基于双层时序网络的学术论文创新度量研究*

陆　伟[1,2]，王玉琦[1,2]，罗卓然[1,2]，于丰畅[1,2]

（1. 武汉大学信息管理学院，武汉，430072；
2. 武汉大学信息检索与知识挖掘研究所，武汉，430072）

摘要：学术论文作为科学研究重要成果之一，在知识传播和创新扩散过程中扮演着关键角色。本文从知识组合和复杂网络的视角出发，提出了“问题—方法”双层时序网络的构建模型，以 ACM（Association for Computing Machinery，美国计算机学会）论文为数据基础，实现了“问题—方法”双层时序网络构建。进一步地，综合双层时序网络中节点的新颖性和传播性特征计算了论文的创新程度，并对结果进行了可视化和案例分析。本文从论文词汇网络角度研究了学术论文创新机制，丰富了论文创新度量研究内涵。

关键词：创新测度；学术论文；双层时序网络

1　引言

作为科研创新评价的重要指标，对学术论文的创新度量是对知识发展脉络的微观考察，有利于挖掘学术论文的价值潜力，推动科学评价正确发展。学术论文作为科学研究的重要成果之一，是学科发展前沿和领域创新的直接体现。随着当前学术科研的深入发展，学术论文体量有了飞跃式增长。论文数量的增长体现了科学研究的发展和推进，但海量文献也造成了相关学科创新脉络梳理的困难，面对良莠不齐的海量文献，如何对论文贡献和新颖性进行有效挖掘分析越来越成为创新度量的关键。

已有的科研成果对单篇文献创新的评价焦点往往集中在论文（论文关键词、参考文献、

* 基金项目：国家自然科学基金重点项目“数智赋能的信息资源与知识管理理论与方法创新”（72234005）。

作者简介：陆伟（1974—），男，辽宁鞍山人，武汉大学信息管理学院教授、博士生导师，主要研究方向：信息检索与可视化、数据智能与创新评价、AI 人机协同等。王玉琦（1996—），女，江西上饶人，武汉大学信息管理学院硕士研究生，主要研究方向：数据挖掘、信息服务。于丰畅（1990—），男，湖北武汉人，武汉大学信息管理学院博士后，主要研究方向：非格式化文档理解、机器视觉、深度学习等。

通讯作者：罗卓然（1993—），女，湖北武汉人，武汉大学信息管理学院博士研究生，主要研究方向：创新评价、数据挖掘。

合作作者等）或外部指标（被引网络、共被引网络、期刊影响因子等）两个方面[1]。前者可以通过对单篇论文的文本内容的深度挖掘概述论文的核心内容和价值，但缺乏横向研究领域维度与纵向时间维度的比较；后者可以通过与外部指标的联系，得出论文在某个研究范围内的相对创新价值，但外部指标是论文传播影响因素的直观体现，难以完全反映文献的整体内容。因此，提出一种综合两种方法、利用论文内部文本要素构建整体网络，从而挖掘判定论文创新的方法，可以在一定程度上弥补上述研究的局限。

在学术论文创新评价方法中，知识网络凭借对多维度知识量化和展示的优势，具象化了论文创新的识别、发展、传播和影响过程，成为学术论文度量的重要研究手段之一。在此背景下，本文对论文词汇功能进行细化拆分，聚焦于形成“问题”和“方法”功能的关键词汇，通过对关键功能词汇的网络联系分析来挖掘论文创新的产生机制，并试图将不同时期的论文词汇网络进行比较，对比论文创新性和影响力之间的相似关系，进而从论文内部文本词汇网络角度分析发掘学术论文的创新产生机制和传播机制，拓展论文创新评价研究方法。

2 相关研究

2.1 学术论文创新性评价研究

学术论文创新性评价是对论文内容质量评价的细化，已有研究主要从内容和影响两个方面评价学术论文的创新性：一是从论文自身的新颖性、突破程度等方面进行评价，即通常意义上的内在指标；二是从论文发表后的传播力、扩散程度等方面进行分析研究，即相对而言的外在指标。

论文的内在指标是论文发表时就具备的特征，一方面是指学术论文的核心内容，包括主题、标题、摘要，以及正文所包含的文本、数字、公式等。基于学术论文主题的创新性评价一般从主题词、关键词入手，与同领域已发表的文章比较词汇组合或语义的差异，Zhang 等[2]基于 TREC①2002 的主题新颖性识别数据，提出了基于词重叠度的新颖性判别方法，认为重叠度越高，新颖性越低。Zhou 等[3]采用词表的方法，对论文的关键词进行规范处理和语义扩展，然后计算文本相似度从而得出新颖性。Yi 和 Tsai[4]基于向量空间模型进行新颖性探测，如果当前论文与历史论文之间向量相似性越大，那么新颖性越小。Kumaran 和 Allan[5]利用文本分类和命名实体识别技术改进了基于向量空间的新颖性探测方法，提高了主题新颖性判别的准确度。He 和 Chen[6]基于时序嵌入和向量余弦值度量了某领域学术论文的创新力，并分析了论文创新力对领域文献数量增长的预测作用。除主题词、关键词外，基于文本内容的论文创新性评价研究已经拓展到句子层面，Tsai 和 Zhang[7]提出了 D2S（document-to-sentence，

① TREC：Text Retrieval Conference，文本检索会议。

篇章到句子）方法用来评价论文的新颖性，即先将文本分割成句，计算每句话的新颖性值再进一步计算单篇论文的新颖性值。Dahi[8]利用学术论文中的创新声明比对原文中相关的句子，抽取和识别分析创新的真实性及其所属的范畴。Ronzano 和 Saggion[9]则通过论文语句的修辞功能进行语义标注，从而自动总结出论文的创新贡献和创新结论。论文内在指标的另一方面是指论文的关联内容，包括论文作者、机构、支撑项目及期刊质量，这些指标可以从不同侧面补充反映论文的情况，Karlovčec 和 Mladenić[10]通过合著网络的图结构和语义相似度分析，讨论了合著作者的跨学科程度与论文创新及相关学科发展速度的关系。Boyack 和 Klavans[11]则讨论了传统的期刊影响因子（impact factor，IF）和论文创新的联系，发现在不同学科中期刊对论文创新和传播影响的程度不同，计算机领域的期刊和论文创新的相关程度较高。钱佳佳等[12]基于词频原则分别提出了科技论文的问题新颖度、方法新颖度、问题—方法组合新颖度计算方法，通过权重赋值计算了论文整体的新颖度。罗卓然等[13]以组合创新理论为基础，从词汇语义的角度出发开展了基于词汇功能的学术论文新颖性度量研究，实现了从语义层面更精细地度量新颖性。

论文外在指标评价即利用论文主题内容以外信息进行的评价，如参考文献、刊载期刊、引用数量等，由于外在评价综合了论文发表后的关注和引用，有较强的客观性和获取性，常常被用于系统评价的手段之一。Li 等[14]引用 Prigogine 的耗散理论，认为学术知识网络是一种有关知识的负熵产品，网络中前期的基础节点输出了足够的信息量传递给后续的节点，从而形成了创新发展的路径。Mukherjee 等[15]基于共被引关系建立“常规性—新颖性”的二维坐标系，将论文划分为四个创新类型。Uzzi 等[16]将参考文献中的同主题期刊对视为知识域，通过对海量文献的共被引关系分析，发现对论文创新影响最大的因素主要来源于先前工作的特殊常规组合，率先提出了基于知识重组的论文创新分析。Boyack 和 Klavans[17]将 Uzzi 研究中的对象由 Web of Science 的期刊对迁移到 Scopus 数据上，并用基于期望标准差的 K50 指标替代了 *Z*-score 指标，结果显示新方法可以在更早的时间领域内探测到同样的结论。

2.2 多层复杂网络相关研究

网络作为对现实世界关系连接的抽象化表示，被广泛应用在社区检测、链路预测等领域。网络的复杂性不仅体现在网络结构的多样化，节点类型也呈现出多样化趋势，仅仅依靠单层网络难以体现相关的作用细节，因此多层复杂网络的概念应势而生[18]。Mucha 等[19]在 2010 年提出了多层复杂网络的概念，他们认为将拥有不同角色关系且存在不同结构的单层网络中的节点实体进行连接，综合多个网络拓扑交互权重可以得到一个多层复杂网络。Buldyrev 等[20]根据层内和层间节点属性定义了多层复杂网络——单个网络内节点同质而网络间节点不同质的 N 层（$N \geqslant 2$）网络，并提出多维型网络（multilayer networks）和依存型网络（interdependent networks）两类多层复杂网络，前者每层网络之间的节点代表相同实体的不同性质，所包含

的网络层数与边类型数目相等，且层与层之间仅有对应实体节点存在边，而后者突破了节点同属相同实体的限制，层间不同实体可以存在连边，整体网络之间存在依赖关系。多层复杂网络的结构研究依旧围绕拓扑统计、中心测度、社区检测、网络传播等方面，同时，由于分层结构的存在，对层间相互作用关系的考量成为研究重点所在，de Arruda 等[21]提出了基于多层网络张量表示的 Pearson 和 Spearman 系数定义，为多层复杂网络的层面加权对比提供了思路；Tu 等[22]则采用 PageRank 的路径研究方法挖掘多层复杂网络的中心节点。在社团检测和传播分析上，Kuncheva 和 Montana[23]将信息传播中的随机游走方法拓展到多层网络上，完成了社团检测的任务并进行了鲁棒性检测；Cheng 等[24]则利用节点在不同层次网络之间的关系作为约束进行社团结构识别和检测。

2.3 基于复杂网络的创新度量研究

论文创新的“新”并不是凭空产生的，一定的历史沿革和发展革新才能创造出新的突破，复杂网络的引入为呈现论文网络发展结构随时间的变化提供了支撑，相关研究按照网络的节点元素构成划分为三类，即基于引用关系的文献网络、基于词汇关系的文献网络和混合类型的文献网络。

引文网络是知识传递的直接表现，利用引用关系矩阵变形而产生的共被引网络和耦合网络也从某些方面反映了文献知识传递。Su 等[25]将 PageRank 的思想应用于引文网络，通过分析网络中论文节点的相关数值对论文进行评价。Fragkiadaki 等[26]区分了直接引用和间接引用的深度，提出论文评价的 F 值。Chen[27]通过网络节点的掩盖探究文献节点对整体结构的影响，利用网络结构变异的三个指标（模块变化率、聚类链接和中心发散）来评价论文的影响力。基于关键词/主题/方法等词共现关系形成的网络从论文内容本身出发评价论文创新性。Foster 等[28]利用 Medline 生物医药论文摘要中的实体构建了实体网络，并将沟通以往未连接的子网络实体桥梁视为“知识跳跃”，借此定义论文创新评价模式。Ma 等[29]利用 LDA（latent Dirichlet allocation，潜在狄利克雷分配）主题模型得到论文主题分布矩阵，再结合引用关系构造了主题网络，然后利用网络结构分析来进行论文评价。随着网络研究的深入，也出现了综合两种或多种研究方法的多元网络论文评价研究。Morris[30]将参考文献网络和论文作者网络进行结合，提出了第一网络主节点和第二网络分区中的次节点交互添加的二分网络构建方法，并将其运用到“作者—引文”网络的构建中。索传军[31]则从知识转移的视角，通过引文网络的知识元变迁来评价论文的老化和创新。谭琳洁和刘向[32]在 LeaderRank 算法的基础上进行了改进，同时考虑节点新鲜度及邻居节点的数量和质量的作用，构建了一个论文现时影响力评价模型。魏瑞斌[33]则利用自引用网络的主路径分析完成了论文创新评价。李康[34]利用研究者协作网络和主题共词网络对生物安全领域文献进行了多层复杂网络构建，并通过层次聚类和社团检测等网络分析方法对该领域未来研究热点进行了预测。

综上，目前已有部分文献将复杂网络方法应用在创新测度与评价研究中，但大多是通过

引文网络和合著网络进行的引用网络分析，缺乏从论文内容的词汇角度开展词汇语义网络创新测度的研究。因此，从词汇层面更细粒度的双层时序网络进行研究，可以在一定范围内为未来文献网络和创新测度研究提供参考和借鉴。

3 双层时序网络构建与新颖性计算

3.1 数据及数据处理

本文研究的主体对象是学术论文的词汇知识单元，需要预先对学术论文的“研究问题”和“研究方法”词汇进行抽取，这里的问题和方法并不是传统关键词层面上的通用词汇，而是能够代表一篇论文核心内容的研究问题及研究方法短语，每一篇论文因内容差异而具备一对独特的“问题—方法”短语。根据以上研究要求，本文利用程齐凯等[35]提出的基于标题生成策略和注意力机制的问题—方法抽取方法，收集了 ACM 数据库中 1951~2017 年的 295 567 篇论文，剔除摘要缺失和内容重复的文献后，抽取了论文的问题—方法词对，最终得到包含 142 114 个问题描述词汇短语和 140 749 个方法描述词汇短语的 200 878 篇学术论文数据集（数据统计如图 1 所示，每篇论文的数据字段类型如表 1 所示）。

表 1 论文数据字段展示

数据字段	含义	类型
art_ido	论文序号	varchar
art_title	论文标题	text
art_abstract	论文摘要	text
art_pub_year	论文发表年份	int
art_pub_date	论文发表日期	varchar
art_au_text	论文作者	text
art_inst_text	论文机构	text
art_au_tags	作者关键词	text
art_method	论文研究方法	text
art_question	论文研究问题	text
citation_count	被引量	int
downloads_12months	近 12 个月的下载量	int
downloads_6weeks	近 6 星期的下载量	int

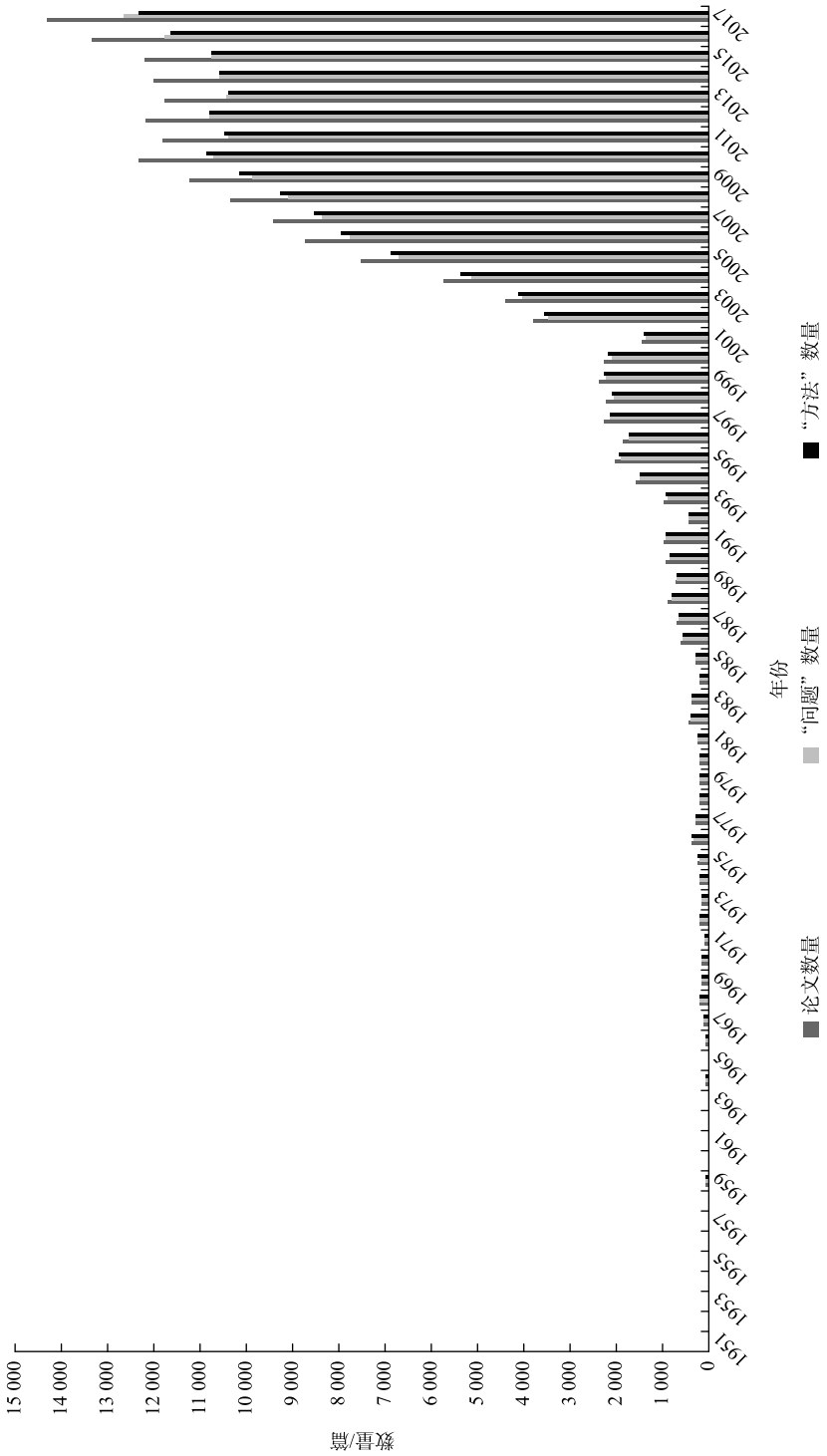


图1　论文及“问题—方法”数据逐年分布图

3.2 双层时序网络数学模型构建

3.2.1 网络数学模型构建

给定一组包含 T 个时间切片的点集合$\{1, 2, \cdots, t, \cdots, T\}$，“问题—方法”双层时序网络可以定义为一个图的序列 $\mu_G=\left(G^{[1]},G^{[2]},\cdots,G^{[t]},\cdots,G^{[T]}\right)$。其中，$G^{[t]}=\left(N^{[t]},E^{t},P^{[t]}\right)$ 为 t 时刻的网络，$N^{[t]}$ 为 t 时刻网络中所有节点的集合，$E^{[t]}$ 为 t 时刻所有网络连边的集合，$P^{[t]}=\left(P_1^{[t]},P_2^{[t]},\cdots,P_L^{[t]}\right)$ 为一系列 t 时刻的子网络层 $\alpha=\{1,2,\cdots,L\}$ 的集合，$P_\alpha^{[t]}=\left(N_\alpha^{[t]},E_\alpha^{[t]}\right)$ 表示 t 时刻多层时序网络中第 α 层所对应的网络结构，$N_\alpha^{[t]}$ 为此时第 α 层上所有节点的集合，$E_\alpha^{[t]}$ 为此时第 α 层上所有连边的集合。因此，存在 $N^{[t]}=N_1^{[t]}+N_2^{[t]}+\cdots+N_\alpha^{[t]}$，$E^{[t]}=E_1^{[t]}+E_2^{[t]}+\cdots+E_\alpha^{[t]}+E_{1\sim2}^{[t]}+\cdots+E_{(\alpha-1)\sim\alpha}^{[t]}$。对于时间集合中的任意 t 时刻，$G^{[t]}$ 可以用一个 $N_t\times N_t$ 的邻接矩阵[①] $W^{[t]}$ 来表示，其中任意节点 v_i 和节点 v_j 的边值表示为 $w_{ij}^{[t]}$，其取值表示如下：

$$w_{ij}^{[t]}=\begin{cases}a_{ij}^{[t]}=\begin{cases}a_{ij} & \left(v_i\text{和}v_j\text{处于同层，且存在语义相似度为}a_{ij}\text{的连边}\right)\\ 0 & \left(v_i\text{和}v_j\text{处于同层，不存在语义相似度关系}\right)\end{cases}\\ 1 \qquad \left(v_i\text{和}v_j\text{处于不同层，且存在文献共现关系}\right)\\ 0 \qquad \left(v_i\text{和}v_j\text{处于不同层，不存在文献共现关系}\right)\end{cases} \tag{1}$$

在本文中，提取出来的“问题—方法”短语作为网络的节点，α 的取值为 2，分别对应“研究问题层”和“研究方法层”。在层内邻接网络中将词汇短语的语义相似度计算值作为连边的取值，在层间邻接网络中将文献共现数目作为连边的取值。

3.2.2 节点语义相似度计算

双层时序网络在构建的时候可以看作单层网络在不同时间点上的集合，因此针对特定时刻，每层网络的构建仍与静态网络相类似。对于研究问题和研究方法，常见的网络构建有频次共现和语义相似度计算两种方法。本文采用的语义短语包含描述性定义，因此单纯的频次共现对内容挖掘不足，故不适宜采用，这里采用语义相似度计算方法。

语义相似度计算方法多样，大体上可以分为基于语料库知识结构、基于相对位置分布相似性和基于篇章连接性分布三类方法[36]。本文采用的是“Glove+ CNN[②]”的语义相似度计算模型（图 2），具体步骤如下：①对所有论文的标题、摘要和关键词字段进行词干抽取、大小写转化、去除标点并分篇存储；②指定时间窗口，将处理好的时间窗口内的文本输入 Glove 算法进行 Embedding（嵌入）计算，得到所有出现过的单词向量预训练结果；③利用预训练结果将组成节点短语的词汇向量组合成多维向量矩阵，输入单层 CNN（卷积核的感受野

① 本文矩阵、向量等字母均用白体表示。

② CNN：convolutional neural network，卷积神经网络。

为 2 × 2）挖掘相似特征后计算最终的语义相似度。该值代表了节点短语在特定时间窗口的相似值。

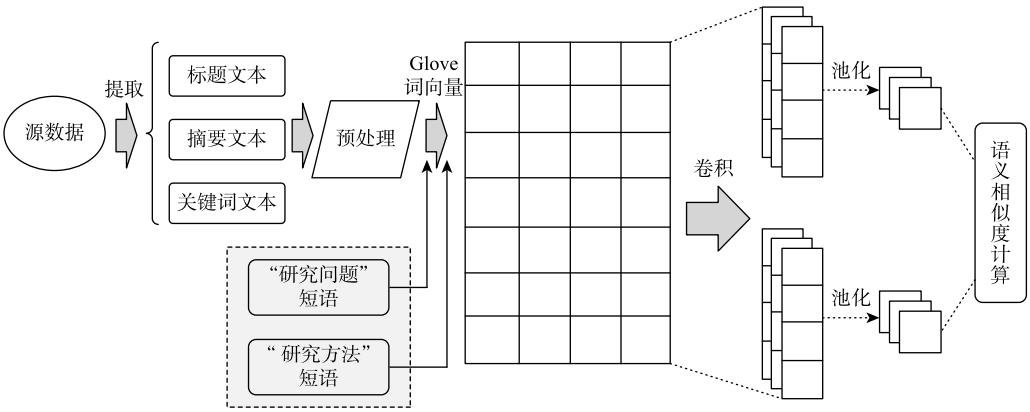


图 2　节点语义相似度计算流程

Glove 算法的全称为 global vectors for word representation，即全局词向量，主要的算法包括以下几点[37]。

（1）根据输入语料构建一个共现矩阵 $\left[\frac{X_{ij}}{d}\right]$，其中 X_{ij} 表示单词 i 和单词 j 在距离为 d 的上下文窗口内的共现次数。

（2）根据共现矩阵和词向量的近似关系，构建损失函数 J。

$$J=\sum_{i,j=1}^{V} f\left(X_{ij}\right)\left(\omega_i^{\mathrm{T}}\tilde{\omega}_j+b_i+\tilde{b}_j-\log\left(X_{ij}\right)\right)^2 \tag{2}$$

（3）损失函数 J 的权重函数是一个两段函数，旨在平衡单词共现次数过小或过大时对损失函数整体的影响。

之所以采用 Glove 算法，是因为它是一个基于全局词频统计的词表征工具，在保留了局部窗口共现信息的前提下，能够更好地利用全局统计量，同时在训练速度和语义效果上有更好的表现。并且，在 Glove 生成词向量后，通过输入短语节点获取多维向量，经过 CNN 层组合（卷积层、激活层、池化层）采用 sigmoid 输出文本节点之间的相似性，也符合语义计算挖掘的一般流程，可以得到较好的结果。

3.2.3　双层时序网络构建

本文构建的是“研究问题”和“研究方法”双层时序网络，针对 t 时刻层内网络采用邻接矩阵的构建方法，由于网络是无向网络，所形成的层内邻接矩阵为对称矩阵（表 2）。其中各个矩阵的数值就是对应的短语文本的语义相似度计算结果，对于对角线元素统一取值为零，即网络中排除自环结构。

表 2 “问题—方法”双层时序网络层内邻接矩阵

0	$v_1^{[t]}$	$v_2^{[t]}$	…	$v_i^{[t]}$	…	$v_j^{[t]}$
$v_1^{[t]}$	0	$w_{12}^{[t]}$	…	$w_{1i}^{[t]}$	…	$w_{1j}^{[t]}$
$v_2^{[t]}$	$w_{12}^{[t]}$	0	…	$w_{2i}^{[t]}$	…	$w_{2j}^{[t]}$
⋮	⋮	⋮	0	⋮	⋮	⋮
$v_i^{[t]}$	$w_{1i}^{[t]}$	$w_{2i}^{[t]}$	…	0	…	$w_{ij}^{[t]}$
⋮	⋮	⋮	⋮	⋮	0	⋮
$v_j^{[t]}$	$w_{1j}^{[t]}$	$w_{2j}^{[t]}$	…	$w_{ij}^{[t]}$	…	0

邻接矩阵是对每个选定时刻 t 进行计算，取值为相应时刻的语义相似度计算值。在本文中，考虑到低语义相似度的节点连边意义不大，且大量存在会对网络分析造成干扰，增加不必要的计算复杂度，故对矩阵每行元素进行如下处理：$w_{ij}^{[t]}$ 等于单词 i 和单词 j 的语义相似度值当且仅当 $w_{ij}^{[t]} \geqslant 0.5$ 或 $w_{ij}^{[t]} \in \text{Top5}\left\{w_{ij}^{[t]} \mid v_i^{[t]}\right\} \cup \text{Top5}\left\{w_{ij}^{[t]} \mid v_j^{[t]}\right\}$ 时成立，其他情况下赋值为零。

层间网络连通体现了问题和方法的共现关系，这里将每篇论文当成层间的一条连线，即单词 i 和单词 j 的层间连线取值为 1，当且仅当单词 i 和单词 j 同时作为一篇论文的研究问题和研究方法时成立，否则不存在相应的连线关系。同时，由于每篇论文的研究问题和研究方法都由独特的短语对组成，在特定时刻 t，层间网络连线代表的实际上是截至时刻 t 已发表的论文数量。

综合层内网络和层间网络，最终形成的“问题—方法”双层时序网络示意图如图 3 所示。

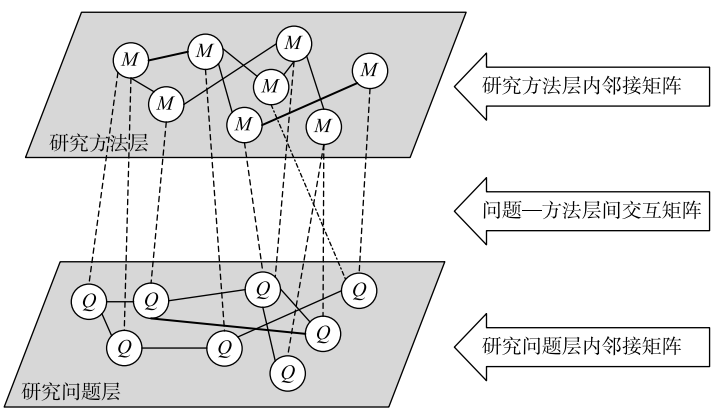


图 3 “问题—方法”双层时序网络示意图

3.3 创新程度得分计算

针对单层网络而言，一个创新节点的引入需要两个前提条件：①从新颖性角度而言，节点的加入让原有的邻居节点产生了之前没有的联系，即计算前一时刻的邻居节点两两间距离到现时刻距离减少的程度；②从传播性角度而言，节点的加入发展了更多的后续相关节点，

即计算后一时刻的中心节点与邻居节点间连边权重相比现时刻权重增加的程度。因此，在无向网络 t 时刻出现的节点 i 的创新程度 $S_i^{[t]}$ 可以用如下的简单线性组合式来衡量：

$$S_i^{[t]} = \mathrm{Nov}_i^{[t]} + \mathrm{Spr}_i^{[t]} \tag{3}$$

节点的创新度量包括新颖性和传播性两个方面。式（3）中，$\mathrm{Nov}_i^{[t]}$ 代表 t 时刻节点 i 的新颖性，这个新颖性是对比 t–1 时刻网络变化的程度得出的；$\mathrm{Spr}_i^{[t]}$ 代表 t 时刻节点 i 的传播性，这个传播性是对比 t+1 时刻网络变化的程度得出的。

$$\mathrm{Nov}_i^{[t]} = \sum_{(x,y)\in \Gamma^{[t]}(i)} \left\{ \| \mathrm{Distance}(x,y)^{[t-1]} - \mathrm{Distance}(x,y)^{[t]} \| \right\} \tag{4}$$

其中，$\mathrm{Nov}_i^{[t]}$ 表示 t 时刻节点 i 的新颖性；$\Gamma^{[t]}(i)$ 表示节点 i 在 t 时刻的邻居节点集合；$\mathrm{Distance}(x,y)^{[t]}$ 表示 t 时刻节点 x 和节点 y 之间的最短路径长度。针对 t 时刻的节点 i 所有的邻居节点，如果在 t–1 时刻任意一对节点两两间不存在直接相连关系，通过计算无权图的节点对最短路径的变化可以得到，节点 i 存在对原有节点对的沟通作用；如果在 t–1 时刻任意一对节点两两间已经存在直接相连关系，通过计算有权图的节点权重路径的变化可以得到，节点 i 存在对原有节点对的加强作用。

$$\mathrm{Spr}_i^{[t]} = \sum_{x\in \Gamma^{[t+1]}(i)} \mathrm{Degree}(x,i)^{[t+1]} - \sum_{y\in \Gamma^{[t]}(i)} \mathrm{Degree}(y,i)^{[t]} \tag{5}$$

其中，$\mathrm{Spr}_i^{[t]}$ 表示 t 时刻节点 i 的传播性；$\Gamma^{[t]}(i)$ 表示节点 i 在 t 时刻的邻居节点集合；$\mathrm{Degree}(y,i)^{[t]}$ 表示 t 时刻节点 y 和节点 i 的连边权重；$\Gamma^{[t+1]}(i)$ 表示节点 i 在 t+1 时刻的邻居节点集合；$\mathrm{Degree}(x,i)^{[t+1]}$ 表示 t+1 时刻节点 x 和节点 i 的连边权重。

本文构建的多层网络是基于语义节点类型的，因此考量的实际上是连接两个不同层节点的连线的中心性排序和计算。这里对学术论文“问题—方法”双层时序网络的中心性计算借鉴 MultiRank 和熵权法[38]的设计思路，即根据网络不同层上的节点创新取值进行加权，同时论文节点的值和两端节点中数值较大的节点关系更加紧密，具体的计算公式如下：

$$S_{(i,j)}^{[t]} = \omega_{(i,\alpha)} S_i^{[t]} + \omega_{(j,\beta)} S_j^{[t]} \tag{6}$$

其中，节点 i 为 α 层上的节点，该层共有 n 个节点；节点 j 为 β 层上的节点，该层共有 m 个节点；$S_i^{[t]}$ 和 $S_j^{[t]}$ 分别为利用式（3）计算得到的 t 时刻节点 i 和节点 j 经过归一化后的创新程度值。

$$\omega_{(i,\alpha)} = \frac{1 + \dfrac{1}{\ln n} \displaystyle\sum_n \left(\frac{S_i^{[t]}}{\sum_n S_n^{[t]}} \times \ln \frac{S_i^{[t]}}{\sum_n S_n^{[t]}} \right)}{n + \displaystyle\sum_n \left\{ 1 + \frac{1}{\ln n} \sum_n \left(\frac{S_i^{[t]}}{\sum_n S_n^{[t]}} \times \ln \frac{S_i^{[t]}}{\sum_n S_n^{[t]}} \right) \right\}} \qquad (i \in \alpha) \tag{7}$$

$$\omega_{(j,\beta)}=\frac{1+\frac{1}{\ln m}\sum_{m}\left(\frac{S_j^{[t]}}{\sum_m S_m^{[t]}}\times\ln\frac{S_j^{[t]}}{\sum_m S_m^{[t]}}\right)}{m+\sum_{m}\left\{1+\frac{1}{\ln m}\sum_{m}\left(\frac{S_j^{[t]}}{\sum_m S_m^{[t]}}\times\ln\frac{S_j^{[t]}}{\sum_m S_m^{[t]}}\right)\right\}}\qquad (j\in\beta)\qquad (8)$$

式（7）和式（8）对不同网络层进行了赋值，采用了 MultiRank 和熵权法的设计思路，即在不同的网络层节点中，若某个单层网络的信息熵越小，表明该网络层代表的创新数值变异程度越大，提供的信息量越多，在综合评价中所能起到的作用也越大，其权重也就越大。

针对特定的一篇论文表现在多层网络中的形式为“一条连接不同网络层间节点对的连线”，对这条连线上所有节点创新值的加权中心性评价即对整篇论文的创新评分，在这个过程中，权重的分配思路与 MultiRank 和熵权法类似，即整体创新得分应更看重得分高、信息量大的节点数值。

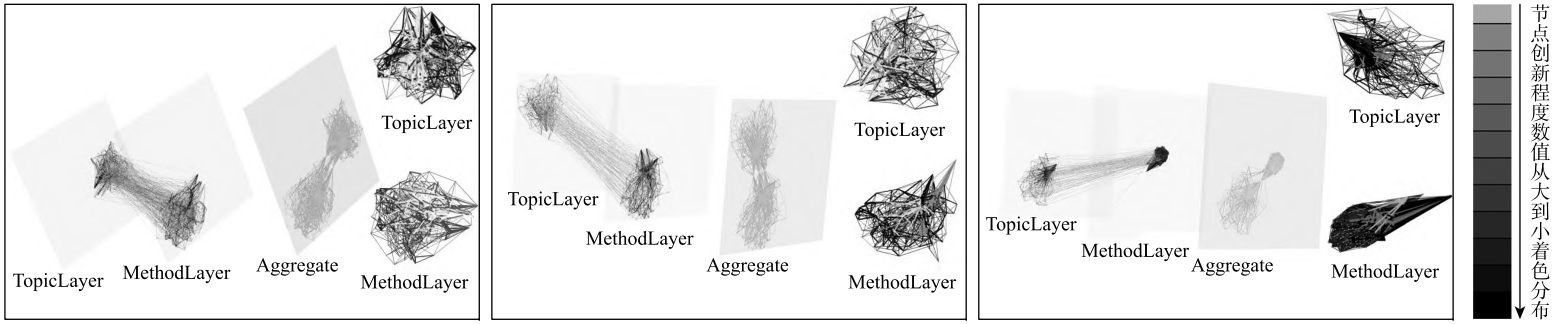
4 实验结果与分析

经过多次实验发现，针对本文数据集中的计算机领域的文献而言，经过 10 年左右的时间一篇文献的主题变化和影响变化将会趋于稳定，能够较好地展示出文献的内容发展变化情况。考虑到数据集文献的时间跨度划分操作均衡情况，本文将数据集按照每隔 11 年一个时间窗口进行划分，共分为 6 个静态时间点，构建每个时间点上的双层网络图，计算每篇文献创新分布的“研究问题”和“研究方法”层上的节点创新程度。为了更好地可视化展示结果并进行分析，这里按照同期被引数量选取占总篇数 10%的 3 个文献样本作为高、中、低文献组，分析不同文献组之间的创新分布关系。

4.1 双层时序网络结果可视化

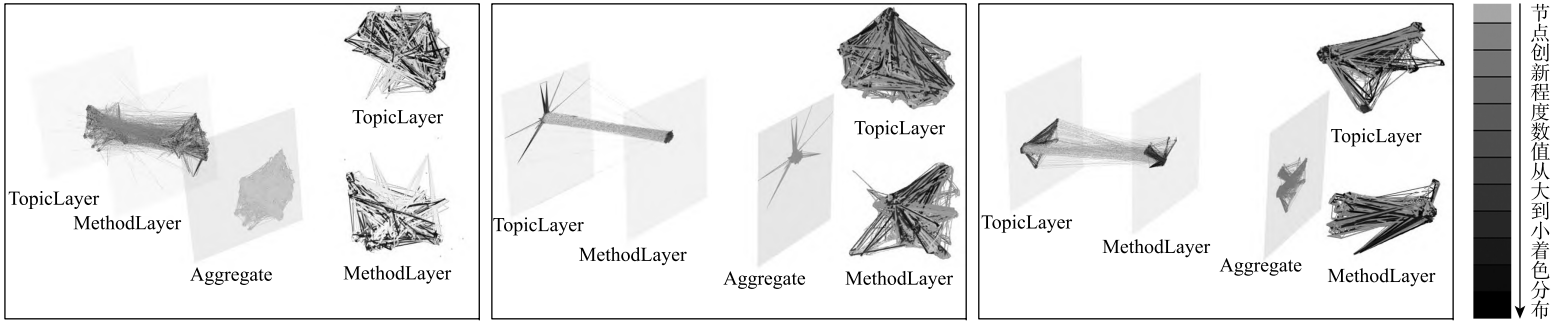
本文利用 MuxViz 绘制了“问题—方法”双层时序网络，并比较了不同被引频次的文献之间“问题—方法”双层时序网络的区别，如图 4 和图 5 所示。其中，TopicLayer 代表“研究问题”层，MethodLayer 代表“研究方法”层，Aggregate 是两个网络的聚合层，也就是一般的多模网络聚合层。网络使用了 Kamada_Kawai、Fruchterman_Reingold 和 Spring[39]三种布局方法融合进行渲染展示，具体表现的是当时间节点 T 取值为 1995 年和 2006 年两个时刻时，前 10% 的高被引、中间 40%~50% 的中被引、后 10% 的低被引文献组的创新权重分布情况，节点 i 按照创新程度 $S_i^{[t]}$ 的大小来着色（颜色 RGB①值依次加大，呈现色度加深的效果），连边的色度深浅由两端节点创新程度取值较大的点决定。

① RGB 是指光学三原色：R 是红色（red），G 是绿色（green），B 是蓝色（blue）。



（a）高被引文献“问题—方法”双层网络图　（b）中被引文献“问题—方法”双层网络图　（c）低被引文献“问题—方法”双层网络图

图4　t = [1995]时高、中、低文献组创新权重分布图



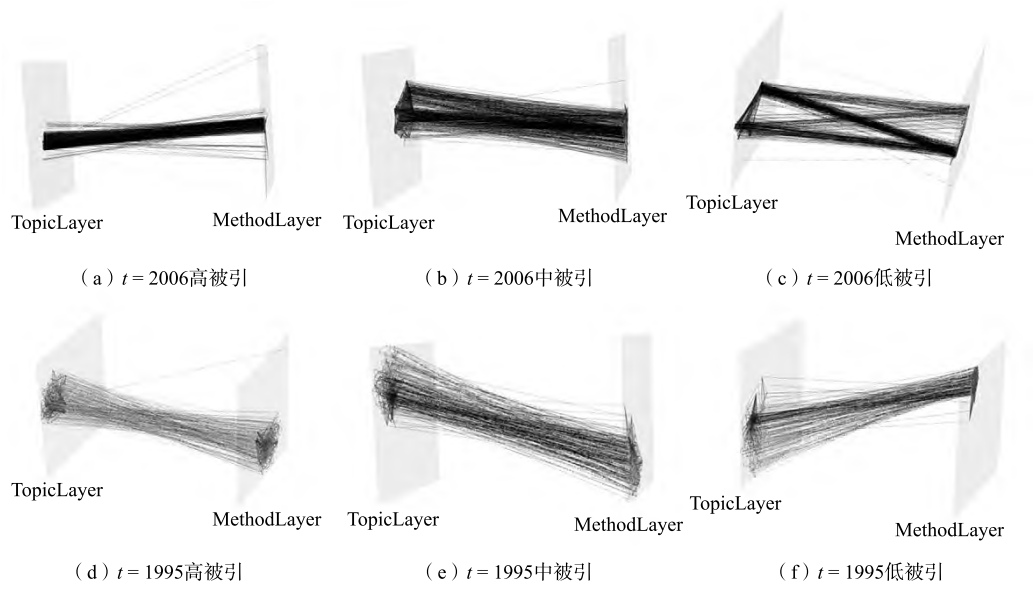
（a）高被引文献“问题—方法”双层网络图　（b）中被引文献“问题—方法”双层网络图　（c）低被引文献“问题—方法”双层网络图

图5　t = [2006]时高、中、低文献组创新权重分布图

从网络分布对比高、中、低被引文献可以发现，高被引文献之间的语义网络联系更加紧密，相互之间的关系呈现团状分布，相比较而言，中、低被引文献的语义网络分布稍散，尤其在低被引文献的语义网络中出现了游离节点；从节点灰度（即创新程度）对比高、中、低被引文献可以发现，被引和创新存在一定的正向关联，高被引文献的语义网络中的节点灰度偏浅，代表创新程度数值大，而随着被引文献的降低网络图整体的创新程度数值在降低；从问题方法区别对比高、中、低被引文献可以发现，高、中、低被引文献与问题创新或者方法创新的来源关联性不大，整体网络均呈现“研究方法”网络层创新程度得分略高于“研究问题”网络层，但相比之下差距并不明显，说明整个数据库中的文献整体方法创新的比例和问题创新的比例在整体上大致趋同，与被引次数关系不大。

从时间来看，t = [2006]的节点的创新程度得分普遍高于 t = [1995]的节点，这主要是由于随着时间的推进，文献的数量增加明显，大量的语义知识网络形成了更紧密的网络关系，也带来了更多的影响性得分，因此呈现了更高的得分。

此外，对 1995 年和 2006 年的高、中、低被引文献的双层时序网络进行对比展示，如图 6 所示。其中，双层网络的层间连线代表不同的文献，粗细代表论文创新得分。从结果可以发现，高、中、低被引文献的层间连线在分布上没有明显的差别，不存在网络中心和网络边际节点的研究问题或者研究方法的偏倚情况。



（a）t = 2006高被引　（b）t = 2006中被引　（c）t = 2006低被引

（d）t = 1995高被引　（e）t = 1995中被引　（f）t = 1995低被引

图 6　不同时间高、中、低被引文献创新得分及分布图

总的来说，双层时序网络的引入更好地从语义层面呈现了学术论文的创新来源，相比异质多模网络而言，分层网络的加入由于划分了不同的内容层次而强调了论文语义创新来源的区别。此外，通过高、中、低三个对照组，也从被引角度说明，创新程度高的文献

大多都有相对完善的背景基础，是从前人的理论知识基础衍生发展而来的，相比较而言更容易获得被引。同时，学术论文语义创新网络中存在“同类引导”的效应，即高被引的创新论文在语义上会呈现和同为高被引的创新论文更密切的关系，辅助证实了创新的传递影响存在。

4.2　结果数据分析

在单层节点创新程度计算的基础上，以每隔 11 年为 1 个时间切片，对 1951~2017 年所有文献构成的 6 个时间网络图上的学术论文计算创新得分并排序，表 3 展示了 1995 年时间点上论文创新得分排序前十的信息。

表 3　*T*=[1995]时论文创新得分前十信息

论文标题	被引次数	问题创新得分	方法创新得分	论文创新得分
Resolving the tension between integrity and security using a theorem prover（https://doi.org/10.1145/50202.50231）	22	secret formal（秘密范式）(138.92)	databas revel（数据库检索）(37.39)	88.34
Pseudo-randomly interleaved memory（https://doi.org/10.1145/115952.115961）	208	polynomi interleav modulo（多项式交互模块）(1.45)	interleav memori（交叉存储器）(134.317)	78.64
A use of drawing surfaces in different collaborative settings（https://doi.org/10.1145/62266.62286）	306	design process（设计过程）(1.45)	collabor draw surfac（合作绘图界面）(134.32)	78.59
Compiler-directed data prefetching in multiprocessors with memory hierarchies（https://doi.org/10.1145/77726.255176）	209	data prefetch（数据预期）(67.51)	memori hierarchi（存储层次）(86.60)	75.81
Scheduling and page migration for multiprocessor compute servers（https://doi.org/10.1145/195473.195485）	211	os schedul and page migrat polici（操作系统调度和页面迁移策略）(5.59)	multiprogram shared-memori multiprocessor（并行程序共享存储式多处理器）(134.53)	73.64
Representing circuits more efficiently in symbolic model checking（https://doi.org/10.1145/127601.127702）	433	partit transit relat（过境关联）(52.01)	symbol verif（符号验证）(89.24)	68.31
Data placement in Bubba（https://doi.org/10.1145/50202.50213）	504	effici data placement（高效数据放置）(48.26)	data-intens applications（数据密集型应用）(82.78)	68.31
A portable platform for distributed event environments（https://doi.org/10.1145/122759.122776）	34	a portabl platform（可移动式平台）(51.91)	distribut event environ（分布式事件环境）(78.6)	64.93
On the representation and querying of sets of possible worlds（https://doi.org/10.1145/38713.38724）	5	incomplet inform databas（不完全信息数据库）(72.91)	a hierarchi（分层结构）(56.07)	63.13

续表

论文标题	被引次数	问题创新得分	方法创新得分	论文创新得分
Augmenting the organizational memory：a field study of answer（https://doi.org/10.1145/192844.193019）	25	a field study（实地研究）（58.11）	organiz memori system（组织记忆系统）（61.13）	59.87

从表 3 可以看出，尽管论文的创新得分和排序与被引次数之间存在一定的正向关系，但并不能在创新和被引之间建立直接的相关关系。例如，表 3 中排名第一的论文被引次数反而偏低，这是因为“Resolving the tension between integrity and security using a theorem prover”提出了一个有关完整性和安全性的定理证明，首创了有关数据知识库的秘密范式（即“secret formal”的研究问题），在此之后 1997 年 Thuraisingham 和 Ford[40]将此问题对应的解决方法进一步拓展并形成专利文件，得到了 112 次的引用，正是因为在引用研究中存在时间上的偏好，所以尽管这篇开创的论文创新得分很高，但并没有太多的引用出现。因此，由于不同于被引的时间累积效果和偏好现象（即经典文献呈现时间被引累积，同方向文献存在近期偏好），一篇论文的创新尽管需要通过影响的监测加以证实，但一定时间后，论文创新得分更多地呈现出一种相对固定的趋势，即研究问题和研究方法节点的网络中心性数值随时间变化减少，更少被后续的内容影响。这也是知识语义网络给科学论文评价带来的优势，可以更好地关注内容自身而减少外部因素的干扰。

5 讨论

本文以学术论文创新测度为研究目标，通过梳理目前国内外已有的创新评价指标和方法，归纳了创新的两个维度特征——新颖性和传播性，利用复杂网络和深度学习的方法，探索了知识网络视角下论文创新测度的相关问题，利用“问题—方法”数据集进行了学术论文创新分析的实证研究。首先，本文构建了“研究问题”和“研究方法”的双层时序网络概念，并给出了相关的数学模型；其次，利用 Glove+CNN 的语义相似度计算得到了网络的边关系矩阵，并通过层内矩阵和层间矩阵的连接形成了双层时序网络；最后，计算每层网络节点创新程度，并对双层时序网络“节点对”进行评分计算及排序，具体的操作主要在于双层时序网络中心节点对的权重分配和计算上，通过加权之后的节点对的得分相加进而对论文进行创新评分。

通过对计算结果和可视化结果的分析，本文将对双层时序网络的发现总结为三点：①有关方法创新和问题创新的分布和被引次数没有直接关系，即不同被引区间的文献都可能存在方法创新或问题创新比例大的文献；②随着科学文献的体量增长，文献创新中的方法创新和问题创新均呈现增长的趋势，更广泛的知识交流在一定程度上促进了整体论文创新的提升；③双层时序网络对科研论文学术创新分布趋势挖掘有一定的作用，比多模网络或超网络的可视化

展现更加清晰，如实反映了网络图中创新节点随时间向中心移动的趋势。总体而言，本文通过引入双层时序网络的概念，对论文创新分布和评价进行了探索性的实验，初步从语义词汇功能的角度揭示了学术论文创新的产生机理，并将相关的研究结论推广到评价应用之上，为后续研究中相关网络方法的拓展提供了一个思路。

参考文献

[1] 罗卓然，王玉琦，钱佳佳，等. 学术论文创新性评价研究综述[J]. 情报学报，2021，40（7）：780-790.

[2] Zhang M，Song R H，Lin C，et al. Expansion-based technologies in finding relevant and new information：THU TREC 2002 novelty track experiments[J]. NIST Special Publication，2003，251：586-590.

[3] Zhou Z L，Wang Y N，Gu J Z. New model of semantic similarity measuring in WordNet[C]. 2008 3rd International Conference on Intelligent System and Knowledge Engineering Xiamen，2008.

[4] Yi Z，Tsai F S. Chinese novelty mining[C]. Conference on Empirical Methods in Natural Language Processing，Singapore，2009.

[5] Kumaran G，Allan J. Text classification and named entities for new event detection[C]. Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval，Sheffield，2004.

[6] He J G，Chen C M. Predictive effects of novelty measured by temporal embeddings on the growth of scientific literature[J/OL]. Frontiers in Research Metrics and Analytics，2018，3：9.

[7] Tsai F S，Zhang Y. D2S：Document-to-sentence framework for novelty detection[J]. Knowledge and Information Systems，2010，29（2）：419-433.

[8] Dahi T. Contributing to the academic conversation：a study of new knowledge claims in economics and linguistics[J]. Journal of Pragmatics，2008，40（7）：1184-1201.

[9] Ronzano F，Saggion H. Knowledge extraction and modeling from scientific publication[C]. Second International Workshop on Semantic，Analytics，Visualization，Montreal，2016.

[10] Karlovčec M，Mladenić D. Interdisciplinarity of scientific fields and its evolution based on graph of project collaboration and co-authoring[J]. Scientometrics，102：433-454，2015.

[11] Boyack K W，Klavans R. Predicting the importance of current papers[C]//Ingwersen P，Larsen B. Proceeding of the 10th International Conference of the International Society for Scientometrics and Informetrics. Stockholm：Karolinska University Press，2005：335-342.

[12] 钱佳佳，罗卓然，陆伟. 基于问题—方法组合的科技论文新颖性度量与创新类型识别[J]. 图书情报工作，2021，（14）：82-89.

[13] 罗卓然，陆伟，蔡乐，等. 学术文本词汇功能识别——在论文新颖性度量上的应用[J]. 情报学报，2022，41（7）：720-732.

[14] Li Y L，Zhang G J，Feng Y Q，et al. An entropy-based social network community detecting method and its application to scientometrics[J]. Scientometrics，2015，102（1）：1003-1017.

[15] Mukherjee S，Uzzi B，Jones B，et al. A new method for identifying recombinations of existing knowledge associated with high-impact innovation[J]. Journal of Product Innovation Management，2016，33（2）：224-236.

[16] Uzzi B，Mukherjee S，Stringer M，et al. Atypical combinations and scientific impact[J]. Science，2013，342（6157）：468-472.

[17] Boyack K W，Klavans R. Atypical combinations are confounded by disciplinary effects[C]. 19th International Conference on Science and Technology Indicators，Leiden，2014.

[18] de Domenico M，Solé-Ribalta A，Cozzo E，et al. Mathematical formulation of multilayer networks[J]. Physical Review，X，2013，3：041022.

[19] Mucha P J，Richardson T，Macon K，et al. Community structure in time-dependent，multiscale，and multiplex networks[J]. Science，2010，328：876-878.

[20] Buldyrev S V，Parshani R，Paul G，et al. Catastrophic cascade of failures in interdependent networks[J]. Nature，2010，464（7291）：1025-1028.

[21] de Arruda G F，Cozzo E，Moreno Y，et al. On degree–degree correlations in multilayer networks[J]. Physica D：Nonlinear Phenomena，2016，（323/324）：5-11.

[22] Tu X，Jiang G P，Song Y R，et al. Novel multiplex pagerank in multilayer networks[J]. IEEE Access，2018，6：12530-12538.

[23] Kuncheva Z，Montana G. Community detection in multiplex networks using locally adaptive random walks[C]. Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining，Paris，2015.

[24] Cheng W，Zhang X，Guo Z S，et al. Flexible and robust co-regularized multi-domain graph clustering[C]. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining，Chicago，2013.

[25] Su C，Pan Y，Zhen Y，et al. PrestigeRank：a new evaluation method for papers and journals[J]. Journal of Informetrics，2011，5（1）：1-13.

[26] Fragkiadaki E，Evangelidis G，Samaras N，et al. F-value：measuring an article's scientific impact[J]. Scientometrics，2010，86（3）：671-686.

[27] Chen C. Predictive effects of structural variation on citation counts[J]. Journal of the American Society for Information ence and Technology，2012，63（3）：431-449.

[28] Foster J G，Rzhentsky A. Evans J A. Tradition and innovation in scientists' research strategies[J]. American Sociological Review，2015，80（5）：875-908.

[29] Ma T C，Li R N，Ou G Y，et al. Topic based research competitiveness evaluation[J]. Scientometrics，2018，112（2）：789-803.

[30] Morris S A. Bipartite yule processes in collections of journal papers[C]//Society for Scientometrics and

Informetrics. Proceedings of the 10th International Conference of the International，2005：316-321.

[31] 索传军. 知识转移视角下的学术论文老化与创新研究[J]. 图书情报工作，2014，58（5）：5-12.

[32] 谭琳洁，刘向. 面向含时引文网络的论文现时影响力评价[J]. 中国科技期刊研究，2020，31（4）：468-473.

[33] 魏瑞斌. 基于自引网络和主路径分析的论文主题创新实证研究[J]. 图书情报工作，2018，62（3）：64-70.

[34] 李康. 基于多层复杂网络的生物安全文献分析研究[D]. 天津大学硕士学位论文，2018.

[35] 程齐凯，李鹏程，张国标，等. 学术文本词汇功能识别——基于标题生成策略和注意力机制的问题方法抽取[J]. 情报学报，2021，40（1）：43-52.

[36] 孙叔琦. 基于统计的词汇级语义相关计算研究[D]. 哈尔滨工业大学博士学位论文，2015.

[37] Pennington J，Socher R，Manning C D. Glove：global vectors for word representation[C]. Conference on Empirical Methods in Natural Language Processing，Doha，2014.

[38] Bianconi G. Multilayer Networks：Structure and Function[M]. London：Oxford University Press，2018.

[39] Komarek A，Jakub P，Vladimir S. Network visualization survey[C]. International Conference on Computational Collective Intelligence，Madrid，2015.

[40] Thuraisingham B M，Ford W. Apparatus and method for the detection of security violations in multilevel secure databases[P]. US，US5694590A，1997.

Research on the Innovation Measurement of Academic Papers Based on Two-Layer Temporal Network

Lu Wei[1, 2]，Wang Yuqi[1, 2]，Luo Zhuoran[1, 2]，Yu Fengchang[1, 2]

（1. School of Information Management，Wuhan University，Wuhan 430072，China；2. Information Retrieval and Knowledge Mining Laboratory，Wuhan University，Wuhan 430072，China）

Abstract：As one of the important outcomes of scientific research，academic papers play a key role in the process of knowledge dissemination and innovation diffusion. From the perspective of knowledge combination and complex networks，this paper proposes a “question-method” two-layer temporal network construction model，and then implements the construction of the network with the data of ACM papers. Furthermore，we calculated the innovation degree of the paper by combining the novelty and spread characteristics of the nodes in the two-layer temporal network，and visualized and analyzed the results with case studies. This paper studies the innovation mechanism of academic papers from the perspective of lexical networks and enriches the methods of paper innovation evaluation.

Keywords：innovation measurement；academic papers；two-layer temporal networks