。例如，专利申请人可能通过自我引用或申请人间的互引行为提升特定专利的影响力；某些引用则可能出于格式要求或审查流程而与实际技术无关[8]。这些因素会干扰基于技术关系的专利网络分析结果。因此，评估专利重要性的模型应当具备良好的鲁棒性与抗噪性，以减少非真实引用关系对排名结果的影响。

进一步地，分析PCDRank和PCSRank对专利间虚假引用行为的稳健性。为了评估虚假引用对不同专利重要性模型的影响，首先在专利网络中随机选取一件未被引用的专利，即这件专利的入度为零，并将其作为通过虚假引用行为提高其影响力的目标专利T。随后，为模拟虚假引用行为，在专利网络中添加m件新专利，并假设每件专利包含r件引用专利。对于m件专利，每件专利都需引用目标

专利，即r件引用专利中有一件是目标专利T，其余$(r-1)$件引用专利则从原专利网络中随机选取。通过这个流程可构建出一个受到蓄意操纵的专利引用网络。然后，基于入度中心性、特征向量中心性、PageRank、PCDRank和PCSRank模型，将目标专利在原专利网络中的排名分别定义为$R_{I_n}(T_0)$、$R_E(T_0)$、$R_{\mathrm{PR}}(T_0)$、$R_{\mathrm{PCD}}(T_0)$和$R_{\mathrm{PCS}}(T_0)$。当第m件专利加入原专利引用网络时，根据不同模型，目标专利T的排名分别定义为$R_{I_n}(T_m)$、$R_E(T_m)$、$R_{PR}(T_m)$、$R_{\mathrm{PCD}}(T_m)$和$R_{\mathrm{PCS}}(T_m)$。接着，用ΔR_{I_n}、ΔR_E、ΔR_{PR}、ΔR_{PCD}和ΔR_{PCS}表示目标专利在原专利网络中添加m件专利后的排名变化，如$\Delta R_{I_n}=R_{I_n}(T_0)-R_{I_n}(T_m)$。

目标专利T的排名变化越小，说明识别模型的稳健性越强。随着新专利加入原专利网络，目标专利在各识别模型中的排名变化如图7所示。由图7可见，基于入度中心性的指标模型最容易受到虚假引用的影响。随着新专利的增加，目标专利的入度中心性排名变化较大。特征向量中心性与入度中心性相比提高了稳健性。相比之下，PageRank和本文所提的PCDRank和PCSRank模型均表现出更强的稳健性。当更多的新专利加入到原专利网络中时，目标专利的排名变化相对较小。此外，图7还验证了各识别模型在不同技术领域中的适用性。由此可以发现，PCDRank和PCSRank模型在不同技术领域中均表现出较强的稳健性。

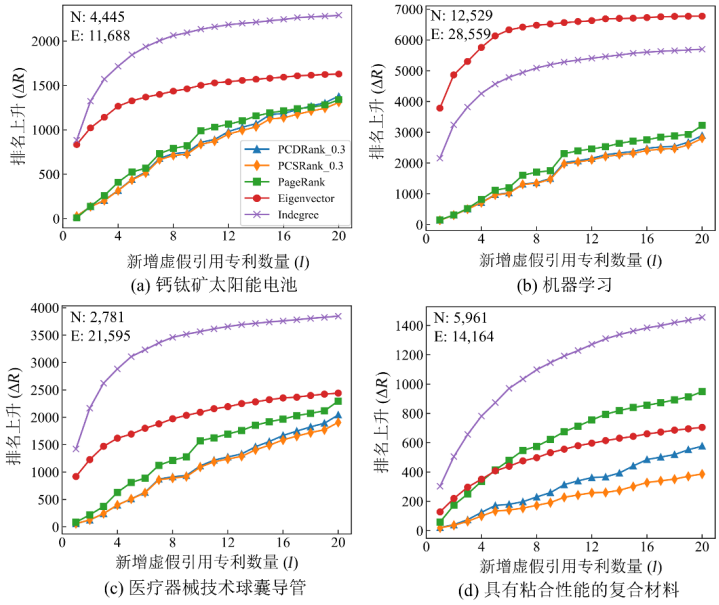


图7　各指标对虚假引用行为的稳健性

Fig. 7　Robustness of each index to false citation behavior

4　结束语

本文探索了基于高阶网络结构的核心专利识别方法，丰富了专利网络分析的理论与方法体系，提升了核心专利识别的精度与稳健性。在理论贡献方面，通过引入高阶网络理论，拓展了专利网络分析的观察尺度，揭示了专利网络中三元及以上节点间的复杂交互模式，深化了对技术知识流动与演化机制的理解。在方法贡献方面，提出了基于前馈循环模体的专利重要性评价模型（PCDRank和PCSRank），通过结合高阶结构特征与全局网络拓扑特性，构建了更加精准的核心专利识别框架，并利用信息传播动力学模型验证其有效性与稳健性。本文从复杂网络科学视角丰富了专利网络分析方法，并得到以下结论。

1）专利引用网络的高阶结构包含了更为丰富的结构信息，有助于增加专利分析信息量，优化识别效果。本文考虑了专利网络中高阶结构对专利重要性的影响，前馈循环模体结构能够捕捉三元及以上专利间的复杂交互模式。实证分析结果表明，高阶网络在结构层面对原始网络进行了“压缩”，剔除了未参与模体的边缘专利，使得保留下来的专利集中于互动密切的知识单元，使得核心专利在网络社区内部呈现出密集分布特征。

2）基于高阶网络结构的核心专利识别方法能获得更为准确的识别结果。本文将专利网络的高阶结构中的前馈循环模体结构特征与专利重要性相关联，并提出了两种基于前馈循环模体的专利重要性评估模型。两种模型既能反映网络的全局拓扑特性，又能捕捉到高阶结构的影响。理论分析和实证结果表明，所提方法能够反映专利技术关系价值，并取得较为准确的识别结果，丰富了专利网络分析中对有向网络的分析方法。

3）专利网络的高阶结构提高了识别模型的稳健性。本文使用复杂网络信息传播动力学模型计算专利的实际传播影响力，进而验证了PCDRank和PCSRank模型的有效性。通过对蓄意操纵的专利引用网络进行仿真建模，揭示了高阶网络结构对专利网络稳健性的影响。实证研究发现，两种模型在不同技术领域中均表现出较强的稳健性。当前专利网络的演化机制研究尚停留在理论层面，本研究尝试分析了专利网络上的动力学过程，有助于理解专利信息复杂系统，并为后续揭示专利网络演变规律奠定理论和方法基础。

同时，本研究仍存在以下局限性与不足：一方面，本文所提的PCDRank和PCSRank模型本质是对PageRank算法的改进，虽能在捕捉高阶结构信息的同时反映专利网络的全局拓扑特性，但也具有原始算法的局限性，如模型易使较老的专利获得更高的排名，从而失去了对较新专利的公平性，未来可结合修正“年龄偏差”的方法进一步优化识别模型；另一方面，本文仅从前馈循环模体探索了高阶网络结构对核心专利识

别的影响，未来可继续探索其他类型的模体或是不同形式表示的高阶网络（如超图或单纯型）对专利网络结构与功能的影响。□

参考文献

[1] 王康，苏盖美，李逸飞，等. 国内和全球创新网络的协同嵌入与企业技术赶超［J］. 经济研究，2025，60（3）：106-123.（WANG Kang，SU Gaimei，LI Yifei，et al. Effects of synergistic embedding of domestic and global innovation networks on enterprise technological Catch-Up［J］. Economic Research Journal，2025，60（3）：106-123.）

[2] 冉征，郑江淮. 技术范式下的创新发展：技术多样化与技术专业化的经济增长效应研究［J］. 管理世界，2024，40（9）：1-20，59，21-24.（RAN Zheng，ZHENG Jianghuai. Innovative development in the technological paradigm：a study of the economic growth effects of technological diversification and technological specialisation［J］. Management World，2024，40（9）：1-20，59，21-24.）

[3] 邓三鸿，郭剑明，施宇杰. 重塑未来：颠覆性技术主题研究进展及未来展望［J］. 科技情报研究，2025，7（2）：1-12.（DENG Sanhong，GUO Jianming，SHI Yujie. Reshaping future：research progress and future prospects of disruptive technology topics［J］. Scientific Information Research，2025，7（2）：1-12.）

[4] 郭剑明，王婧怡，袁润. 综合属性指标和引用关系的核心专利识别方法研究［J］. 情报学报，2024，43（5）：575-587.（GUO Jianming，WANG Jingyi，RUAN Run. Core patent identification method integrating attribute indicators and citation relationships［J］. Journal of the China Society for Scientific and Technical Information，2024，43（5）：575-587.）

[5] 郭剑明，王婧怡，袁润. 综合属性和关系的专利网络社区发现和核心专利识别研究［J］. 情报科学，2024，42（11）：120-127.（GUO Jianming，WANG Jingyi，RUAN Run. Patent network community discovery and core patent recognition based on comprehensive attribute and relationship［J］. Information Science，2024，42（11）：120-127.）

[6] 郭剑明，王婧怡，周云峰，等. 基于专利网络节点重要性的核心专利识别方法研究［J］. 情报杂志，2023，42（5）：162-168，191.（GUO Jianming，WANG Jingyi，ZHOU Yunfeng，et al. Research on the core patents identification method based on node importance of patents network［J］. Journal of Intelligence，2023，42（5）：162-168，191.）

[7] 关鹏，王曰芬. 国内外专利网络研究进展［J］. 数据分析与知识发现，2020，4（1）：26-39.（GUAN Peng，WANG Yuefen. Advances in patent network［J］. Data Analysis and Knowledge Discovery，2020，4（1）：26-39.）

[8] 杨冠灿，陈亮，张静，等. 专利引用关系形成的解释框架：一个指数随机图模型视角［J］. 图书情报工作，2019，63（5）：100-109.（YANG Guancan，CHENG Liang，ZHANG Jing，et al. Framework for explanations of patent citation formation：an exponential random graph model perspective［J］. Library and Information Service，2019，63（5）：100-109.）

[9] YANG Rongmei，LIU Bo，LÜ Linyuan. Simplicial motif predictor method for higher-order link prediction［J］. Expert Systems with Applications，2025，269：126284.

[10] 肖沪卫，顾震宇. 专利地图方法与应用［M］. 上海：上海交通大学出版社，2011：148.（XIAO Huwei，GU Zhenyu. Patent map methods and applications［M］. Shanghai：Shanghai Jiao Tong University Press，2011：148.）

[11] 李治东，熊焰，方曦. 基于熵权层次分析法的核心专利识别应用研究［J］. 情报学报，2016，35（10）：1101-1109.（LI Zhidong，XIONG Yan，FANG Xi. Application research of identifying core patent base on entropy weight of analytic hierarchy process method［J］. Journal of the China Society for Scientific and Technical Information，2016，35（10）：1101-1109.）

[12] 李姝影，张鑫，许轶，等. 核心专利集筛选及专利技术主题识别影响［J］. 情报学报，2019，38（1）：13-20.（LI Shuying，ZHANG Xin，XU Yi，et al. Patent screening of core documents and impact of patent technology subject identification［J］. Journal of the China Society for Scientific and Technical Information，2019，38（1）：13-20.）

[13] 温迎强，杨伯军，万鹏，等. 基于灰色模糊理论的核心专利识别方法研究［J］. 科技管理研究，2022，42（2）：172-183.（WEN Yingqiang，YANG Bojun，WAN Peng，et al. Research on the identification method of core patents based on grey fuzzy theory［J］. Science and Technology Management Research，2022，42（2）：172-183.）

[14] 巩永强，王超，王锐，等. 复杂网络视角下的核心专利识别研究［J］. 情报理论与实践，2022，45（10）：103-113.（GONG Yongqiang，WANG Chao，WANG Rui，et al. Study of identification of core patents in the perspective of complex network［J］. Information Studies：Theory & Application，2022，45（10）：103-113.）

[15] 陈祥，冯佳，穆晓敏，等. 技术知识扩散视角下核心专利识别方法研究［J］. 情报理论与实践，2022，45（10）：132-138.（CHEN Xiang，FENG Jia，MU Xiaomin，et al. Study of identification of core patent in the perspective of diffusion of technology knowledge［J］. Information Studies：Theory & Application，2022，45（10）：132-138.）

[16] BATTISTON F，AMICO E，BARRAT A，et al. The physics of higher-order interactions in complex systems［J］. Nature Physics，2021，17（10）：1093-1098.

[17] BIANCONI G. Higher-order networks［M］. Cambridge：Cambridge University Press，2021.

[18] 李江，刘影，王伟，等. 识别高阶网络传播中最有影响力的节点［J］. 物理学报，2024，73（4）：48901.（LI Jiang，LIU Ying，WANG Wei，et al. Identifying influential nodes in spreading process in higher-order networks［J］. Acta Physica Sinica，2024，73（4）：48901.）

（下转第149页）

velopment trend of key technologies for lithium extraction from salt lakes ［J］. Modern Chemical Industry，2024，44（4）：6-12，18.）

［28］蒋晨啸，陈秉伦，张东钰，等. 我国盐湖锂资源分离提取进展［J］. 化工学报，2022，73（2）：481-503.（JIANG Chenxiao，CHEN Binglun，ZHANG Dongyu，et al. Progress in the separation and extraction of lithium resources from salt lakes in China ［J］. CIESC Journal，2022，73（2）：481-503.）

［29］国务院关于印发《中国制造2025》的通知［EB/OL］.［2025-08-03］. http：//www. gov. cn/zhengce/content/2015-05/19/content_9784. htm.

［30］邓同兴，王敏丽，陈彦锋，等. 氯化锂—匹罗卡品诱发癫痫持续状态小鼠海马星形胶质细胞改变［J］. 中国药物经济学，2023，18（3）：106-108.（DENG Tongxing，WANG Minli，CHEN Yanfeng，et al. Changes in hippocampal astrocytes in mice with status epilepticus induced by lithium chloride-pilocarpine ［J］. Chinese Journal of Pharmaceutical Economics，2023，18（3）：106-108.）

［31］廖明生，肖峰，王淑坤，等. 碳酸锂对癫痫性精神病患者攻击性的疗效研究［J］. 药品评价，2022，19（24）：1504-1506.（LIAO Mingsheng，XIAO Feng，WANG Shukun，et al. Efficacy study of lithium carbonate on aggression in patients with epileptic psychosis ［J］. Drug Evaluation，2022，19（24）：1504-1506.）

［32］工业和信息化部　国家发展和改革委员会　科学技术部　商务部　国家卫生健康委员会　应急管理部　国家医疗保障局　国家药品监督管理局　国家中医药管理局关于印发“十四五”医药工业发展规划的通知［EB/OL］.［2025-08-03］. https：//www. gov. cn/zhengce/zhengceku/2022-01/31/content_5671480. htm.

作者简介：蔡鑫，男，1996年生，硕士生。研究方向：技术预见。**靳军宝**（通信作者，Email：jinjb@llas.ac.cn），男，1986年生，博士，副研究馆员，硕士生导师。研究方向：产业技术情报与知识产权分析，科技人才流动与评价。**王政媛**，女，2000年生，硕士生。研究方向：技术预见。**郑玉荣**，男，1974年生，博士，副研究馆员。研究方向：产业技术分析与区域发展咨询。**白光祖**，男，1984年生，博士，研究员，博士生导师。研究方向：技术情报挖掘，产业发展战略与区域创新体系。**曹琨**，男，1993年生，博士，助理研究员。研究方向：知识发现与技术识别。**孙凡雅**，女，2003年生，硕士生。研究方向：技术预见。

作者贡献声明：蔡鑫，开展调研，方法设计，开展实验，论文撰写。**靳军宝**，论文选题与设计，指导论文写作、修改、定稿。**王政媛**，论文资料收集与整理。**郑玉荣**，论文思路指导。**白光祖**，论文思路指导。**曹琨**，论文方法指导。**孙凡雅**，前期调研。

录用日期：2025-08-25

（上接第114页）

［19］刘波，曾钰洁，杨荣湄，等. 高阶网络统计指标综述［J］. 物理学报，2024，73（12）：128901.（LIU Bo，ZENG Yujie，YANG Rongmei，et al. Fundamental statistics of higher-order networks：a survey［J］. Acta Physica Sinica，2024，73（12）：128901.）

［20］郭剑明，王婧怡，袁润. 基于网络快照的核心专利预测方法研究［J］. 情报理论与实践，2024，47（6）：166-174.（GUO Jianming，WANG Jingyi，YUAN Run. Research on core patent prediction method based on network snapshot［J］. Information Studies：Theory & Application，2024，47（6）：166-174.）

［21］FRANCESCHET M，COLAVIZZA G. Quantifying the higher-order influence of scientific publications ［J］. Scientometrics，2020，125（2）：951-963.

［22］BICK C，GROSS E，HARRINGTON H A，et al. What are higher-order networks? ［J］. SIAM Review，2023，65（3）：686-731.

［23］LI Meng，WANG Zhigang，ZENG An，et al. Higher-order structure based node importance evaluation in directed networks［J］. Information Processing & Management，2025，62（1）：103948.

［24］FELMLEE D，MCMILLAN C，WHITAKER R. Dyads，triads，and tetrads：a multivariate simulation approach to uncovering network motifs in social graphs ［J］. Applied Network Science，2021，6（1）：63.

［25］杨冠灿，李纲，刘彤. 基于关系融合的专利网络结构分析研究［M］. 北京：科学技术文献出版社，2019.（YANG Guancan，LI Gang，LIU Tong. Analysis and research of patent network structure based on relationship fusion［M］. Beijing：Scientific and Technical Documentation Press，2019.）

［26］MORI F，OKADA T. Information-transfer characteristics in network motifs ［J］. Physical Review Research，2023，5（1）：13037.

［27］NEWMAN M. Networks［M］. Oxford：Oxford University Press，2018.

［28］NEWMAN M. Message passing methods on complex networks ［J］. Proceedings of the Royal Society A，2023，479（2270）：20220774.

作者简介：郭剑明（ORCID：0000-0001-6343-6547），博士生。研究方向：信息分析与复杂网络。**杨杰**（ORCID：0000-0002-1276-0483），博士生。研究方向：科学评价与计量。**方静**，馆员。研究方向：信息管理与科学评价。**邓三鸿**（ORCID：0000-0002-6910-3935，通信作者，Email：sanhong@nju.edu.cn），博士，教授，博士生导师。研究方向：科学评价与信息检索。

作者贡献声明：郭剑明，分析数据，撰写与修改论文。**杨杰**，处理数据，论文修改。**方静**，清洗数据，参与提纲讨论。**邓三鸿**，提出研究思路，指导论文写作。

录用日期：2025-08-18