基于主题契合度的专家推荐模型研究*

靳健[1]，杨海慈[1]，李凝[2]，耿骞[1]
（1.北京师范大学政府管理学院，北京 100875；2.中央财经大学外国语学院，北京 100081）

摘要：评审专家遴选是会议、期刊等投稿论文评审的一项重要步骤。本研究首先依据主题在学科中出现频率，分析待评审论文及候选专家论文的主题重要性；其次，综合考虑主题重要性，构造整数优化模型以高效地为多篇待审文稿推荐评审专家，并加入专家专长与投稿论文的匹配度、评审专家权威性、评审专家工作量分配等因素以满足专家推荐的现实需求；最后，从平均覆盖率、人均审核率和重要性匹配度三个角度对本模型的有效性进行验证。结果表明，提出的优化模型可以较好地完成专家推荐任务。

关键词：评审专家推荐；主题契合度；整数优化
中图分类号：TP301.6
DOI：10.3772/j.issn.1673-2286.2017.04.007

1 引言

投稿论文的评审是学术会议、学术期刊组织者的一项重要工作。论文评审组通常需先对投稿论文进行分类，其次邀请合适学者担任评委完成论文评审。在此过程中，如何高效、准确地遴选出合适的评审专家成为一个值得讨论的问题。

评审专家的遴选标准包括专家权威度[1-2]、专家知识结构与待评审论文主题的匹配程度[3-4]、专家学术关系网络等[5-6]。针对专家权威度，部分研究应用现有评价指标（如H指数等），部分研究综合多方面因素提出新的评价指标[2]。在主题提取方面，有的研究利用主题模型分析论文主题与专家知识结构的相关性[7]，有的研究利用模糊语义分析专家主题[8]，也有部分研究利用社会网络模型等实现专家推荐[9-10]。

但是，在以主题表示论文或专家知识结构的相关研究中，较少考虑不同学科背景下的主题重要性差别。例如，在特定学科内，部分主题较基础但研究者较多；相反，在同学科中，可能还存在部分主题关注度较低的情况。假设某一待评审论文重点探讨了这种关注度相对较低的主题，为保证评审质量，对于该待评审论文，应尽可能推荐对该主题相对较了解的评审专家。为此，本研究将首先分析主题在学科中的出现频率，从而形成一种考虑“非均匀主题”的评审专家推荐优化模型。

本研究提出的方法将可应用于有关学术论文和学术项目审批的专家推荐场景。本研究的创新点在于：结合主题在学科背景中的出现频率，对不同主题的重要性做出标识，使主题向量更好地表达待评审论文特征，进而优化评审专家推荐效果。

2 国内外研究现状

2.1 专家发现与专家推荐

专家发现问题最早出现在问答平台及社交媒体的相关研究中。与专家推荐相类似，专家发现主要考虑如何利用专家的活动来定义和选择专家。例如，Hwang等假设专家发现的目的是帮助用户跟随专家的选择[11]，因此，将专家定义为对某一类别产品给出较多评论的用户；Zhou等对基于链接分析专家的方法提出质疑，并根据用户在某一主题上的专业性和声誉来评定专家[12]；Wang等提出仅在小圈子活动的专家可能地位虚高，并从内容和社区地位角度考虑专家的衡量标准[13]；梁凯春等对专家和知识设定不同级别的中心度和权

* 本研究得到教育部人文社会科学研究青年基金项目“面向论文评审专家推荐的兴趣变化挖掘与回避机制生成的研究”（编号：16YJC870006）和ISTIC-EBSCO文献大数据发现服务联合实验室基金项目“融合异构科研数据的评审专家推荐研究”资助。

威度，并假设专家的中心级别由其推荐内容的权威度决定，而知识的权威度由推荐人的中心度决定[14]。还有许多不同的方法对专家发现问题做出分析，并利用主题模型对其展开分析，如Daud等分析潜在狄利克雷分配（Latent Dirichlet Allocation）、时间-作者-主题（Temporal-Author-Topic）等模型间的差异，并提出一种考虑时间因素的时间-专家-主题模型（Temporal-Expert-Topic）[7]。

社交问答平台中的专家发现与评审中的专家推荐在问题定义、模型算法等方面存在很多相似之处。例如，社交平台中用户的提问可以类比评审环境中的待审论文；网站评论、转发类比论文中的引用关系；用户发帖量类比学者发文量。同时，这两种方式中的研究问题各有其特殊性。例如，社交网络环境中的专家发现还需要考虑论坛话题、用户与专家的互动，而在评审环境中则需要考虑期刊、评审历史等因素。

在评审专家推荐方面，不少研究主要依据专家权威度、待评审论文主题、专家社会关系等指标。Sun等关注专家权威度、出版业绩、项目绩效、历史评价绩效和同行评审等方面[8]，其研究中的专业水平和历史绩效通过专家自行填写表格确定，因而该推荐结果在一定程度上可能会受主观因素的影响；Li等利用某一领域的引用关系计算待审论文和审稿人的相似性[15]，通过论文数量、论文质量和时间变量描述专家的权威度；Han等分析专家论文的发表历史、与待评审论文的主题匹配度、专家与投稿者的社会关系、专家权威度四个因素，并从准确率和召回率评价实验结果[16]；Sun等通过论文与专家的相关性、专家发文质量以及社区连通性等多个因素实现专家推荐，并借助MapReduce提高运行效率[6]；Yukawa等利用余弦相似度描述作者间、作者与论文以及论文间的关系[17]。也有研究融合多种不同指标并赋予相应权重，以计算专家得分，最终根据得分排名实现评审专家的推荐。例如，Zheng等提取关于专家和文档的多组特征，并利用基于排序学习的算法推荐评审专家[2]；Han等通过单独评价各指标的准确率和召回率，并根据各指标的表现来确定权重[16]。有的研究使用层次分析法确定各指标权重[6,8,15]，但该方法需人工参与，结果易受主观因素影响。为克服上述方法存在的弊端，还有研究通过图论和网络模型法均衡各指标权重，以期实现专家推荐效果最优化[1,18,19]。

专家发现与专家推荐的相关研究重点分析了专家权威度、待评审论文与专家论文间的匹配等多个因素，并将评审专家推荐理解为一种检索问题，而相对较少考虑评审活动的实际现状，即通常需要对多篇而非单篇论文作出推荐专家等。因此，本研究将综合考虑多方面因素，并对多篇论文的专家组推荐展开研究。

2.2 评审分配问题

评审分配问题大多考虑多篇论文需邀请多名专家评审等实际问题，并分析许多较实际的限制条件。Xu等强调评审分配中论文数量的重要性[3,20]，即为每位评审专家分配合适数量的论文，以保证审稿质量；Fan等运用知识规则最大限度地分析论文的分组问题，以使论文分组的组间相似性较小而组内相似度较大[4]；Wang等将评审分配问题定义为“组对组”的任务分配[21]，即在优化分配前对专家和论文分组，以控制待评审论文的评审者数量，减少评审者负担。

整合待评审论文与评审专家的约束条件挑战。Karimzadehgan等将专家与待评审论文的匹配问题构建为一个整数优化模型[22-23]，首先利用主题模型分析专家与待评审论文的相关性，并将被分配论文量、专家兴趣、待评审论文主题以及专家主题匹配度等因素纳入优化模型中；在此基础上，Wang等将模型进一步扩展为具有多个自变量的优化模型[21]，以实现专家推荐；Tang等利用优化模型实现专家分配，并将结果应用于学生选择导师等情境[24]；还有研究将作者和待评审论文抽象为网络结构中的节点，利用随机游走模型实现待评审论文的最优化分配[9]。

以上评审分配问题的相关研究将评审专家的推荐理解为一种任务分配题，综合考虑论文评审活动中多个实际因素（如专家工作量、待评审论文的最少评审专家数量、专家组主题多样性及专家组主题对评审论文主题的覆盖等），最终建立优化模型为多篇论文推荐合适专家。但相关研究相对较少地考虑同一学科中不同主题重要性对专家组推荐的影响。

3 学科背景下主题重要性的专家推荐优化模型

3.1 问题定义

假设待评审论文n篇，专家库R中包含专家m人，即$R=\{r_i\}_{i=1}^{m}$。专家r_i发表的论文集为d_i，所有专家发表的论

文共同组成论文集合$D=\{d_i\}_{i=1}^{m}$。假设专家论文集d_i中包含h个主题，构成向量$\tau=\{t_i\}_{i=1}^{h}$；并且，假设各主题与专家的相关性为C_{ij}（$i\in(1,m)$，$j\in(1,h)$），主题与待评审论文的相关性为B_{ij}（$i\in(1,n)$，$j\in(1,h)$）。同时，专家遴选系统还需考虑真实场景的限制条件。例如，每位专家评审的论文数量不超过p，每篇待评审论文的评审专家人数不低于q等。本研究各符号含义如表1所示。

表 1 符号定义

m	专家库总人数
n	投稿论文数量
h	专家论文及待评审论文中提取的主题个数
p	每篇论文评审组的专家数量
q	每位专家可以评审的最多论文数量
$R=\{r_i\}_{i=1}^{m}$	专家库R
$D=\{d_i\}_{i=1}^{m}$	专家论文共同组成的论文集
$\tau=\{t_i\}_{i=1}^{h}$	h个主题组成的向量以表示专家论文集
B	$n\times h$矩阵，B_{ij}表示主题j在论文i中出现的概率
C	$m\times h$矩阵，C_{ij}表示主题j在专家i中出现的概率
$F=\{f_i\}_{i=1}^{h}$	h个主题在特定学科中的概率分布
S	$n\times h$矩阵，表示考虑文档环境后论文与主题的相关性
Y	$m\times n$的0—1矩阵，表示评审组员和待评审论文的分配关系

3.2 建模过程

3.2.1 整体思路

在模型构建中，本研究以专家已发表论文为基础数据，对待评审论文的主题重要性作出分析，并通过以下4个步骤实现专家组的推荐，具体流程如图1所示。

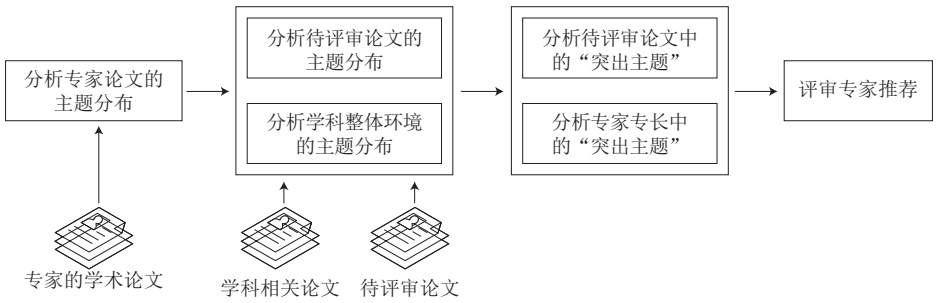

图 1 专家推荐过程

（1）专家论文主题分析：获取候选专家已发表论文，分析这些论文的主题分布。（2）待评审论文及学科论文主题分析：以候选专家库中专家发表的论文主题分布为依据，分析待评审论文的主题及学科整体环境中的主题分布。（3）突出主题分析：以待评审论文的主题分布及学科整体环境中的主题分布为依据，分析各主题在待评审论文中是否为突出主题。（4）专家分配与推荐：在推荐评审专家时，既要保证同一位专家不能分配过多论文，也要保证每篇论文的审稿专家数量。本文将据此分别建立优化框架，融合主题重要性，得到专家推荐列表。

3.2.2 主题提取

（1）文档主题生成模型简介。本研究拟用文档主题生成模型（Latent Dirichlet Allocation，LDA）模型提取分析专家发表论文及待评审论文的主题分布。LDA是一种层次贝叶斯模型[25]。模型假设一篇文档是由一系列多项式分布的主题表示，而主题由一系列多项式分布的单词来表示。

在LDA模型中，存在两个先验假设，一种是文档对主题的先验假设$\theta_d\sim Dir(\bullet|\alpha)$，另一种是主题对单词的先验假设$Dir(\bullet|\beta)$。由于存在后验概率无法直接求解的问题，因此主题模型一般采用近似求解的方式求得参数。如Blei等采用基于变分贝叶斯的方式[25]，Griffiths等采用基于吉布斯采样的方式[26]，Minka利用基于期望推进等方式获取LDA模型的参数[27]。一般来说，吉布斯采样由于简单直观且效果较好，因此成为一种较常用的参数求解方式。

（2）主题提取过程。评审专家推荐首先要保证所推荐的专家熟悉待评审论文的研究领域。本研究假设待评审论文内容的最小单元是“主题”，且论文内容可用主题向量表示。为实现待评审论文内容与专家研究方向的匹配，本研究从待评审论文及专家论文集D中提

取主题。

假设专家r_i的全部论文可以拼接成一篇文档d_i，以分析专家r_i的研究领域。其中，θ_{di}为d_i中h个主题的概率分布。m位审稿专家的论文形成集合$D=\{d_1, d_2, \cdots, d_m\}$。本研究利用LDA模型估计集合$D$的主题概率分布。根据LDA模型思想，集合$D$的似然估计可表示为公式（1）。

$$\log p(D|\alpha, \beta) = \sum_{di\epsilon D}\sum_{w\epsilon di} c(w, d_i) \log\left(\sum_{Z=1}^{h} p(w|z, \beta)\, p(z|d_i, \theta_{di})\right) \quad (1)$$

其中，α和β为LDA的先验参数，$c(w, d_i)$是论文d_i中词语w的数量，$p(w|z, \beta)$是主题z生成词语w的概率，$p(z|d_i, \theta_{di})$是论文d_i包含主题z的概率。本研究利用吉布斯采样算法估计论文d_i的主题分布，并输出h个主题在d_i中的分布。

以所有候选评审专家的论文集合作为样本，LDA模型可帮助获得h个主题及这些主题在集合D中的概率分布C。随后，以待评审论文作为测试集，分析h个主题在待评审论文中的分布B及在学科整体环境中的分布F_t。其中，“学科整体环境”由学术数据库中与该学科相关论文模拟。

3.2.3 主题重要性建模

在某一特定学科范围内，出现频率较高的主题可能是较热门或较基础的内容。对于待评审论文来说，若某一主题在学科中出现频率低，但在该论文中出现频率较高，则说明该主题与待评审论文关系密切，且可能在论文中较重要。因此，为保证待评审论文的评审质量，需要为“小众主题”安排更专业的评审专家。

在一般情况下，若两个主题在待评审论文i中出现概率相同，则在学科范围内，出现概率相对较小者标识该论文的能力较强。根据该假设，在学科中出现频率较高的主题应赋予较小权重，反之赋予较大的权重。结合主题提取的结果，利用TFIDF思想，待评审论文i与主题j的相关性可定义为公式（2）。

$$\forall_i \in [1, n],\ \forall_j \in [1, h] P_{ij} = B_{ij} \times log\left(\frac{1}{fj}\right) \quad (2)$$

同时，对于某特定待评审论文或专家论文，根据论文与主题的相关性矩阵，利用聚类算法分析h个主题与论文的相关度s。然后，选取s值最低的一类主题为待评审论文中的“不突出主题”，其他主题为“突出主题”。

3.3 评审专家推荐的优化模型

评审专家的推荐应包括限制条件为：每一篇待评审论文有p位专家参与评审，每一位专家评审的论文数量不超过q，专家与待评审论文重要主题的相关性强，专家应尽可能多地覆盖待评审论文的各主题，专家在待评审论文重要主题的覆盖度应高于次重要主题。

优化方案的目标是$\max\sum C_{mh} S_{nh}^{T} Y_{mn}^{T}$，即推荐专家在待评审论文主题上的权威度之和最大。该优化问题的输入是主题在专家中出现的概率矩阵C、考虑文档环境后论文与主题的相关性S、每篇待评审论文的评审人数及每位专家的评审负荷，输出为待评审论文推荐专家的任务分配矩阵Y_{mn}。y_{ij}=1表示专家i被分配担任待评审论文j的评审专家。根据以上假设，评审专家推荐的优化方案可表示为公式（3）。

$$\max\sum C_{mh} S_{nh}^{T} Y_{mn}^{T} \quad (3)$$

$$\text{s.t.:}\quad C1{:}\forall_j \in [1, n], \sum_{i=1}^{m} y_{ij} = p$$

$$C2{:}\forall_i \in [1, m], \sum_{j=1}^{n} y_{ij} \leqslant q$$

$$C3{:}\forall_i \in [1, m],\ \forall_j \in [1, n], y_{ij} \in \{0, 1\}$$

4 实验设计与分析

实验需要两部分数据：一是包括专家个人信息、发文全文等的专家库数据，二是待审论文集。由于受现实条件限制，本研究采用中文学术论文数据库中的采样数据完成模型实验。

4.1 数据集

本研究的实验数据来自万方数据库。数据对象为信息管理与信息系统领域下的5 226名学者及其75 880篇论文。学者信息包括姓名、单位、发文、被引、H指数、发表期刊、学科、关注热点等，论文信息包括题名、发表时间、被引频次、发文期刊、摘要和关键词等。信息管理与信息系统领域下的分支较多，涉及内容广泛，包括经济学、哲学、教育学、医学、水利工程学等众多学科。在现实中，期刊和会议通常有较明确的学科范围。因此，依据万方数据库的学科标签，本研究选择经济学类目下的相关论文，构建数据集A（100篇论文，200名专家）和数据集B（300篇论文，400名专家）。

4.2 实验结果

本研究借助JAVA库中的JGibbLDA工具实现专家主题的遴选。JGibbLDA工具适用于在JAVA环境中运用吉布斯采样技术完成LDA主题模型中参数估计和预测问题。下文所展示的试验示例中模型参数设置为：主题（ntopics）个数为10，吉布斯采样迭代次（niters）为1 000，主题中最可能出现的词语（twords）的返回个数为10，LDA模型中的参数和采用默认值，即取值为50/K（K为预测所得的主题数），取值为0.1。经过该模型分析，获得专家论文集中的10个主题，每个主题的概念可由返回的最有可能出现的词语及其出现概率表示。表2展示了其中两个主题及其特征词语。

本研究通过数据集*A*和数据集*B*分析考虑主题重要性的优化模型效果。其中，表3为设定每篇待评审论文需要有4名专家作出评审，每位专家的最大工作负荷为5，且论文主题个数为10时为某一篇待评审论文推荐专家的结果。表3中展示了待评审论文和评审专家与10个主题的相关性。其中，4名评审专家的突出主题基本覆盖了待评审论文的所有突出主题。表4所示为推荐结果中专家评审负荷的人数分布。其中，每篇论文分配的专家数为*p*，且每位专家分配的论文数量不超过*q*。

表 2 专家论文集预测所得主题示例

主题例1		主题例2	
主题特征词	词语出现概率	主题特征词	词语出现概率
市场需求	0.038 627	数据挖掘	0.029 647
电子商务	0.026 787	数据仓库	0.011 381
知识管理	0.014 947	时间序列	0.008 571
信息资源	0.009 028	客户关系管理	0.008 571
指标体系	0.009 028	聚类分析	0.007 166
层次分析法	0.009 028	信息熵	0.005 761
多属性决策	0.006 068	分类	0.005 761
信息管理	0.006 068	专家系统	0.005 761
企业信息化	0.006 068	关联规则	0.004 356
知识流程外包	0.004 588	个性化推荐	0.004 356

表 3 主题相关性

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10
待评论文	0.02	**0.05**	0.02	**0.20**	0.02	**0.05**	0.02	**0.62**	0.02	0.02
评审专家1	0.02	**0.25**	0.02	**0.13**	0.02	**0.17**	0.02	**0.33**	0.02	0.02
评审专家2	0.06	0.02	0.06	0.02	0.02	**0.10**	0.02	**0.66**	0.02	0.02
评审专家3	**0.10**	0.02	0.02	0.02	0.02	0.02	0.10	**0.66**	0.02	0.02
评审专家4	0.03	0.03	0.03	0.03	0.08	0.03	0.03	**0.58**	0.03	**0.18**

注：粗体代表突出主题。

表 4 评审专家负荷量对比

		评审5篇文章人数	评审4篇文章人数	评审3篇文章人数	评审2篇文章人数	评审1篇文章人数	评审专家总人数
数据集*A*	*p*=3；*q*=5	20	8	16	39	39	125
	p=5；*q*=5	50	11	38	29	29	162
数据集*B*	*p*=3；*q*=5	85	29	41	72	72	319
	p=5；*q*=5	222	32	46	42	42	382

4.3 实验结果检验及评价

对于大部分机器学习问题，人工评判可能是模型评估较有效的方法。对于评审推荐问题，现有研究中人工评估模型可靠性的方法主要包括两类：一是人工阅读论文后匹配合适专家，并以此作为标准结果，计算推荐结果的召回率和准确率[22]；二是请相关专家或用户根据已有评价标准对算法推荐结果作出评估[8]。然而，人工选择和认可的“最优”结果受限于标记专家的能力和知识范围。因此，人工标记的数据能否作为标准结果存在一定争议。此外，召回率和准确率不能体现实验结果中不同重要程度的主题差异。为此，本研究提出通过以下3种评价指标来定性评估和比较不同模型。

4.3.1 评价指标定义

平均覆盖率（avc）：专家擅长的主题占评审论文所有主题的比例。该指标可用来衡量模型推荐的评审组对该论文的总体评审能力，见公式（4）。

$$avc=\frac{h_A}{h} \tag{4}$$

h_A为该待评审论文被覆盖的主题数量，h为论文主题总数。如果一篇待评审论文的“突出主题”包含在所有评审专家的“突出主题”中，则认为该主题已被覆盖。

人均审核率（avn）：一篇文章中，平均每个主题被覆盖的频次占评审组总人数的比例。该指标从主题角度评估论文被审核的效果。理想结果为待评审论文重要主题的人均审核率高，甚至接近1。这意味待评审论文的所有评审专家在重要主题上都具有评审能力，见公式（5）。

$$avn=\frac{\sum n_{A'}}{n\times h\times p} \tag{5}$$

$n_{A'}$为某一待评审论文中所有主题被覆盖的总频次，n为待评审论文的总数，h为论文主题数量，p为每篇待评审论文的评审组人数。

重要性匹配度（cor_s）：评审专家在各主题上的专业水平与待评审论文主题重要性的斯皮尔曼相关系数。该指标用于衡量待评审论文中较重要的主题是否得到优先分配，见公式（6）和（7）。

$$cor_s=\frac{\sum r_s}{m\times n} \tag{6}$$

$$r_s=1-\frac{6\sum_{i=1}^{h} d_i^2}{h\left(h^2-1\right)} \tag{7}$$

r_s为待评审论文与评审专家关于主题分布的斯皮尔曼相关系数，d为待评审论文和评审专家的主题向量秩次差。cor_s值高则说明待评审论文的重要主题相对于次重要主题可分配到更多资源。

4.3.2 对比算法

本研究选取未考虑主题重要性差别的均匀主题优化模型和考虑主题重要性差别的非均匀主题贪心算法作为对比模型，检验不同模型针对同一数据集的推荐效果。均匀主题模型采用优化算法实现专家和待评审论文的匹配，但不区分同一待评审论文中不同主题的重要性。而贪心算法推荐模型利用贪心算法的原理完成专家推荐。贪心算法具体步骤如下。

（1）利用主题模型提取专家及待评审论文的主题分布B_{nt}^{greedy}和R_{mt}^{greedy}；

（2）依据非均匀主题模型的计算主题重要性并形成新的专家及待评审论文主题分布矩阵$B_{nt}^{greedy'}$和$R_{mt}^{greedy'}$；

（3）建立循环，计算第i（$i\in[1,n]$）篇待评审论文与全部m位候选专家的相关度；

（4）将m位专家按与第i篇待评审论文的相关度排序，为每篇待评审论文选择出相关度最高的前p个专家，同时将被选择的专家的负担值加1；

（5）每完成一篇待评审论文的专家分配后，检验专家集中所有专家的负担值是否小于q，若小于q，则将此专家从专家集中去除；

（6）n次循环后完成m位专家对n篇待评审论文的分配。

4.3.3 实验结果及分析

（1）非均匀主题模型与均匀主题模型的比较分析。图2和图3为当p=3，q=5时，数据集A的3种模型方式分配评审专家的实验结果。由图2和图3可见，非均匀主题的两个算法较均匀主题的算法在平均覆盖率和人均审核率两个指标上具有显著优势。由此得出，经过主题非均匀化处理的模型可以更好地完成评审专家对待评审论文的分配任务。

（2）模型中的变量对实验结果的影响分析。主题提取过程中需要提前设定主题数量h。从图2和图3可见，两种非均匀算法均在h=30时取得最佳结果，但主题数量的改变对实验结果的影响不显著。也有研究主

题数量对实验结果的影响，但结果表明表现最好的主题数量不尽相同。因此，本研究认为主题数量对模型结果的影响可能与数据集的选择有关。

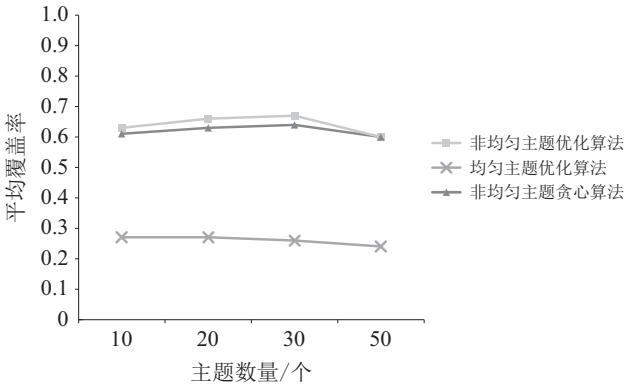

图 2 主题数量对平均覆盖率的影响

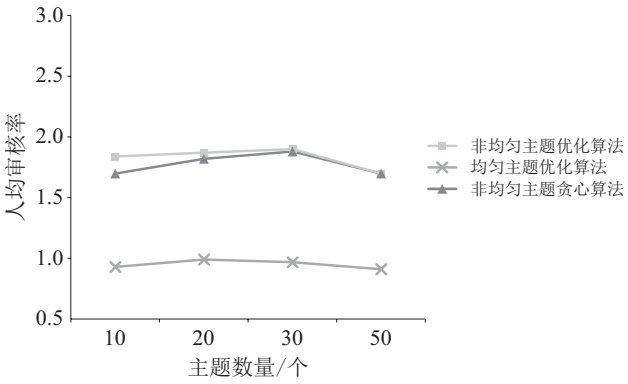

图 3 主题数量对人均审核率的影响

如图4和图5分别展示了3种分配方式在数据集B中的实验结果。当p=5，q=5时平均覆盖率和人均审核率最高。这是因为评审专家数量增加可能影响评审质量，此时每篇待评审论文中有更多主题被审核，并且每个主题接受的审核人次也需要相对增多。

（3）优化算法与贪心算法的比较分析。图2和图3中非均匀主题的优化算法与贪心算法在覆盖率、审核率两项指标上差距不大。其主要原因是数据集中专家人数相对于评审需求较充足。因此，即使贪心算法的每轮循环都“贪心”地选出当前最合适的专家，也不会对整体结果产生显著影响。然而，从图4和图5可见，随着待评审论文对评审专家人数需求的增加，当p=5，q=5时，优化算法的优势得以凸显。为突出各个算法的差异，在数据集B中增加一种分配方式p=2，q=2进行实验。从图4、图5可见，在该条件中，两种算法的差距进一步加大。这说明，当专家库规模有限时，优化算法相较贪心算法更有利于实现资源的最优配置。

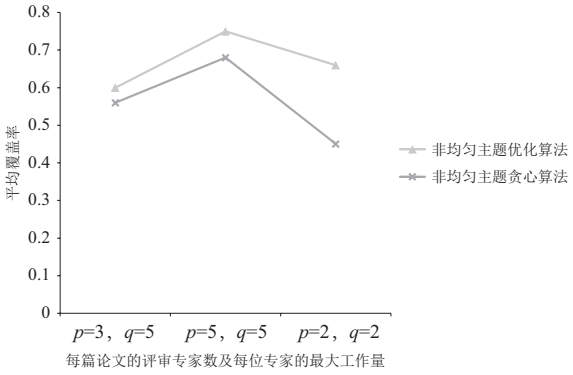

图 4 分配方式对平均覆盖率的影响

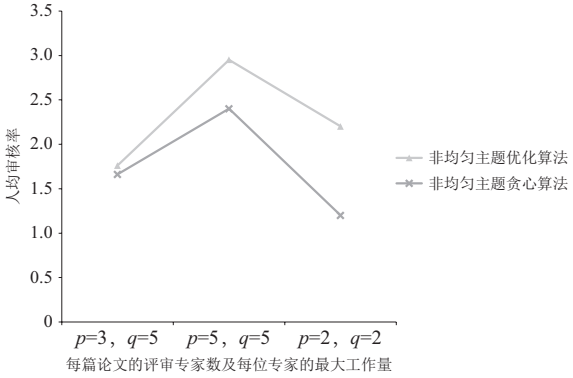

图 5 分配方式对人均审核率的影响

从图6可见，两种算法在重要性匹配度cor_s这一指标上的差距较大，且系数变化主要受数据集影响，而与分配方案关系不大。在数据集A中，优化算法推荐出的专家主题向量与待评审论文主题向量的相关系数为0.40，而贪心算法的相关系数为0.12。此外，由于贪心算法在每次循环中直接选择适合评审当前论文的最优专家，因此论文的循环顺序也会对推荐结果产生一定影响。这导致贪心算法存在较大不稳定性，一定程度上限制了其广泛应用。

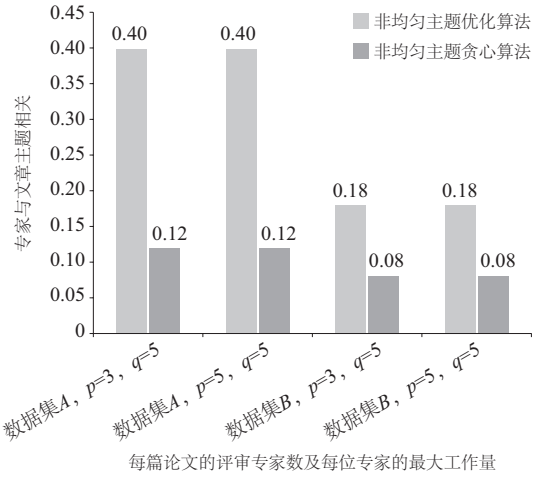

图 6 不同算法与数据集在重要性匹配度上的比较

5 结论与展望

本研究分析如何为多篇待审论文推荐评审专家的问题，并提出非均匀主题重要度的专家推荐优化模型。该模型考虑不同学科环境下待评审论文的主题重要度差异，并在优化模型中综合考虑专家与待评审论文主题的相关程度、待评审论文的被评审人数、专家评审负荷等因素。最终，针对每篇待评审论文形成评审专家组。

通过与其他两种模型推荐结果的比较，验证本检验模型的有效性。结果发现，经过非均匀化处理的模型在覆盖率和人均审核率上均有明显优势。当专家资源较充足时，优化算法与贪心算法差别不大；但随着待审核论文专家数量增加，优化算法的优势逐渐凸显。在相关系数方面，优化算法始终表现出较高相关性，而贪心算法表现较差。

鉴于模型存在的问题，后续研究可对数据集进行更精细设置。如结合学科现状对网上抓取的学者信息进行人工筛选和过滤，形成高质量专家集，更真实地模拟评审环境下的专家推荐，使模型推荐结果更易于评价。

参考文献

[1] 韩晓,郭嘉丰,杜攀,等.一种面向权威度和多样性的自动学术调研框架[J].计算机学报,2015(2):365-373.

[2] ZHENG H T, LI Q,JIANG Y,et al.Exploiting multiple features for learning to rank in expert finding[C]//9th International Conference on Advanced Data Mining and Applications,December 14-16,Hangzhou,[s.n],2013:219-230.

[3] SUN Y H,MA J,FAN Z P,et al.A hybrid knowledge and model approach for reviewer assignment[J].Expert Systems with Applications,2008,34(2):817-824.

[4] FAN Z P,CHEN Y,MA J,et al.Decision support for proposal grouping:a hybrid approach using knowledge rule and genetic algorithm[J].Expert Systems with Applications,2009,36(2):1004-1013.

[5] 李春英,汤庸,陈国华,等.面向学术社区的专家推荐模型[J].智能系统学报,2012,7(4):365-369.

[6] SUN J,XU W,MAJ,etal.Leverage RAF to find domain experts on research social network services:a big data analytics methodology with MapReduce framework[J].International Journal of Production Economics,2015,165:185-193.

[7] DAUD A,LI J,ZHOU L,et al.Temporal expert finding through generalized time topic modeling[J].Knowledge-Based Systems,2010,23(6):615-625.

[8] SUN Y H,MA J,FAN Z P,et al.A group decision support approach to evaluate experts for R&D project selection[J].IEEE Transactions on Engineering Management,2008,55(1):158-170.

[9] TANG J,JIN R,ZHANG J.A topic modeling approach and its integration into the random walk framework for academic search[C]//2008 Eighth International Conference on Data Mining,December 15-19,2008,Washington.New York:IEEE,2008:1055-1060.

[10] Yang C,Ma J,SILVA T,et al.A multilevel information mining approach for expert recommendation in online scientific communities[J].The Computer Journal,2015,58(9):1921-1936.

[11] HWANG W S,LEE H J,KIM S W,et al.Efficient recommendation methods using category experts for a large dataset[J].Information Fusion,2016,28(C):75-82.

[12] ZHOU G,ZHAO J,HE T,et al.An empirical study of topic-sensitive probabilistic model for expert finding in question answer communities[J].Knowledge-Based Systems,2014,66(9):136-145.

[13] WANG G A,JIAO J,ABRAHAMS A S,et al.ExpertRank:a topic-aware expert finding algorithm for online knowledge communities[J].Decision Support Systems,2013,54(3):1442-1451.

[14] 梁凯春,蔡淑琴,林森.基于分众分类的专家推荐算法研究[J].武汉理工大学学报(信息与管理工程版),2007,29(4):87-90.

[15] LI X,WATANABE T.Automatic paper-to-reviewer assignment,based on the matching degree of the reviewers[J].Procedia Computer Science,2013(22):633-642.

[16] HAN S G,JIANG J P,YUE Z,et al.Recommending program committee candidates for academic conferences[C]//Proceedings of the 2013 Workshop on Computational Scientometrics:Theory & Applications,October 28,2013,San Francisco.New York:ACM,2013:1-6.

[17] YUKAWA T,KASAHARA K,KATO T,et al.An expert recommendation system using concept-based relevance discernment[C]//Proceedings of the 13th International Conference on Tools with Artificial Intelligence,New York:IEEE,2001:257-264.

[18] LIU X,SUEL T,MEMON N.A robust model for paper reviewer assignment[C]//Proceedings of the 8th ACM Conference on Recommender Systems.New York:ACM,2014:25-32.

[19] 巩军,刘鲁.基于个人知识地图的专家推荐[J].管理学报,2011,8(9):1365-1371.

[20] XU Y H,MA J,SUN Y H,et al.A decision support approach for assigning reviewers to proposals[J].Expert Systems with Applications,2010,37(10):6948-6956.

[21] WANG F,ZHOUS R,SHI N.Group-to-group reviewer assignment problem[J].Computers &Operations Research,2013,40(5):1351-1362.

[22] KARIMZADEHGAN M,ZHAI C X.Integer linear programming for

constrained multi-aspect committee review assignment[J].Information Processing & Management,2012,48(4):725-740.

[23] KARIMZADEHGAN M,ZHAI C X,BELFORD G.Multi-aspect expertise matching for review assignment[C]//Proceedings of the 17th ACM Conference on Information and Knowledge Management.New York: ACM,2008:1113-1122.

[24] TANG W,TANG J,LEI T,et al.On optimization of expertise matching with various constraints[J].Neurocomputing,2012,76(1):71-83.

[25] BLEI DM,NG A Y,JORDAN M I.Latent Dirichlet Allocation[J].Journal of Machine Learning Research,2003(3):993-1022.

[26] GRIFFITHS T L,STEYVERS M.Finding scientific topics[J]. Proceedings of the National Academy of Sciences,2004,10(11):5228-5235.

[27] MINKA T P.Expectation propagation for approximate Bayesian inference[C]// Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence.[S.l.]:Morgan Kaufmann Publishers Inc.,2013(17):362-369.

作者简介

靳健，男，1982年生，博士，讲师，研究方向：文本挖掘、学术数据挖掘、客户关系管理，E-mail：jinjian.jay@bnu.edu.cn。
杨海慈，女，1995年生，研究方向：学术数据挖掘。
李凝，女，1981年生，博士，讲师，研究方向：文本分析、语用学。
耿骞，男，1965年生，博士，教授，研究方向：信息检索、管理信息系统。

A Topic Relevance Aware Model for Reviewer Recommendation

JIN Jian[1],YANG HaiCi[1], LI Ning[2], GENG Qian[1]
(1.School of Government, Beijing Normal University, Beijing 100875, China;
2.School of Foreign Studies, Central University of Finance and Economics, Beijing 100081, China)

Abstract: Reviewer assignment is an important step in the phase of paper evaluation for conference organizers and journal editors. In this study, the importance of a topic is initially estimated by its occurrence frequency within a specific research area, which helps to express submissions and reviewers' expertise. Next, in consideration of topic importance in submissions, an integer optimization model is formulated to recommend a reviewer group. Also, different practical constraints are reckoned in the optimization model, which includes the affinity between reviewers and submissions, reviewers' expertise, the burden assignment of each reviewer, etc. To evaluate the effectiveness, the proposed approach is benchmarked with two baseline algorithms in terms of coverage, average number of reviewers, relevance between reviewers and topics, etc. Comparative experiment results show that the proposed approach is capable to recommend reviewers effectively.

Keywords: Reviewer Assignment; Topic Relevance; Integer Optimization

（收稿日期：2017-03-30）