

技术评审专家遴选方法 在颠覆性技术专家预判平台上的应用

林毅^{1,2} 张均胜^{1,2} 刘志辉^{1,2} 王唯滢³

(1. 中国科学技术信息研究所, 北京 100038; 2. 富媒体数字出版内容组织与知识服务重点实验室, 北京 100038; 3. 中国科普研究所, 北京 100081)

摘要: 评审专家遴选是技术评审中的关键环节。鉴于颠覆性技术专家预判平台预判系统对时效性和智能型的要求, 专家遴选对预判结果具有决定性影响。通过学术专长匹配和专业遴选来选择符合要求的专家, 可以降低成本, 提高推荐效率与准确度, 完成颠覆性技术的预测任务。基于学术网络表示学习的方法既可以避免大量特征工程, 又可以方便不同类型的特征进行融合。利用异质网络表示学习方法和标签排序的学术专长画像方法构建专家库, 并使用融合专家综合评价指标特征的匹配方法对待预判的颠覆性技术和专家专长进行匹配, 为专家遴选提供一份专业背景匹配的候选专家列表。这种方法在Academic Social Network数据集上进行模拟实验。实验结果表明, 这种方法能提升项目评审专家学术专长匹配, 在加入综合指标特征后, 专家的综合指标特征能有效地反馈到实验结果中, 从而提高评审系统的时效性和智能性。

关键词: 专家遴选; 标签排序; 特征融合; 颠覆性技术

DOI: 10.3772/j.issn.1674-1544.2024.02.007

CSTR: 15994.14.issn.1674.1544.2024.02.007

中图分类号: TP181; TP391.1

文献标识码: A

Application of Technical Review Expert Selection Method on Disruptive Technical Expert Prediction Platform

LIN Yi^{1,2}, ZHANG Junsheng^{1,2}, LIU Zhihui^{1,2}, WANG Weiyong³

(1. Institute of Scientific and Technical Information of China, Beijing 100038; 2. the Key Laboratory of Rich-Media Knowledge Organization and Service of Digital Publishing Content Institute of China, Beijing 100038; 3. China Research Institute for Science Popularization, Beijing 100081)

Abstract: The selection of review experts is a key step in technical reviews. Given the requirements for timeliness and intelligence of the predictive platform for disruptive technology experts, expert selection has a decisive impact on the prediction results. By matching academic expertise and conducting professional selection to select experts who meet the requirements, costs can be reduced while improving recommendation efficiency and accuracy, thus completing the prediction task of disruptive technologies. Based on academic

作者简介: 林毅 (1989—), 男, 博士, 中国科学技术信息研究所助理研究员, 研究方向为情报分析; 张均胜 (1980—), 男, 博士, 中国科学技术信息研究所研究员, 研究方向为情报分析 (通信作者); 刘志辉 (1979—), 男, 博士, 中国科学技术信息研究所研究员, 研究方向为竞争情报; 王唯滢 (1988—), 女, 博士, 中国科普研究所助理研究员, 研究方向为科技传播、科普理论与政策。

基金项目: 中国科学技术信息研究所创新基金青年项目“颠覆性技术预判专家画像构建及遴选方法研究”(QN2022-11); 中国科学技术信息研究所创新基金青年项目“战略性新兴产业集群中的技术创新主体竞合网络构建及测度方法研究”(QN2023-07); 中国科学技术信息研究所创新基金青年项目“基于技术特异性的产业链核心企业识别方法研究”(QN2023-06)。

收稿时间: 2023年6月13日。

network representation can not only avoid a lot of feature engineering, but also facilitate the fusion of different types of features. This paper proposes to build an expert database and use a matching method integrating the characteristics of experts' comprehensive evaluation indicators to match the predicted subversive technology and expert expertise on the basis of heterogeneous network representation learning methods and precise profiling methods of academic expertise in label ordering, so as to provide a list of candidate experts related to professional background for expert selection. The method proposed in this paper has been tested on the Academic Social Network dataset, and the simulation results show that the method can improve the academic expertise matching of technology evaluation experts. After adding the comprehensive index feature, the expert's comprehensive index feature can be effectively fed back into the experimental results, thus improving the timeliness and intelligence of the evaluation system.

Keywords: expert selection, label rank, feature fusion, disruptive technology

0 引言

科学技术的重要突破对于推动经济发展和社会进步至关重要。随着经济的不断增长和技术的快速发展，人们越来越需要应对技术革新所带来的挑战。因此，跨学科融合以及信息技术的发展已成为当今世界最为重要的趋势之一，这种趋势正在改变着我们的生活和工作方式。在这样的背景下，颠覆性技术的发展正成为创新活动的焦点，并对军事、工业、科技和社会产生深远的影响。

为学术成果、项目、学者等进行评价等活动寻找合适的专家是一项具有挑战性的任务，因为专业知识是一个定义松散的概念，很难形式化。特别是专家库的构建和遴选合适的专家对技术进行预判成为颠覆性技术专家预判平台中的关键环节，对预判结果具有决定性影响。匹配颠覆性技术预判系统所需的专家，是影响评议质量关键的环节。构建专家库，利用融合专家综合评价指标特征的匹配方法，可以为专家遴选提供与专业背景相关的候选专家列表。除了要考虑待评估项目与专家的学术专长匹配情况外，还要考虑专家所处的位置和工作量，兼顾专家的可用性和可中断性。本文针对颠覆性技术专家预判平台评估中专家遴选这一重要环节，提出一种基于异质网络匹配专家学术专长与设置定性评估专家综合指标相结合的遴选方式，以更好地遴选出合适的专家对待评估技术进行评估。这种方法将各种因素考虑

在内，以提供候选专家的最终排名。此方法不仅考虑了项目与专家学术专长的匹配程度，还考虑了专家的基本情况、道德修养、行业影响力和评审业绩等其他特征，且每个特征的影响可以由自动机器学习、系统管理员设置，或者由特定项目的明确偏好来决定。

1 相关研究

1.1 学术专长画像的相关研究

学术专长是专家在特定研究领域的专业知识和技能^[1]的体现，专业领域和研究方向是专家学术专长的有力反映^[2]。通过分析专家的学术成果、学术行为等信息，可以生成精准有效的标签来描述其学术专长。目前主要的学术专长画像方法有文本挖掘方法和网络表示学习等。

文本挖掘方法的本质在于关键词提取。人们越来越对设计提供更多相关关键词的关键词提取器感兴趣^[3]。提取更有意义的关键词对大数据的分类、数据挖掘、索引和文本分析等都很重要。此外，提取更多相关关键词基本上改进了随后使用的机器学习方法。目前流行方法有TextRank方法^[4]和TF-IDF方法^[5]等。每个关键字提取器为给定文档生成一组候选关键字的方式在每种情况下都不同，从而为集成方法生成更广泛的候选关键字池。在应用过滤启发式算法并重新计算候选关键字得分之后，集成方法使用动态阈值函数为给定文档选择一组关键字。在此基础上，监督学习的应用范围很广，包括最近邻分类器、决策

树、基于规则的分类器和概率分类器等^[6]。聚类和主题建模是在文本数据上下文中常用的两种常用无监督学习算法。在主题模型中，每个主题可以表示为单词的概率分布，每个文档可以表示为主题的概率分布^[7]。

近年来，已经有一些与图嵌入和表示学习相关的研究。图分析任务可以大致分为节点分类、链接预测、聚类和可视化4类。对于节点分类，大致有两种方法的类别，即使用随机游走传播标签的方法^[8-9]和从节点中提取特征并对其应用分类器的方法^[10-11]。对于链接预测方法，包括基于相似性的方法^[12]、最大似然模型^[13]和概率模型^[14]。对于聚类方法，包括基于属性的模型^[15]和直接最大化集群间距离的方法^[16]。

1.2 专家专长匹配方法的相关研究

机器学习方法应用于研究专家遴选方法的问题已受到广泛的关注，且获得了更好效果，节省了成本。机器学习方法利用了专家信息和评审文档的语义相关性，以提高推荐专家的学术专长匹配精度。目前常用的方法是基于知识嵌入的方法和基于深度学习模型的语义匹配方法。

基于知识嵌入的方法，提供了在连续语义空间中建模主题和主题相关性的可能性。基于单词向量和单词的分布式表示，已被证明能够有效地捕获语言中的语义规律^[7]。具有相似语义和句法属性的单词往往会被投影到向量空间的附近区域。将LDA中原有的离散单词类型替换为连续词嵌入，高斯LDA^[18]表明可以在主题模型中纳入词嵌入中的附加语义，并进一步增强性能。

基于深度学习模型的语义匹配方法，提高了语义间的联系，使模型能够准确推荐。何柔莹等^[19]基于卷积神经网络框架，将注意力机制引入模型的输入层，以增强特征提取的能力。Duan等^[20]采用基于句子对建模的评审专家的匹配方法，通过监督信息来建模句子对之间的关系。在卷积神经网络的基础上监督信息，使模型能够更准确地学习评审专家和文档之间的特征。为提高颠覆性技术预判系统中专家遴选的时效性和准确性，本文提出了一种专家学术专长与评审综合指标结合的方法。

2 颠覆性技术专家预判平台专家遴选方法

本文提出的颠覆性技术专家预判平台专家遴选方法的具体流程如图1所示。首先，对专家的学术成果文档集合进行分析，提取出专家的学术标签特征并进行排序；然后，对待预判的颠覆性技术文本文档提取出关键词并进行多特征学术专长语义匹配，遴选出适合的评审专家。

2.1 学术专长匹配方法

基于颠覆性技术专家预判平台的专家学术专长匹配方法如图2所示。首先，在专家专长特征提取时，分析专家的学术成果集合，提取出专家的候选专长标签并进行排序，从而有效地减小了标签规模；然后，同评议特征表示进行匹配度计算，筛选出匹配度高的候选专家。

2.1.1 概念定义

专家表示提取由异质语义表示集合和学术成果文档集合两部分构成。专家异质语义表示集

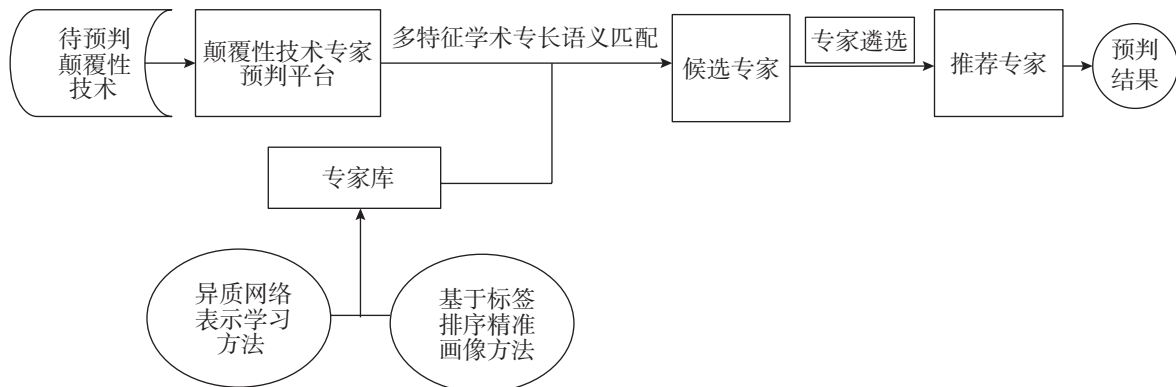


图1 颠覆性技术专家预判平台专家遴选框架

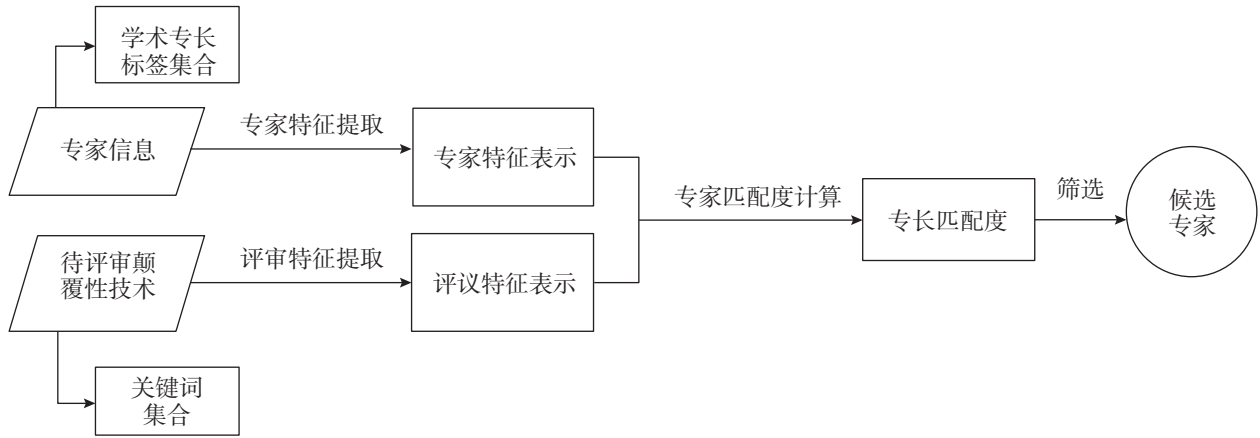


图2 学术专长匹配流程

合 H_E 提取的具体步骤是：专家集合 E 包括专家库中所有专家，即 $E = \{e_i | 1 \leq i \leq |E|\}$ 。其中， e_i 表示第 i 个专家， $|E|$ 表示库中专家总数。异质语义表示集合 $H_E = \{h_{e,i} | 1 \leq i \leq |E|\}$ 。其中， $h_{e,i}$ 为专家 e_i 的异质语义表示。专家 e_i 的学术成果文档由 d_i 表示，用论文、专著和研究报告等来描述专家学术成果。学术成果文档集合 $D_i = \{d_i | 1 \leq d \leq |D_i|\}$ 。其中， $|D_i|$ 表示专家 e_i 的学术成果文档数。学术专长标签用 $l_{i,k}$ 表示专家 e_i 的第 l 个学术专长标签。 $L_{e,i} = \{l_{i,k} | 1 \leq l \leq |L_{e,i}|\}$ 表示学术专长标签集合。其中， $|L_{i,k}|$ 表示专家 e_i 的学术专长标签数。专长匹配度 $score_{i,f} \in [0,1]$ 表示学术专长 $l_{i,k}$ 与专家 e_i 匹配度的数值，数值越大匹配度越高。

2.1.2 专家专长标签生成排序

对待评估专长标签 $l_{i,k}$ 使用 TextCNN 模型进行向量化，得到待评估专长表示 $X_{i,k}$ 。对专长匹配进行得分计算，首先利用异质语义表示和专长表示向量乘法获取，将匹配特征输入全连接层，对神经网络进行训练并输出专长匹配度，并对标签进行排序。用 t_i 表示给定专家的候选标签专长的异质语义表示，专家 e_i 在专长标签 $l_{i,k}$ 上的专长匹配得分计算公式为：

$$X_{\max,n} = [\psi(X_{t,1;s}), \dots, \psi(X_{t,c;c+s}), \dots, \psi(X_{t,d;d+s})] \quad (1)$$

$$X_{i,l} = [\max(X_{m,1}), \dots, \max(X_{m,n}), \dots, \max(X_{m,d})] \quad (2)$$

式中， ψ 表示卷积函数，给定卷积核 s 和卷积数量 d 对词向量矩阵 X_i 进行卷积后，生成专长标签 $l_{i,k}$ 的特征映射矩阵 X_m ，然后利用最大池化函

数对特征映射矩阵进行池化，生成专长标签 $l_{i,k}$ 向量表示 $X_{i,k}$ 。

其次，对专长匹配得分 $score_{i,j}$ 进行计算。其计算公式为：

$$X_{prf,i,l} = \tanh(X_{e,i} \otimes X_{i,l}) \quad (3)$$

$$score_{i,l} = \sigma(W_{prf} \cdot X_{prf,i,l} + b_{prf}) \quad (4)$$

式中， \otimes 表示向量乘法； \tanh 表示双曲正切函数； W 和 b 分别表示权重和偏置； σ 为 sigmoid 函数。

基于标签排序的学术专长使用 LabelRank 算法筛选排名 topK 专长画像输出结果 $TOP_i = \{L_{e,i} | 1 \leq i \leq |E|\}$ ，对每位专家的学术专长画像进行迭代，返回学术专长画像结果。利用 GetElement 函数获取专家集合 E 中每个专家的异质语义表示 e_i ；使用 LabelMining 和 LabelVector 函数对文档集合 D 进行候选专长挖掘，生成专长召回和标签表示。根据式 (3) 和式 (4) 计算专长匹配度计算得分对匹配的专家专长标签进行降序排序，将专家 e_i 的学术专长标签集合 $l_{e,i}$ 加入专家学术专长画像中。

2.2 专家遴选综合评价指标

在综合指标体系^[22]基础上，本文将专家基本情况指标分为年龄、职称、学历和学术研究领域；将专家评议业绩指标分为累计数、命中率、成功率和综合离散率；将专家道德修养指标分为科学道德、科学态度和工作作风。从而建立了一个综合指标体系来评估专家的综合能力和素质。这些指标的具体含义见表 1。

表1 专家遴选指标体系表

一级指标	二级指标	准则
A1 基本情况指标	B1 年龄	<45、45 ~ 60、>60
	B2 职称	高级职称、副高、中级
	B3 学历	全日制、自考、函授
	B4 学术	任职国际、国家、省部
	B5 研究领域	单学科、交叉学科、边缘学科
A2 评议业绩指标	B10 累计数	高、中、低
	B11 命中率	>90%、70% ~ 80%、60% ~ 70%、<60%
	B12 成功率	>90%、70% ~ 80%、60% ~ 70%、<60%
	B13 综合离散率	高、中、低
A3 道德修养指标	B14 科学道德	优、良、中、可、差
	B15 科学态度	优、良、中、可、差
	B16 工作作风	优、良、中、可、差
A4 行业影响力	B17 社会工作	多、适量、少
	B18 社会头衔	多样、一般、少
	B19 社会活动	广泛、一般、少

评审专家的基本情况是评价的基础，道德修养是评议业绩的支撑与保障。因此，基本情况、道德修养、行业影响力和评审业绩构成了评审专家评价指标体系中其他不可或缺的4个重要组成部分，由于不同领域专家遴选的侧重点不同，指标的权重分配也不同。

2.2.2 综合指标特征提取

本文采用熵权法来定义专家的综合评价指标。其基本评价模型为：待评方案为 m ，评价指标项为 n ，形成指标数据矩阵 $X = (X_{ij})_{m \times n}$ 。对于指标 X_j ，其信息熵为 $E_j = -k \sum_{i=1}^m p_{ij} \ln p_{ij}$ ，其中 $p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}}$ 。由此可以看出，对于指标 x_j ，指标值 x_{ij} 的差距越大，信息熵越小，指标的变化程度就越高，提供的信息量就越大，从而导致这项指标的权重越大；相反，这项指标的权重也就越小。本文将评议业绩指标 (A1)、评议业绩指标 (A2)、道德修养指标 (A3)、行业影响力 (A4) 这4维综合评价指标特征与专长匹配特征进行拼接，作为训练模型的输入特征。

3 实证结果分析

3.1 实验数据

本文利用开源学术数据集 Academic Social Network^[21]对专家的学术专长进行画像，并验证

融合特征标签排序方法的有效性和准确性。这个学术网络是由AMiner学术知识库构建而成的，其中包含了1 712 432位学者、2 092 356篇论文以及论文的标题、摘要和作者信息。本文对数据集进行预处理，提取了包含34 278位专家、93 462篇论文、6 785个出版物数据的子集，包含专家、文献标题、文献摘要和文献引用等信息。此处选择80%数据集作为训练集，另外20%的数据集作为测试集，训练集中包含27 422位专家，测试集中包含6 856位专家，旨在训练一个学术专长画像模型并对其识别性能进行评估。

3.2 实验结果与分析

本文使用精确率、召回率和F1评估指标作为评价标准，并使用论文本挖掘方法中常用的TFIDF、TextRank和LDA模型进行对比实验。top5标签的实验结果见表2。

从表2可以看出，使用文本挖掘方法中的前5种方法进行召回可以包含大部分专家的学术专长标签。这表明在保证能够充分召回学术专长标签的前提下，选择文本挖掘方法对LabelRank方法的最终性能影响较小。然而，这些方法无法很好地区分有效的学术专长标签。在所使用的TextRank、LDA和TFIDF文本挖掘方法中，TextRank方法的精度指标略好于LDA和TFIDF方法。但是，这3种方法的精确率和召回率均不高，说明它们识别出的专长标签中包含大量噪声

标签，难以进行区分。

实验的对比结果如图 3 所示。其中，实线表示精确率，虚线表示召回率。从图 3 中可以看出，LabelRank 方法可以有效地对学术专长标签进行筛选，从而提高学术专长画像的精确度。实验结果表明，相较于文本挖掘方法，LabelRank 方法在精确率和召回率方面表现更优，说明它能够有效区分文本挖掘方法召回的专长标签，并进行合理的打分和筛选。此外，随着学术专长标签集合的增大，精确率会逐渐降低，但召回率会逐渐增加。因此，在确定学术专长标签集合大小时，需要根据具体的精确率和召回率需求进行权衡。根据是更侧重于精确率还是更侧重于召回率的不同需求，选择更适合的学术专长标签集合的大小。

专家#1257961 的学术专长识别示例如表 3 所示。从结果中可以看出，文本挖掘方法在区分正例学术专长标签方面表现不佳，即使在得分较

高的结果中也仍然存在负例学术专长标签。相比之下，在本文使用的 LabelRank 方法中，得分较高的正例学术专长标签能够在学术专长画像结果中排名更加靠前。具体而言，TextRank 打分中将负例学术专长标签的“optimization model”排在第 1 位；TFIDF 打分的前 5 个结果中包含两个负例学术专长标签。同时，LDA 打分中也包括一个负例标签“overall throughput”。与其他方法相比，使用 LabelRank 方法可以将被标注为正例学术专长标签能够更准确地识别，本文使用的 LabelRank 方法有效性得到了验证。

在颠覆性技术专家预测平台系统中，对评审综合指标在全连接层进行特征融合。在实验中，对 A1、A2、A3、A4 进行 0 到 1 之间的随机取值，并选取了有对比度的几组结果在表 4 中展示。

从表 4 可以看出，对评审指标特征进行随机

表 2 top 5 对比实验结果

模型	精确率	召回率	F1
TFIDF	0.787	0.543	0.643
TextRank	0.854	0.585	0.694
LDA	0.805	0.658	0.724
LabelRank	0.905	0.822	0.862

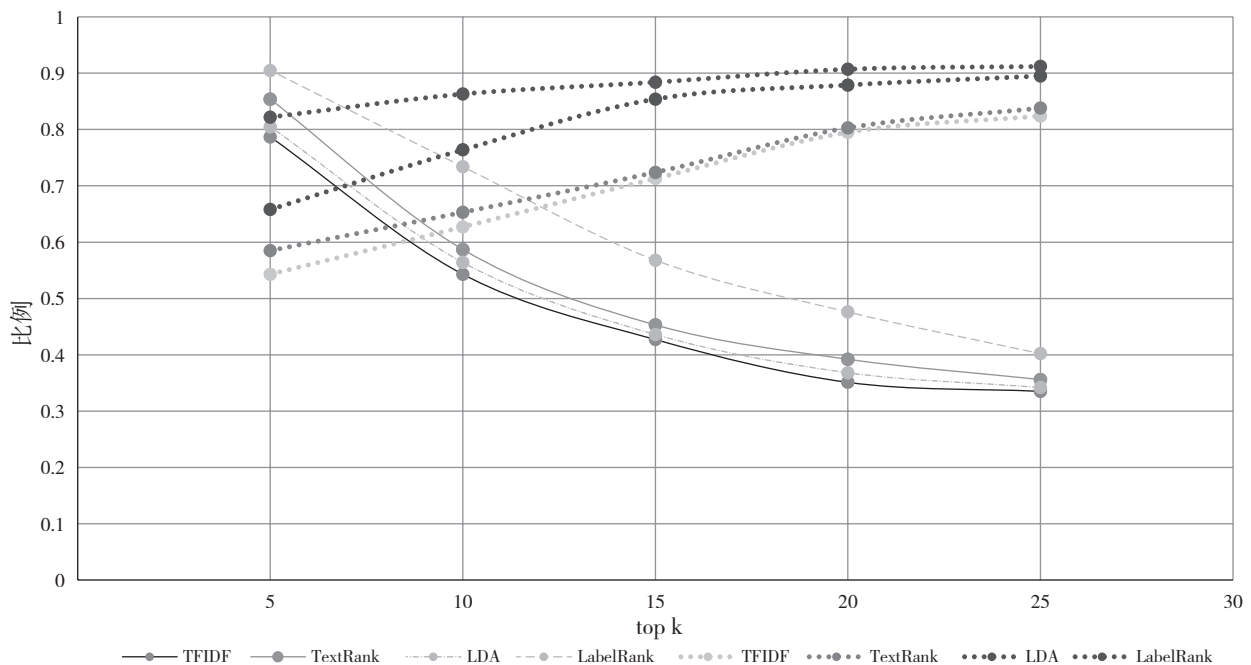


图 3 学术专长实验结果对比

表3 专家#1257961的学术专长标签匹配实例

模型	标签标注结果
TFIDF	load balancing(1);temporary table(0);independent task(1);insulin infusion rate(0);task allocation(1)
TextRank	optimization model(0);task allocation(1);load balancing(1);static task graph(1);task scheduling(1)
LDA	load balancing(1);dynamic link algorithm(1);overall throughput(0);task allocation(1);initial task(1)
LabelRank	independent task(1);load balancing(1);task allocation(1);initial task(1);static task graph(1)

表4 融合特征结果对比

方法	A1	A2	A3	A4	标签打分结果
LabelRank	—	—	—	—	independent task(0.743); load balancing(0.742); task allocation(0.705); initial task(0.684); static task graph(0.673)
LabelRank+ 评审指标特征 1	0.54	0.46	0.73	0.46	independent task(0.745); load balancing(0.744); task allocation(0.713); initial task(0.701); static task graph(0.691)
LabelRank+ 评审指标特征 2	0.87	0.92	0.68	0.77	independent task(0.824); load balancing(0.822); task allocation(0.803); initial task(0.787); static task graph(0.773)
LabelRank+ 评审指标特征 3	0.13	0.41	0.21	0.13	independent task(0.623); load balancing(0.622); task allocation(0.565); initial task(0.547); static task graph(0.543)

赋值后，分别选择了特征数值适中（评审指标特征1）、较大（评审指标特征2）、较小（评审指标特征3）的3组特征进行结果对比。在融合评审指标特征后可以看出，评审特征对打分结果会造成正相关的影响，并且在较高的评审特征中对原本打分越低的标签结果越高，而在较低的评审特征中对原本打分越高的标签则其影响越高。同时，也能明显地看到，融合评审指标特征会影响标签打分的最终结果，可以使综合指标（基本情况、道德修养、行业影响力和评审业绩）较高的专家在标签匹配上获得更高的得分，具有更高的被推荐评审项目的可能。

4 总结与展望

为了确保颠覆性技术预判系统的评议质量，选择合适的专家是至关重要的。为此，本文建立

一个专家库，采用融合专家综合评价指标特征的匹配方法，为专家遴选提供与专业背景相关的候选专家列表。在考虑待评估项目与专家的学术专长匹配情况的同时，还考虑专家所处的位置和工作量，以确保专家的可用性和可中断性。针对颠覆性技术专家预判平台中对专家遴选的准确性和时效性的要求，本文鉴于学术网络表示学习的方法既可以避免大量特征工程，也方便不同类型的特征进行融合，提出了一种基于异质网络的融合特征方法。这种方法提取专家学术专长标签和专家综合指标，输入全连接层进行训练，将专家的科研能力与综合指标相结合，进一步为预判平台遴选出合适的专家人选。

本文在 Academic Social Network 数据集进行实验并取得了很好的效果。在与其他模型的对比实验中，这种方法的精确率达到了 90.5%，召回

率达到 82.2%。实验证明, 此方法在学术专长识别性能上均优于其他方法。结果表明, 本文提出的方法可以提升技术预判专家学术专长匹配。在加入专家综合指标特征后, 专家的综合指标特征能有效地反馈到实验结果中, 从而提高评审系统的时效性和智能性。

本文研究的不足之处是暂时没有考虑到冷启动的问题。在今后的研究工作中, 将着重于冷启动和实时推荐等问题, 进一步围绕评估项目与专家的学术专长匹配的专家遴选问题展开研究, 并应用于颠覆性技术专家预判平台的专家推荐中。

参考文献

- [1] BEDARD J. Expertise and its relation to audit decision quality[J]. *Contemporary accounting research*, 1991, 8(1): 198–222.
- [2] 陈孙, 李楠, 梁冰, 等. 基于成果特征的学者学术专长识别方法[J]. *图书情报工作*, 2019, 63(20): 96–103.
- [3] BHARTI S K, BABU K S, PRADHAN A, et al. Automatic keyword extraction for text summarization in multi-document e-newspapers articles[J]. *European journal of advances in engineering and technology*, 2017, 4(6): 410–427.
- [4] RADA M, PAUL T. TextRank: bringing order into text[C]//*Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*. Stroudsburg: ACL, 2004: 404–411
- [5] PAGE L, BRIN S, MOTWANI R, et al. The pagerank citation ranking: bring order to the web[R/OL].(2008–12–28)[2021–04–12].<http://ilpubs.stanford.edu:8090/422>.
- [6] SEBASTIANI F. Machine learning in automated text categorization[EB/OL].[2023–06–13].<https://courses.ischool.berkeley.edu/i256/f06/papers/sebastiani02.pdf>.
- [7] AGGARWAL, CHARU C, ZHAI Chengxiang. An introduction to text mining[EB/OL].[2023–06–13].https://link.springer.com/chapter/10.1007/978-1-4614-3223-4_1.
- [8] AZRAN A. The rendezvous algorithm: multi-class semisupervised learning with Markov random walks[EB/OL].[2023–06–13].<https://www.gatsby.ucl.ac.uk/~azran/ICML07.pdf>.
- [9] BALUJA S, SETH R, SIVAKUMAR D, et al. Video suggestion and discovery for youtube: taking random walks through the view graph[EB/OL].[2023–06–13].https://www.researchgate.net/publication/221023074_Video_suggestion_and_discovery_for_youtube_Taking_random_walks_through_the_view_graph.
- [10] BHAGAT S, ROZENBAUM I, CORMODE G. Applying link-based classification to label blogs [EB/OL]. [2023–06–13].<https://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=0E9060C47D0EB5A98B244AABC FB4E198?doi=10.1.1.160.7259&rep=rep1&type=pdf>.
- [11] LU Q, GETOOR L. Link-based classification[EB/OL]. [2023–06–13].https://www.researchgate.net/publication/221345326_Link-based_Classification.
- [12] HABERMEHL S, SCHLESINGER C, PRILL D. Comparison and evaluation of pair distribution functions, using a similarity measure based on cross-correlation functions[J]. *Journal of applied crystallography*, 2021, 54(2): 612–623.
- [13] CLAUSET A, MOORE C, NEWMAN M E. Hierarchical structure and the prediction of missing links in networks[J]. *Nature*, 2008, 453 (7191): 98–101.
- [14] FRIEDMAN N, GETOOR L, KOLLER D, et al. Learning probabilistic relational models[EB/OL].[2023–06–13].<https://max.book118.com/html/2017/0724/123956190.shtm>.
- [15] ZHOU Y, CHENG H, YU J X. Graph clustering based on structural/attribute similarities[EB/OL]. [2023–06–13].https://www.researchgate.net/profile/Yang-Zhou-27/publication/220538529_Graph_Clustering_Based_on_StructuralAttribute_Similarities/links/0046353c0210f50c3c000000/Graph-Clustering-Based-on-Structural-Attribute-Similarities.pdf.
- [16] SHI J, MALIK J M. Normalized cuts and image segmentation[EB/OL].[2023–06–13].<https://www.semanticscholar.org/paper/Normalized-cuts-and-image-segmentation-Shi-Malik/b94c7ff9532ab-26c3aedbee3988ec4c7a237c173>.
- [17] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [EB/OL].[2023–06–13].<https://www.semanticscholar.org/paper/Distributed-Representations-of-Words-and-Phrases-Mikolov-Sutskever/87f40e6f3022adbc1f1905e3e506abad05a9964f>.
- [18] RAJARSHI D, MANZIL Z, CHRIS D. Gaussian lda for topic models with word embedding[C]//*Proceedings of the 53rd Annual Meeting of the Association for Com-*

- putational Linguistics. 2015.
- [19] 何柔莹, 徐建. 基于注意力卷积神经网络的工作票专家推荐方法[J]. 南京理工大学学报(自然科学版), 2019, 43(1): 17-25.
- [20] DUAN Z, TAN S, ZHAO S, et al. Reviewer assignment based on sentence pair modeling[EB/OL].[2023-06-13].https://www.nstl.gov.cn/paper_detail.html?id=97a09b2da40664c058cadf5af687d769.
- [21] TANG J, ZHANG J, YAO L M, et al. ArnetMiner: extraction and mining of academic social networks[EB/OL].[2023-06-13].https://www.researchgate.net/profile/Zhong-Su-3/publication/51986580_ArnetMiner_Extraction_and_Mining_of_Academic_Social_Networks/links/00463529d7dd7ece84000000/ArnetMiner-Extraction-and-Mining-of-Academic-Social-Networks.pdf.
- [22] 高琢玉. 基于多目标决策的专家遴选算法的研究[D]. 长沙: 中南大学, 2024.

(上接第26页)

- [13] 陈洁. 科技报告质量管理评价体系研究[J]. 中国科技资源导刊, 2019, 51(2): 55-60.
- [14] 陈传夫, 代钰珠, 曾建勋. 科技报告开发利用与知识产权问题研究[J]. 情报学报, 2014(8): 793-799.
- [15] 许燕, 张爱霞, 麻思蓓. 科技报告服务中的知识产权平衡机制[J]. 科技管理研究, 2018(3): 193-197.
- [16] 任亮, 许燕, 郑彦宁. 科技报告全周期服务著作权风险研究: 以国家科技报告服务系统(NSTRS)为例[J]. 情报科学, 2022(2): 148-154.
- [17] 陈洁, 韩非, 云杉, 等. 地方科技报告制度建设探究[J]. 中国科技资源导刊, 2017(2): 1-7.
- [18] 何静, 蔡利超. 广东省科技报告制度建设现状及对策研究[J]. 中国科技资源导刊, 2017(3): 61-66.
- [19] 王星, 赵捷. 国家科技报告服务系统构建研究[J]. 中国科技资源导刊, 2015(5): 26-34.
- [20] 汪芸辉, 薛钦文. 国家科技报告服务系统用户满意度模型研究[J]. 中国科技资源导刊, 2019(4): 82-87.
- [21] 陆丽娜, 袁芳. 关系网络范式下科研合作选择研究: 以纳米技术科技报告为例[J]. 科技管理研究, 2020(23): 115-123.
- [22] 沈思, 李沁宇, 叶媛, 等. 基于TWE模型的医学科技报告主题挖掘及演化分析研究[J]. 数据分析与知识发现, 2021(3): 35-44.
- [23] 剧晓红, 毛平. 创新驱动下的科技报告服务影响因素研究[J]. 现代情报, 2021(2): 107-114.
- [24] 毛刚, 贾志雷, 侯人华. 情报学视角下的科技报告研究[J]. 情报杂志, 2013, 32(12): 62-66.
- [25] 汪雪峰, 付芸, 邱鹏君, 等. 关于我国国家科技报告服务模式探索[J]. 科技管理研究, 2016, 36(7): 190-195.
- [26] 应向伟. 科技报告服务模式及在科技管理中的探索研究[J]. 科技管理研究, 2018(2): 34-38.
- [27] 陈峰. 面向企业用户的科技报告增值服务方式探讨[J]. 中国科技资源导刊, 2017(4): 51-54.
- [28] NTIS[EB/OL]. [2023-07-31]. <https://www.ntis.gov/>.
- [29] 张爱霞, 杜薇薇. 美国联邦政府科技信息管理及对我国的启示[J]. 中国科技资源导刊, 2022(3): 21-30.
- [30] 杨贺晴, 徐枫, 周秀霞, 等. 5R视角下CASHL学科用户调研及营销策略研究: 以东北师范大学为例[J]. 大学图书馆学报, 2020(6): 19-27.
- [31] 徐建华, 路锦怡, 汪汉清. 2018年图书馆学期刊问卷调查法文章分析[J]. 图书馆学研究, 2020(8): 2-9.
- [32] 夏立新, 白阳, 康雨培, 等. 基于BIBFRAME的科技报告语义服务研究[J]. 情报学报, 2017, 36(7): 695-705.