情报学报　2026 年 8 月　第 45 卷　第 8 期
Journal of the China Society for Scientific and Technical Information, Aug. 2026, 45(8): 1092-1105

DOI: 10.3772/j.issn.1000-0135.2026.08.002

基于单纯形的专利合作高阶拓扑结构测度及创新产出关系研究

郭剑明[1,2]，杨　杰[1,2]，宋欣雨[1,2]，刘　倩[1,2]，邓三鸿[1,2]

（1. 南京大学信息管理学院，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），南京　210023）

摘　要　在团队化与协作化日益加深的创新环境下，如何从结构层面刻画专利合作网络并揭示其对创新产出的影响，已成为创新情报研究中的一项重要议题。现有研究通常基于二元关系视角，难以刻画专利合作中普遍存在的群体协作结构。本文引入代数拓扑中的高阶网络表示方法，系统刻画专利合作网络中的多主体互动结构，并探究其高阶结构特征与发明人创新产出间的关系。基于PatentsView数据库中700余万项专利数据，本文构建了专利合作网络的单纯形模型，并采用拓扑量化方法测度网络的高阶结构特征。其中，互不连通组件数量β_0用于表征合作网络的基础连通性，高阶环路β_1用于刻画无法由低阶协作填充的群体协作模式。在此基础上，构建“发明人-申请年份”面板数据，采用负二项固定效应回归模型，系统检验高阶结构特征与技术生产率、专利被引数量、专利5年被引数量及专利权利要求数等多维创新产出指标间的关系。研究结果表明，专利合作网络中，互不连通组件数量β_0与4种创新产出指标之间呈现倒U形关系，即在合理区间内，组件数量的增加与创新产出水平呈现正向关联，但当组件数量处于较高水平时，其与创新产出的关联强度呈现减弱趋势；高阶环路β_1与4种创新产出指标均呈现显著正相关关系，表明高阶协作结构与知识整合能力及技术创新表现之间存在密切联系。稳健性分析结果显示，上述结论在不同技术发展阶段及不同技术领域中均保持稳定，具有较强的普适性。

关键词　专利合作；高阶协作结构；创新产出；单纯形；贝蒂数

Research on Measurement of High-order Topology Structure of Patent Cooperation Based on Simplex and Its Association with Innovation Output

Guo Jianming[1,2], Yang Alex Jie[1,2], Song Xinyu[1,2], Liu Qian[1,2] and Deng Sanhong[1,2]

(1. *School of Information Management, Nanjing University, Nanjing*　210023; 2. *Key Laboratory of Data Engineering and Knowledge Services in Provincial Universities (Nanjing University), Nanjing*　210023)

Abstract: In an innovation environment where teamwork and collaboration are deepening, describing the patent cooperation network from the structural level and revealing its influence on innovation output have become important topics in innovation intelligence research. Existing research is mostly based on the perspective of binary relationships; therefore, it is difficult to describe the common group cooperation structure in patent cooperation. This paper introduces a high-order net-

收稿日期：2025-12-17；**修回日期**：2026-03-23

基金项目：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范式变革预测研究”（25ATQ008）。

作者简介：郭剑明，1996年生，博士研究生，主要研究领域为专利情报、复杂网络；杨杰，2000年生，博士，博士后，主要研究领域为科学计量、创新评价、复杂网络分析；宋欣雨，1997年生，博士研究生，主要研究领域为科学计量与学术评价；刘倩，1996年生，博士研究生，主要研究领域为科学计量、信息分析；邓三鸿，通信作者，1975年生，博士，教授，博士生导师，主要研究领域为领域知识图谱、科学计量、信息分析与科学评价，E-mail：sanhong@nju.edu.cn。

work representation method in algebraic topology, systematically describes the multi-agent interaction structure in the patent cooperation network, and examines the association between its higher-order structural features and inventors' innovative outputs. Based on more than 7 million patent records in the PatentsView database, this paper constructs a simplex complex model of the patent cooperation network and uses a topological quantization method to measure the high-order structural characteristics of the network. Among them, the number of disconnected components β_0 is used to characterize the basic connectivity of the cooperative network, and the high-order loop β_1 is used to characterize the group cooperation mode that cannot be filled by low-order cooperation. On this basis, this study constructs an inventor−application year panel dataset and employs negative binomial fixed-effects regression models to systematically assess how higher-order structural features are associated with multiple dimensions of innovative output, including technological productivity, total patent citations, five-year patent citations, and the number of patent claims. The results show that there is an inverted U-shaped relationship between the number of disconnected components β_0 in the patent cooperation network and the four innovation output indicators; that is, within a reasonable range, an increase in the number of components is positively correlated with the level of innovation output, but when the number of components is high, the correlation with innovation output shows a weakening trend. The high-order loop β_1 is positively correlated with the four innovation output indicators, suggesting that higher-order collaborative structures are closely linked to knowledge integration capabilities and technological innovation performance. The robustness analysis shows that the above conclusions are stable in different technical development stages and different technical fields and have strong universality.

Keywords: patent cooperation; high-order collaborative structure; innovation output; simplex; Betti number

0　引　言

随着现代科学技术的迅猛发展，创新活动的复杂性与研发成本持续攀升，标志着“大科学”时代正式来临[1]。知识积累的指数级增长带来了沉重的“知识负载”，单一主体难以独自掌握跨学科的全面知识，科研团队合作由此成为应对挑战、实现技术突破的核心形式[2]，合作网络也由此成为解析团队结构与科学协作的重要工具[3]。创新主体之间的互动正由简单的成对合作，逐步演化为多主体参与、跨团队协同的群体协作模式[4]。如何从结构层面理解这种复杂协作关系，并揭示其对创新绩效的作用机制，不仅是情报学研究中关于科学合作与知识生产的重要理论议题[5]，也是优化创新团队组织、提升科技创新效率的现实政策关切[6]。专利合作网络不仅能够精准呈现创新主体间的关联形态，更为揭示合作结构与创新绩效的内在逻辑、解析技术创新的演化规律提供了关键支撑[7]。因此，基于专利合作网络的创新产出关系研究，不仅有助于解构团队结构的深层特征，更是揭示技术协作机制、预测并提升创新绩效的核心途径。

现有研究在合作网络的拓扑特性（如无标度、小世界性）及其与创新产出间的关系方面取得了丰硕成果，但仍局限于基于图论的二元（pairwise）视角[8]。这种二元互动框架难以充分刻画真实创新活动中普遍存在的群体协作关系[9]。以多名发明人共同参与同一专利为例，传统成对关系表示会将高阶协作关系简化为若干二元连接，进而丢失团队内部蕴含的高维拓扑信息。近年来，网络科学的发展催生了如单纯形（simplex）[10]和超图（hypergraph）[11]等高阶表示方法，在捕捉多元主体互动方面具有独特优势，并在多个领域得到成功应用[12]。尽管此类方法已初步引入科学合作分析[3,13]，但在系统考察高阶结构与个体创新产出关系的研究层面仍显不足，且已有结论的普适性有待验证。高阶网络不仅能够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[14]。然而，现有研究鲜有将这一前沿方法应用于专利合作网络，导致对高阶拓扑结构与技术创新产出之间的关系缺乏深入的量化认知。

基于上述背景，本文旨在探索基于高阶网络结构的专利合作网络测度方法，并检验高阶结构特征与创新产出间的关系，以期丰富现有专利合作网络分析的理论与方法。在理论贡献方面，本文深化了创新网络研究的结构视角，将高阶拓扑引入专利合作研究，拓展了对创新协作机制的理解边界，深化了对群体协作关系的理解。在方法贡献方面，本文构建并应用了基于代数拓扑的高阶网络表示与量化方法，将专利合作网络抽象为单纯复形结构，并结合负二项回归模型，检验高阶拓扑特征与发明人创新产出间的关系，同时开展多维稳健性分析以验证结论的稳定性。从政策价值与意义来看，本文揭示了高阶协作结构与创新产出之间的正向关联，为研发团队构建、创新资源配置与科技政策制定提供实

证依据和新的决策参考。

1 相关研究

1.1 合作网络结构及创新产出影响

随着创新活动的复杂性不断提高、技术分工进一步细化，合作已成为推动技术进步的重要组织方式。团队科学研究的兴起，推动了对合作网络结构及其影响机制的深入探索。在宏观层面，已有研究系统地揭示了合作网络普遍具有的无标度、小世界及强社区结构等拓扑特征[15]。与此同时，分析工具的进步使得网络模型从基础的二元结构，不断向加权[16]、多层[17]、动态[2]等更精细的维度拓展，从而能够对协作模式进行更细微的刻画。相关研究表明，个体或团队在网络中的结构位置对其创新表现具有显著影响[18]。节点中心性[19]、连接强度[5]、结构洞[20]等传统网络指标被证实与科研人员（或发明人）的生产力、论文或专利影响力密切相关。在创新研究领域，针对专利数据的分析表明，占据协作网络结构洞的发明人能增加创新绩效，而邻域度中心性具有负向作用[21]；连接状态与连接强度对专利颠覆性有显著负向影响，而对专利影响力和专利市场价值则有正向效应[5]。在微观层面，相关研究进一步深入到团队内部构成与互动特征对创新产出的具体影响。团队的新鲜度，如新成员或新合作关系的引入，被证实能够有效预测成果的跨学科性与颠覆性[22]。团队的知识交叉程度对科研主体下一阶段的科研产出数量呈现倒U形的影响关系[23]。团队多样性（包括年龄、性别、机构、学科、地理距离等方面）被发现是提升论文或专利影响力的重要因素[2,24-26]。这些研究从不同角度揭示了细粒度的团队结构属性是驱动高质量创新产出的微观因素。

1.2 科学结构分析中的高阶网络表示

尽管上述研究取得了丰硕成果，但传统合作网络建模主要基于二元互动，将合作关系简化为个体间的成对连接，未能充分刻画现实中普遍存在的、涉及3个及以上主体的“群体互动”[3]。这种局限可能忽略了对系统功能至关重要的高阶协作模式[9]。为此，网络科学与代数拓扑领域提出了一系列高阶网络表示形式，其中最具代表性的包括单纯形[10]与超图[11]。这些高阶表示能够显式刻画参与同一合作事件的多主体关系，捕捉传统二元网络无法捕捉的复杂结构[27-28]。在科学研究领域，已有学者尝试将高阶网络表示引入结构分析。相关研究考虑网络中高阶结构对节点重要性的影响，构建了基于高阶结构的有向网络中节点重要性的评估框架[29]。还有学者探索使用单纯形描述科学合作网络的形状，证明了合作网络中强单纯形闭包的存在[13]。超图模型也被用来分析科研影响力和颠覆性创新之间的作用关系[30]。此外，相关研究使用单纯复形捕捉科学家之间的群体互动，探究了局部协作网络中高阶结构对科学家生产力的影响[3]。这些研究为理解群体协作的复杂性、识别合作中的结构性模式提供了新的理论与方法工具。

综上，当前研究充分探讨了宏观与微观层面团队结构特征与创新产出间的关系。然而，当前合作网络研究主要基于合作者间的二元关系。尽管少数研究涉及高阶结构特征，但关于专利合作网络拓扑结构（如高阶环路）与创新产出的具体关系仍缺乏系统性的分析论证。鉴于此，本文旨在引入高阶网络方法测度专利合作网络，探究高阶结构特征与创新产出的关联。

2 研究设计

2.1 专利合作网络构建及高阶结构测度

专利合作网络可分为专利权人合作网络、发明人合作网络、申请人合作网络、机构合作网络、地域或国家合作网络等[31]，本文以发明人为节点，合作者间的协作关系为边构建专利合作网络。引入代数拓扑中的单纯形理论，以克服传统图论方法在描述群体协作关系时的局限性，并利用拓扑学中的贝蒂数（Betti number，β_k）量化并测度合作网络的高阶结构特征。

2.1.1 专利合作网络与单纯形

单纯形网络是常用的高阶网络表示方式，它基于代数拓扑的单纯复形（simplicial complex）思想，用单纯形来表示高阶网络中的节点交互关系。在传统图论中，合作关系被抽象为两个节点之间的连边，仅能表示简单成对关系[18]。而单纯形理论允许将3个及以上主体的共同协作关系表示为一个高维单形[13]。例如，3位发明人共同申请一项专利可表示为一个2-单形（三角形），4位发明人则对应一个3-单形（四面体）。与表示成对关系的简单连边相比，这种表示方式能够保留群体协作的整体性。

本文构建专利合作网络并抽象为单纯形的整体过程如图 1 所示。为聚焦对发明人创新产出具有最直接影响的核心协作环境，本文构建并分析局部专利合作网络。首先，基于目标发明人在前 5 年，即 $(t-5)$至$(t-1)$年的合作记录构建其局部专利合作网络，描绘其在 t 年的协作环境。通过将构建时间窗口设定在产出年份之前，保证了回归分析的时序逻辑。图 1a 展示了发明人在 t 年的直接合作者集合；在此基础上，追溯这些合作者在时间窗口内是否存在共同的专利申请，如图 1b 所示。具体而言，识别发明人在 t 年的所有直接合作者，并追溯这些合作者在窗口期内的共同专利申请，从而形成仅包含该发明人“一阶邻域”的合作结构。这样的“局部网络”能够捕捉直接影响发明人知识流动与团队组织的微观互动关系[2]。

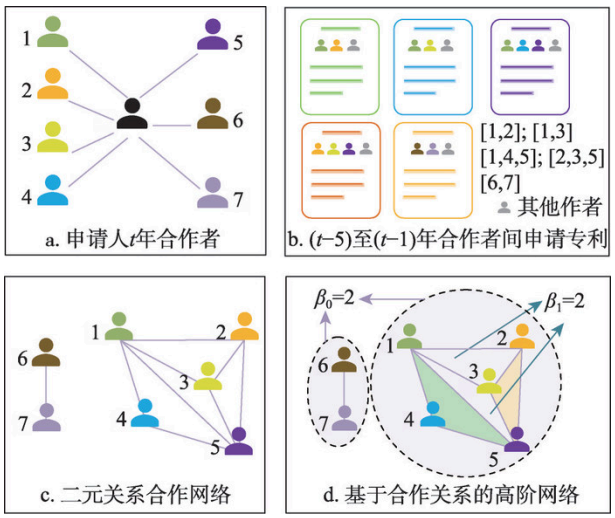


图 1　专利合作网络高阶结构测度示意

在获得这些合作关系后，可进一步将合作者之间的合作关系构建为图 1c 所示的二元关系合作网络。为准确表征专利合作中普遍存在的多人合作关系，使用单纯形抽象化表示局部专利合作网络，通过将 $(d+1)$ 个共同合作的发明人表示为 d-单形，使得图 1d 所示的高阶合作结构能够被完整保留。传统网络仅能以 1-单形（边）表示两个发明人之间的二元合作，而单纯复形则能利用 d-单形表示 $(d+1)$ 个主体之间的共同合作，从而保留了群体协作的完整信息[3]。单纯复形 γ 的定义要求其满足“面封闭性”，即 $\forall \alpha \in \gamma$，其所有面 $\alpha' \subset \alpha$ 均有 $\alpha' \in \gamma$。这种严格的数学定义确保了网络中低维（如两两合作）与高维（如 3 人共同合作）协作关系的层级嵌套性[32]。更重要的是，单纯复形能够明确区分“仅有两两合作”与“多人共同合作”这两种不同的拓扑结构，从而编码出传统方法无法识别的“高阶孔洞”[33]，这对于量化协作网络的冗余性与知识整合程度十分重要。

2.1.2　专利合作网络的高阶结构测度

在构建单纯复形后，使用贝蒂数（β_k）对局部专利合作网络的高阶拓扑特征进行量化。在现有科学合作高阶网络[3]与同调理论[13]相关研究的基础上，本文将贝蒂数这一高阶拓扑测度指标引入专利合作网络研究。贝蒂数是一种拓扑不变量，直观地反映了单纯复形在不同维度上“孔洞”（holes）结构[34]。贝蒂数是基于同调理论（homology theory）计算得出的，其中，k 维同调群 $H_k(K)$ 的秩为 β_k，它精确地量化了不能被更高维结构填充的独立 k-维孔洞的数量[35]。使用贝蒂数能够超越传统网络指标，直接捕捉由大于两位发明人协作关系所构成的 $k \geqslant 1$ 维高阶拓扑结构。

设 K 为一个有限单纯复形，$C_k(K)$ 表示其第 k 阶链空间，$\partial_k: C_k(K) \to C_{k-1}(K)$ 为对应的边界算子。第 k 阶同调群为 $H_k(K) = \ker\partial_k / \mathrm{im}\partial_{k+1}$，第 k 阶贝蒂数 β_k 定义为该同调群的秩，其计算公式为

$$\beta_k = \mathrm{rank}(H_k(K)) = \dim(\ker\partial_k) - \dim(\mathrm{im}\partial_{k+1}) \quad (1)$$

在具体计算中，上式可等价写为

$$\beta_k = m_k - r_k - r_{k+1} \quad (2)$$

其中，m_k 表示 k-单纯形的总数；$r_k = \mathrm{rank}(\partial_k)$ 表示第 k 阶边界算子的秩；$r_{k+1} = \mathrm{rank}(\partial_{k+1})$ 表示第 $(k+1)$ 阶边界算子的秩。r_k 描述了由 k-单纯形边界所张成的 $(k-1)$ 维子空间的维数，而 r_{k+1} 则刻画了能够由 $(k+1)$-单纯形所填充的 k-维循环的数量。因此，β_k 衡量了在去除可被填充的循环之后，复形中本质上独立的 k-维拓扑结构数量。

合作网络结构中拓扑不变量的存在较为稀疏，即 k 值越高，β_k 存在的可能性越低，故本文考虑使用 β_0 和 β_1 两个常用的贝蒂数。

零阶贝蒂数（β_0）：表示单纯复形中互不连通组件的数量，即合作网络被划分为多少个彼此独立的子结构。在局部专利合作网络中，较高的 β_0 意味着合作者之间整体连通性较弱，合作关系呈现碎片化特征；较低的 β_0 则表明合作子群之间存在更紧密的连接。图 1d 中虚线包围的部分体现了两个不同连通组件的存在。

一阶贝蒂数（β_1）：表示单纯复形中 1-维孔洞

的数量，即由边围成但未被更高维单形填充的闭合路径。在局部专利合作网络中，这类结构通常对应若干发明人之间存在循环式合作连接，但未形成完全重叠的多人共同合作关系。从协作组织结构角度来看，β_1反映了合作网络中跨合作子群的循环连接程度。较高的β_1意味着网络中存在多路径协作闭环，有助于知识在不同团队之间流动与整合，从而增强网络的结构弹性与信息整合能力。图1d右侧所示结构中未被着色的两个三角形结构存在典型的1-维孔洞。

高阶结构已被证实在复杂系统中具有重要功能意义，类似的机制在专利合作网络中同样可能存在。因此，通过β_0和β_1对专利合作网络进行高阶拓扑度量，有助于揭示发明人合作环境的深层结构特征及其与创新活动的潜在关系，从而深化对合作网络“结构-功能”的理解。

2.2 发明人创新产出及关系分析

为分析局部专利合作网络的高阶结构特征与发明人创新产出之间的关系，构建基于“发明人-申请年份”的面板数据集，并通过负二项回归模型估计高阶拓扑结构与创新产出之间的关系。

2.2.1 回归分析变量

技术创新具有多维属性，仅衡量创新产出的数量往往难以全面反映发明人的真实绩效。与现有研究[18]一致，本文从“产出数量”与“产出质量”两个维度选取因变量指标，如表1所示。

表1 回归模型的变量描述

类型	变量	符号表达	变量描述
因变量	技术生产率☆	productivity	发明人在观测年份t申请的专利总量
	被引数量	citation	发明人t年申请专利被引总次数
	5年被引数量	citation5	发明人t年申请专利在公开后5年内被引次数
	权利要求数	claim	发明人t年申请所有专利的权利要求项总和
自变量	零阶贝蒂数	β_0	网络中互不连通组件的数量
	一阶贝蒂数	β_1	网络中高阶环路(无法被群体合作填充的空心协作结构)的存在性;二元变量
控制变量	网络密度	density	反映合作者间整体关联紧密程度
	平均连接强度	AvgTie	发明人与每位直接合作者在(t−5)至(t−1)年的平均合作专利数
	协作强度	CollabStr	(t−5)至(t−1)年合作者之间的合作专利数与所有合作者同期专利总数之比
	职业年龄	career age	发明人从首次申请专利年份至观测年份t的时间跨度
	滞后产出	lagged productivity	发明人上次有专利申请记录年份的技术生产率
	网络规模	network size	专利合作网络中的发明人数量

注:“☆”表示核心因变量。受篇幅影响,表4和表5仅罗列核心因变量(技术生产率)的回归模型及其结果,其他因变量结果如图5所示。

首先，将技术生产率（productivity）作为核心因变量，定义为发明人在观测年份t申请的专利总量，以直观衡量其技术创新活跃度。其次，为捕捉创新的质量与技术价值，引入3个补充指标：专利被引数量（citation，发明人t年申请专利的被引总次数）、专利5年被引数量（citation5，t年申请专利在未来5年内的被引次数）以及专利权利要求数（claim，发明人t年申请专利的权利要求总数）。既往研究表明，被引次数衡量专利在产业界与技术领域的知识扩散程度，而权利要求数则被广泛视为专利价值、技术复杂度与法律强度的稳定指标，能够表征专利的质量[5,18,21]。通过构建数量与质量并重的4个因变量，旨在更全面地识别和检验专利合作网络高阶拓扑结构对发明人创新产出的多维影响。

在自变量方面，使用贝蒂数刻画局部合作网络的高阶结构特征（表1）。零阶贝蒂数β_0用于衡量网络中互不连通组件的数量，体现合作网络的碎片化程度；一阶贝蒂数β_1表征网络中高阶环路的存在性，反映无法由团队合作填充的空心协作结构。鉴于实际观测中复杂高阶孔洞极为稀疏，本文将β_1简化为二元变量，用于识别高阶协作闭环的边际效应。

其次，为确保高阶结构效应不被传统网络特征或个体差异混淆，本文纳入多类控制变量（表1）。在网络层面，包括网络密度（network density）、平均连接强度（average tie strength）和协作强度（collaborative strength）。在发明人i的局部协作网络中，网络密度用于刻画其合作者之间整体关联的紧密程度，计算公式为

$$\text{density}_t = \frac{2E_t}{N_t(N_t - 1)} \tag{3}$$

其中，N_t表示发明人在$(t-5)$至$(t-1)$时间窗口内的直接合作者数量；E_t表示该合作者子网络中实际协作数量。该指标取值介于0和1之间，数值越大，表明合作者之间的联系越紧密。

平均连接强度用于衡量发明人与其合作者之间合作关系的平均深度，定义为

$$\mathrm{AvgTie}_t = \frac{1}{|\mathcal{N}_t|}\sum_{j\in\mathcal{N}_t} w_{ij}^{(t-5,t-1)} \tag{4}$$

其中，$\mathcal{N}_t$表示发明人在$(t-5)$至$(t-1)$时间窗口内的合作者集合；$w_{ij}^{(t-5,t-1)}$表示发明人i与合作者j在该时间窗口内的合作专利数量。该指标反映了发明人与其直接合作者之间平均合作频次，数值越大，表明合作关系越稳定，协作强度越高。

协作强度用于刻画发明人合作活动在其合作者整体创新产出中的相对重要性，定义为

$$\mathrm{CollabStr}_t = \frac{\sum_{j\in\mathcal{N}_t} w_{ij}^{(t-5,t-1)}}{\sum_{j\in\mathcal{N}_t} p_j^{(t-5,t-1)}} \tag{5}$$

其中，$p_j^{(t-5,t-1)}$表示合作者j在同一时间窗口内的专利总产出。该指标刻画了发明人i的合作行为在其合作者创新活动中的占比，用于控制整体协作强度和合作者创新活跃度的影响。

在个体层面，控制职业年龄（career age），即发明人从首次申请专利年份至观察年份t的职业时间跨度；控制滞后产出（lagged productivity），即发明人上次有专利申请记录年份的技术生产率，以控制个人产出的持续性与路径依赖。这些指标已被许多创新合作网络研究证明与创新效率及知识流动密切相关[2-3,25]。

最后，加入年份和技术领域的固定效应，本文针对1980—2017年38个年份的13个阶段（每3年划分为一个阶段）和9个CPC（cooperative patent classification）“部”所属技术领域分别生成哑变量，将其作为固定效应加入模型，以控制专利在不同时间维度、技术领域的异质性。

2.2.2　回归模型

由于所选创新产出指标（如专利数量、专利被引数量、5年被引数量及权利要求数）均为非负计数变量，且在样本中呈现明显的离散特征，本文采用负二项回归模型估计高阶结构特征与发明人创新产出的关系。相较于泊松回归模型，负二项模型允许条件方差大于条件均值，更适用于厚尾分布的计数型数据，因此，在创新研究与科学计量研究中被广泛采用[18]。

基于“发明人-申请年份”的面板数据，设$Y_{i,t}$表示发明人i在年份t的创新产出。为系统识别不同高阶结构特征的作用机制，本文采用逐步回归策略，构建如下两个回归模型。

模型1. 零阶贝蒂数高阶结构的基准效应：

$$\mathbb{E}(Y_{i,t}|\cdot) = \exp(\alpha + \gamma_0\beta_{0,i,t} + \gamma_1\beta_{0,i,t}^2 + \delta^{\mathrm{T}}\boldsymbol{X}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}) \tag{6}$$

其中，$\beta_{0,i,t}$表示发明人i在年份t对应的零阶贝蒂数高阶结构特征指标；$\boldsymbol{X}_{i,t}$为控制变量向量，涵盖网络密度、平均连接强度、协作强度、职业年龄以及滞后产出等因素；μ_i为个体固定效应，用于剔除发明人天赋、性别等不随时间变化的个体异质性；τ_t为年份固定效应，用于控制宏观技术周期、政策环境变化等共同时间冲击；$\varepsilon_{i,t}$为随机扰动项；系数γ_0反映零阶贝蒂数高阶结构特征与创新产出之间的统计关联；α为常数项；γ_1用于刻画零阶贝蒂数与创新产出之间的非线性关系；δ^{T}表示各控制变量及其回归系数的线性组合。

模型2. 在模型1的基础上，引入一阶贝蒂数$\beta_{1,i,t}$：

$$\mathbb{E}(Y_{i,t}|\cdot) = \exp(\alpha + \gamma_0\beta_{0,i,t} + \gamma_1\beta_{0,i,t}^2 + \gamma_2\beta_{1,i,t} + \delta^{\mathrm{T}}\boldsymbol{X}_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t}) \tag{7}$$

其中，$\beta_{1,i,t}$表示发明人i在年份t的一阶贝蒂数高阶结构特征指标。系数γ_1衡量在控制零阶贝蒂数高阶结构特征后，一阶贝蒂数高阶结构与创新产出的边际关联。模型1与模型2构成嵌套关系，有助于区分不同高阶结构特征在解释发明人创新产出差异中的相对作用。

上述高阶结构特征与网络控制变量均基于$\Delta t = (t-5, t-1)$的滚动时间窗口构建，以刻画发明人在当前创新产出之前的历史协作结构，从而在一定程度上缓解可能存在的内生性或反向关联问题[2]。

3　实证分析

3.1　研究数据

本文所用数据来源于美国专利商标局（United States Patent and Trademark Office，USPTO）提供的专利全文与引文记录，以及由USPTO首席经济学家办公室支持的专利数据共享平台PatentsView。USPTO专利数据库覆盖了1976—2022年7566869项专利及117308691条引用关系，构成了结构完善、质量可靠的大规模专利数据集。每项专利均包含两

个申请日期与发布日期。与既有研究一致，本文聚焦于实用新型专利（utility patent），因其具备更严格的审查程序、更高的技术新颖性与更强的信息披露规范，更适用于分析创新产出[36]。

回归分析基于按“发明人-申请年份”构造的面板数据集。若某位发明人在多个年份与不同合作者共同申请专利，则其记录将在数据中多次出现，可以反映其跨年度的协作行为。为确保面板数据分析的稳健性，本文剔除了总申请专利数少于5项且参与申请年份不足3年的发明人，从而避免使用过短或过稀疏的科研生涯数据。随后，对于每位发明人，基于其与合作者的年度交互，构建对应年份的局部合作网络。经过筛选与处理，最终回归面板数据集包含82926位发明人以及774389条年度观测记录。

3.2 描述性统计分析

表2展示了不同CPC技术领域的发明人数量及对应的观测数。从整体来看，技术领域间的发明活动呈现显著的不均衡分布。其中，“物理（G）”与“电学（H）”领域规模最大，分别包含25294与22976名发明人，显示出高度集中的创新活跃度。相对而言，“纺织；造纸（D）”与“新兴/交叉技术（Y）”的发明人规模明显较小，反映出这些领域的技术创新较为局部化。总体上，该分布揭示了技术领域在创新产出规模上的结构性差异。

表2 各CPC“部”技术领域的发明人数量

CPC	中文名称	发明人数量	发明人-申请年份观测数
A	人类生活必需	10982	93646
B	作业;运输	8917	78059
C	化学;冶金	10131	103936
D	纺织;造纸	475	3870
E	固定建筑	1248	10191
F	机械工程;照明;加热;武器;爆破	4154	34713
G	物理	23781	222714
H	电学	22012	214857
Y	新兴/交叉技术	1226	12403
	合计	82926	774389

表3汇总了主要变量的描述性统计结果。从结果可见，创新产出类变量（包括专利数量、被引数量及权利要求数）均呈现明显的高度右偏与过度离散特征，即中位数显著低于均值，且最大值远高于典型样本的产出水平。这一分布特征意味着数据存在“幂律分布”的结构性不平衡，并伴随着显著的长尾现象。此外，各项产出变量的方差明显大于其均值，表明数据不满足泊松分布关于均值与方差相等的基本假设，故本文选择负二项回归模型进行估计较为合适。该模型不限制均值与方差的数量分布，并通过额外的离散参数有效吸收数据中的过度离散部分，因而能够在厚尾计数分布条件下提供更稳健的估计结果。

表3 各变量描述性统计分析

类型	变量	样本量	均值	中位数	最小值	最大值	标准差
因变量	技术生产率	82926	3.0	2	1	760	5.07
	被引数量	82926	64.6	15	0	105000	388
	5年被引数量	82926	15.9	4	0	75800	155
	权利要求数	82926	51.2	26	1	9240	107
自变量	零阶贝蒂数	82926	1.6	1	0	61	1.97
	一阶贝蒂数	82926	<0.1	0	0	1	0.10
控制变量	网络密度	82926	0.2	0.05	0	1	0.34
	平均连接强度	82926	0.6	0.24	0	50	1.02
	协作强度	82926	0.1	0.05	0	1	0.21
	职业年龄	82926	11.6	10	0	41	9.54
	滞后产出	82926	2.7	1	0	760	4.81
	网络规模	82926	5.6	4	0	252	6.73

进一步检验解释变量之间是否存在潜在的多重共线性问题。图2所示为自变量与控制变量间的Pearson相关系数。为避免因变量对相关结构的干扰，相关性分析主要针对自变量与控制变量展开。总体来看，大多数变量之间的相关系数处于较低水平，绝对值普遍低于0.30，表明变量间线性相关性较弱。其中，零阶贝蒂数与网络规模之间呈现一定程度的正相关（$r = 0.67$），β_0^2 与 β_0 之间存在较高相

关性（$r = 0.80$），这一结果符合二次项变量的构造逻辑，属于结构性相关关系，并非模型设定问题。

在此基础上，本文计算了方差膨胀因子（variance inflation factor，VIF）以检验多重共线性问题。结果显示，各解释变量的 VIF 值均处于合理区间(1.03,2.90)，显著低于通常采用的临界阈值 10，表明模型不存在严重的多重共线性问题。需要说明的是，针对包含二次项的变量β_0，本文采用中心化处理后计算 VIF 值，以避免结构性相关导致的虚假高 VIF 估计。变量间的相关系数与 VIF 结果表明，本文的回归模型设定具有良好的稳定性与解释有效性。

图 3 对局部专利合作网络的拓扑指标β_0与β_1的统计特征进行整体刻画。总体来看，β_0呈现明显的右偏分布（图 3a），大部分发明人的合作网络仅包含少量连通组件，而β_0取值较高的情形较少。β_1的分布则高度集中于 0（图 3b），$\beta_1 = 1$的比例约为 1%，表明一维高阶环路结构在整体样本中较少存在。

从时间维度来看，β_0的年度均值自 20 世纪 80 年代起持续上升（图 3c），反映合作网络逐步趋于由更多组件构成；然而$\beta_1 = 1$的年度概率在整个时期内始终维持低位，未随合作关系扩张而增加，显示高阶环路结构的稀缺性。同时，P(Obs)表示各年份观测数占全部观测数的比例，其总体上升趋势说明样本观测数随时间增加，为后期年份的估计提供了相对充分的数据基础。在技术领域层面（图 3d），不同 CPC 下β_0存在显著差异，如“物理（G）”“电学（H）”等领域的β_0较高，而β_1在所有技术类别中均保持近似为 0。这表明虽然不同领域的网络连通性存在结构性差别，但高阶协作环路在各技术领域均较为稀少。

3.3　回归模型结果分析

图 4 展示了高阶拓扑指标与发明人技术生产率之间的关系。首先，从描述性结果来看（图 4a），β_0与生产率之间呈现明显的非线性趋势。随着β_0的提升，发明人平均生产率总体上升，但在较高区间增速逐渐放缓。二次多项式拟合结果显示，二次项系数为负（−0.004），表明β_0与生产率之间可能存在倒 U 形关系，即过高的连通组件数量未必持续带来生产率提升。相比之下，β_1的结果呈现较为明确的单调差异（图 4b）。当$\beta_1 = 1$时，发明人的平均生产率显著高于$\beta_1 = 0$的情形，表明存在一维高阶环路结构的合作网络往往伴随更高的创新产出水平。

在控制网络密度、平均连接强度、协作强度、职业年龄、滞后产出以及网络规模后，负二项回归的边际效应估计进一步验证了这一非线性结构（图 4c）。估计结果显示，β_0的一阶项为正，而二次项为负，且达到统计显著性水平。根据回归系数计算得到的转折点约为β_0=20.88。在该水平之前，连通组件数量的增加显著提升预期生产率；超过该阈值后，边际效应逐步减弱，表明过度分散的网络结构可能降低协作效率，从而抑制创新产出。进一步地，图 4d 展示了不同 CPC 技术领域中的边际效应差异。总体来看，不同技术领域均呈现倒 U 形趋势，但转折点位置存在明显差异。例如，“物理（G）”领域的转折点约为β_0=40.2，意味着在该领

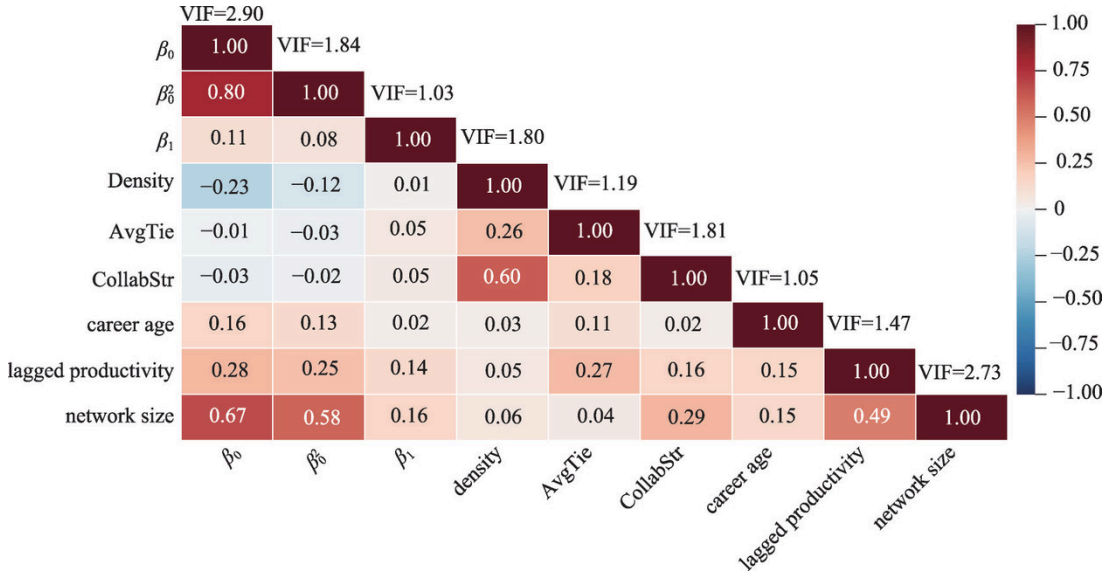


图 2　变量间 Pearson 相关系数矩阵及 VIF 结果（彩图请见 https://qbxb.istic.ac.cn）

域中，网络连通组件数量达到较高水平时，创新产出仍维持在较高区间；而“人类生活必需（A）”领域的转折点约为β_0=12.4，表明其在较低连通组件水平即达到产出峰值。该结果说明，高阶网络结构与创新绩效的关联关系具有显著的技术领域异质性。

β_0和技术生产率间的回归模型结果如表4所示。分析显示，β_0与发明人生产率呈现显著正相关（模型1~模型5均显著，p<0.001）。在加入β_0的平方项后（模型2~模型5），β_0^2呈现显著负系数，表明β_0

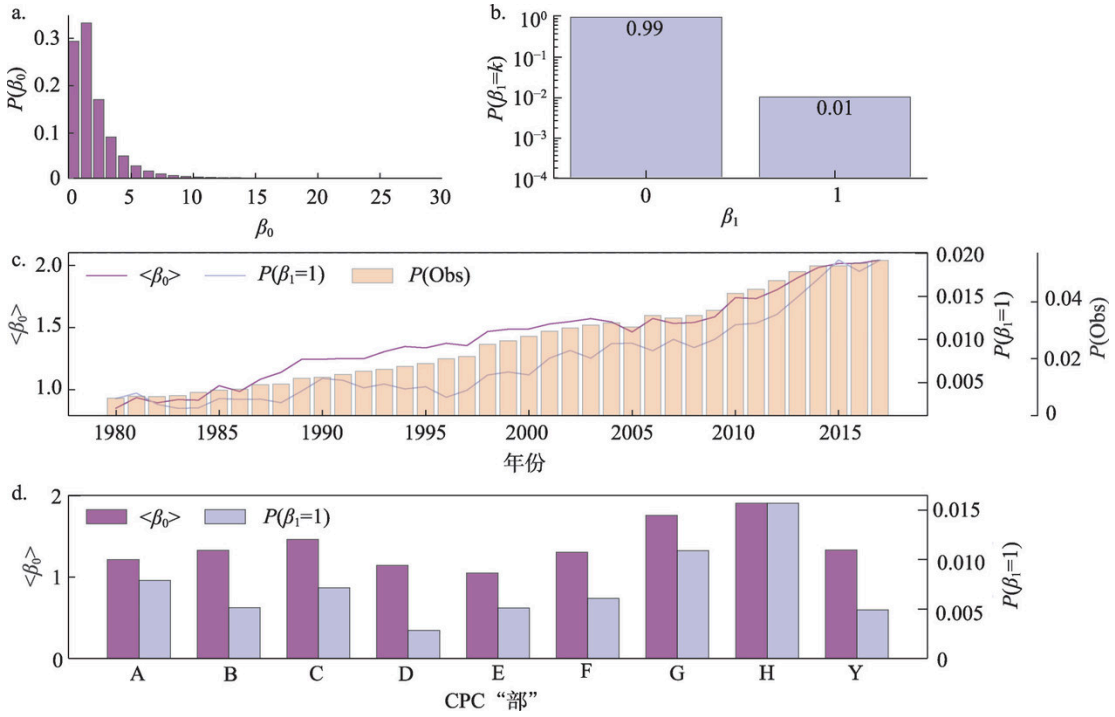


图3 β_0与β_1的分布及趋势特征

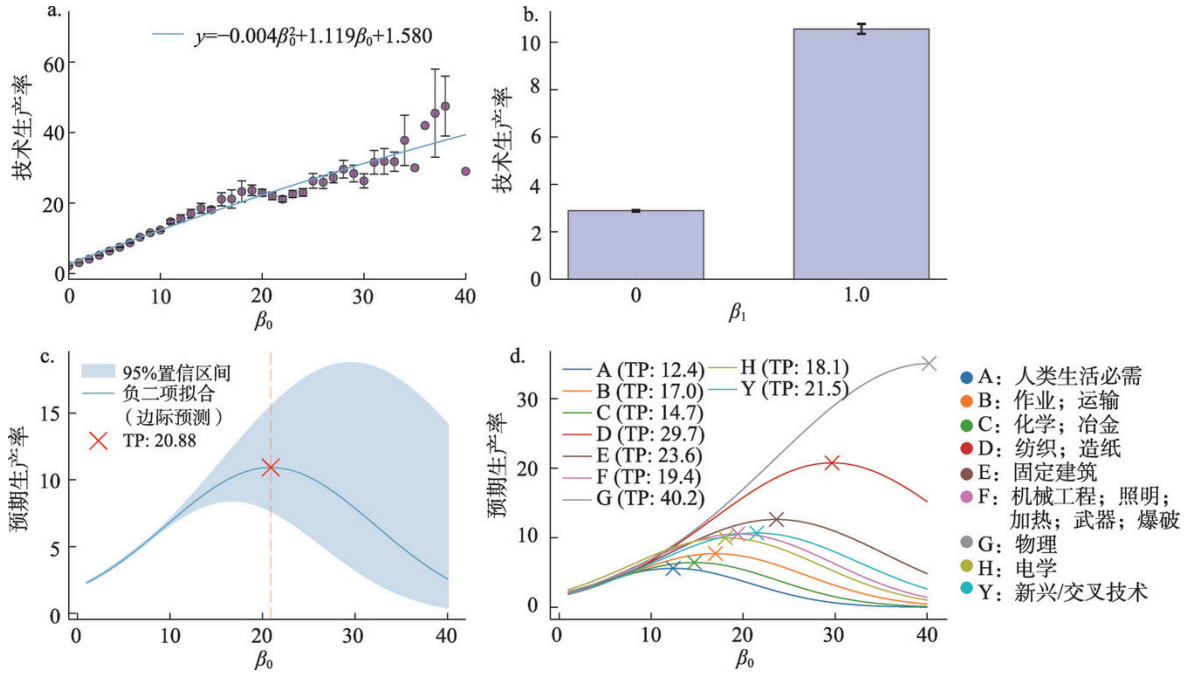


图4 高阶拓扑指标与发明人技术生产率之间的关系（彩图请见https://qbxb.istic.ac.cn）

误差棒表示95%置信区间，TP是转折点（turning point）。

表 4　β_0 和技术生产率间的倒 U 形关系

变量	模型 1	模型 2	模型 3	模型 4	模型 5
β_0	0.2178***	0.2750***	0.2720***	0.1045***	0.1120***
	(0.0021)	(0.0057)	(0.0053)	(0.0029)	(0.0028)
β_0^2		−0.0062***	−0.0059***	−0.0024***	−0.0023***
		(0.0006)	(0.0005)	(0.0002)	(0.0002)
density（对数值）			0.1701***	−0.3143***	−0.2057***
			(0.0070)	(0.0056)	(0.0057)
AvgTie（对数值）			0.4339***	0.0513***	−0.0197***
			(0.0062)	(0.0028)	(0.0025)
CollabStr（对数值）				0.4847***	0.5495***
				(0.0070)	(0.0066)
career age（对数值）				−0.0652***	−0.0426***
				(0.0010)	(0.0020)
lagged productivity（对数值）				0.4752***	0.2834***
				(0.0028)	(0.0022)
network_size（对数值）				0.7123***	0.8142***
				(0.0065)	(0.0057)
发明人 FE	不包含	不包含	不包含	不包含	包含
申请年份 FE	不包含	不包含	不包含	不包含	包含
观测数	774389	774389	774389	774389	770919
R^2	0.072	0.075	0.105	0.208	0.254

注：所有模型均基于负二项回归，括号内数据表示在发明人层次聚类的稳健标准误（standard error，S.E.）；*表示 p<0.05，**表示 p<0.01，***表示 p<0.001；FE 表示固定效应（fixed effects）。

与生产率之间存在倒 U 形关系，即网络中连通组件数的增加在初期显著提升生产率，但达到较高水平后边际增速减缓。加入网络密度、平均连接强度、协作强度、职业年龄、滞后产出和网络规模等控制变量，以及控制发明人固定效应与年份固定效应后（模型 5），模型拟合度 R^2 提升至 0.254，进一步支持其与生产率之间的稳健关联。需要说明的是，在负二项回归等计数模型中，R^2 指标通常低于线性模型，其数值需结合模型设定加以理解。

β_1 和技术生产率间的回归模型结果如表 5 所示。结果显示，其对生产率的影响显著且正向（模型 1~模型 5 均显著，p<0.001）。加入 β_0 及其他网络特征后（模型 2~模型 5），β_1 系数逐步下降，但仍保持高度显著，表明一维高阶环路的存在对创新产出具有独立的正向作用。控制发明人和年份固定效应后（模型 5），β_1 与生产率的关联仍保持显著，R^2 为 0.254。

3.4　稳健性分析

为检验核心结论在不同创新产出指标下的稳健性，以技术生产率、专利被引数量、专利 5 年被引数量及专利权利要求数 4 类指标为因变量，重新估计包含 β_0 与 β_1 的回归模型（图 5）。结果显示，在 4 类创新产出指标中，β_0 的一次项系数均为正且在 0.11~0.17 范围内波动，二次项系数均为负并集中在 −0.005~−0.002（p<0.001）。这表明，合作网络中互不连通组件数量与不同创新产出指标间呈现稳定的倒 U 形关系，即适度增加网络组件数量与更高水平的创新产出相关，但当组件数量过高时，其与创新产出的关联强度呈现减弱趋势。

高阶环路结构 β_1 在 4 种创新产出指标对应的回归模型中均呈现显著正向影响，其回归系数介于 0.14~0.20（p<0.001）。回归结果中，β_1 的影响方向与显著性保持一致，表明一维高阶协作环路对创新产出的正向关联不依赖于具体的产出度量方式，具有较强的稳健性。

单纯复形中的零阶贝蒂数 β_0 与二元关系视角下的连通分量（子图）数量在结构含义上具有高度对应性。为检验该结构指标与创新产出之间关联结果的稳健性，本文将合作网络转化为二元图表示形式，并计算其连通分量数量，在保持控制变量与固定效应设定一致的条件下，重新估计创新产出回归模型（图 6）。结果显示，二元网络子图数的一次项与二次项系数方向及显著性与 β_0 模型结果高度一

表5 β_1和技术生产率间的正向关系

变量	模型1	模型2	模型3	模型4	模型5
β_1	1.293***	0.7768***	0.5011***	0.1784***	0.1461***
	(0.0089)	(0.0077)	(0.0068)	(0.0052)	(0.0052)
β_0		0.2672***	0.2412***	0.1034***	0.1116***
		(0.0007)	(0.0007)	(0.0006)	(0.0007)
β_0^2		−0.0061***	−0.0051***	−0.0024***	−0.0023***
		(4.93×10^{-5})	(4.53×10^{-5})	(3.46×10^{-5})	(3.60×10^{-5})
density（对数值）			−0.3838***	−0.3132***	−0.2036***
			(0.0039)	(0.0034)	(0.0040)
AvgTie（对数值）			0.3813***	0.0499***	−0.0203***
			(0.0015)	(0.0015)	(0.0018)
CollabStr（对数值）			1.1940***	0.4826***	0.5449***
			(0.0049)	(0.0045)	(0.0054)
career age（对数值）				−0.0644***	−0.0420***
				(0.0006)	(0.0016)
lagged productivity（对数值）				0.4723***	0.2817***
				(0.0009)	(0.0011)
network_size（对数值）				0.7102***	0.8122***
				(0.0035)	(0.0040)
发明人FE	不包含	不包含	不包含	不包含	包含
申请年份FE	不包含	不包含	不包含	不包含	包含
观测数	774389	774389	774389	774389	770919
R^2	0.008	0.078	0.124	0.208	0.254

注：所有模型均基于负二项回归，括号内数据表示在发明人层次聚类的稳健标准误；*表示p<0.05，**表示p<0.01，***表示p<0.001；FE表示固定效应（fixed effects）。

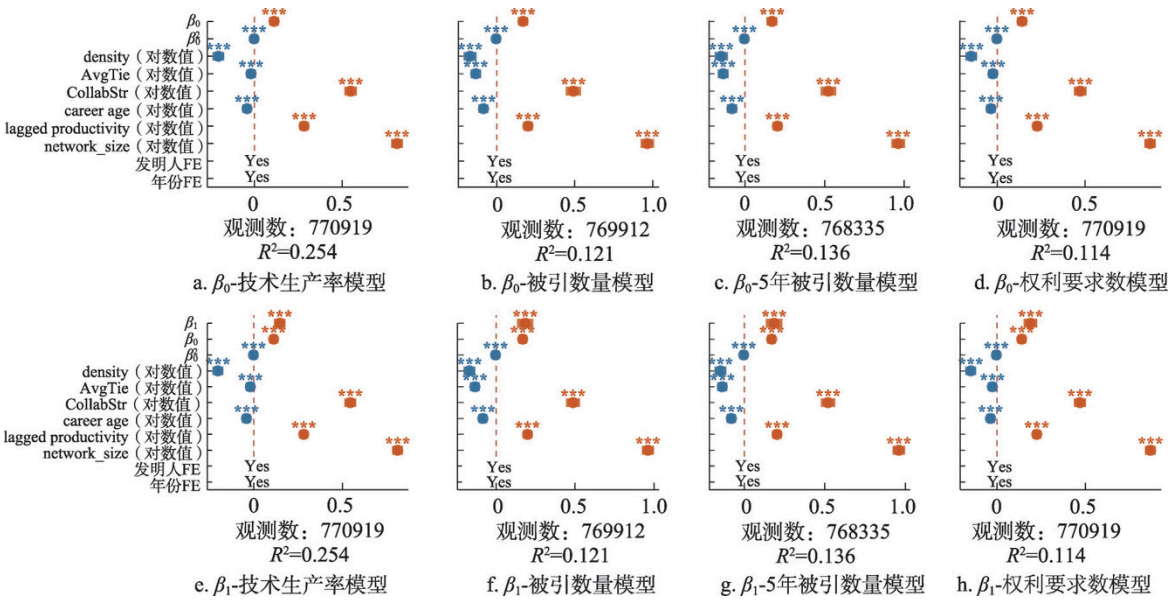


图5 β_0与β_1对4个创新产出指标的回归结果（彩图请见https://qbxb.istic.ac.cn）

回归系数的正负由散点颜色表示，橙色代表回归系数为正，蓝色代表回归系数为负；

误差棒表示95%置信区间；*表示p<0.05，**表示p<0.01，***表示p<0.001。

致，仍呈现稳定的倒U形关系。这表明网络连通结构与创新产出之间的非线性关联在不同结构表达框架下均保持稳定。

然后，考察上述统计关联在不同技术周期下是

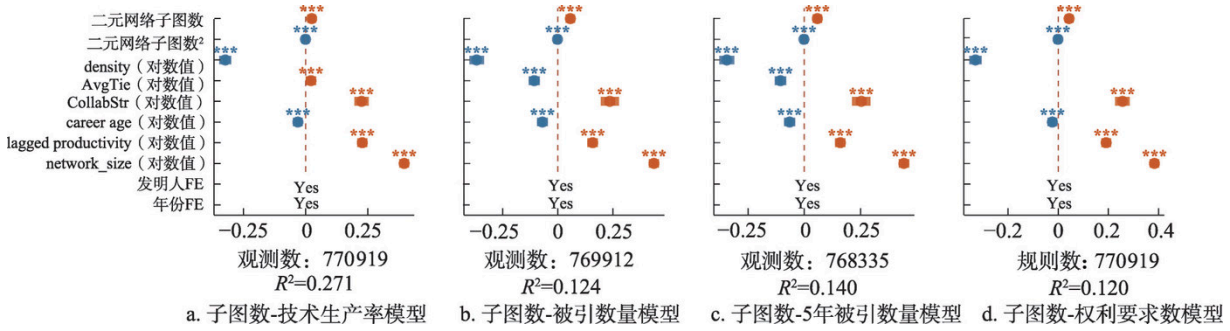


图 6　二元网络子图数对 4 个创新产出指标的回归结果（彩图请见 https://qbxb.istic.ac.cn）

回归系数的正负由散点颜色表示，橙色代表回归系数为正，蓝色代表回归系数为负；
误差棒表示 95% 置信区间；*表示 p<0.05，**表示 p<0.01，***表示 p<0.001。

否随时间保持稳定。为此，基于 1980—2017 年的数据样本，将研究区间划分为 13 个连续区间，分别估计 β_0、β_0^2 与 β_1 对技术生产率的影响。分时期回归结果显示（图 7），β_0 的一次项系数在各时期均为正，其取值主要分布在 0.08~0.12，除个别时期样本受限外，在 13 个时期中有 12 个时期呈现显著正向影响（12/13）；同时，β_0^2 的系数在 −0.005~−0.001 范围内波动，且在 13 个时期中有 11 个时期显著为负（11/13）。上述结果表明，合作网络连通组件数量与技术生产率之间的倒 U 形关系，在长期时间维度上具有高度稳定性，其关联方向与数量级并未因技术范式或创新环境变化而发生偏移。

在图 7 中，高阶环路结构 β_1 在多数时期中对技术生产率表现出显著的正向影响，其回归系数主要集中在 0.06~0.26。在全部 13 个分时期回归结果中，有 11 个时期 β_1 显著为正（11/13），仅在样本规模较小或技术体系尚未稳定的早期阶段，其估计结果未达到常规显著性水平。总体来看，β_1 的估计值随时间呈现平缓波动。

最后，本文进一步检验高阶拓扑结构对发明人技术生产率的影响是否在不同技术领域中保持稳定。基于 CPC 一级分类标准，分别在“人类生活必需（A）”“作业；运输（B）”“化学；冶金（C）”“纺织；造纸（D）”“固定建筑（E）”“机械工程；照明；加热；武器；爆破（F）”“物理（G）”“电学（H）”以及“新兴/交叉技术（Y）”9 个技术领域内，独立估计 β_0、β_0^2 与 β_1 的回归模型，如图 8 所示。

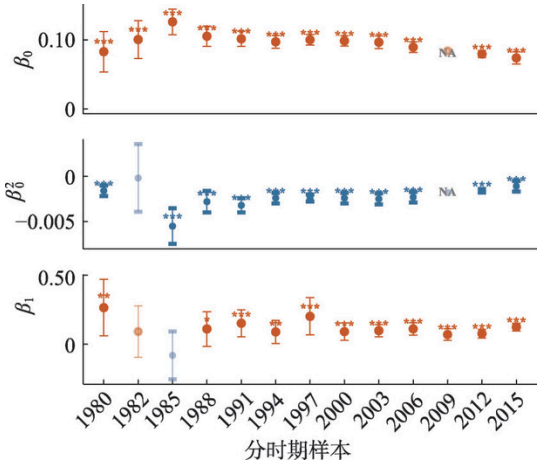


图 7　β_0、β_0^2 与 β_1 对技术生产率的分时期回归系数趋势（彩图请见 https://qbxb.istic.ac.cn）

回归系数的正负由散点颜色表示，橙色代表回归系数为正，蓝色代表回归系数为负；颜色较深表示系数达到统计显著性水平，颜色较浅表示未达到统计显著性水平，NA 表示该时期模型未能获得有效估计值。
误差棒表示 95% 置信区间；*表示 p<0.05，
表示 p<0.01，*表示 p<0.001。

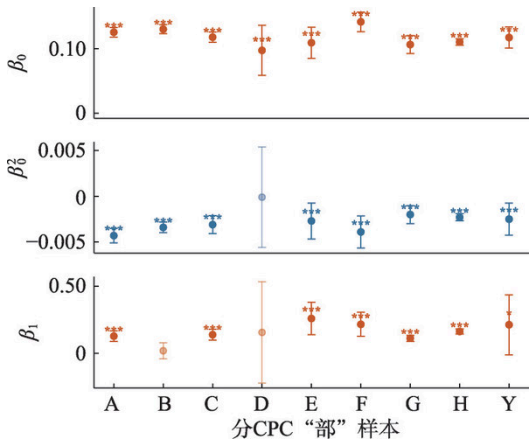


图 8　β_0、β_0^2 与 β_1 对技术生产率的分领域回归系数趋势（彩图请见 https://qbxb.istic.ac.cn）

回归系数的正负由散点颜色表示，橙色代表回归系数为正，蓝色代表回归系数为负；颜色较深表示系数达到统计显著性水平，颜色较浅表示未达到统计显著性水平，NA 表示该时期模型未能获得有效估计值。
误差棒表示 95% 置信区间；*表示 p<0.05，
表示 p<0.01，*表示 p<0.001。

从网络连通组件数量的影响来看，β_0在所有技术领域中均呈现显著正向关联（p<0.001），其回归系数主要分布在0.106~0.141。其中，“机械工程；照明；加热；武器；爆破（F）”领域（γ_0=0.1441，S.E.=0.0770）与“作业；运输（B）”领域（γ_0=0.1299，S.E.=0.0035）的正向关联相对较强，而“纺织；造纸（D）”领域（γ_0=0.0972，S.E.=0.0198）与“电学（H）”（γ_0=0.1102，S.E.=0.0025）领域的估计值相对较低。同时，除了“纺织；造纸（D）”技术领域外，β_0^2在其他领域中均显著为负（p<0.001），其系数范围集中于−0.005~−0.002，表明网络连通组件数量与技术生产率之间的倒U形关系具有明显的跨领域一致性。该统计关联在“人类生活必需（A）”领域（γ_1=−0.0043，S.E.=0.0004）与“机械工程；照明；加热；武器；爆破（F）”领域（γ_1=−0.0039，S.E.=0.0009）中较为显著，说明在以工程实践为特征的技术领域中，合作规模扩张的边际递减效应更为突出。

考察高阶协作结构的作用可以发现，β_1在所有技术领域中均表现出显著的正向影响，其回归系数主要位于0.019~0.259范围。其中，“固定建筑（E）”领域（γ_2=0.2592，S.E.=0.0619）、“机械工程；照明；加热；武器；爆破（F）”领域（γ_2=0.2160，S.E.=0.0456）以及“新兴/交叉技术（Y）”领域（γ_2=0.2128，S.E.=0.1138）的效应量相对较高；相比之下，“作业；运输（B）”领域（γ_2=0.0195，S.E.=0.0303）的估计值略低。上述结果验证了高阶拓扑特征对发明人生产率的跨领域稳健性。

4 结 论

本文面向专利合作网络中多主体协作的真实特征，引入代数拓扑视角，构建单纯形模型，以克服传统二元网络在刻画群体协作关系时的局限性，系统量化了合作网络的高阶结构特征，并基于大规模“发明人-申请年份”面板数据检验其与创新产出间的关系。在理论贡献方面，本文将高阶拓扑方法引入创新网络分析框架，拓展了合作网络研究从二元互动向群体协作的结构视角；在方法贡献方面，本文构建并应用了基于代数拓扑的高阶网络表示与量化方法，将专利合作网络抽象为单纯复形结构，并结合负二项回归模型进行关联性检验。通过多创新产出指标、跨时间、跨技术领域的系统实证，为理解合作结构与创新表现之间的深层关联提供了新的证据。本文的研究结果表明，高阶拓扑结构为理解合作网络结构特征与创新产出之间的关联提供了重要维度，为创新团队构建、专利合作组织方式优化以及创新政策制定提供了新的实证基础。研究结论主要包括以下3点。

（1）专利合作网络的基础连通性（β_0）对发明人创新产出呈现一定的倒U形关系。适度数量的连通组件有助于保持合作资源的多元性，与较高的专利产出水平相关；当网络过度碎片化时，合作关系被分割为多个孤立子群，知识整合效率下降，产出水平的增长效率随之降低。

（2）高阶协作结构（β_1）对发明人创新产出具有稳定的正向促进作用。当合作网络中存在无法由3人及以上协作完全填充的高阶环路结构时，代表团队之间存在跨团组的复杂协作闭环，与更为活跃的信息流动结构特征相关，从而提升个体的创新能力。从实践角度来看，提升β_1并非刻意制造“结构孔洞”，而是通过促进跨领域合作、鼓励跨团队联合项目等方式增加网络中的多路径协作闭环。

（3）上述统计关联在不同创新产出指标、不同技术周期与技术领域中保持稳健。技术生产率、专利被引数量、专利5年被引数量及专利权利要求数4种创新产出指标结果稳定。此外，无论是1980年以来的不同创新阶段，还是在不同CPC技术领域内部，β_0的倒U形作用与β_1的正向作用均显著成立，表明高阶拓扑特征与创新产出之间的关联在不同时间与技术领域中保持稳健。

本文仍存在若干局限。首先，本文基于观测性面板数据与计量模型分析，结果反映变量之间的统计关联关系，并不构成严格的因果识别。随着科研问题日益综合化与复杂化，科技合作及学科交叉不断深化[23]，虽然本文使用面板模型与固定效应控制了大量潜在混杂因素，但高阶拓扑结构的形成仍可能受到未观测团队偏好或组织策略的影响。其次，本文主要关注β_0与β_1两类高阶拓扑特征，而更高维的同调结构（如β_2及更高维孔洞）在大规模合作网络中的潜在作用尚未展开讨论，未来可结合持久同调等工具进一步拓宽高阶结构的测度维度。此外，本文未涉及高阶结构对创新质量异质性的作用机制，如对颠覆性创新或跨领域创新的影响[37]。未来可进一步结合跨国数据、科研团队组织结构或其他类型的创新合作网络，为高阶网络理论在科技创新领域的应用奠定坚实基础。

参 考 文 献

[1] Fortunato S, Bergstrom C T, Börner K, et al. Science of science [J]. Science, 2018, 359(6379): eaao0185.

[2] Yang A J, Guo J M, Shi Y J, et al. Repeat collaboration and scientific innovation: evidence from dynamic ego networks of Nobel laureates[J]. Humanities and Social Sciences Communications, 2025, 12: Article No.1620.

[3] Yang W L, Wang Y. Higher-order structures of local collaboration networks are associated with individual scientific productivity[J]. EPJ Data Science, 2024, 13: Article No.15.

[4] 卢超, 李梦婷, 周宸宇. 基于合贡献者网络的学者学术影响力评估研究[J]. 情报学报, 2025, 44(11): 1444-1457.

[5] 吴柯烨, 孙建军, 张子涵. 基于专利合作网络的弱链接效应研究[J]. 情报学报, 2024, 43(11): 1257-1268.

[6] 曾建勋, 卢春江, 林鑫, 等. 创新情报: 时代背景、概念内涵与功效作用[J]. 情报学报, 2025, 44(9): 1075-1082.

[7] 刘洪愧, 李希晨, 姜丽, 等. 全球创新网络下的中国企业创新——跨国合作与质量提升[J]. 管理世界, 2025, 41(8): 56-83.

[8] 冷萱, 何佳鑫. 科学家集聚、合作网络与创新效率[J]. 世界经济, 2025, 48(10): 192-224.

[9] 詹姆斯·索罗维基. 群体的智慧——如何做出最聪明的决策[M]. 王宝泉, 译. 北京: 中信出版社, 2010: 183-186.

[10] Shi D H, Chen G R. Simplicial networks: a powerful tool for characterizing higher-order interactions[J]. National Science Review, 2022, 9(5): nwac038.

[11] Zhang Y Z, Lucas M, Battiston F. Higher-order interactions shape collective dynamics differently in hypergraphs and simplicial complexes[J]. Nature Communications, 2023, 14: Article No. 1605.

[12] Bianconi G. Higher-order networks: an introduction to simplicial complexes[M]. Cambridge: Cambridge University Press, 2021.

[13] Patania A, Petri G, Vaccarino F. The shape of collaborations[J]. EPJ Data Science, 2017, 6(1): Article No.18.

[14] 郭剑明, 杨杰, 方静, 等. 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.

[15] Newman M E J. The structure of scientific collaboration networks [J]. Proceedings of the National Academy of Sciences of the United States of America, 2001, 98(2): 404-409.

[16] Ke Q, Ahn Y Y. Tie strength distribution in scientific collaboration networks[J]. Physical Review E, 2014, 90(3): 032804.

[17] Wang F F, Dong J X, Lu W Z, et al. Collaboration prediction based on multilayer all-author tripartite citation networks: a case study of gene editing[J]. Journal of Informetrics, 2023, 17(1): 101374.

[18] 关鹏, 王曰芬, 傅柱, 等. 基于专利合作网络的研发团队识别及创新产出影响研究[J]. 数据分析与知识发现, 2022, 6(5): 99-111.

[19] 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.

[20] Guan J C, Pang L X. Bidirectional relationship between network position and knowledge creation in Scientometrics[J]. Scientometrics, 2018, 115(1): 201-222.

[21] 邵桂兰, 许杰, 李晨. 合作网络结构洞、邻域中心性与发明人创新绩效[J]. 科技管理研究, 2021, 41(4): 168-174.

[22] Zeng A, Fan Y, Di Z R, et al. Fresh teams are associated with original and multidisciplinary research[J]. Nature Human Behaviour, 2021, 5(10): 1314-1322.

[23] 冯小东, 黄雨杭. 基于SciBERT的科研合作知识交叉测度及其对科研主体持续科研产出的因果效应[J]. 情报学报, 2025, 44(7): 869-891.

[24] Yang A J, Ding Y, Liu M J. Female-led teams produce more innovative ideas yet receive less scientific impact[J]. Quantitative Science Studies, 2024, 5(4): 861-881.

[25] Yang A J, Xu H M, Ding Y, et al. Unveiling the dynamics of team age structure and its impact on scientific innovation[J]. Scientometrics, 2024, 129(10): 6127-6148.

[26] Lin Y L, Frey C B, Wu L F. Remote collaboration fuses fewer breakthrough ideas[J]. Nature, 2023, 623(7989): 987-991.

[27] 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): 048901.

[28] Vasilyeva E, Kozlov A, Alfaro-Bittner K, et al. Multilayer representation of collaboration networks with higher-order interactions [J]. Scientific Reports, 2021, 11: Article No.5666.

[29] Li M, Wang Z G, Zeng A, et al. Higher-order structure based node importance evaluation in directed networks[J]. Information Processing & Management, 2025, 62(1): 103948.

[30] 李欣哲, 鲁晓. 新颖组合与团队协作的双重奏——基于超图模型的科学基金青年资助项目的创新影响作用分析[J]. 科学学研究, 2026, 44(1): 204-215.

[31] 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.

[32] Battiston F, Cencetti G, Iacopini I, et al. Networks beyond pairwise interactions: structure and dynamics[J]. Physics Reports, 2020, 874: 1-92.

[33] Landry N W, Young J G, Eikmeier N. The simpliciality of higher-order networks[J]. EPJ Data Science, 2024, 13(1): Article No.17.

[34] Krebs J, Polonik W. On the asymptotic normality of persistent Betti numbers[J]. Advances in Applied Probability, 2025, 57(2): 492-523.

[35] 刘波, 曾钰洁, 杨荣渭, 等. 高阶网络统计指标综述[J]. 物理学报, 2024, 73(12): 128901.

[36] Park M, Leahey E, Funk R J. Papers and patents are becoming less disruptive over time[J]. Nature, 2023, 613(7942): 138-144.

[37] 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.

（责任编辑　魏瑞斌）