《人工智能研究杂志》76（2023年）761-827 投稿于2022年10月；发表于2023年3月

审稿人分配问题：文献系统综述

Meltem Aksoy yontarm@itu.edu.tr
Seda Yanik sedayanik@itu.edu.tr
伊斯坦布尔技术大学，工业工程系，
Maçka，伊斯坦布尔，土耳其

Mehmet Fatih Amasyali amasyali@yildiz.edu.tr
耶尔德兹技术大学，计算机工程系，
Davutpaşa，伊斯坦布尔，土耳其

摘要

恰当的评审人分配会显著影响提案评审的质量，因为只有将提案分配给相关评审人，才能确保评审准确且公平。将评审人分配给已提交提案这一关键任务，是评审流程的起点，也称为审稿人分配问题（RAP）。由于人工分配存在明显局限，期刊编辑、会议组织者和资助项目管理人员都需要自动审稿人分配方法。自1992年以来，许多研究针对自动化流程的需求提出了分配方案。本综述论文旨在为学者和从业者全面介绍有关RAP的现有研究。为实现这一目标，本文对过去三十年发表的、可在 Web of Science、Scopus、ScienceDirect、Google Scholar 和 Semantic Scholar 等数据库中查阅的 103 篇审稿人分配领域出版物进行了深入的系统综述。本文将RAP方法按其底层技术分为两大类及多个子类，并对其进行了分类和讨论。此外，本文还提出了各类别潜在的未来研究方向。本综述表明，RAP研究正变得愈发重要，同时仍需投入更多努力来开发新方法和框架。

1. 引言

专家的识别与分配是一个新兴的研究主题，近年来受到许多研究人员的深入探讨。专家推荐旨在帮助人们更轻松地找到合适的人来回答特定问题。此项任务可能涉及识别或分配在某一主题领域具备特定专业知识的专家。本文聚焦于审稿人分配，这是专家分配中的一个独特领域。

同行评审是用于评估科学研究论文和项目的质量与可行性的常规专业流程。将专家评审人分配给项目提案或研究论文，这一任务称为审稿人分配问题（RAP），是学术会议、期刊和资助机构同行评审流程中的一项关键且具有挑战性的工作。

为提高评审质量并确保评估公平、公正，分配评审人时应考虑以下重要事项：(1) 应根据评审人的专业领域为其分配提案

© 2023 AI Access Foundation. All rights reserved.

；(2) 提案应由各自主题领域中最具资质的审稿人进行评估；(3) 评审人应积极主动、热衷于审查提案；(4) 任何评审人审查的提案数量都不应超过约定数量；(5) 每位评审人审查的提案数量应相同；(6) 每份提案都应由一定数量的评审人审查；(7) 应避免申请者与评审人之间可能损害评审流程公平性的关系。

传统上，审稿人分配由单个决策者（如资助项目经理）或小型委员会负责，所有这些任务均由人工完成。然而，人工分配流程非常耗时，而且委员会的判断带有主观性或偏见。此外，由于无法妥善考虑所有要求，因此很难优化分配方案。而且，由于许多申请都在公布的截止日期临近时才提交，审稿人分配必须在时间极其有限的情况下完成。在当前的同行评审流程中，评审人和提案的数量很多，甚至可达数千。这么多评审人和提案会拖慢人工分配，使同行评审成为会议主席、期刊编辑和资助项目管理人员面临的难题。因此，人工方法极其费力，也无法始终取得最优结果。由此，自动分配评审人的方法已成为学术界的重要主题。最早探讨审稿人分配问题（RAP）的研究由 Dumais 和 Nielsen（1992年）提出，介绍了基于潜在语义索引（LSI）的相似度分数计算方法。为满足自动化流程的需求，自 Dumais 和 Nielsen（1992年）的开创性研究以来，许多研究提出了分配方案。

关于 RAP 方法的多篇综述文章已经发表，但它们未能反映该主题的最新研究进展。这些研究大多只关注 RAP 的某个特定方面，没有全面探讨这一问题，因此未能提供整体研究视角。例如，Goldsmith 和 Sloan（2007年）进行了一项调查，讨论了若干优化准则，以及程序委员会主席和现有会议管理工具采用的一些技术。

在另一项调查中，Kalmukov 和 Rachev（2010年）考察了用于界定论文和审稿人胜任能力的现有方法。Kolasa 和 Krol（2011年）研究了论文审稿人指派问题的启发式算法。首篇聚焦 RAP 各个方面的文献综述由 Wang 等人（2008年）提出，并引用了 19 项研究。2010年，Wang 等人又推出了该研究的更新版，内容也更加全面。这项研究引用了 36 项作为 RAP 问题解决方案提出的研究。由于这些调查发表于十多年前，因此未能反映该领域最新的科学进展。Patil 和 Mahalle（2020年）的近期研究探讨了 RAP 通常采用的绩效指标。此外，迄今为止开展的所有调查都只涉及会议审稿人分配问题，而没有考虑期刊和资助机构中审稿人的分配研究。

此外，也有一些综述论文聚焦于整个评审流程，并将 RAP 视为该流程的一个组成部分。尽管这些研究设有专门章节，介绍如何为提案遴选审稿人，但这些调查不足以从研究视角全面探讨 RAP。例如，Price 和 Flash（2017年）概述了如何利用机器学习和人工智能工具将同行评审流程中的各个步骤自动化。近年来，Shah（2022年）发表了一篇全面的综述论文，探讨同行评审流程中的挑战及其提出的解决方案。这项研究通过分析整个评估过程中因不诚实行为、校准失误和主观性等问题而产生的评审

流程中的不平等，对同行评审相关文献进行了总体概述。在这项研究中，RAP 被视为由审稿人专业知识不匹配所导致的不公正问题，并且仅从一个维度进行了浅显考察。

为了解决由以往调查的不足所造成的研究空白，本研究旨在建立RAP相关研究的最新基线，并呈现一份涵盖RAP各个方面的全面系统性文献综述。本文深入评述了1992年至2022年期间发表、并可在Web of Science、Scopus、ScienceDirect、Google Scholar和Semantic Scholar数据库中获取的103项研究。

本综述的主要贡献是让读者全面了解 RAP 及当前的先进方法。此外，这篇文献综述提出了一个基础框架，用于对解决 RAP 的现有方法进行分类。研究人员可以利用本出版物思考文献中的空白，并确定未来潜在的研究主题。此外，资助机构、科学期刊和会议可以更好地了解如何更高效地管理审稿人分配流程。

本研究的组织结构如下：第 2 节介绍开展系统性文献综述所采用的方法。第 3 节介绍 RAP 的不同应用领域。第 4 节介绍审稿人分配流程的各个阶段。第 5、6 和 7 节按 RAP 的各个阶段考察并总结相关出版物。第 8 节总结并提出未来研究的推荐。

2. 综述方法

本文所呈现的系统性文献综述遵循 Snyder（2019）推荐的指南，该指南包含三个主要阶段：（i）设计综述，（ii）开展综述，以及（iii）报告综述。图 1 展示了开展这一全面文献综述的计划。

在设计阶段，研究问题已确定，并且已考虑总体综述方法，包括选择检索词和合适的数据库。需要回答的主要研究问题包括：什么是 RAP，研究领域有哪些？分别使用哪些信息来构建审稿人简介和提案简介？提出了哪些不同的工具和技术来解决 RAP？如何对为解决 RAP 而提出的各种工具和方法进行分类?

检索在五个电子数据库中进行：Web of Science、Scopus、ScienceDirect、Google Scholar 和 Semantic Scholar。选择这些数据库是因为它们收录了有关 RAP 的高影响力出版物和会议论文集。检索词为“审稿人分配”和“审稿人分配问题”。本调查考察了可在数字数据库中检索到的英语研究论文（来自期刊和会议），其出版时间为 1992年至2022年。为确保所选出版物与本调查相关，我们确定了纳入与排除标准，用于分析每项潜在的主要研究。本次综述的标准如图 1 所示。研究必须满足所有纳入标准且不符合任何排除标准，才具备纳入资格。

在审查阶段，我们采用了检索策略，全面综述 RAP 文献。检索策略步骤如下：1. 查看出版物的标题、关键词和摘要，并选出符合纳入标准的出版物。2. 阅读上一步中未被排除的出版物，以

确定是否应根据排除标准将其从调查中剔除。3. 查阅符合条件的论文的参考文献目录，以发现新的研究。如果发现这些研究，请注意它们必须符合纳入与排除标准。

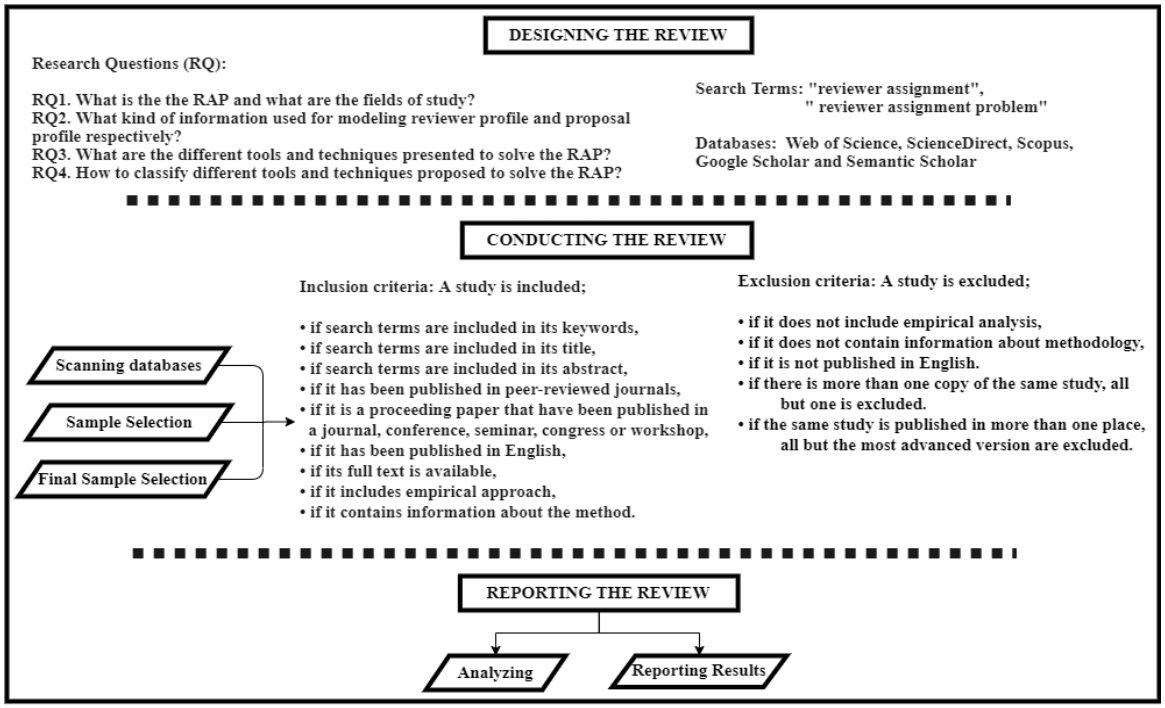


图 1：系统性文献综述的计划。

2022年8月，对数据库进行了检索，并定期更新检索结果。通过使用预先确定的检索词，共检索到404篇论文。使用 Mendeley 删除重复研究后，剩余334篇论文。确定涉及 RAP 的潜在主要研究后，对出版物按年份和研究领域的分布进行了考察。如图2所示，总体分布表明，近年来涉及 RAP 的研究数量有所增加。

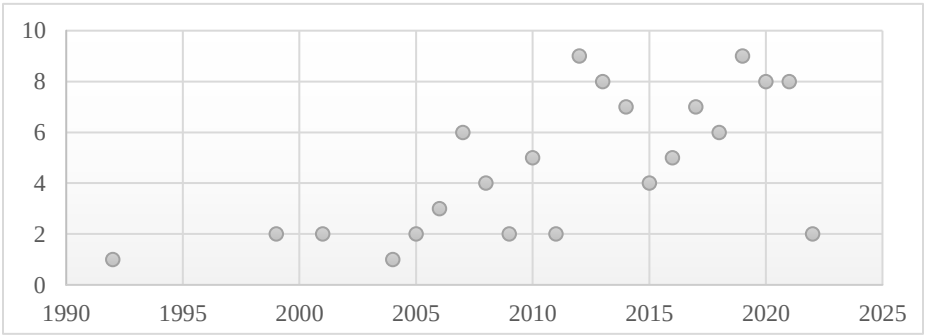


图 2：各年份研究的分布。

利用 Web of Science 类别对出版物进行分类，以确定其所属研究领域。表1展示了各研究领域的研究分布。鉴于研究在广泛的研究领域中分布近乎均衡，可以得出结论：RAP 是多个研究领域中的重要主题，且用于研究 RAP 的方法多种多样，具有跨学科特点。如表1所示，计算机科学与信息系统、教育研究和人工智能是排名前三的研究领域。

类别	记录数	占334的百分比
计算机科学与信息系统	44	13,17
教育 教育研究	40	11,98
计算机科学 人工智能	36	10,78
工程 电气与电子	24	7,19
计算机科学理论与方法	22	6,59
计算机科学跨学科应用	20	5,99
教育学科学学科	19	5,69
信息科学与图书馆学	13	3,89
外科学	13	3,89
计算机科学软件工程	12	3,59
运筹学与管理科学	11	3,29
儿科	11	3,29
卫生保健科学与服务	10	2,99
普通内科	10	2,99
临床神经病学	9	2,69
电信	9	2,69
多学科工程	7	2,10
公共、环境与职业健康	6	1,80
放射学、核医学与医学影像	6	1,80
生物学	5	1,50
计算机科学与控制论	4	1,20
骨科	3	0,90

表1：各研究领域的研究分布。

随后，将纳入与排除标准应用于每篇文章的标题、关键词和摘要。在这一步中，排除了五篇以英语以外语言撰写的论文，另有三篇文章因无法获取全文而被排除。经过这一筛选流程，共有 74 项研究被认定值得阅读全文。经过深入分析，65 项研究符合标准并被认为符合纳入条件。

通过对主要研究进行参考文献与引文检索，共识别出 1768 篇文章，并依据排除标准和纳入标准对其进行了审查。随后，发现其中 60 项研究可能与本调查主题相关。对这 60 项研究的全文进行了深入审查，最终确认其中 38 项符合资格。共有 103 篇论文满足标准，并被纳入系统综述。图 3 展示了这一流程的综合结果。

3. RAP 的应用领域

研究人员采用了多种方法来响应自动化审稿人分配任务的需求。为了评估所提出方法的性能，他们使用了来自不同应用领域的各种数据集。本调查所分析研究的应用领域列于附录 B 和 C 的第二列中。此外，图 4 展示了这些研究按提案类型划分的分布情况。可以看出，大多数研究关注为会议论文分配适当的审稿人（54%）。研究最多的第二个应用领域是为项目提案分配审稿人（25%），第三个是为期刊论文分配审稿人（21%）。

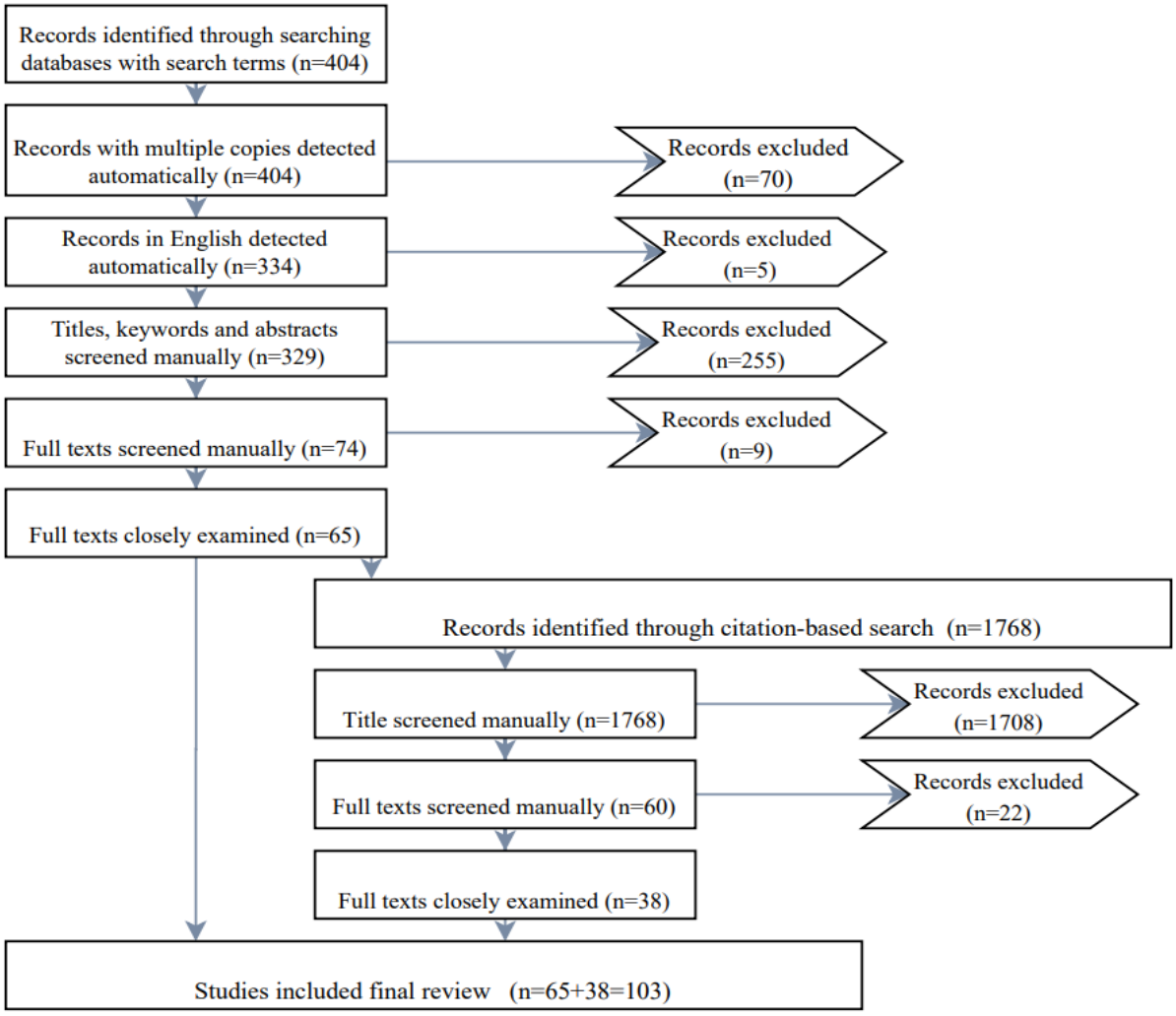


图 3：研究筛选流程。

尽管不同应用领域已有许多关于 RAP 的实证研究，但很难找到公开可用的真实审稿人分配信息。主要原因是，学术期刊、会议和资助机构出于隐私

方面的顾虑，不会共享评审人和提案信息。因此，一些学者创建了各种数据集，用于评估所提出模型的性能，并作为各类研究的基准数据集。可获取的数据集列于附录 A。

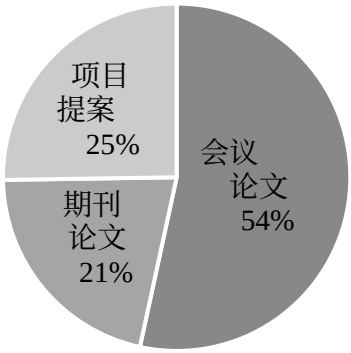


图 4：各提案类型研究的分布。

3.1 会议论文的 RAP

许多会议广泛采用同行评审流程，将符合条件的论文与不符合条件的论文区分开来，并选出高质量的会议论文。评审流程从论文在预定截止日期前投稿开始。投稿阶段会收集论文的各种数据，以创建论文档案。投稿截止后，会议程序委员会主席必须从审稿人库中选出专家评审人来评估已投稿的论文。通常会要求评审人对自己愿意阅读的论文进行“竞价”。根据竞标率以及评审人在注册流程中提供的信息（如专业领域、所属机构和资历），将合适的评审人分配给相应论文。评审人依据会议的评审表评估论文，并将评审意见提交给程序委员会主席。通常，程序委员会主席会根据评审人的意见，最终决定是否接收每篇论文。作者通常会收到评审人评审意见的匿名摘录，以便据此改进论文。最后，已接收论文的作者可以提交经修订的作品版本，即所谓的“终稿格式”。即使是分配过程中的一个小错误，也可能导致严重的误判，显著损害会议的声誉以及作者对会议的信心。因此，遴选评审人是最关键的阶段，因为这会直接影响整个评审流程的质量（Shah，2022年）。

在会议同行评审流程中，必须同时为论文分配评审人，并且必须在规定时间内完成评审人的分配任务。论文和评审人数量较少的小型会议可以人工将论文分配给评审人。然而，大多数会议的投稿论文和评审人数量都相当多，而且论文通常在会议截止日期临近时才提交。在这种情况下，要为每篇论文人工分配最合适的评审人颇具挑战。此外，由于需要考虑许多约束条件（准确率、负载均衡、利益冲突等），大型会议的人工分配准确率也越来越低。

为克服会议评审流程中人工操作的种种困难，自动化已成为迫切需求。与人工流程相比，自动化审稿人分配不仅有望简化任务，还能实现偏差更少、更系统化的论文与审稿人匹配。关于

如何为学术会议的论文自动分配审稿人，已引起众多学者的关注。他们提出了许多自动审稿人分配方法，其中一些已在实际会议中得到应用。面向会议实际应用、最为知名的尝试是多伦多论文匹配系统（TPMS）（Charlin & Zemel，2013年）。自2010年以来，该系统已用于机器学习和计算机视觉领域的会议，例如 NIPS、ICCV 和 ICML。此外，该系统还被整合到微软研究院会议管理工具包中，这是一个用于管理稿件投稿和评审流程的网络系统。其他知名系统，如 GRAPE（Di Mauro & Ferilli，2005年）和 Erie（Li & Hou，2016年），也用于 NIPS、ICML 和 IEEE INFOCOM 等大型会议的审稿分配流程。

3.2 期刊论文的RAP

对论文进行恰当且准确的评审，对于确保科学期刊的质量起着至关重要的作用。与会议论文的情况一样，科学论文应依据具体标准进行评估，并分配给审稿人，以便在客观评估后发表。在论文投稿和评估方面，会议管理与期刊管理有许多共同点。然而，两者之间的关键区别在于向出版物分配审稿人。期刊管理系统应在每个时间点为给定论文确定最合适的审稿人，而不应像会议管理系统那样将分配视为一个全局优化问题。此外，会议必须进行一次性分配，而期刊则必须进行顺序分配。会议设有论文投稿截止日期。相比之下，期刊通常没有截止日期，各篇论文彼此独立地处理。尽管期刊投稿没有截止日期，但审稿人仍应在规定时间内完成评估。

对于期刊论文，编辑和副编辑通常会定期手动分配审稿人。人工分配在很大程度上取决于编辑对该主题最新学术进展的了解。由于投稿数量不断增加，期刊编辑需要在约束条件下寻找合适的论文—审稿人配对，这已成为一项具有挑战性且耗时的任务。系统需要实现自动化，以便进行更加公平、客观和高效的评审分配。尽管有此需求，针对期刊论文RAP提交的实证研究数量（Biswas & Hasan，2007年；Karimzadehgan & Zhai，2009年；Daud 等人，2010年；Andrade-Navarro 等人，2012年；Wang等，2013；Protasiewicz，2014年；Yin 等人，2016年；Peng 等，2017年；Jin 等人，2018a、2018b年；Zhao 等人，2018年；Duan 等，2019年；Chughtai 等人，2019年；Tan 等人，2021年；Hoang 等人，2021年）少于针对会议论文的实证研究数量。

3.3 项目提案的 RAP

由于竞争激烈且资金来源有限，为项目提供资助并选择合适的项目，对资助机构至关重要（Tian 等人，2002）。项目选择是一项复杂且耗时的工作，包含许多子任务。在这一复杂过程中，一项重要且具有挑战性的任务是为项目提案分配评审人。评审人的意见会影响项目的潜在价值，因此他们在项目选择中发挥着至关重要的作用。将最合格的专家分配给项目提案进行评审的方法，可能会对项目选择的质量以及资助机构的投资回报产生重大影响。

用于项目提案评审的 RAP 已引起研究人员的关注。不过，基于真实数据开展实证分析的研究数量仍然不足。美国国家科学基金会

（NSF）（Janak 等人，2005年；Hettich 与 Pazzani）和中国国家自然科学基金委员会（NSFC）（Xu 等，2010年；Mai 等人，2012年；Li & Watanabe，2013年；Silva 等人，2014年；Liu 等人，2016年；Yue 等人，2017年）是两个尤为突出的机构，其数据集被广泛用于实证分析。

在本研究的其余部分中，术语“提案”将同时指论文提案和项目提案。

4. 审稿人分配流程

审稿人分配流程及相关方法的

无论审稿人分配采用何种结构（人工或自动），Wang 等（2010）都将该流程分为三个步骤：（1）搜索候选审稿人；（2）计算每份提案与每位候选审稿人之间的相似度分数；（3）在满足约束条件的情况下优化分配，以最大化相似度分数。Yin 等（2016）将审稿人分配任务视为专家检索问题，并将候选审稿人与提案之间的主题相关性作为主要关注点。从这一视角出发，他们将 RAP 流程概括如下：首先，收集潜在候选审稿人的信息，以表征其专业知识和知识储备。其次，将提案建模为查询。最后，根据审稿人的专业知识与提案内容之间的关系选出审稿人。

本系统综述研究结合了 RAP 的两种流程定义，一种由 Wang 等（2010）提出，另一种由 Yin 等（2016）提出，并将 RAP 流程划分为三个阶段。使用不同信息源创建审稿人与提案画像，以表示候选审稿人与提案，被视为阶段（1）；本调查将其称为审稿人与提案的表示。阶段（2）的主要目标是根据候选审稿人与提案的画像，计算每个可能的提案—评审人配对的相似度分数。任意提案与评审人之间的相似度分数，是反映评审人与提案之间专业知识匹配程度的数值（Shah，2022）。相似度分数越高，表示预期的评审质量越高。根据计算出的相似度分数，在阶段（3）中可以确立最优分配，同时考虑工作量均衡、提案需求和评审人偏好等实际约束条件。

为了创建审稿人与提案的画像，所有研究人员在计算提案与评审人之间的相关性或对其进行优化之前，都分析了可能表示评审人或提案的各种信息。这意味着所综述的所有研究都应用了第一阶段。然而，阶段（2）和阶段（3）未必如此。已有许多用于解决 RAP 的方法。一些研究仅关注计算评审人与提案之间的相似度分数（例如 Rodriguez & Bollen，2008年；Daud 等人，2010年；Xu 与 Du，2013年；Protasiewicz，2014年；Mittal 等人，2020年等），而另一些研究则尝试在一定约束条件下优化目标，以解决分配问题（Cook 等，2005年；Li 等人，2008年；Tang 等人，2010年；Tang 等人，2012年；Daş 与 Gökçen，2014年；Guo 等人，2018年；Boehmer 等人，2021年等）。此外，还有一些研究涵盖 RAP 的所有阶段，并将其设计为审稿人推荐系统或决策支持系统（例如 Tian 等人，2005年；Janak 等人，2006年；Sun 等人，2007年；Conry 等人，2009年；Xu 等人，2010年；Silva 等人，2014年；Liu 等人，2016年；Hoang 等人，2021年）。

正如许多研究人员所认同的那样（Wang 等，2008年；Xu 等，2010年；Kou 等，2015a；Peng 等，2017年；Duan 等，2019年；Yang 等，2020年），针对第2阶段和第3阶段所提出的方法分为两大类：i）基于信息检索的审稿人分配问题（IRRAP）；ii）基于优化的审稿人分配问题（ORAP）。103篇出版物在这两大类别中的分布如图5所示。如图5所示，一些研究仅关注其中一种方法，而另一些研究则同时关注两种方法。在这项系统综述研究中，除对审稿人和提案进行表示外，

还根据为解决审稿人分配问题（RAP）而提出的方法，将研究分为两个部分进行考察。涉及两类方法的研究，则根据所采用的方法，分别在各类别下单独讨论。

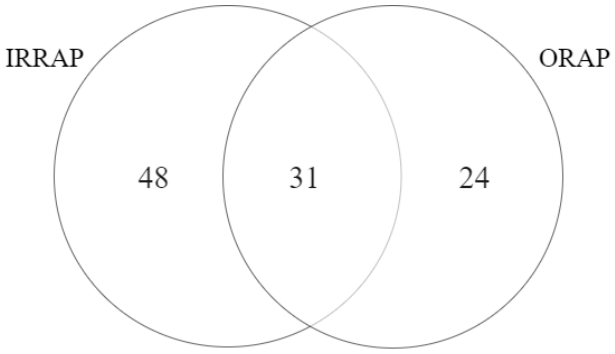


图5：两个主要类别中各研究的分布。

IRRAP 聚焦于 RAP 的第二阶段，主要目标是计算提案—审稿人相似度评分（Dumais & Nielsen，1992年；Mimno 和 Mccallum，2007年）。RAP 被视为一个信息检索问题，而基于信息检索的方法将每个提案视为查询，将审稿人视为待检索的文档，以获取适当的审稿人。在本调查中，根据不同类型信息的使用方式，IRRAP 分为三组：基于文本/基于内容的方法、基于网络的方法以及基于人工信息的方法。基于人工信息的方法利用审稿人明确提供的额外信息，将审稿人分配给提案；例如，收集审稿人的兴趣或对提案的评分，允许审稿人提供有关自身的主题等。这种方法通常需要收集人工信息。基于文本/基于内容的方法通常使用自然语言处理（NLP）技术，以统计方式学习文档中的词共现信息，例如词与主题之间的信息（主题模型）、词与文档之间的信息（语言模型），以及词与词之间的信息（词嵌入模型）。基于文本/内容的方法还可分为两类相关方法：基于语义的方法和基于词的方法。基于网络的方法通常利用网络关系，例如引文网络、共同作者网络、协作网络，以及审稿人与审稿人之间或审稿人与提案之间的其他相关网络。

ORAP 专注于 RAP 的第三阶段，将 IRRAP 转化为优化问题，并试图从运筹学的角度解决这一问题（Benferhat 和 Lang，2001年；Cook 等，2005年；Xu 等，2010年；Kolasa & Krol，2011年）。本研究考察了 ORAP 所采用的数学模型及用于求解这些模型的算法。

本调查考察的所有研究，依据 RAP 流程的各个阶段以及 RAP 的求解方式，归纳为三个主要标题。第 5 节讨论评审人和提案的表示方式。第 6 节讨论用于计算提案与评审人之间相似度分数的信息检索方法；第 7 节则讨论用于优化审稿人分配的数学模型和算法。图 6 展示了 RAP 方法的整体分类。如图 6 所示，每个类别都以一个标题编号标注。本文按各标题综述相关研究。

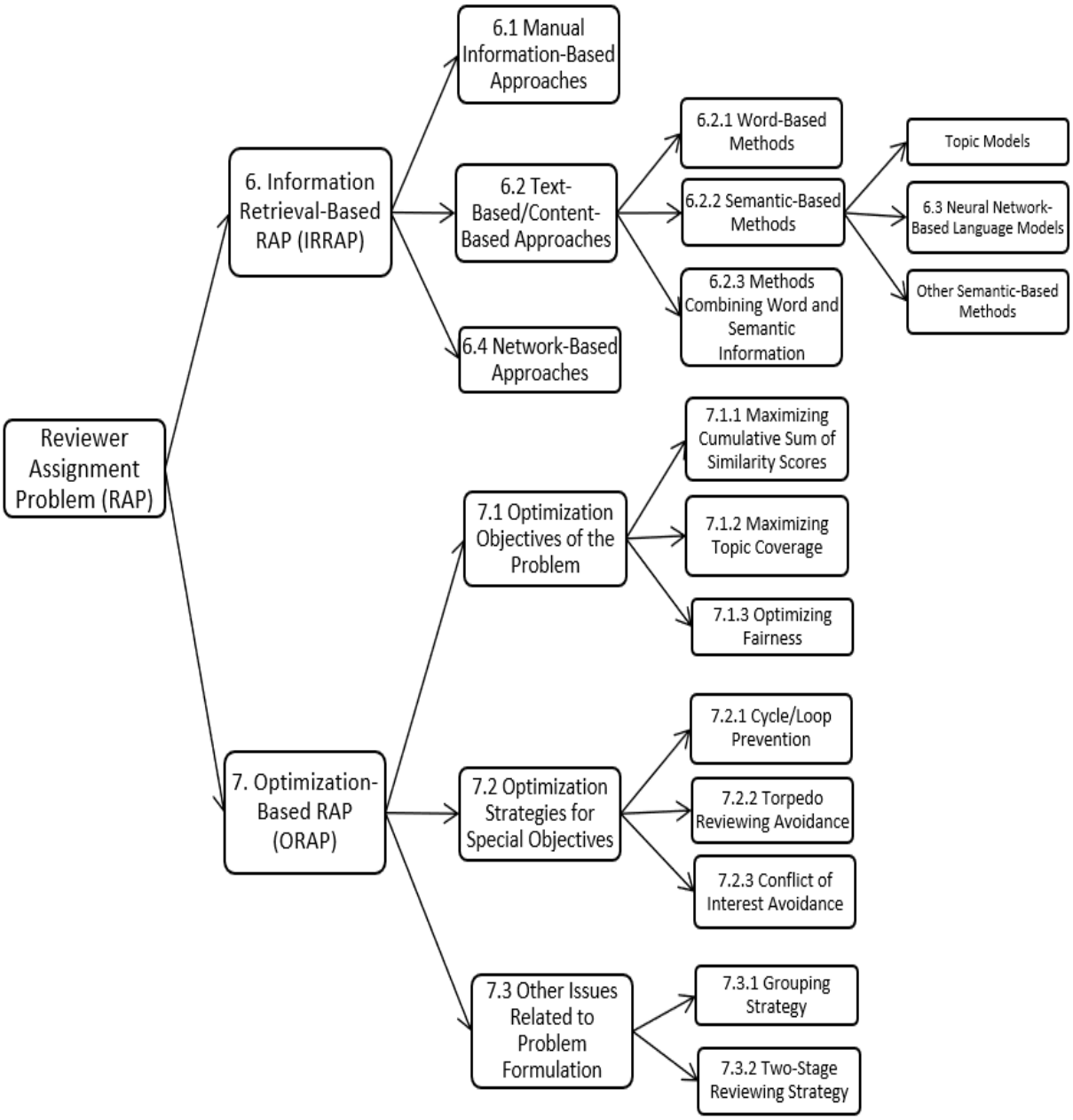


图 6：RAP 方法分类。

5. 第一阶段：审稿人与提案的表示

在计算提案与评审人之间的相似度分数之前，通常会先定义两项内容：评审人档案和提案档案。研究人员会考察可用于表示评审人和提案的多种信息。通常会利用评审人已发表论文中的信息来创建评审人档案，例如过往出版物的标题、摘要、关键词、全文、参考文献列表、引文列表和元数据。除这些信息外，档案还可以包括任何

评审人的个人信息，例如其姓名和姓氏、所属机构、职称、兴趣领域、专业领域、h指数、引文网络、合著者信息、其他关系/联系信息（例如学术关系、同事、竞争关系、学生—导师关系、项目合作者、合作关系等）、声誉与资历、过往评审经验、提案申请经验、项目管理经验（项目数量、项目质量）等。

此外，自我申报的关键词和评审人对提案的偏好（每位评审人对每个提案的自我评价、每位评审人对每个提案给出的意愿度，以及对提案的兴趣）也可用于表示评审人。

图7和图8展示了根据用于创建评审人简介的信息对出版物进行的分布情况。根据图7，最常用于描述审稿人的三种个人信息是专业知识/研究领域、合著信息和所属机构。选择这些信息的主要原因是它们易于获取。候选审稿人在申请成为审稿人时通常会提供此类信息。然而，仅凭审稿人的专业知识/研究领域、合著信息或所属机构，可能无法充分体现他们的情况。此外，候选审稿人常常由于主观性以及可能产生误解，而无法恰当地选择自己的专业领域或研究领域。因此，利用审稿人已发表论文所包含的丰富信息，可以获得准确得多的结果。

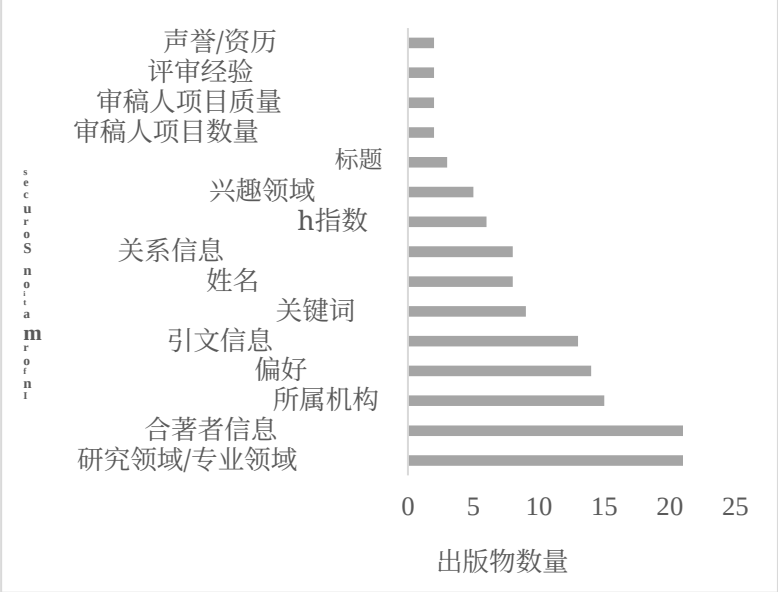


图7：与审稿人相关的各信息来源的出版物数量。

如图8所示，关于评审人已发表文章，最受青睐的三种信息源是全文、摘要和标题。虽然只需一种信息即可建立审稿人档案，但通常会使用多种信息来准确描述评审人。例如，审稿人以往发表的论文的摘要和标题经常会结合起来，用于构建其档案。

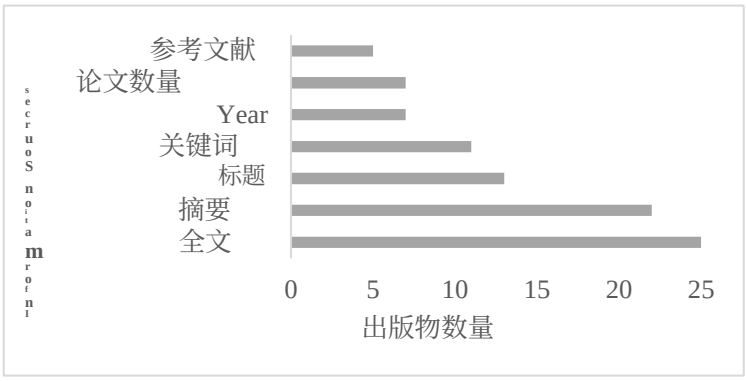


图8：与审稿人已发表的论文相关的各信息来源出版物数量。

创建提案档案时，可以使用与提案相关的信息，例如标题、关键词、摘要、全文、主题、研究/专长领域、参考文献、引文信息、合著者信息，以及申请人所属机构、学科/专业等信息。此外，还可以使用申请人在提交阶段指定的关键词。图9展示了根据用于创建提案档案的信息对研究进行的分类。最常用的三类信息是提案的摘要、关键词和全文。2014年之后的研究中，全文的使用更为普遍（73%）。这主要是因为自然语言处理中出现了新一代高性能文本/词表示技术。

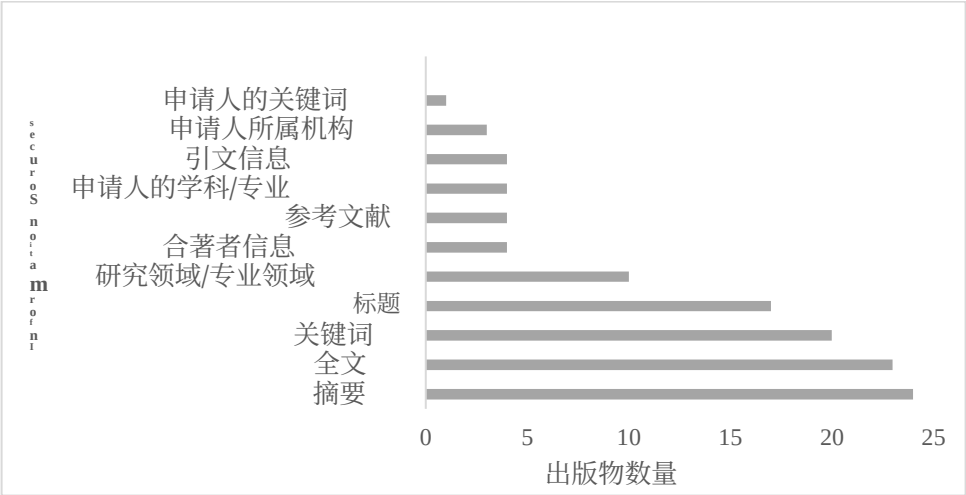


图9：与提案相关的各信息来源出版物数量。

在提案与审稿人画像过程中，检测提案与评审之间潜在的利益冲突（COI）至关重要。双方之间的任何合作或竞争都可能导致不公。因此，应在分配任务中检测并避免潜在的利益冲突。基于此，后续章节将讨论关于检测提案与审稿人之间潜在利益冲突的研究。

5.1 利益冲突检测

检测利益冲突对于确保提案评审过程不偏不倚、公平公正至关重要。在审稿人分配任务中，应检测并避免各种类型的利益冲突。Wu 等（2018）将利益冲突分为两类：确定性利益冲突和潜在利益冲突。确定性利益冲突与明确的关系有关，提案方和审稿人可以按照预先设定的申报规则声明这些关系。确定性利益冲突包括共同署名、同事关系、学生—导师关系、项目合作者关系以及竞争关系。另一方面，潜在利益冲突涉及根据申报规则无需声明的关系，必须根据明确的关系推断出来，例如亲密朋友、学术合作者、亲属等。在当前知名的会议管理系统中，例如 EasyChair 和 Microsoft Conference Management Toolkit，使用确定性利益冲突十分常见。这些系统通常要求审稿人和申请人查看一长串潜在评审人名单，并披露任何利益冲突。在某些情况下，这份名单可能包含数百名审稿人。此外，作者可能有意或无意地忽略一些潜在的利益冲突案例。因此，自我申报流程既耗时，也可能不完整。 尽管会议的程序委员会主席、期刊的编辑或资助项目管理人员（在资助 机构）可以自行检查潜在的利益冲突，但这一过程耗时、效果有限且容易出错。因此，需要有效且高效的方法来自动检测利益冲突。

在某些情况下，由于缺乏必要信息，很难检测利益冲突。然而，在许多其他情况下，社交网络中存在隐式或显式信息，例如科研社交网络（Scholarmate）、共同作者网络（如 ArnetMiner、DBLP）和职业网络（如 LinkedIn）。为减少提案与审稿人之间以及审稿人彼此之间的潜在利益冲突，Yin 等（2016）将概率主题模型和激活扩散模型整合到研究人员的共同署名网络中。此外，Pradhan 等（2020）使用共同署名图检测利益冲突。他们从 ArnetMiner 中提取了每位审稿人的合著者名单，并将其与审稿人和作者自行声明的合著者关系进行交叉核对。Silva 等（2014）通过挖掘 ScholarMate 数据构建协作网络，以识别审稿人之间的潜在利益冲突。Yan 等（2017）利用学术活动数据构建了研究人员及其所属机构的学术网络。研究人员之间的关系包括同事关系、合著者关系和学生—导师关系，而机构之间的关系则包括合作关系和同事关系。他们利用这些学术网络中提案与审稿人之间的路径距离检测潜在利益冲突。

早期研究主张利用加权关系自动检测潜在利益冲突。由于缺乏对利益冲突的通用定义，在涉及利益冲突的情境中，真实标签并不容易获取。Wu 等（2018）提出了一套名为 PISTIS（利益冲突感知审稿人推荐平台）的半自动系统，以解决利益冲突申报与检测问题。PISTIS 从各种公开来源检索异构关联，包括共同署名（来自 DBLP）和共同工作时长（从所属机构信息中提取）。该系统还使用因子图模型提取导师—被指导者关系（Wang 等人，2010）。这些异构信息用于构建异构出版物网络，作为系统的主要知识库。为简化申报流程，该系统通过监督排序模型推荐潜在利益冲突，并可利用历史申报数据迭代更新该模型。

6. 第 二阶段：提案与评审人之间的相似度分数计算 s

RAP 的第二阶段，即计算提案与审稿人之间的相似度分数，在文献中通常称为基于信息检索的审稿人分配问题（IRRAP）。基于信息检索的方法将每项提案视为查询，以检索合适的审稿人。本文根据所用信息的不同类型，将现有基于检索的审稿人分配研究分为以下三类：(i) 基于人工信息的方法，(ii) 基于文本/内容的方法，以及 (iii) 基于网络的方法。图10展示了这些研究在三类中的分布。如图10所示，一些研究只关注其中一类，而另一些研究则同时涉及多类。本文随后将从三个方面介绍所考察的出版物。

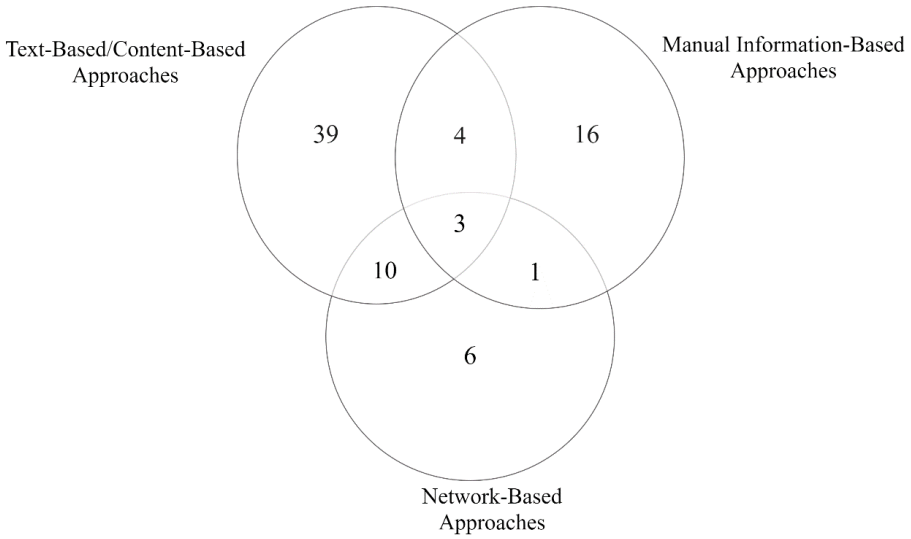


图10：各子类别中 IRRAP 方法的分布。

6.1 基于人工信息的方法

基于人工信息的方法利用评审人提供的额外信息，将评审人分配给提案。竞价，也称为评分，是基于人工的系统中最常见的策略；在这一过程中，评审人会参与表达自己的偏好（Benferhat & Lang, 2001年；Basu 等人, 2001年；Janak 等人, 2005年；Goldsmith & Sloan, 2007年；Papagelis & Plexousakis, 2008年；Xu 等人, 2010年；Garg 等人, 2010年；Kolasa & Krol, 2011年；Chen 等人, 2012年；Daş & Gökçen, 2014年；Tayal等人, 2014年；Liu 等人, 2015年；Liu 等人, 2016年；Yeşilçimen & Yıldırım, 2019年；Jecmen 等人, 2020年；Fiez 等人, 2020年）。在竞标方法中，评审人必须阅读提案摘要或全文，并为每份提案评分，以表示自己愿意评审的程度。通常，竞标等级量表包含3、4或5个等级（Huang 等人, 2010年；Wang 等人, 2010年）。例如，在 ICDM'07 中，竞价范围为1至4，分别表示以下偏好：4 = “非常愿意评审”，3 = “愿意评审”，2 = “不太愿意评审”，1 =“不愿意评审”。竞标阶段结束后，每位评审人都会得到一份涵盖所有提案的偏好列表，该列表体现了评审人与提案之间的相似度分数。竞价值越高，表示相似度分数越高。专家最有能力判断自己是否具备评审某份提案所需的专业能力。根据预期用途可以看出，由公平且负责任的评审人实施竞价，有助于提高评估的质量和可靠性，

并确保评审人的满意度（Cabanac & Preuss, 2013年）。然而，竞标方法存在明显缺点。首先，提案数量增加时，竞标会变得复杂且耗时过长。其次，许多会议中，大多数申请都缺少足够的竞价。例如，JCDL 2005年的264份提案中，有146份没有收到正向竞价（Rodriquez & Bollen, 2007年）；而在 IMC 2010，68%的提案没有收到任何竞价（Beverly & Allman, 2012年）。如果获得竞价的提案数量不足，最终形成的提案—评审人竞价矩阵就会稀疏，从而大幅降低审稿人分配方法的准确率。第三，评审人倾向于根据当前的研究兴趣竞价，而非依据累积专长。这种情况可能降低分配质量，也会影响对提案进行评估的准确率。最后，也是最重要的一项缺点是，这些方法容易受到恶意行为和利益冲突的影响，进而损害评估的公平性。采用这些方法时，评审人可以通过策略性竞价，大幅提高自己被分配到其意图针对的提案的机会（Jecmen 等人, 2020年；Wu 等人, 2021年）。

审稿人和申请人提供的专业关键词或主题是分配流程中使用的其他人工信息（Hartvigsen 等人, 1999年；Xu 等人，2012年；Charlin 等人，2012年；Liu 等人，2014年；Silva 等人，2014年；Liu 等人，2015年；Liu 等人，2016年；Pradhan 等人，2020年）。通常要求申请人在投稿期间提供一份简短的关键词列表，以代表其提案。同样，在注册流程中，要求审稿人选择与其专业领域相对应的关键词。关键词列表可以通过从预先确定的词典（例如 ACM计算机分类系统）中进行选择来创建，也可以以自由文本的形式确定。每个提案和每位审稿人都由一组关键词表示。计算两组关键词之间相似度分数的一种直接方法，是统计它们共有的关键词数量（Hartvigsen 等人，1999年；Xu 等人，2012年）。此外，还可以使用基于共有关键词的杰卡德相似度和戴斯相似度测量来计算相似度分数（Silva 等人，2014年；Liu 等人，2015年；Liu 等人，2016年）。一些研究人员同时依据审稿人的偏好和审稿人提供的关键词来确定分配结果（Di Mauro 等人，2005年；Conry 等人, 2009年；Huang 等人，2010年）。但是，自行声明的关键词具有主观性，就像竞价一样，可能导致误解和利益冲突。

与基于人工信息的方法的缺点相对应，当前采用人工方法的研究少于旨在实现 RAP 自动化的研究，例如基于文本的方法。不过，在一些研究中，为了测试所提出的技术，会将基于人工信息的信息检索方法作为一种比较方式（例如，比较基于竞价和基于 LSI（Dumais & Nielsen, 1992）的提案与审稿人匹配、竞价与相对有序的粒子群算法（Rodriguez & Bollen, 2008），以及评审人提供的专家主题与通过 PLSA 获得的信息（Karimzadehgan & Zhai, 2012））。

6.2 基于文本/基于内容的方法

基于文本或基于内容的方法是一种信息检索技术，通过比较评审人的内容属性与提案所关注的内容属性来进行推荐（Davoodi 等人，2013年）。这些方法属于隐式方法，无需用户采取任何额外操作，即可识别提案和评审人的能力，并利用文本挖掘技术计算审稿人与提案之间的相似度分数。

本节参考 Tan 等人（2021），根据所使用的信息类型将基于文本/基于内容的方法分为两类：基于词的方法和基于语义的方法。尽管大多数研究单独关注基于语义的方法或基于词的方法，但也有一些研究同时采用这两类方法。本节后文将分为三个标题，分别介绍采用基于词的方法、基于语义的方法以及结合语义信息和词语信息的方法的研究。此外，附录 B 对这些研究进行了详细总结。

6.2.1 基于词的方法

基于词的方法利用词频统计信息来描述评审人与提案之间的关系。基于词的方法分为以下两类：向量空间模型和统计语言模型。

向量空间模型（VSM）：在向量空间模型（VSM）中，每篇提案文档和每篇审稿人文档都表示为一个词袋向量，其中各词的权重由该文档中词语的词频-逆文档频率（TF/IDF）分数确定。词向量是指将每个词表示为一个数值向量，该向量反映了该词的使用方式或含义。审稿人与提案之间的相似度分数根据词语信息的向量表示计算得出。

Yarowsky 和 Florian（1990）首次使用 VSM，并通过余弦相似度计算评审人与提案之间的相似度分数。Basu 等人（2001）提出了三种基于 VSM 的相似度分数计算方法：使用文档全文的 TF/IDF 表示；通过基于机器学习的关键词提取，将提案和评审人都表示为关键词向量；以及通过基于领域本体的主题提取，在主题向量上精确定义提案和评审人。三种方法的结果表明，使用的信息越多，效果越好。Hettich 和 Pazzani（2006）提出了一种基于 TF/IDF 加权向量空间的推荐方法，用于为美国国家科学基金会（NSF）的资助申请推荐评审小组。Caldera 等人（2014）提出了另一种方法，该方法使用 TF/IDF 对投稿论文进行分类，并与潜在评审人的现有出版物进行比对，以实现评审人适配度评分的自动化。在大多数情况下，使用自由文本并辅以关键词提取和本体驱动的主题映射来自动分配提案与评审人，能够获得相关性更高的提案—审稿人匹配。为证明这一点，Biswas 和 Hasan（2007）通过利用研究人员以往发表的论文和领域本体，使用 TF/IDF 创建评审人简介。与其他研究不同，Shon 等人（2016）并未在计算评审人与提案之间的相似度分数时直接使用 TF/IDF 值，而是将根据评审人简介和提案简介中的词频得出的 TF/IDF 值用于模糊加权关键词中。

尽管研究人员已经证明，基于向量空间模型的技术可以在无人为干预的情况下将投稿分配给评审人，但向量空间模型存在一个根本缺陷，可能导致提案—评审人匹配不相关。向量空间模型是一种基于词汇匹配的技术，其假设是词语彼此独立。这种不切实际的假设会导致语境和相邻词语完整性的丢失，并忽略句子的语法结构。n 元语法可用于克服向量空间模型的局限性。然而，随着 n 的增加，由于词元组合的数量不断增加，向量的大小会呈指数级增长。

统计语言模型（SLM）：统计语言模型（SLM）使用多种概率和统计技术来预测序列中语言元素的概率分布，例如

词语、句子等。在评审人分配问题（RAP）中，SLM 分析提案和评审人的文本语料，并预测二者之间的相似度。Mimno 和 Mccallum（2007年）提出了采用狄利克雷平滑的 SLM，用于衡量评审人与提案之间的亲和度。Mai 等（2012年）开发了一种基于概率语言模型的新系统，用于为每份提案选择合适的评审人。对于广泛使用的 TPMS，Charlin 和 Zemel（2013年）采用 SLM 计算每个提案—评审人对的相似度分数。同样，Tang 等（2010年）和 Tang 等（2012年）使用基于 SLM 的检索方法计算成对相似度分数。Karimzadehgan 等（2008年）引入了一种混合统计语言模型，用于对评审人的多个专业知识方面进行建模。

尽管采用平滑方法的 SLM 在实践中具有可用性，但数据稀疏性和维度灾难仍是亟待解决的问题。语言建模采用数据驱动的方法来捕捉自然语言文本序列的关键统计特性，可用于预测序列中的后续词语，或在相关任务中执行槽位填充（Minaee等人，2022年）。最简单的统计语言模型——n 元语法模型，能够捕捉相邻词元之间的联系。另一方面，这些模型往往无法考虑词元之间的长距离依赖关系，而这些词元承载着语义关联。

6.2.2 基于语义的方法

当前基于词的方法存在的问题是，用户若不使用信息编入索引时所采用的词语，就无法获取自己正在寻找的信息。这是由自然语言特有的若干复杂特征造成的。其中最常见的一种是同义关系，即同一概念可以通过多种方式（词语）表达。另一种是多义性，即许多词语具有多种含义，因而用户查询中的短语可能会匹配到文档中的某些术语，而这些术语在语义上与用户并不相关。简言之，尽管基于词的方法能够刻画提案在词语层面的频率特征，却无法描述语义信息。因此，在某些情况下，这些模型计算出的相似度分数可能存在较大偏差。为解决这些问题，人们提出了基于语义的方法。下文将基于语义的方法分为两类介绍，即主题模型和其他基于语义的模型。

主题模型

基于语义信息的研究通常聚焦于主题模型。主题模型是一类统计模型，用于捕捉每位评审人的出版物和提交提案中的主题分布，这有助于了解提案的集中程度以及评审人的专业知识（Kim和Lee，2018年）。许多主题建模技术都用于提取评审人和提案的特定主题分布，例如潜在语义索引（Dumais & Nielsen，1992年）、概率潜在语义分析（Karimzadehgan 等人，2008年；Mirzaei 等人，2019年）、潜在狄利克雷分配（Blei等人，2003年）、作者主题模型（Karimzadehgan & Zhai，2009年；Tang 等人，2012年；Liu 等人，2014年；Kou 等，2015a，2015b）以及作者人物主题模型（Mimno 和 Mccallum，2007年）。

潜在语义索引/潜在语义分析（LSI/LSA）：潜在语义索引（LSI）或潜在语义分析（LSA）（Deerwester 等人，1990）是一种常用的降维方法，用于评估两篇文档之间的相似度。LSI 基于一种名为奇异值分解（SVD）的数学方法，揭示词与其底层概念之间的关系。Dumais 和 Nielsen（1992年）提出了首项使用 LSI 方法解决 RAP 的研究。他们将审稿人和提案的主题表示提取为 k维 LSI 空间中的点。然后

每个提案—评审人配对之间的相似度分数通过这两个点的余弦距离进行测量。Ferilli 等人（2006）开发了一个专家系统 GRAPE（全球审稿分配处理引擎），用于自动化 RAP。该系统以提案的主题以及审稿人的专业知识和偏好作为输入。Li 和 Hou（2016年）为 IEEE INFOCOM 2015年和 INFOCOM 2016年创建并实现了一个审稿分配系统（Erie）。在 Erie 中，提案与候选审稿人的匹配分数使用 LSI 进行计算。他们利用 LSI 方法提取提案的主题以及评审人的专业知识。在 Chughtai 等人（2019年）开发的基于本体的推荐系统中，使用概率方法提取并赋予本体权重，然后基于 LSI 对其进行比较，以计算相关性指数。

LSI 与语言无关，采用纯数学方法；因此，其计算方法相对直接、高效且易于使用。然而，它也有几个主要缺点：(i) LSI 需要大量文档语料库和词汇，才能得出准确的结果。由于数据量庞大，因此需要额外的存储空间和计算时间。(ii) 使用 LSI 获得的主题信息缺乏明确的概率，因此难以解释。(iii) 很难确定最佳主题数量。(iv) SVD 的计算开销很大，而且随着新数据的到来，难以更新。(v) LSI 向量占用的空间过大。

概率潜在语义分析（PLSA）：为克服 LSI 的缺点，Hofmann（1999年）提出了概率潜在语义分析（PLSA）。PLSA 可视为对 LSA 的概率解释，旨在根据涉及的概率，确定哪些文本和词语对应哪些主题。与 LSI 相比，PLSA 能更好地解释潜在空间中的维度。此外，研究表明，与 LSI 相比，PLSA 在捕捉多义性和同义关系方面能够提供更好的模型（Farahat 和 Chen，2006年）。

尽管概率潜在语义分析优于潜在语义分析，但它在审稿人分配问题解决方案中的应用并不广泛。Karimzadehgan 等（2008年）首次使用该方法提取审稿人专业知识的不同方面，然后根据这些方面将审稿人与提案进行匹配。在 Mirzaei 等人（2019年）的一项近期研究中，概率潜在语义分析被用于从审稿人以往发表的论文和提案中获取潜在主题。由于该模型存在若干缺陷，概率潜在语义分析在审稿人分配问题中的应用受到很大限制。概率潜在语义分析并非一种定义明确的主题模型，无法用于生成新文档。此外，它使用期望最大化（EM）算法进行模型参数训练。期望最大化算法由基于观测数据生成的似然函数驱动，并通过迭代最大化这些似然函数来估计参数。因此，训练后的模型性能差异很大，且在很大程度上取决于初始值。另一个问题是，参数数量会随文档数量直接增加，导致模型过拟合。

潜在狄利克雷分配（LDA）： 潜在狄利克雷分配（LDA）的创建旨在解决概率潜在语义分析（PLSA）的主要缺陷，即缺乏针对新文档的定义明确的生成模型（Blei等人，2003年）。LDA是一种主题模型，将文档表示为由词识别出的潜在主题的随机混合。LDA利用狄利克雷分布来确定文本和短语属于特定主题的概率。该模型通过狄利克雷分布捕捉词和文本的可交换性，从而实现连贯的测试数据生成过程。研究发现，LDA被广泛应用于从评审人已发表的研究和提交的提案中提取主题表示（Mimno 和 Mccallum，2007年；Tang 等人，2010年；Tang 等人，2012年；Charlin 和 Zemel，2013年；Liu 等人，2014年；Xu 和 Zuo，

2016年；Peng 等，2017年；Yan 等人，2017年；Nguyen 等人，2018年；Medakene 等人，2019年；Pradhan 等人，2020年；Yang 等，2020年；Patil 和 Mahalle，2021年；Tan 等人，2021年；Hoang 等人，2021年）。

作为最先进的主题模型之一，潜在狄利克雷分配方法成功识别出RAP研究主题。然而，由于其假设，该方法存在两个根本性缺陷。第一个假设是，主题数量必须固定且预先已知。通过使用旨在确定最合适主题数量的算法，可以部分克服这一限制。第二个假设涉及文档的主题独立性。语料库中的一个文档可能包含多个主题，或者同一主题可能出现在多个文档中。然而，在潜在狄利克雷分配方法中，论文中的主题比率是服从狄利克雷分布的随机变量，并且必须相互独立。由于这一假设，潜在狄利克雷分配无法利用主题内词语之间潜在的语义关系。

作者-主题模型（ATM）：与其他主题模型相比，传统的潜在狄利克雷分配在提案主题提取以及评审人与提案—评审人相似度评分计算方面取得了显著改进。然而，潜在狄利克雷分配没有考虑作者信息。鉴于潜在狄利克雷分配在建模作者兴趣方面存在不足，有人提出了一些作为其扩展的基线方法。作者-主题模型（ATM）由Rosen-Zvi 等人于2004年提出，是最广泛用于提案与审稿人匹配的潜在狄利克雷分配变体。

假设一篇文档有多位已知作者，且每位作者都被视为多个主题的混合体，作者-主题模型（ATM）旨在确定文档中的某个词由哪位作者负责。Kou 等人（2015a，2015b）使用 ATM 确定提案和审稿人发表成果的主题分布。随后，他们根据不同的主题权重，为每份提案提供专家组推荐。除了审稿人的专业能力和提案与审稿人的相似度等通常关注的方面外，Jin 等人（2017）还考虑了审稿人的研究兴趣，并使用 ATM模型和期望最大化算法估算提案与审稿人出版物的主题分布。

通过将不同类型的信息纳入 ATM，人们提出了几种面向 RAP 的新型基于 LDA 的主题模型。Tang 等人（2008年）通过纳入出版场所改进了 ATM，并提出了作者-会议-主题模型（ACT）。虽然给定主题的含义基本保持不变，但主题（语义相关短语）的出现情况及其相关因素会随时间变化。基于这一想法，Daud 等（2010年）提出了时间-专家-主题（TET）模型，该模型能够识别不同时期特定主题领域的专家，并展示专家的兴趣和关系随时间发生的变化。

为了确定评审人与提案之间的主题距离，Mimno 和 Mccallum（2007年）提出了作者-角色-主题（APT）模型。APT 模型在 ATM 中引入了角色层。APT 不再将某位作者的所有论文汇总为单一的主题分布，而是允许将每位作者的文档划分为一个或多个组。这些组代表同一位作者以不同“角色”撰写的内容，每个组都由独立的主题分布表示。Li 和 Watanabe（2013年）通过在模型中添加时间成分，扩展了 APT 方法。

在 ATM 中，每位作者的主题分布仅依据其文档确定，而不同作者的文档之间的相似度则被忽略。这些局限性会影响作者兴趣建模的整体性能，尤其是在分析跨学科研究中的文档时。Jin 等人（2018a）提出了作者-主题-主题（AST）模型，以克服在分配专家时遇到的困难。

到跨学科提案。AST模型通过添加主题层扩展了ATM模型。这样可以控制层次主题的创建，并允许作者之间共享主题。

尽管有人声称潜在空间中的主题在语义上连贯，但主题模型只考察文档语料库中的词语共现分布，而不考察其中包含的语义信息（Ekinci 和 Omurca，2020年）。此外，概率主题模型需要大型文档语料库，才能准确识别主题及其分布。在应用于摘要等短文档时，这可能会成为一个问题（Anjum 等人，2019年）。此外，现有的概率主题模型假定词语彼此独立、互不相关。虽然这一假设提高了处理效率，却消除了词语之间丰富的关系。由于存在这些局限，仅靠主题模型无法生成语义相关、有意义且连贯的主题。

其他基于语义的方法

除主题模型外，研究中还提出了其他基于语义的技术。与其他研究中采用的审稿人和提案表示方法不同，Li 和 Chen（2011年）基于知识集理论，开发了提案与审稿人之间的相似度计算方法。Müngen 和 Kaya（2018年）认为，仅查看候选审稿人的研究背景信息和元数据不足以搜寻审稿人。因此，他们提出了一种名为主题亲和传播（II-TAP）的新统计方法，该方法依据多种特征确定审稿人的专业领域，例如审稿人的论文、这些论文的元数据、审稿人随时间变化的兴趣，以及审稿人合著者的专业知识。此外，他们还制定了一项策略，以弥补应用中所用论文的摘要和关键词存在的不足。首先，使用 TextRank算法和 Zemberek-NLP3库获取论文的摘要和关键词。然后，使用频率查找模块补全缺失的关键词。

一些研究人员认为，无法准确确定评审人在特定领域的专业水平，而且评审人的专业能力会因领域和子领域而异（Xue 等人，2012年；Daş 与 Gökçen，2014年；Tayal等人，2014年；Mittal 等人，2019年）。因此，将专业水平描述为清晰集可能导致提案与评审人匹配不准确。由于难以用清晰值表示专业水平，一些研究使用语言变量定义相似度分数，以表示每位专家对每项提案的专业知识。Daş 和 Gökçen（2014年）使用语言变量确定相似度分数，以表示每位评审人针对每项提案的专业知识。在该研究中，语言变量用三角模糊数表示。另一项研究中，Tayal 等（2014年）使用二型模糊集估算评审人在不同领域的专业水平。此外，研究人员使用最能代表提案的三个术语构建了提案的模糊集。借助模糊函数，计算评审人与提案的二型模糊集之间的相等性。

有时，评审人无法手动提供其各个专业领域的详细信息。因此，确定涵盖所有相关学科的完整范围可能会很困难，从而降低为提案分配合适评审人的可能性。利用模糊逻辑从评审人过往文章中提取关键词，或许是一种可行的选择。Andrade-Navarro 等（2012年）设计了一款名为 Peer2ref 的网络工具，帮助期刊编辑识别合格的评审人。该工具使用模糊二元关系法，从提案的标题和摘要中提取重要概念作为关键词。随后，它会在 MEDLINE 的术语索引档案中搜索这些关键词；这些档案根据作者的书目资料编制而成。如果有多位作者同名，就可能难以找到合适的

评审人候选者。为解决姓名歧义问题，Peer2ref 会利用作者的合著信息自动评估档案。Mittal 等人（2019年）提出了一种基于模糊集和扩张概念的策略，以解决 RAP 中的歧义问题。他们首先提出了一种自动关键词提取方法，用于确定提案和评审人出版物中的关键术语。随后，他们构建了模糊词语关系图。通过应用模糊图中心性指标（即模糊度中心性、模糊介数中心性和模糊接近中心性），为每个提案和评审人构建模糊集，以表示相关术语及其权重。接着，采用模糊扩张原理，计算每个提案—评审人配对之间的相似度。为利用语境知识丰富文本表示，可将一些本体作为外部来源，用于向评审人和提案的文本文档中嵌入背景知识。例如，WordNet（Miller，1995）揭示了词语之间的许多同义关系。在 Mittal 等人（2019年）的论文中，WordNet 被用于寻找评审人与提案的模糊集之间的差异。Xue 等（2012）提出了区间模糊本体方法，通过使用学科本体和区间值模糊集理论对提案及评审人的研究主题进行建模，为 RAP 添加语义匹配。他们基于 WordNet 计算提案与评审人之间的关系。然而，由于这些本体是人工构建的，其覆盖范围有限，维护工作也极其繁重。因此，作为一种更可行的解，Davoodi 等（2013年）利用维基百科资料库中的背景知识构建了语义核，该资料库通过连接相关概念提供知识图谱。语义核用于丰富评审人的简介。

6.2.3 结合词语与语义信息的方法

一些学者主张结合词语和语义信息，为目标提案寻找最合适的审稿人。为了更准确地计算审稿人与提案之间的相似度分数，Tang 等（2010年）、Tang 等（2012年）和 Tan 等人（2021年）同时使用了潜在狄利克雷分配模型和 SLM。同样，在广为人知的审稿人分配系统 TPMS 中，Charlin 和 Zemel（2013年）使用 SLM 和潜在狄利克雷分配对提案及审稿人已发表的出版物进行建模。Peng 等人（2017年）指出，审稿人的科学兴趣往往是多方面的，并且可能随时间发生变化。他们表示，较新的出版物更能反映审稿人当前的研究兴趣。因此，作者提出了一种时间感知且基于主题的分配模型。该模型使用潜在狄利克雷分配和词频-逆文档频率来表示审稿人和提案，从而更准确地计算相似度分数。Pradhan 等（2020年）也使用潜在狄利克雷分配和词频-逆文档频率对审稿人的研究兴趣进行建模。Silva 等（2014年）通过整合“精确匹配”和“语义匹配”分数，提出了一种新颖的相似度分数测量方法。精确匹配分数是审稿人的专业知识出现在提案中的预测概率，使用 SLM 进行计算。语义匹配则由二分图的加权边得出；该二分图是基于本体的提案与审稿人语义网络的简化形式。

表2展示了聚焦于评审人和提案语义信息的方法分布。如表2所示，LSI、概率潜在语义分析和潜在狄利克雷分配等主题建模方法，是广泛采用的基于语义的方法，主要聚焦于RAP的第二阶段。由于潜在狄利克雷分配模型相较于其他主题模型具有优势，因此被广泛用于识别代表评审人和提案的主题。

基于神经网络的语言模型在多种自然语言处理下游任务中取得了卓越成果，包括文档分类、命名实体识别和机器翻译（Romanov and Khusainova，2019年；Kalyan 等人，2021年；Kapočiūtė-Dzikienė 等人，2021年）。尽管使用这类新型

生成语言模型计算相似度分数并不常见，但截至2018年，已有少数研究采用了这种方法。将基于神经网络的模型用于计算相似度分数的研究表明，这类模型比其他基于语义的方法更有效（Ogunleye 等人，2018年；Duan 等，2019年；Yong 等人，2019年）。基于神经网络的语言模型实际上属于基于语义的方法类别。不过，考虑到这类方法在计算相似度分数方面已带来的显著优势，以及未来可能带来的益处，本文将其单独列为一节进行介绍。

6.3 基于神经网络的语言模型

针对 RAP 提出的基于词的模型和主题模型主要采用基于频率的技术。这类技术将文档视为术语集合，虽然每个术语的频率至关重要，却忽略了术语的顺序。这些方法最显著的缺点是无法编码词语之间的语义和句法关联，因此生成的表示在语义上较弱。此外，基于频率的技术在处理大型数据集时，常会生成高度稀疏的向量，其中每个文档的向量表示的 n 维度等于该文档词汇表的大小 n。这种维度问题通常称为稀疏性诅咒。由于词缀的影响，土耳其语、中文和捷克语等形态丰富的语言中，数据稀疏性尤为普遍。

方法	2991	7002	8002	0102	2102	3102	4102	5102	6102	7102	8102	9102	0202	1202
LSI	✓								✓			✓		
PLSA			✓									✓		
LDA				✓	✓	✓	✓		✓	✓✓	✓	✓	✓✓	✓✓✓
ATM			✓	✓				✓✓		✓				
APT		✓				✓								
AST											✓			
Word2Vec											✓✓	✓✓		✓
BERT												✓		
CNN												✓		
双向长短期记忆网络												✓		
SciBERT												✓		
Doc2Vec													✓	
WordNet					✓							✓		
语义核						✓								
常见主题模型												✓		

表2：按年份划分的基于语义的方法研究分布。

基于深度学习和神经网络的语言模型已被开发出来，用以解决基于频率的技术中的此类问题，并自动学习文档的句法和语义特征。基于神经网络的语言模型利用词嵌入来学习词序列分布和文档的语义表示。词嵌入可用于计算连续值

的文档分布式表示，并将其置于低维空间中。基于神经网络的词嵌入方法分为两类进行考察。第一类称为静态词嵌入方法或基于预测的词嵌入方法，能够编码词语的形态和语义特征，并考虑子词元素（字符 n 元语法）。Word2Vec（Mikolov 等人，2013年）、GloVe（Pennington 等人，2014年）和 FastText（Bojanowski 等人，2017年）是最常见的静态词嵌入方法。第二类是基于上下文的词嵌入方法，包括采用深层架构的模型，如 ULMFiT（Howard & Ruder，2018年）、BERT（Devlin 等人，2019年）、GPT-2（Radfort 等人，2019年）和 DistilBERT（Sanh 等人，2019年）。静态词嵌入方法将词表示为稠密向量，而基于上下文的表示方法顾名思义，会将上下文信息添加到表示中。借助非常深的架构，这些方法能够从语境中学习词语的含义，并将其编码为向量。

近年来，基于神经网络的词嵌入方法已用于许多自然语言处理任务，包括词预测、翻译、自动语音识别和文本相似度计算。对于审稿人分配任务，提案—审稿人相似度因此可以通过其文档稠密向量表示的相似度来衡量。Ogunleye 等人（2018年）使用 Word2Vec 计算了评审人与提案之间的相似度分数。此外，他们还使用词频-逆文档频率、LSI 和潜在狄利克雷分配等基于词频的模型来测量相似度分数。对研究结果的比较表明，使用 Word2Vec 得到了最佳的提案—审稿人匹配。Zhao 等人（2018年）开发了一种名为词移距离—构造性覆盖算法（WMD–CCA）的新分类方法，将 RAP 作为分类问题来解决。该方法基于这样一种观点：利用审稿人和提案的研究领域标签信息可以提高分配质量。首先，使用 Word2Vec 表示提案和审稿人。然后，使用词移距离算法衡量提案与审稿人之间的最小文本距离。最后，根据领域关系预测候选审稿人和提案的领域标签。这种方法能够通过统计学习来学习领域特征。然而，它无法让机器从正确的视角学习领域信息并借鉴人类知识。为帮助机器从恰当的视角学习领域关系，Duan 等人（2019年）提出了一种句对建模方法。该方法利用提案和审稿人标题与摘要之间的领域关联作为监督信息。他们使用六种不同的基于神经网络的语言模型对句对进行建模，并通过训练集学习提案及审稿人论文的领域特征，具体包括：BERT-卷积神经网络、Word2Vec-CNN、BERT-bilSTM、Word2Vec-CNN、随机词嵌入-CNN 和随机词嵌入-bilSTM。最后，他们通过训练模型预测审稿人与提案之间的相似度分数。Yong 等人（2021年）还利用 Word2Vec 提取关键词的层次化语义表示。随后，利用知识图谱和规则匹配建立规则引擎，根据审稿人与每个提案的相关性生成有序的审稿人列表。

Anjum 等人（2019年）声称，包括向量空间模型、SLM 和概率主题模型在内的现有方法，无法充分解决提案与审稿人的专业领域之间的词汇不匹配和部分主题重叠问题。为处理提案与评审人之间的词语不匹配和部分主题重叠，作者提出了基于 Word2Vec 的共同主题模型方法。

实际上，审稿人或提案通常属于多个研究领域。因此，RAP也被表述为一个多标签分类问题，并根据多个估计标签分配审稿人（Zhang等，2019年）。此外，他们还引入了层次化透明表示（即Hiepar），用于

表示审稿人和提案的语义信息。Hiepar通过一个基于注意力机制的双层双向门控循环单元网络进行学习。该模型能够捕捉审稿人和提案的层次结构，以及它们各自独特的语义内容。

Cohan 等人（2020年）提出了 SPECTER 模型，利用在多种自然语言处理任务上预训练的 Transformer 语言模型当前取得的成功，学习学术出版物的独立表示。BERT 等 Transformer 语言模型只考虑文档内部的语境，不利用文档之间的任何信息。这限制了它们学习最优文档表示的能力。SPECTER嵌入方法结合预训练的 SciBERT（Beltagy 等人，2019年）模型和从引文图中提取的关系信号，生成论文嵌入。与目前可用的预训练语言模型不同，SPECTER 无需针对特定任务进行微调，即可用于下游应用。Hoang 等人（2021年）开发了一种决策支持系统，可自动为提案选定一组审稿人。在该系统中，审稿人和提案的简介均使用 Doc2Vec（Le 和 Mikolov，2014年）表示为向量。

6.4 基于网络的方法

基于文本/内容的方法仅限于词与元素描述之间由词典定义的关系，无法发现隐式关系。隐式关系指评审人与评审人之间，或提案与评审人之间的社交或学术网络，类似于人们建立的社交网络。 分析此类关系可以增强评审人的推荐效果。在本调查中，能够利用社交和学术关系信息的方法称为基于网络的方法。

社交网络分析是一种特殊的基于网络的方法，允许用户利用非正式互动中的社会结构，作为额外的信息来源。社交网络分析中的网络信息来源通常包括引文网络、协作网络、合著网络，以及评审人与评审人之间、提案与评审人之间众多相互关联的网络。

引文网络是一种描绘论文及其引文关系（即被引文章与施引文章之间连接）的有向图。该图能够识别文档之间的直接和间接联系。一些引文相关特征，例如文档中的引用次数，以及被引文章与施引文章的出版日期之间的时间间隔，也可用于区分文档间连接的权重，并估算它们的相似程度。Jin 等人（2017年）和 Jin 等人（2020年）利用引文网络辅助识别潜在评审人。所提方法基于以下假设：如果某个评审人候选人受到某一领域许多合格专家的引用，那么该候选人在该领域就具有较高的权威性。研究人员将引用行为组织到主题权威图中，以分析每位评审人在各项提案中的相对权威性。他们采用主题PageRank，为给定主题确定具有较高权威性的评审人。在另一项研究中，Tran 等人（2017年）旨在为某一特定会议的程序委员会换届推荐候选人。据作者称，该会议的关键词数量有限、较为笼统，有时还不是英语。使用这些关键词在文献数据库中搜索潜在评审人，会得到很多噪声结果。因此，这项研究提出了一种不单纯依赖文本的技术。在所提方法中，研究人员构建了一个三模网络，用于刻画专家候选人在特定会议中的知识，而候选人与会议的

接近程度则根据其各种关系的权重来确定，例如会议委员会成员身份、共同署名、引文网络等。

Xu 和 Du（2013年）指出，针对 RAP 且侧重于基于文本的信息检索技术的研究，未能整合相关信息。为解决这一问题，他们提出了一种结合语义网络、社交网络和引文网络的三层网络模型。在构建该网络模型时，评审人和出版物被设为节点，两者之间的关系则用连接表示。语义分析用于捕捉研究领域之间的语义关系；引文分析则用于分析文章之间的引文关系。粒子群算法利用三层模型获取的三类信息，提供审稿人推荐。Rodriguez 和 Bollen（2008年）提出了一种基于共同署名网络的协同过滤系统，该网络由包含顶点、边和权重的三元图定义。每个顶点代表网络中的一位作者，每条边代表两位作者合作发表的一篇出版物，每个权重则表示任意两位合作者之间关系的强度。他们采用较为先进的粒子群技术计算相似度分数。Silva 等（2014年）利用社交网络构建连通矩阵，根据审稿人的知识共享行为识别志趣相投的审稿人。

7. 第三阶段：基于优化的审稿人分配问题

RAP的第三阶段，即基于优化的审稿人分配问题（ORAP），将IRRAP转化为一个优化问题。在这一阶段，RAP根据不同的优化目标和约束条件进行建模，然后采用各种优化技术求解。本节将讨论优化审稿人分配的数学模型和算法。

一般而言，ORAP方法采用两阶段框架。第一阶段，在提案集与审稿集之间构建加权二分图。通常，提案与其审稿人之间边的权重表示两者的相关性。第二阶段，根据二分图，在各种约束条件下优化适当的目标函数，从而确定最终的提案—审稿人分配（Pradhan 等人，2020年）。RAP可视为广义指派问题（GAP）的一种进阶形式：两个规模庞大的对象集必须进行匹配，使一个集合中的每个对象都分配给另一个集合中数量有限的对象。自1999年以来，已有多种数学模型和算法被提出，用于将一组审稿人与一批提案进行匹配，以最大化整体优化目标。表3按年份展示了提出RAP数学模型的现有研究分布。所提出的数学模型在优化目标、约束条件和策略方面各不相同。在ORAP方法中，优化目标通常表述为最大化相似度分数的累积和、最大化主题覆盖率以及优化公平性。然而，仅仅建立一个能够优化目标的提案与审稿人匹配方案还不够。分配方案还必须考虑诸多实际约束条件。表4列出了RAP数学模型中考虑的约束条件及其定义。根据表4，尽管不同研究考虑了不同的约束条件，但通常都力求满足以下两项约束条件：

(1) 每位评审人都应分配到预先确定数量的提案（评审人工作量）；(2) 每份提案都应由一定数量的评审人进行评审（提案需求）。

给定提案集合 $P = \{1, 2, \dots, n\}$ 和可用评审人集合 $R = \{1, 2, \dots, m\}$。结合这两个常用的约束条件以及最大化相似度分数总和的目标，RAP 在其最一般的形式下可表述如下：

$$\max \sum_{i=1}^{n} \sum_{j=1}^{m} s_{ij} x_{ij} \tag{7-1}$$

$$s.t. \sum_{j=1}^{m} x_{ij} = d, \quad \forall i = 1, \dots, n \tag{7-2}$$

$$\sum_{i=1}^{n} x_{ij} = w, \quad \forall j = 1, \dots, m \tag{7-3}$$

其中 $x_{ij} = \begin{cases} 1 & if\ reviewer\ j\ is\ assigned\ to\ proposal\ i \\ 0 & ot\ \ erwise \end{cases}$, , h

在方程 7-1 中，s_{ij} 表示提案 i 与评审人 j之间的相似度分数。方程 7-2 和 7-3 分别表示对每项提案所需的评审数量（d）和评审人工作负载（w）的约束条件。

本调查在以下三个小节中讨论 ORAP 方法。第一小节讨论该问题的不同优化目标。第二小节介绍针对特殊目标的优化策略。最后一小节提及与问题形式化相关的其他问题。此外，附录 C 从优化角度详细总结了研究 RAP 的研究。在附录 C 中，你可以找到所提出的数学模型以及用于求解这些模型的算法。你还可以找到用于衡量所提出方法效果的不同性能比较策略和度量指标。

7.1 问题的优化目标

从优化视角研究 RAP 的研究提出了不同的数学模型和算法，以实现既定的优化目标。根据优化目标，本调查将研究分为三类：最大化相似度得分累积和、最大化主题覆盖和优化公平性。以下各节将分别讨论每一类。

7.1.1 最大化相似度得分累积和

研究中最常见的优化目标，是最大化所有已匹配提案—评审人配对的相似度得分累积和（Cook 等，2005年；Goldsmith & Sloan，2007年；Taylor，2008年；Kolasa & Krol，2011年；Charlin 等人，2012年；Charlin & Zemel，2013年；Long等人，2013年；Li & Hou，2016年）。这些研究假定已有已提交提案和评审人的列表，以及提案—评审人相似度得分矩阵。他们认为，当总相似度得分达到最大值时，每份提案都会分配到最合适的评审人。一些知名的会议管理工具，如 TPMS（Charlin & Zemel，2013年）、EasyChair（easychair.org）和 HotCRP（hotcrp.com），都实现了最大化总相似度的优化策略。

为使相似度分数之和最大化，大多数研究人员将 RAP 构建为整数线性规划（ILP）模型（Guerv´os & Valdivieso，2004年；Janak 等人，2006年；Li 等人，2008年；Taylor 等人，2008年；Conry 等人，2009年；Xu 等人，2010年；Chen 等人，2012年；Charlin 等人，2012年；Li 等人，2013年；Charlin &

Zemel，2013年；Silva 等人，2014年；Li & Hou，2016年；Ogunleyu 等人，2017年；Jin 等人，2018b年；Yang 等人，2020年；Kat，2021年）。

一些研究人员提出了一种称为混合整数规划（MIP）的数学建模方法，该方法在目标函数中引入了一组额外的连续变量，以实现约束松弛。（Janak 等人，2006年；Garg 等人，2010年；Wang等，2013；Lian 等人，2018年；Yeşilçimen 和 Yıldırım，2019年；Leyton-Brown 等人，2022年）。

当审稿人和提案的数量较少时，精确算法或典型的线性规划求解器（例如 CPLEX）能够生成提案—审稿人分配方案，以优化累积相似度得分（Janak 等人，2006年；Taylor 等人，2008年；Charlin 等人，2012年；Ogunleyu 等人，2017年；Jin 等人，2018b年）。由于提案和审稿人数量庞大以及时间有限这双重困难，精确算法无法在可接受的时间和内存范围内高效解决该问题。因此，众多研究人员对使用近似解算法或启发式算法来解决审稿人分配问题进行了研究，例如贪心算法（Cook 等，2005年；Karimzadehgan & Zhai，2009年，2012年；Long等人，2013年；Guo 等人，2018年；Boehmer 等人，2021年）、两阶段随机偏置贪心算法（Wang 等人，2013年）、阶段加深贪心算法（Kou 等，2015a，2015b；Mirzaei 等人，2019年）、基于加权矩阵分解的贪心算法（Pradhan 等人，2020年）、评审人贪心轮询算法（Payan & Zick，2022年）、蚁群算法（Li 等人，2008年）、随机划分（Jecmen 等人，2021年）、行列生成（Leyton-Brown 等人，2022年）、遗传算法（Li 等人，2007年；Xu等人，2010年；Chen 等人，2012年；Li 等人，2013年）、粒子群优化算法（Yang 等，2020年）、爬山算法（Hettich 与 Pazzani，2006年）、OBH算法（Yeşilçimen & Yıldırım，2019年）、PeerReview4All（Stelmakh 等人，2019年）以及 FairFlow（Kobren 等人，2019年）。一些研究人员提出了混合算法，即结合一种以上的算法，以消除单一算法的缺点（Yang 等，2020年）。例如，Guervós 和 Valdivieso（2004年）尝试通过结合贪心算法和进化算法来解决 RAP。与单独使用贪心算法相比，所提出的混合算法能够在更短的时间内实现更加精确的提案—评审人匹配。Xu等人（2010年）提出了一种基于贪心随机自适应搜索过程和遗传算法的混合算法，通过将遗传算法与局部搜索策略相结合，克服遗传算法的局限性并提高其有效性。Kolasa 和 Krol（2011年）将遗传算法与蚁群优化（遗传算法与蚁群优化）相结合，以获得可行的分配方案并防止过早收敛。同一项研究还提出了禁忌搜索与遗传算法（禁忌搜索与遗传算法）混合算法，以实现类似的目标。实验表明，尽管混合算法能够提供比单独使用遗传算法或蚁群优化更加精确的基于适应度的分配方案，但其时间性能较差。

当提案和审稿人被表示为二分图中的节点时，网络流模型非常适用于 RAP。图中的每条边都对应着成本和容量。成本由每位审稿人与每项提案之间的相关性决定。通过添加源节点和汇节点，并定义每条弧的作用，可以满足提案数量和审稿人数量等容量限制。研究人员基于网络流理论提出了不同方法，例如最大费用流问题（Goldsmith & Sloan，2007年）、最小费用流问题（Huang 等人，2010年）、最小凸费用流问题（Tang 等人，2010年；Tang 等人，2012年）以及等价最小凸费用流问题（Yan 等人，2017年）。

Method / Year	1991	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
广义分配问题（整数线性规划）		✓		✓	✓ ✓	✓	✓	✓		✓ ✓ ✓ ✓	✓ ✓	✓ ✓	✓		✓ ✓ ✓		✓ ✓	✓ ✓	✓ ✓	
混合整数线性规划				✓							✓					✓	✓			✓
最小值凸成本Flow问题								✓	✓											
成对受约束问题											✓							✓		
网络流理论					✓							✓								
受约束多方面委员会评审分配(CMACRA)							✓										✓			
加权覆盖率基于小组评审人分配问题													✓ ✓							
容量约束瓶颈转运问题	✓																			
集合覆盖问题			✓																	
最小值成本流问题								✓												
最大值主题覆盖率论文评审人分配（MaxTC-PRA)											✓									
模糊线性规划												✓								
有容量限制的运输问题														✓						
等价的最小值凸成本Flow问题															✓					
均衡多任务分配																		✓		
策略证明																		✓		✓

via 分区 分配																			
加权 z- 无环 Peer 评审															✓		✓		
两阶段 论文 分配 问题																		✓	

表3：各年份数学模型的分布。

Cook 等人（2005）为 RAP 提出了一种替代策略。他们将该问题建模为集合覆盖问题，其主要目标是使评审每一对提案的共同评审人数量之和最大化。Lian 等（2018）将 RAP 视为双边匹配问题，其中一方（评审人）对另一方（提案）具有偏好。考虑到双方都有容量约束，该问题被建模为混合整数规划。与其他研究不同，该研究为双方都设置了容量上限和下限。受多准则决策和社会选择理论研究的启发，本研究在所提出的 RAP 流网络算法中使用有序加权平均数（OWA）。Pradhan 等（2020）将 RAP 建模为多任务分配问题的均衡模型。在此案例中，任务对应提案，工作者对应评审人。他们旨在通过提高主题相似度并降低利益冲突，在评审人的均衡工作量分配下最大化总利润。为求解该模型，他们提出了一种基于加权矩阵分解的贪心算法，该算法通过搜索分配方案，使每位评审人的最大与最小总利润之差达到最高，从而逐步提高总利润。

一些常用的决策科学技术，例如模糊优化和层次分析法（AHP），也可用于促进RAP的实施。例如，Daş 和 Gökçen（2014）从固有不确定性的角度处理了RAP。他们提出了一个模糊整数线性规划模型，旨在通过清晰约束最大化提案与评审人组之间的相似度分数，同时最小化组建评审小组的成本。评审小组成本由支付给评审人进行评估的费用，以及评审人产生的差旅和餐饮等开支构成。他们在模型中增加了一个约束条件，以确保成本不超过特定预算。Tayal 等（2014）提出了另一种用于求解RAP的模糊优化方法，该方法在一组相关约束条件下，优化评审人的二型模糊专长集与提案的模糊关键词集之间的相等度得分。这项研究提出了一个涵盖专长建模各个方面的集成解。Xue 等（2012）采用区间模糊本体方法对Tang 等（2012）的数学模型进行了修改，以计算相似度分数。Li 和 Watanabe（2013）建立了一个用于评审人推荐的集成模型，同时考虑评审人的相关性和专业知识。专业水平和相关性得分采用基于AHP的方法进行计算。

7.1.2 最大化主题覆盖率

分配给提案的所有评审人的累积专业知识应覆盖提案的所有方面。若将提案视为单一主题而忽略其多个方面，就会导致审稿人分配不充分，尤其是对于包含多个子主题的提案。因此，一些研究人员

将最大化主题覆盖率视为全局目标，并提出了不同的方法以实现全面覆盖（Karimzadehgan & Zhai，2009年、2012年；Tang 等人，2010年；Xue 等人，2012年；Tang 等人，2012年；Long 等人，2013年；Tayal 等人，2014年；Kale 等人，2015年）。他们将评审人的专业知识和提案内容表示为两组主题，或假设这两组主题已知。这些主题集可能源自会议、期刊或资助机构通常提供的学科关键词列表，也可能通过主题模型自动发现。随后，他们使用不同的主题覆盖率评估评审人分配方案的质量。

Long 等人（2013）提出了最大主题覆盖论文审稿人分配（MaxTC-PRA）问题，其中，根据集合覆盖率衡量将一组审稿人分配给一份提案的质量。该比率是审稿人组覆盖的提案主题数量与提案主题总数之比。Tang 等人（2010年）和 Tang 等（2012年）提出了一种基于约束的优化问题，用于考虑“多主题覆盖”的提案与审稿人匹配。他们将两种主题覆盖度量纳入所提出的模型，作为约束条件。第一种是覆盖率，要求分配给每份提案的审稿人组尽可能多地覆盖提案的各个方面。覆盖率的公式见公式 7-4。第二种是置信度，即分配给某份提案的每位评审人的专业能力应尽可能涵盖提案的各个方面。置信度可通过公式 7-5 计算。为求解所提出的优化框架，他们将该问题转化为凸费用网络流问题，并提供了一种能够保证获得最优解的高效算法。

$$Coverage(q_i) = \frac{\left|T(q_i) \cap \mathrm{U}_{v_j \in V(q_i)} T(v_j)\right|}{|T(q_i)|} \tag{7-4}$$

$$Confidence(q_i) = \frac{1}{m} \sum_{v_j \in V(q_i)} \frac{\left|T(q_i) \cap T(v_j)\right|}{\left|T(q_i) \cap \mathrm{U}_{v_j \in V(q_i)} T(v_j)\right|} \tag{7-5}$$

其中，提案（q_i）和评审人（v_j）的相关主题分别表示为 $T(q_i)$ 和 $T(v_j)$，展示了与提案和评审人各自最相关的方面。$V(q_i)$ 是候选审稿人集合。

这些提出的策略提高了评审分配的各方面覆盖率。然而，在这些方法中，审稿人被独立分配给每份提案，并未考虑整个委员会。

这使得在一组评审人之间平衡评审工作量变得困难，并可能导致将过多的提案分配给在热门领域具备专业能力的评审人。

此外，尽管一些提案由知识渊博的审稿人评估，但其他提案可能不可避免地会被分配给专业知识有限的审稿人。与独立地将审稿人分配给提案的方法不同，一些学者定义了一种多方面分组分配。

Karimzadehgan 和 Zhai（2009年、2012年）提出了受约束的多方面委员会评审分配（CMACRA）问题，其目标是在评审人工作量和提案需求约束下，基于多方面的提案与审稿人匹配，将一组评审人分配给一组提案。他们将 CMACRA 问题建模为整数线性规划（ILP）。他们采用覆盖率和置信度指标来确定分配质量。寻找覆盖率和置信度的最优值通常是 NP难问题（Zhai 等人，2003年）；因此，他们采用贪心技术来获得近似解。

基于加权覆盖度，他们定义了基于加权覆盖的分组式审稿人分配问题（WGRAP）模型，以最大化每个主题的总加权覆盖评分。此外，Kou 等人（2015a、2015b）提出了一种名为阶段加深贪心算法的近似算法，实现了 1/2 的近似比。Mirzaei 等人（2019年）还提出了另一种基于分组的多方面提案—审稿人分配公式，称为“基于潜在研究领域的多方面评审团队分配”（MARTA-LRA）。该方法使用潜在研究领域（通过拼接审稿人发表成果中的术语来表示），以确定提案和审稿人各主题的重要性。

Kale 等人（2015年）通过纳入审稿人的时效性，为最大主题覆盖优化策略增添了一个新的维度。研究人员利用评审人论文发表所在期刊的影响因子来确定评审人的时效性。为确保审稿人尽可能评阅主题各异的提案，研究人员提出了基于贪心策略的最大主题覆盖论文—审稿人分配算法。

7.1.3 优化公平性

尽管最大化相似度分数总和或优化主题覆盖通常能够实现准确的审稿人分配，但这可能会导致某些提案受到不公平对待。主要原因在于，为了优化全局目标，可能会将更多合格的审稿人分配给一个提案，而牺牲其他提案。在这种方法下，某些提案可能会被分配给缺乏相关主题领域专业知识的审稿人，而另一些提案则可能被分配给胜任的审稿人。 Kobren 等人（2019年）分析了 CVPR'17、18 会议上数千篇论文的数据。分析表明，采用这种方法得出的分配结果中，至少有一个提案对所有审稿人的相似度分数均为零。为了实现公平，不应让任何审稿人以牺牲其他人为代价获益。

Hartvigsen 等（1999）通过最大化总专业知识，部分解决了公平性问题，同时施加约束，确保每个提案至少有一位专业知识超过特定阈值的评审人。该方法为每个提案分配了一位高度专业化的审稿人，但未能避免在分配其余审稿人时歧视某些提案。近期研究尝试通过以下方式克服这些问题：（a）对有效提案—评审人匹配的最低相似度施加严格约束；或（b）最大化最坏情况下提案相似度之和。

为实现公平性，Garg 等人（2010）旨在最大化处境最差的提案的效用。考虑到评审人的字典序偏好，他们基于最大最小公平性目标提出了字典序最小分配和秩最大分配的概念，如公式 7-6 所示。他们认为，如果在覆盖率和工作量约束条件下，所有提案的相似度之和达到最小值，那么这种分配就是公平的。他们证明，当等价类超过两个时，求解字典序最小分配或最优平等主义分配是 NP难的。

$$(g,p) = max_M \, min_j \, \sigma_j(M) \tag{7-6}$$

其中，$\sigma_j(M)$ 是分配 M下评审人 j 的签名。签名通过加权偏好或字典序排名进行比较。

Stelmakh 等人（2019年）研究中讨论的公平性目标，是最大限度地提高处于最不利地位的提案的评审质量，而不是最大限度地提高所有提案的总评审质量。受最大最小公平性概念的启发，他们旨在针对给定的相似度矩阵 S，找到可行分配 A，以最大化公平性目标 Γ^S 。$R_A(i)$ 表示在分配 A下评审提案 i 的评审人集合。公平性目标的公式见方程 7-7。

$$\Gamma^S(A) = min_{i\in[n]} \sum_{\mathrm{j}\in R_A(i)} s_{ij} \tag{7-7}$$

与针对 RAP 的特定确定性目标优化分配的各种方法不同，本研究还考察了在统计准确性方面的分配。他们提出了基于增量最大流技术的 PeerReview4All 算法，以满足公平性和统计准确性这两项核心要求。该算法在最大值-最小值公平性目标方面实现了常数因子内的最优解。此外，PeerReview4All 算法为正确识别应予接受的最佳提案提供了稳健的统计保证。

所提出的算法通过最大化处境最差的提案的相似度之和来生成更公平的分配，但它们存在一些缺点：(i) 限制允许的最小值相似度通常会使问题不可行，因为可能不存在任何匹配，能够在满足阈值的前提下为所有提案提供足够的覆盖率。(ii) 这些算法不会在审稿人之间强制实现负载平衡，因此可能产生审稿人被分配到数量差异极大的提案的匹配结果。Kobren 等人（2019年）提出了一种考虑局部公平性的全新模型，以消除这些问题。该模型被表述为一个整数线性规划，用于优化全局目标，同时包含用于平衡审稿人工作负载的上限和下限约束，并包含局部公平性约束，以确保每个提案被共同分配给一组具备足够专业知识的审稿人。如公式 7-8 所示，局部公平性约束可以表述为：每个提案的总体相似度分数必须至少为 T。

$$\sum_{i=1}^{n} x_{ij}\, s_{ij} \geq T, \quad \forall j = 1,2,\dots,m \tag{7-8}$$

由于所提出的模型是NP难问题，因此在求解时引入了两种算法。第一种是FairIr，它能提供一种公平的分配，在满足局部公平性和工作量约束的同时最大化全局目标。第二种算法FairFlow不保证生成公平匹配，但其速度可能快于FairIR和PeerReview4All。

Payan 和 Zick（2021年）将 RAP 构想为一种公平分配问题，旨在将评审人的质量公平地分配给提案。在这项研究中，公平性的核心标准是嫉妒：如果某个提案认为分配给另一个提案的审稿人比自己的更好，就会嫉妒对方。由于不可分割物品无法实现无嫉妒分配，他们将重点放在至多一件物品无嫉妒（EF1）这一标准上。在传统的公平分配情境中，广为人知的轮流选择法通过为代理人确定顺序，并允许他们每次选择一件物品，来生成 EF1 分配。然而，由于同一位审稿人不能重复审阅同一提案的限制，轮流选择法无法实现 EF1。为克服这些障碍，他们提出了审稿人轮流选择机制，这是对传统轮流选择法的扩展。

约束条件	说明		参考文献*
评审人工作量	每位评审人获得的任务不应超过 一定数量的提案。		1-3、5-21、23-30、32、34-37, 39-52
提案需求	每份提案都应分配给一定数量的 审稿人。		1-2、4-11、12-17、19、21-23、25-30、34-37、39、41-46, 48-52
利益冲突回避	不应将评审人分配给一位 提案，若存在任何形式的利益冲突，则、】【 他们之间的利益关系。	所属机构 关系	5、19、29–30、42、52
		共同作者 关系	5、12-14、19、26、29、37、45, 51, 52
		导师—被指导者 关系	15, 19, 21, 29
		竞争者 关系	19、30、46
		项目级 合作	26
策略无关性	任何评审人都无法改变 与其存在利益冲突的提案 。		42, 44
鱼雷式评审	审稿人分配应确保 提案不太可能已经 遭到鱼雷式评审。“鱼雷式 评审”指的是这样一种情况： 评审人故意试图被分配到 自己不喜欢并想要否决的提案。		35
无环 约束	一项评审分配若不包含以下 类型的串通环，则应位于：一个 评审人循环，其中每位评审人评审一份提案 由下一位评审人提供，从而形成一个 评审周期，在此期间每位评审人都给出 良好的评分。		35, 39, 42, 45
评审人的专业知识	分配给提案的评审人应 具备评估该提案所需的知识。		16, 26, 40
权限平衡	评审人的权威性应当 针对每项提案保持平衡性。		13-14
评审人去- 匿名化	无法确定任何审稿人身份， 为每项提案提供较高的确定性。		35
资历	应合理分配经验丰富的评审人。 并尽量保持均衡		45
评审小组成本	评审小组的成本不应超过预先设定的 确定的预算。		25
评审小组规模	评审小组的规模不应超过预先 确定的数量。		25
地域多样性	分配给同一提案的审稿人 不得来自同一地理 区域。		45
评审角色	如果可以为每个提案分配多个评审人的评审角色 ，(i) 每个提案 必须由每个评审角色评估且仅评估一次， 并且（ii）每位评审人必须担任相同的 次数，且在每一种不同的 评审角色中。		4

* (1) Hartvigsen 等人，1999年 (2) Guerv´os & Valdivieso，2004年 (3) Cook 等，2005年 (4) Janak 等人，2006年；(5) Sun 等人，2007年 (6) Li 等人，2007年 (7) Li 等人，2008年 (8) Karimzadehgan & Zhai，2009年 (9) Conry 等人，2009年 (10) Xu 等，2010年 (11) Garg 等人，2010年 (12) Liu 和 Hong，2010年 (13) Tang 等人，2010年 (14) Tang 等人，2012年 (15) Chen 等人，2012年 (16) Karimzadehgan & Zhai，2012年 (17) Xue 等人，2012年 (18) Charlin 等人，2012年 (19) Long 等人，2013年 (20) Tayal 等人，2014年 (21) Wang 等，2013年 (22) Li 等人，2013年 (23) Charlin & Zemel，2013年 (24) Cechlarova 等人，2014年 (25) Daş 与 Gökçen，2014年 (26) Silva 等人，2014年 (27) Kale 等人，2015年 (28) Kou 等，2015a，2015b (29) Liu 等人，2016年 (30) Yue 等人，2017年 (31) Ogunleyu 等人，2017年 (32) Yeşilçimen & Yıldırım，2019年 (33) Mirzaei 等人，2019年 (34) Yang 等，2020年 (35) Jecmen 等人，2020年 (36) Jin 等人，2018b年 (37) Pradhan 等人，2020年 (38) Kat，2021年 (39) Boehmer 等人，2021年 (40) Kobren 等人，2019年 (41) Dhull 等人，2022年 (42) Guo 等人，2018年 (43) Xu 等人，2019年 (44) Lian 等人，2018年 (45) Leyton-Brown 等人，2022年 (46) Yan 等人，2017年 (47) Jecmen 等人，2021年 (48) Stelmakh 等人，2019年 (49) Payan & Zick，2021年，(50) Jin 等人，2017年，(51) Li & Hou，2016年，(52) Hoang 等人，2021年

表4：RAP数学模型中涉及的约束条件及其定义。

7.2 特殊目标的优化策略

在文献中，已有研究采用不同的优化策略来避免恶意行为。本节将讨论有关避免恶意行为的研究。同行评审是一种有效且广泛使用的方法，用于评估提案的可行性和质量。它也是一种竞争性程序，这意味着某项提案的评估结果会受到其他提案评估结果的影响。由于被分配的评审人直接产生和/或参与评审结果的形成，因此，审稿人分配一直是评审流程中最关键的任务之一。由于竞争十分激烈，同行评审容易受到审稿人策略性操纵行为的影响。

在分配审稿人的过程中，通常会询问审稿人的专业知识和兴趣。这为操纵行为打开了方便之门，因为审稿人可能会谎报自己的知识和兴趣。例如，审稿人通常会在运行分配算法、确定自己将被分配评审哪份提案之前，申报自己对哪些提案有偏好。在这种情况下，审稿人可能会故意设置自己的偏好，以便被分配到某份特定提案。操纵行为并不限于竞标系统，因为用于确定论文分配的许多信息都可能被操纵，例如审稿人申报的专业领域或审稿人的出版物清单。此外，不道德的审稿人可能会与同一会议或资助机构的其他提案提交者私下串通。基于这种约定，这些审稿人会设法互相评审对方的提案，并以给予好评作为交换，获取某种外部报酬。

正如 Thurner 和 Hanel（2011年）所述，即使只有少数以自我为中心、采取策略性行为的审稿人，也会显著降低评审流程的质量。因此，许多研究人员都致力于寻找公平的审稿人分配方案，确保不歧视任何提案。本节将在以下三个标题中讨论影响审稿人分配公平性的恶意行为，以及避免这些行为的各种解法。

7.2.1 环路预防

许多人在会议和资助机构中担任双重角色。他们既是提交提案的申请人，也是被指定评估提案的审稿人。因此，在相关领域工作的申请者和评审者更有可能存在超出利益冲突的学术关系。在最坏的情况下，两个不同提案的申请人会被分配为审稿人，互相评审对方的提案。他们可能共同合作，并给出高于应有水平的评分。一些研究人员将这种行为称为“串通圈”，即一群不道德地相互评估和支持的审稿人（Littmann，2021年）。此类审稿人与提案环路违反匿名性，威胁同行评审流程的公平性和正当性，因此应在自动审稿人分配中予以避免。

Guo 等人（2018）尝试通过定义环约束来避免串通圈，从而确保分配中不包含 k-环评审者。以二环为例：如果评审人 A 被分配去评审评审人 B 的提案，那么评审人 B 就不能被分配为评审人，去评审 A 的提案。他们开发了一种环检测算法。该算法利用拓扑排序和深度优先搜索算法检测循环。为在以下三项约束条件下最大化提案与评审人之间的相似度分数总和，他们将 RAP 表述如下：提案需求（方程 7-10）、评审人工作量（方程 7-11）以及 k-循环自由约束（方程 7-12）（方程 7-9）。他们声称，环的长度越长，协作作弊的难度就越大。因此，在所提出的模型中，他们仅允许长度大于预定义阈值 k 的长环。

$$max_{x(i,j)} \sum_j \sum_i s_{ij} x_{ij} \tag{7-9}$$

$$s.t.\ \sum_j x_{ij} = d,\ \ \forall i\ \in B \cup C \tag{7-10}$$

$$\sum_i x_{ij} \leq w,\ \ \forall j\ \in A \cup B \tag{7-11}$$

$$l(g) \geq k,\ \forall g \in G \tag{7-12}$$

$$x_{ij} \in \{0,1\},\ \forall j \in\ A \cup B,\ \forall i\ \in B \cup C \tag{7-13}$$

其中，s_{ij} 是评审人 j 与提案 i之间的相似度分数。如公式 7-13 x_{ij} 所示，13 x_{ij} 是一个二元分配变量；如果评审人 j 与提案 i存在利益冲突，则$x_{ij} = 0$ 。w 是评审者工作负载容量上限， d 是每个提案所需的评审人数。他们用有向图 $G(V, E)$表示审稿人分配。在该有向图中，(i) 节点（V）分为三类：A、 B和C。节点 A 是入度为零的仅评审者节点。节点 B 表示同时是申请人和评审人的人员。节点 C 是出度为零的仅申请者节点。(ii) 边（E）表示评审人节点与申请者节点之间的分配。(iii) $\{r, p\}$表示由 p 提交的提案由 r评审。(iv) $l(g)$表示子图 g 中的环路长度，其中 $g \in G$。

与 Guo 等人（2018）的设定更为接近的是，Boehmer 等人（2021）也考虑了无环评审分配的计算问题。他们提出了加权无环评审者分配问题，并开发了贪心算法来解决这一问题。他们对所提出的算法进行了理论分析，并证明该启发式方法总能找到一个质量损失最小的无环分配。

Leyton-Brown 等人（2022）引入了“无二元循环”约束，以防止出现两位评审人对彼此的提案竞价表示赞同，且被分配来评审这些提案的情况。他们为每个竞标环提出了式 7-14 中所示的松弛变量 $s_{jj'ii'}^{cy} \geq 0$ ，并为与该竞标环相关的每项审稿人分配添加了式 7-15 中所示的惩罚变量 p^{cy} 。

$$x_{ij} \leq s_{jj'ii'}^{cy} \quad \forall j, j', i, i' \in CY \tag{7-14}$$

$$Objective^{cy} = \sum_{j,j',i,i' \in CY} p^{cy} s^{cy}_{jj'ii'} \quad (7\text{-}15)$$

其中，评审人 j 对评审人 j' 提交的提案 i 竞价表示赞同，而评审人 j' 对评审人 j 提交的提案 i' 竞价表示赞同。二元变量 x_{ij} 表示提案 i 是否分配给评审人 j。CY 表示一个循环集合。

这些拟议的无循环分配方法可以避免同一会议中的串通圈。然而，这种防范措施可能会被不道德的人规避，他们会避免形成循环，例如，某位评审人在一次会议上帮助了一名申请人，而该申请人则在其他场合给予回报。

7.2.2 规避鱼雷式评审

分配还可能通过“鱼雷式评审”受到操纵，即评审人故意希望被分配到某项提案，以便否决该提案，这可能是因为该提案属于竞争提案，也可能是因为评审人不喜欢该提案的主题领域。例如，评审人可能会给竞争提案较低的评分，希望自己的提案能够获得接受，从而降低竞争提案被选中的可能性。

为了避免或减轻此类操纵行为，Dhull 等人（2022）提出了一种最大相似度策略无关评审人分配方法，在该方法中，任何评审人都无法通过改变自己的评估来改善其自身提案的结果。他们采用了 Alon 等人（2011）提出的分区方法，以确保策略无关性。通过这种分区方法，提案被划分为若干子组，且不会将任何评审人分配到与其自身提案属于同一子组的提案。随后，针对每个子组，分别汇总各评审人的评估结果，因此，任何评审人的评分都不会影响其自身提案的最终结果。他们提出了多项式时间算法，并从理论上保证了其算法的分配质量。Xu 等人（2020）也研究了在策略无关性约束条件下，如何将评审人分配给能够最大化相似度分数的提案。Dhull 等人（2022）假设每位评审人都提交一份提案，且每份投稿均由一位评审人撰写。相比之下，Xu 等人（2020）考虑了任意作者署名的情况，即每份提案可以有多位作者，而每位评审人也可以提交多份提案。此外，他们开发了一种基于分区方法的算法，但该算法无法提供关于分配质量的理论保证。

这三项研究仅考察了评审人希望自己的提案获批的动机。Jecmen 等人（2020）还设想了评审人给出不诚实评审的其他可能动机，例如不喜欢某个研究领域，或申请人与评审人之间的合作所带来的抑制因素。Jecmen 等人（2020）提出，通过最大化相似度总和，同时加入一定的随机性，来随机分配评审人与论文，以防止鱼雷式评审、不诚实的正面评审以及评审人身份被反向识别。该方法旨在确保无论评审人偏好和相似度如何，提案—评审人配对被分配的概率都很低。

7.2.3 利益冲突规避

另一项重要要求是，“不得将评审人分配给与其存在任何利益冲突的提案”。通常，在 RAP 程序的第二阶段，会检测提案与评审人之间可能存在的利益冲突。在第三阶段，会纳入避免这些已发现的利益冲突的策略。

纳入 RAP 数学模型。本调查在“5.1 利益冲突检测”一节中讨论了利益冲突检测方法。本节介绍了如何将利益冲突规避策略整合到 RAP 模型中。为了确保分配中不存在利益冲突，通常会在目标函数中将二元决策变量 x_{ij} 设为零（Guerv´os & Valdivieso, 2004年；Janak 等人，2006年；Karimzadehgan & Zhai, 2009年、2012年；Tang 等人，2010年；Tang 等人，2012年；Chen 等人，2012年；Xue 等人，2012年）。Sun 等人（2008年）通过将存在利益冲突的评审人的专业水平设为零来避免不公平现象。Long 等人（2013年）创建了一个利益冲突集合 $\mathcal{C}$，其中包含提案与评审人之间所有潜在的冲突。该研究所提出的算法搜索一个集合 A，该集合包含提案—评审人分配对，且与集合 $\mathcal{C}$ 不相交。为了消除利益冲突，Silva 等人（2014年）定义了一个指示变量 c_{ij}，并将其添加到目标函数中。他们将 $c_{ij} = 0$ 设为某个值，以表示该分配不可能实现；否则，将 c_{ij} 设为 1。Wang 等人（2013年）也使用了指示变量 c_{ij}，并基于该变量生成了一个约束函数，以确保不存在利益冲突。

Yan 等（2017）提出了一种称为最小利益冲突论文评审人分配（MinCOI-PRA）的新公式，旨在最大化主题相似度并最小化利益冲突。他们将每种分配的COI（A）定义为公式 7-16 所示。在这一公式中，研究人员之间的利益冲突（作者到评审人）表示为COI_V，研究人员所属机构之间的利益冲突表示为COI_D。

$$COI(A) = \gamma \sum_{i\epsilon P, j\epsilon R} COI_V\,(F(i), j) + (1-\gamma) \sum_{i\epsilon P, j\epsilon R} COI_D\,(D\big(F(i)\big), D(_j)\big) \qquad (7\text{-}16)$$

其中，$\gamma\epsilon$ [0,1] 是用于平衡COI_V 和COI_D的耦合参数。每个提案的申请人表示为F（i），其所属机构表示为D（F（i））。每位评审人的相关机构表示为D（j）。

Pradhan 等（2020）的最新研究将 RAP 公式化为一种最大化型均衡多任务分配问题，通过最大化主题相关性并最小化每位评审人的利益冲突，逐步优化收益。为满足评审人与提案之间利益冲突的最低要求，他们提出了与共同作者距离成反比的利益冲突值。

7.3 与问题构式相关的其他问题

除了在 ORAP 方法中建立并优化数学模型之外，一些研究还关注问题的结构。本调查将这些研究归为两个主题：采用分组策略，而非将提案逐一分配给评审人；以及采用两阶段评估策略，在收到大量提案时逐步完成分配流程。

7.3.1 分组策略

现有的大多数方法都采用相同的策略，旨在为每份提案找出合适的评审人（如图 11A 所示）。由于分配给主题相似的提案的评审人具有不同的学科领域经验，这种策略可能导致评审结果不公平（Wang 等，2013年）。例如，一份关于可持续供应链管理的提案，可能由两位环境科学专家和四位供应链管理专家评审。然而，另一份主题相同的提案可能由三位生态科学专家和两位供应链管理专家评估。此外，当提交的提案很多且有大量合格评审人可供选择时，为每份提案挑选合适的评审人会耗费大量时间（Xu 等，2010年）。

为使相似的评审人评估主题相似的提案，Xu等人（2010年）基于“分组”理念提出了一种新的评审人分配策略（如图 11B 所示）。根据这一策略，首先按研究领域对提案进行分组，然后为组内每份提案分配一组评审人。受分组理念启发，Daş 和 Gökçen（2014）提出了一种新颖的分配策略：他们不再将提案组分配给评审人，而是针对组内每份提案，从潜在评审人中确定最合适的评审人。为在严格约束条件下，最大限度提高提案组与评审人组之间的相似度分数，同时最大限度降低生成评审人组的成本，他们开发了一个模糊整数线性规划模型。Wang等人（2013年）定义了组间分配策略（如图 11C 所示），即根据相似度将提案和评审人分别划分为不同的组，然后将每个提案组分配给一个独立的评审人组。研究人员将RAP构建为一个多目标混合整数规划问题，并提出了一种两阶段随机偏置贪心算法来求解。

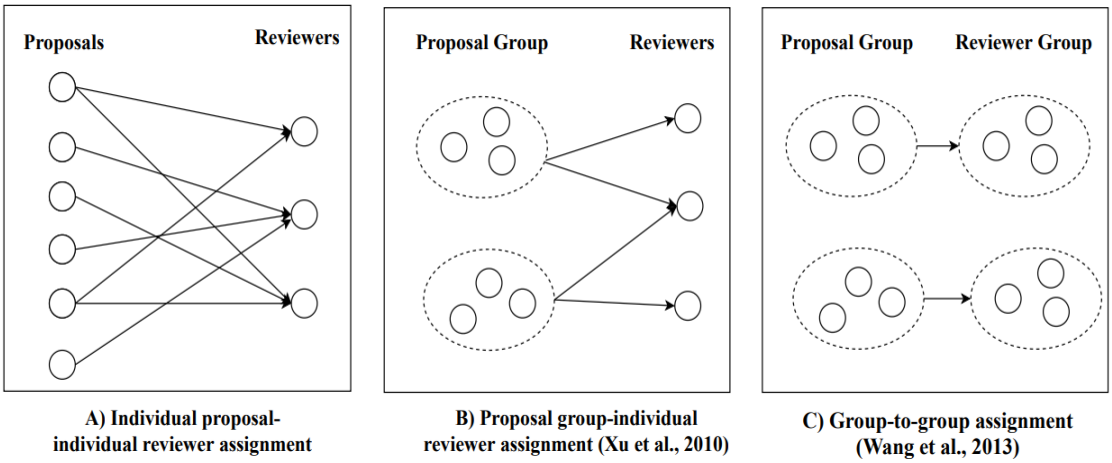


图 11：分配的分组策略。

7.3.2 两阶段评审策略

在大型会议中，识别合适的提案—评审人配对，以及应对妨碍评估质量的恶意行为，变得十分困难。AAAI'21-22 和 IJCAI'22 等会议采用了两阶段评审流程，以便在时间限制下应对这些挑战。在这一系统中，完成首轮评估后，会将部分提案分配给额外评审人进行第二轮评审。将评审流程划分为两个阶段的主要原因是更有效地利用评审人资源，并提高评估质量。在这一总体目标下，两阶段评审可以采用不同的方式，以满足不同的需求。例如，正如 AAAI'21 所实施的那样，只有在第一阶段获得足够高分数且有机会被接受的提案，才会在第二阶段接受评审。两阶段评审还可以帮助评估处于批准与拒绝边缘的“临界提案”（Leyton-Brown，2022年）。此外，它还有助于向第一阶段被拒绝的提案申请人提供有意义的反馈。另外，两阶段评审还可用于弥补以下评审人的不足：在第一阶段未作出回应或仅提供最低限度回应的评审人、因个人生活中的问题而无法进行评估的评审人，以及由于与分配给自己的提案存在利益冲突（COIs）而拒绝评估提案的评审人。

在两阶段评审中，需要额外审查的提案无法预先得知。在时间限制很严格的情况下，很难确定第二阶段应由哪些评审人负责哪些提案。Jecmen 等人（2021年）提出了一种简单的“随机拆分”策略，以同等概率随机选取一部分评审人分配到第二阶段。给定一个由 m 个提案组成的集合 P，以及一个由 n 名评审人组成的集合 R，一部分提案 $\mathrm{P}_2 \subseteq P$ 需要额外评估。他们在未审查 P_2 的情况下，选取一部分评审人 $\mathrm{R}_2 \subseteq \mathrm{R}$ 负责第二阶段，将评审人 $R_1 = R\backslash R_2$ 留给第一阶段的提案分配。他们建立了两阶段 RAP 模型，以最大化两个阶段中评审人分配的相似度分数总和。他们使用最优分配的平均相似度，以便轻松解读结果。公式 7-17 展示了如何在约束条件（公式 7-18、7-19、7-20）下计算第二阶段最优分配的平均相似度。

$$Q(R_2, P_2) = \frac{1}{d(1)n + d(2)\beta n}\left[max_{x \in M(R\backslash R_2, P; w; d(1)} \sum_{i \epsilon P,\ j \epsilon R\backslash R_2} x_{ij} s_{ij} + max_{y \in M(R_2, P_2; w; d(2)} \sum_{i \epsilon P_2}\sum_{j \epsilon R_2} y_{ij} s_{ij}\right] \quad (7\text{-}17)$$

$$s.t.\ \ \sum_{i \epsilon P} x_{ij} + y_{ij} \leq w \quad \forall j \in R \quad (7\text{-}18)$$

$$w|\mathrm{R}_2| \geq d^{(2)}\beta n \quad (7\text{-}19)$$

$$w(m - |\mathrm{R}_2|) \geq d^{(1)}\beta n \quad (7\text{-}20)$$

其中，x_{ij} 是一个矩阵；若评审人j 在第一阶段被分配给提案i，则$x_{ij} = 1$；否则为$x_{ij} = 0$。y_{ij} 是第二阶段用于相同目的的矩阵。s_{ij} 是评审人与提案之间的相似度分数。$d^{(1)}$ 表示第一阶段的提案需求，$d^{(2)}$ 表示第二阶段的提案需求。w 表示评审人工作量上限。β 是提案比例的固定值，而$\beta \in \{\frac{1}{n}, \ldots, \frac{n}{n}\}$。

Leyton-Brown 等人（2022）提出了一种新型两阶段评审系统，通过将较少的评审人分配给接受概率较低的提案，更有效地利用现有评审人资源。两阶段 RAP 被表述为一个混合整数规划问题，其目标是最大化总体加权匹配分数，同时满足提案和评审人容量约束，并避免利益冲突。此外，该研究还考虑了其他软性期望指标，例如评审人相对于提案的资历、提案评审人之间的共同作者距离、分配给同一提案的评审人之间的地域多样性，以及两周期无冲突分配。为确保所得混合整数规划表述能在可接受的时间和内存范围内求解，他们采用了行列生成，并分别求解每个阶段。

8. 结论与未来研究

同行评审是评估和确定提交给学术会议、期刊或资助机构的提案质量时最广泛采用的程序。同行评审流程包含多个步骤。为提案分配合适的评审人，即审稿人分配问题（RAP），是确保评审流程准确、公正的首要且最关键任务。在常规同行评审流程中，会议

主席、期刊编辑或资助项目管理人员必须审阅所有提案，寻找合适的评审人，并在考虑某些限制条件的情况下，手动将评审人与提案进行匹配。随着提案数量不断增加，逐一为每份提案手动分配评审人的传统方式已不再适用，因为这种方式耗时且容易出错。鉴于人工分配存在明显局限，有必要设计自动审稿人分配机制。因此，自动审稿人分配作为一个研究课题和实际问题，已引起学界的广泛关注。

过去 30 年来，来自不同领域的学者开发了审稿人分配方法，以应对 RAP。本研究对文献中提出的自动化流程进行了系统综述。研究人员根据各种方法的优缺点，对大量方法进行了分类和讨论。正如本文综述所示，将评审人分配给提案是一个已取得诸多进展的领域，近年来有关这一主题的研究数量也有所增加。然而，审稿人分配处理仍有改进空间。本节将未来研究可以关注的问题归纳为以下四个标题。

8.1 审稿人与提案画像的未来方向

本文献综述的目标之一，是考察有关构成审稿人与提案画像的变量及其确定方式的现有研究。尽管已有研究考虑了许多不同的变量来表征评审人和提案，但仍有一些方面值得进一步探讨。

- 大多数 RAP 文献都假设审稿人候选数据库已经存在，并且
  它们并未描述数据库构建流程。此外，还应考察审稿人候选数据库的创建过程。现有数据库中的审稿人也被假定始终有资格评审提案，且有时间参与评审。由于工作负载过重或日程安排紧张，即使审稿人被认为已准备好进行评审，也可能不接受这项任务。此外，即便能够确保潜在评审人的资历，这种方法仍可能导致特定研究领域的专业知识短缺。因此，应尽可能扩大审稿人候选池。人们认为，利用信息检索技术从外部来源招募新审稿人，不仅有助于扩大候选池，也有助于获取有关审稿人的客观且最新的信息。这也值得深入研究。

- 衡量审稿人的历史评审表现，并在审稿人分配任务中将其纳入考量，可以提高分配的准确率和质量。为了在审稿人分配任务中考虑审稿人的历史评审表现，可以开展实证研究，以确定评估审稿人评审能力的度量指标，例如评审准确性、质量等。

- 为了建立提案简介模型，一些研究考察了申请人的若干方面，例如申请人所属机构、专业领域和合著信息。未来的研究应寻找并使用更多申请人特征，例如能够体现申请人专业知识的教育背景、工作经验、研究项目和专利。这类信息有助于进行相关性判定和审稿人关联性分析。

- 建立全面可靠的评审人和提案数据库，对于确保准确的提案—评审人匹配至关重要。在这一领域，使用由人工收集且不再更新的信息构成的静态数据库十分常见。此外，也有动态数据库，例如 OpenReview 和

DBLP，它们包含来自多个来源的结构化或非结构化数据。然而，要提高动态数据库的准确率和完整性，仍需付出更多努力。

8.2 相似度分数计算的未来方向

第二阶段的主要目标是根据候选审稿人和提案的简介，为每一组可能的提案—评审人配对计算相似度分数。在这篇综述论文中，现有的相似度分数计算研究根据所使用的信息类型归纳为三类。尽管许多相似度计算系统依赖人工方法，但研究表明，自然语言处理技术中的最先进工具能够显著实现相似度分数计算的自动化并提升其效果。除了基于文本的方法，许多研究还提出了基于隐式关系的网络模型。对这类关系进行分析，有助于加强审稿人推荐，并避免提案与审稿人之间可能存在的利益冲突。目前的研究已在计算提案与审稿人之间的相似度分数方面进行了许多尝试，但仍需开展更多工作。

- 基于神经网络的语言模型在多种自然语言处理下游任务中表现出色，包括文档分类、命名实体识别和机器翻译（Romanov and Khusainova，2019年；Kalyan 等人，2021年；Kapočiūtė-Dzikienė 等人，2021年）。少数关于将基于神经网络的模型应用于相似度分数计算的研究表明，这类模型比其他基于语义的方法更有效（Ogunleye 等人，2018年；Duan 等，2019年；Yong 等人，2019年）。然而，学界对于基于神经网络的语言模型在 RAP 领域的优势尚未达成共识。因此，还需要更多研究来分析新一代嵌入模型在 RAP 中的应用效果。人们认为，尤其是基于变换器的预训练模型，例如 XLNet（Yang 等人，2019年）、RoBERTa（Liu 等人，2019年）和 T5（Raffel 等人，2020年），具有强大的特征提取能力，有助于提高提案与审稿人之间相似度分数的准确率。

- 预训练语言模型可以基于大量关于提案和审稿人的数据，直接用于任何领域的审稿人分配。然而，此类学习模型所生成结果的质量取决于是否拥有大量领域专用训练数据。因此，研究人员和负责安排同行评审的人员必须合作创建并分发大型训练数据集。

- 一些研究人员提出了一组相似度分数计算方法，其中使用了词汇和语义信息。与仅基于词汇或语义信息计算的相似度分数相比，结果令人满意。未来必须开发更多用于计算相似度分数的混合方法。此外，为了改进相似度分数计算以及审稿人分配的速度和公平性，未来的研究可以探讨整合基于文本、基于网络和基于人工的信息。

- 以往的研究通常提出使用加权连接（例如共同作者和引文网络。然而，其他社交网络，例如学术关系、学生—导师关系、竞争关系和项目合作者，也可能影响评审人的分配。例如，具有相同导师但从未共同发表过论文的两名研究人员之间可能存在利益冲突。申请者和评审人的这些潜在利益冲突由本人声明，并且可以进行主观评估，但它们可能会影响同行评审流程的质量。为了

自动检测此类利益冲突，可以使用监督学习模型来学习用户之间的利益冲突分数。但是，缺乏高质量的真实标注数据集，严重阻碍了监督式利益冲突学习的发展。未来的研究应集中于创建真实标注数据集。

- 为了准确检测利益冲突，需要构建交互式系统，例如 PISTIS（Wu 等人，2018年）。这类系统能够自动从开放来源中提取某些形式的潜在利益冲突，并包含用于主观利益冲突类别的声明模块。

- 如果评审人与提案之间存在利益冲突，人们通常认为，
  无论利益冲突的类型或程度如何，都不应将评审人分配给该提案。然而，这种一刀切的方法可能导致缺乏具备评估资格的评审人，尤其是在合格评审人稀缺的领域。与其采用这种一刀切的方法，不如考虑利益冲突的不同类型和程度，制定灵活的利益冲突规避策略。目前，RAP 文献中的研究主要侧重于识别和规避潜在的利益冲突，但很少关注如何衡量可能导致利益冲突的关系强度。未来的研究可以着重于开发用于衡量关系程度的测量指标。例如，假设提案的评审人与申请人是两个独立的节点。两个节点之间的距离可用于估计利益冲突的程度；也就是说，评审人与申请人之间的关系越密切，双方存在的利益冲突就越多。

- 必须从各种来源获取并整合社交网络，例如被指导者—
  指导者网络以及所属机构/同事网络。因此，必须收集反映不同类型关系的各种数据源，并对由此产生的大型数据集进行分析。

8.3 优化分配的未来方向

本系统综述表明，已有大量研究致力于构建数学模型，以便在不同约束条件下，通过优化不同的目标函数来解决 RAP。此外，针对这一问题，研究人员还提出了许多求解算法。但未来仍有一些关键问题需要考虑。

- 尽管在针对 RAP 提出的数学模型中，最大化总体相似度这一目标函数相当常见，但它可能会对某些提案造成不公平。这种方法可能导致提案被分配给缺乏某些提案所属主题领域专业知识的特定评审人。为确保公平，必须保证每个提案都分配给一组具备最低可接受专业水平的评审人。近期研究（Kobren 等人，2019年；Stelmakh 等人，2019年）试图通过对有效提案—评审人匹配的最低接近度施加严格约束，或优化提案在最坏情况下的接近度总和，来克服这些问题。然而，由于约束条件的复杂性，很难找到理想的解。未来研究应着重提高所提方法的计算效率。

- 重新优化策略也值得关注。如果发现评审人与提案之间存在未知的利益冲突，或者评审人严重不具备评审资格，就必须确定新的评审人，并实施新的决策程序。在时间受约束的语境下，开发重新优化方法，而不是重复整个流程，将更为有利。

- 据信，接收评审人的反馈有助于提高审稿人的满意度，并实现更准确的评审人与提案匹配。另一个未来研究方向是如何动态优化分配，包括审稿人对分配结果的反馈。

8.4 其他问题的未来方向

- 本调查显示，大多数 RAP 研究聚焦于一个或两个阶段，而用于解决不同阶段问题的技术很少得到协调。不过，各阶段之间存在内在关系。例如，阶段（3）中算法的分配结果，很大程度上取决于阶段（2）中相似度评分计算的准确性。未来，研究人员应开发能够兼顾全部三个阶段的新方法，而不是只尝试在每个阶段取得最佳结果。

- 相似度测量与评审人分配问题能否通过协作方式得到解决，将会很值得关注；这或许可以纳入以往会议或期刊的反馈，或资助项目先前轮次的反馈。

- 首先，应努力为各个领域制定标准化评审人分配标准和流程，以确保分配的准确性和公平性。

- 在大多数研究中，研究人员使用通过随机分配或人工分配获得的真实数据来评估和测试所提方法。这些数据集包含敏感信息，因此无法公开获取。因此，无法进行直接比较。缺乏 RAP 的标准基准，妨碍了对各种方法性能的比较，进而阻碍了研究的发展。因此，应建立一套经过精心设计的基准数据集。

本篇文献综述为读者全面概述了 RAP 及其前沿的 RAP 相关策略。此外，本文为 RAP 的各个阶段以及针对 RAP 的最新研究分类提供了基础。本文系统地考察了该领域近期的研究进展，并指出了有待进一步研究的领域。基于本文，学者可以考虑文献中的空白，并确定未来研究的潜在主题。此外，资助机构、科学期刊和会议可以更好地了解如何更高效地管理审稿人分配流程。我们最大的愿望是，本研究能够成为考虑开展 RAP 研究的研究人员的入门文档。

附录 A：可用数据集

作者	案例实施	可用数据集	数据集内容
Karimzadehgan 等人（2008年）	会议审稿人分配：ACM, SIGIR	http://sifaka.cs.uiuc.edu/ir/data/review.html	73篇论文，189位审稿人
Karimzadehgan & Zhai（2009）	会议审稿人分配：SIGIR	http://timan.cs.uiuc.edu/data/review.html	73篇论文，30位审稿人
Tang 等人（2010年）	会议审稿人分配：KDD'08, '09 和 ICDM'09	http://www.arnetminer.org/expertisematching	338篇论文，354名审稿人
Daud 等人（2010年）	期刊审稿人分配：DBLP	http://www.informatik.uni-trier.de/~ley/db/	90124 篇论文，112317 作者
Tang 等人（2012年）	会议审稿人分配：	http://www.arnetminer.org/expertisematching	338篇论文，354名审稿人

作者	案例实施	可用数据集	数据集内容
	KDD'08 ,'09 和 ICDM'09		
Karimzadehgan & Zhai (2012)	会议审稿人分配：ACM	http://timan.cs.uiuc.edu/data/review.html http://www.acm.org/dl.	73篇论文，189位审稿人
Kou 等人 (2015a, 2015b)	会议审稿人分配：SIGKDD'08 , '09, ICDM'09 SDM'08 , '09 CIKM'08 ,'09	http://degroup.cis.umac.mo/reviewerassignment/	第一数据集：545 篇论文，203 名评审人；第二数据集：648 篇论文，154 名审稿人；第三数据集：617 篇论文，105 名评审人；第四数据集：513 篇论文，90 名评审人
Nguyen 等人 (2018)	会议审稿人分配：CCIA'14, 2015年，2016年	aminer.org、3tdx.cat、researchgate.net, dblp.uni-trier.de	106 篇论文，96 位审稿人
Duan 等人 (2019年)	期刊审稿人审稿人分配：Arxiv	http://sifaka.cs.uiuc.edu/ir/data/review.html https://www.aminer.cn/expertisematching ftp://3lib.org//oai_dc/arxiv	685 篇论文，1885 审稿人
Mirzaei 等人（2019 年）	会议审稿人分配：SIGIR, PubMed	http://timan.cs.uiuc.edu/data/review.html, http://isl.ce.sharif.edu/pubmed/dataset/, http://www.ncbi.nlm.nih.gov/pubmed	SIGIR- 73 篇论文，189 审稿人 PubMed- 231 篇论文
Kobren 等人（2019 年）	会议审稿人分配：MIDL, CVPR	https://github.com/iesl/fair-matching	MIDL：118 篇论文，117 位审稿人 CVPR：2623 篇论文，1737 位审稿人
Kalmukov（2020 年）	会议审稿人分配：CompSysTech	http://www.compsystech.org/	1484篇论文
Jecmen 等人（2020）	会议审稿人分配：ICLR'18	https://github.com/theryanl/mitigating_manipulation_via_randomized_reviewer_assignment	911篇论文，2435 审稿人
Kreutz & Schenkel（2020年）	会议审稿人分配：MOL'17, BTW'17，ECIR'17	https://doi.org/10.5281/zenodo.4071874	–
Pradhan 等人（2020年）	会议审稿人分配：ICBIM'16	https://easychair.org/my/conference.cgi?welcome=1;conf=icbim2016	59 篇论文，40 位审稿人
Xu 等人（2020）	会议审稿人分配：ICLR'18	https://github.com/xycforgithub/StrategyProof_Conference_Review	911 篇论文，2435 审稿人
Tan 等人（2021 年）	期刊审稿人分配：arXiv, GitHub	https://github.com/aitsc/WSIM/tree/main/datasets	100 篇论文，400 位审稿人
Hoang 等人（2021 年）	期刊审稿人分配：DBLP	https://dblp.org/xml/, https://github.com/Lucaweihs/impact-prediction#getting-data	479篇论文，1000名审稿人
Patil（2021年）	会议审稿人分配：NIPS'19	https://www.kaggle.com/abolihpatil/nips2019-paper-for-reviewer-assignment-ahp-pnm	150篇论文

附录 B：基于文本/基于内容的方法

作者	**Case** 实施 **on**	评审人简介	提案简介	方法	性能指标已使用	比较
Dumais & Nielsen（1992 年）	会议 评审人分配：Hypertext'91	摘要提供方审稿人	标题, 提案的摘要 提案	LSI	平均值相关性, 精确率	与 reviewers' rel_{evance} 判断

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 指标所用	比较
Yarowsky & Florian (1990)	会议 评审人 分配: ACL'99	全文, 摘要、标题, 关键词 $revi_{ewers}$ ' 先前 发表 论文, 书目信息 的 审稿人（共同 作者身份, 引文）	全文, 摘要, 标题, 关键词 of 提案; 参考文献 hic 信息 与 作者	向量空间模型、朴素贝叶斯 分类器	准确率, 平均值 评审角色，One-of-best-2	1. 推荐模型 与人工进行比较 分配, 2. 比较 采用……进行分配 文本、书目、文本 + 书目信息, 3. 完整……的案例 审稿人的 出版物文本与 相同权重、不同 权重以及人工 分配比较, 4. 仅使用分配 关键词与分配 使用关键词、标题, 摘要比较
Basu 等人（2001 年）	会议 评审人 分配: AAAI-98	全文 of $revi_{ewers}$ ' 此前 已发表 论文, 审稿人 姓名, 所属机构, 偏好	摘要 of 提案	词频-逆文档频率, KNN, 扩展 直接 贝叶斯	召回率	1. 推荐 算法比较 彼此之间 2. 与……比较 随机分配和 协同过滤
Hettich 与 Pazzani （2006 年）	项目 提案 评审人 分配: NSF	全文 r评审人rs' 此前 已发表 论文	全文 of 提案	词频-逆文档频率, Hill 爬升 算法	总和 残差项 权重	比较 备选审稿人集合 与提案集
Ferilli 等人（2006年）	会议 评审人 分配: IEA/AIE	摘要和 标题的 $revi_{ewers}$ ' 先前 发表的 论文	标题, 摘要 提案	LSI	准确率	NA
Biswas & Hasan (2007年)	期刊 评审人 分配: 模拟的 data	全文 $revi_{ewers}$ ' 先前 发表的 论文, 领域 本体的 审稿人	全文 of 提案	向量空间模型、KEA	相关性	自由比较 文本，KEA，本体-驱动的主题推断
Zhao 等人 (2018)	期刊 评审人 分配: 模拟 data	研究 兴趣 的审稿人	关键词 of 提案	Word Mover's 距离, 构造性 覆盖 算法, LDA–KNN, LDA–CCA, LSI–KNN, LSI–CCA,	精确率, 准确率, 召回率	1. 合成数据集与 公开数据集的 比较 2. 潜在狄利克雷分配–K近邻、LDA–CCA、LSI–K近邻、LSI–CCA、词移距离–K近邻、LDA–GaussianNB, 线性判别分析–支持向量机、LSI–高斯朴素贝叶斯

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 指标所用	比较
				WMD–KNN, LDA–高斯朴素贝叶斯，潜在狄利克雷分配–支持向量机，LSI–高斯朴素贝叶斯		
Mimno 和 Mccallum（2007年）	会议 评审人 分配: NIPS'06	全文的 $revi_{ewers}$ ' 先前 发表的 论文	全文 of 提案	APT	准确率, 精确率	1. 推荐模型 与人工 分配, 2. 比较不同 狄利克雷平滑 语言模型、APT 和 ATM
Karimzad ehgan 等人（2008年）	会议 评审人 分配: ACM, SIGIR	摘要 $revi_{ewers}$ ' 此前 已发表 论文	摘要, 全文的 提案	概率潜在语义分析，k-均值	置信度, 覆盖率, 平均值 置信度, F分数	1. 评审人方面 建模（最大 边际相关性与概率潜在语义分析）比较，与 论文方面建模 （基于互信息的 聚类与简单 话语分析）, 2. 摘要的比较 与全文的使用
Tang 等人（2008年）	会议 评审人 分配: Open .....的数据库 出版物	.....的全文 $revi_{ewers}$ ' 先前 发表的 论文, 书目 信息 审稿人（共同 作者身份, 引文）	NA	ACT	平均精度均值 精确率	采用三种策略的 ACR 与 SLM 进行比较, 潜在狄利克雷分配、ATM
Conry 等人 等人（2009年）	会议 评审人 分配: ICDM'07	合著 的 出版物数量, 研究领域, 偏好 审稿人，以及 .....的摘要 审稿人的 先前 发表的 论文	摘要 and 研究 领域 提案	VSM	评估 评分	比较 所提出的模型与 Taylor 的（2008年）模型
Tang 等人（2010年）	会议 评审人 分配: KDD'08 ，'09 和 ICDM'09	Reviewers' 此前 已发表 论文，共 同作者, 所属机构	NA	语言 模型, LDA	精确率	所提出的算法与 贪心算法的 比较
Daud 等人（2010年）	期刊 评审人 分配: DBLP	$Revi_{ewers}$ ' 先前 发表 论文	NA	时间 专家 主题 模型	主题熵 主题	所提出的 时间- 专家-主题方法的比较 与两个基线

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 所用指标	比较
						方法，时间序列— 作者-主题和作者- 会议-主题
Andrade-Navarro 等人 等人（2012 年）	期刊 评审人 分配: MEDLINE	摘要和 参考文献 reviewers ’ 先前 发表 论文，合著 作者身份	标题和 摘要的 提案	模糊 二元 关系	平均分 相似度，的 平均值 论文数（每位 评审人 简介）	1. Peer2ref 比较 与人工评审人 筛选结果 2. Peer2ref、eTBlast 和 Jane 网络工具比较
Mai 等人 （2012 年）	项目 提案 评审人 分配: NSFC	研究领域 代码, 评估 经验, 项目 申请 经验	标题, 摘要, and 关键词 of 提案	概率 语言 模型	竞价	两个数据集的比较 提案集合：热门和 冷门主题
Xue等人 （2012 年）	会议 评审人 分配: KDD’08 ，‘09 和 ICDM’09	全文 $revi_{ewers}$ ’ 先前 发表的 论文, 审稿人 姓名, 所属机构	学科 本体, 关键词 of 提案	区间 模糊 本体, 语义 匹配	匹配 度，负载 方差	主题数量比较
Charlin 等人（2012年）	会议 评审人 分配: NIPS’09 ，10	关键词 提供者 审稿人, 引文，以及 经验 审稿人	关键词 and 引文 of 提案	语言 模型, 线性 回归, 贝叶斯 概率 矩阵 分解	RMSE	比较 语言模型、线性 回归, 贝叶斯概率 矩阵分解
Li & Watanabe （2013 年）	研究 项目 提案 评审人 分配: NSFC	出版物数量 出版物, 质量和 时效性 出版物, 期刊 排名, 参考文献的 审稿人	参考文献 s of 提案	APT	NA	NA
Long 等 等人（2013 年）	会议 评审人 分配: KDD’06 ,10	NA	NA	词频-逆文 档频率	分配 公平性	按不同类型进行分配比较 不同类型的冲突 利益冲突
Charlin & Zemel (2013年)	会议 评审人 分配: NIPS’10 ， ICML’12	全文 $revi_{ewers}$ ’ 先前 发表 论文, 专业知识 评分 由 审稿人	全文 of 提案	语言 模型, LDA, 线性 回归, 贝叶斯 概率 矩阵 分解	相关性 得分、均方根误差	1. 语言模型与 潜在狄利克雷分配比较, 2. 线性回归与..... 贝叶斯概率 矩阵分解 比较

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 指标所用	比较
Caldera 等 等人（2014 年）	会议 评审人 分配: Eurographics' 14	全文 r审稿人 ' 先前 发表的 论文	全文 of 提案	词频-逆文 档频率	偏好 审稿人	NA
Liu 等人 （2014 年）	会议 评审人 分配: CIKM'08, NIPS'06	专业知识 主题，共 评审人的作者身份 审稿人, 评审人的全文 评审人的 先前 发表 论文	全文, co- 作者身份 关系 p of 提案	LDA	精确率, 匹配 度, 权威性, 多样性	1. NIPS 数据集的比较 与 CIKM 数据集 2. RWR、 主题相似度、文本 相似度，以及带有 稀疏性约束的 RWR
Protasiewi cz（2014年）	期刊 评审人 分配: Open 数据库 出版物	全文, 摘要、标题, ...的关键词 revi$_{ewers}$ ' 先前 发表的 论文；姓名- 姓氏, 所属机构, 研究领域 共同署名 的评审人	全文 of 提案	多项式 朴素 贝叶斯, 支持向量机，快速 自动 关键词 提取 算法	NA	NA
Kou 等人 (2015a, 2015b)	会议 评审人 分配: SIGKDD'08 ， '09, ICDM'09 SDM'08 ，'09 CIKM'08 ,'09	摘要 r审稿人 ' 先前 发表 论文	摘要 of 提案	ATM	NA	NA
Protasiewi cz 等人 (2016年)	项目 提案 评审人 分配: 国家 中心 研究与 发展 波兰的	全文, 摘要、标题、 关键词 revi$_{ewers}$ ' 先前 发表 论文；姓名- 姓氏, 所属机构, 专业领域, 共同署名 审稿人	标题, 摘要, 以及全文的 文本 提案	朴素 贝叶斯, 层次 凝聚 ve 聚类,	精确率, 准确率，F- 分数	1. 波兰语关键词 提取、关键短语 关键词短语 快速提取算法 自动关键词 提取 2. 最大概率 与.....进行模型比较 多语言分类
Shon等人 （2016 年）	项目 提案 评审人 分配: 韩国国家科学 基金会	摘要、标题, 和关键词 of reviewe rs' 此前 已发表 论文, 审稿人所属机构 审稿人	标题, 摘要, and 关键词 of 提案; 关键词 提供 by 申请人	词频-逆文 档频率	执行时间	NA

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 指标所用	比较
Yin 等人 （2016 年）	期刊 评审人 分配: DBLP	全文 reviewers' 先前 已发表 论文，h- 索引 审稿人	全文 of 提案	主题 模型	多样性，F- 分数, 覆盖率, 权威性	比较 所提出的模型与模型 未考虑 ..... 审稿人因素重要性 与评审人 方面模型
Jin 等人 （2017 年）	期刊 评审人 分配: 万方, ArnetMiner	NA	全文 of 提案	ATM、EM	相关性, 兴趣趋势, 权威性	1. ATM 的比较, SLM 与向量空间模型, 2. 两个 数据集：万方和 ArnetMiner
Peng 等人 （2017 年）	期刊 评审人 分配: 《计算机科学前沿》 计算机 《科学》 期刊	全文, 摘要、标题, 和时间 信息 reviewers' 先前 发表的 论文	标题, 摘要, 以及完整的 文本 提案	潜在狄利克雷分配、TF- IDF	匹配 度	比较 所提方法与 人工分配
Ogunleye 等人 (2017年)	会议 评审人 分配: NIPS	全文的 $revi_{ewers}$ ' 先前 发表 论文	全文 of 提案	Word2Vec, 词频-逆文档频率, LSI，潜在狄 利克雷分配	适用性 评分	.....之间的比较 Word2Vec、词频-逆文档频率、LSI, LDA
Yan 等人 （2017 年）	会议 评审人 分配: ICDM'15, SIGKDD'13,14,15	全文的 $revi_{ewers}$ ' 先前 发表 论文	摘要 of 作者的 先前 发表的 论文	LDA	平均值 匹配 适应度, 平均值 避免 效应	比较 不同耦合 参数值
Jin 等人 (2018a)	期刊 评审人 分配: 万方, ArnetMiner	摘要 $revi_{ewers}$ ' 此前 已发表 论文，共同 作者身份 审稿人	NA	AST	困惑度, Kullback– 莱布勒 散度, 覆盖率	两种不同类型 AST 的比较 不同类型的 AST 和 ATM
Li & Hou (2018)	会议 评审人 分配: IEEE INFOCOM 2015 ,16	全文 $revi_{ewers}$ ' 先前 发表的 论文	全文 of 提案	LSI	NA	NA
Nguyen 等人 等人 (2018)	会议 评审人 分配: CCIA'14 ，'15, '16	研究 主题, 兴趣, 时效性，以及 评审人质量 审稿人	全文 of 提案	LDA	质量指标, 总体 匹配	比较 所提方法和 随机分配
Anjum 等人 （2019年）	会议 评审人 分配: NIPS，第一等级	摘要 reviewers' 先前 发表的 论文	摘要 of 提案	Word2Vec, 常见 主题 模型	精确率	APT 比较, 单文档模型, 层次狄利克雷 过程，带重启的随机游走 ，潜在狄利克雷分配，词

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 指标	比较
						Mover距离 ，隐 主题模型、Doc2Vec, 以及所提出的模型
Medakene 等人 （2019 年）	会议 评审人 分配: ICA2IT'19	参考文献 revi$_{ewers}$ ， 先前 发表 论文, 教学方面的 .....的方面 审稿人, 研究领域	全文 and 参考文献 of 提案	LDA	置信度 度，主题 相似度, 相关性 得分，均方误差, 均方根误差，平均绝对误差, MAPE	有无.....情况下的比较 有和没有的情况 对.....的考量 教学简介
Duan 等人 （2019年）	期刊 评审人 分配: 合成与 arXiv	标题与 摘要 revi$_{ewers}$ ， 先前 发表 论文	提案标题和 摘要 提案	词频-逆文档频率, BERT, CNN, 双向长短期 记忆网络	精确率, 召回率, 宏平均 F1 得分，平均值 平均值, 精确率, 规范化 折扣 累积 收益，二元 偏好	1. SPM-比较 RA、潜在狄利克雷分配、SLM、LDA-LM、TATB、关键词 余弦相似度、BBA, 词移距离算法 2. 比较 合成数据集与 公开数据集 3. 六种神经网络结构之间的比较 神经网络结构 对于句对 模型：Word2Vec-bilSTM、Word2Vec-CNN、Word2Vec-bilSTM-CNN、bilSTM、卷积神经网络、bilSTM-CNN
Mirzaei 等 等人（2019 年）	会议 评审人 分配: AGM SIGIR, PubMed	标题和 摘要的 revi$_{ewers}$ ， 先前 发表的 论文	NA	概率潜在语义分析、潜在语义分析	覆盖率, 平均值 置信度	整数线性规划与 概率潜在语义分析、CFLA-PLSA, WCGRA-PLSA, CMARTA-LRA-ILP, CMARTA-LRA
Mittal 等人 等人（2019 年）	会议 评审人 分配: SemEVAL'10	全文 revi$_{ewers}$ ， 此前 已发表 论文	全文 of 提案	WordNet, 模糊 扩展 原则 模糊 图 连通性 度量	匹配 度	模糊比较 模糊度中心性，模糊 介数中心性, 模糊接近中心性, PageRank，HIT
Chughtai 等人 （2019 年）	期刊 评审人 分配	全文 r审稿人 ， 先前 发表 论文	标题, 摘要, 关键词 of 提案	LSA, 词频-逆文档频率, 本体-基于 知识-领域	熵, 相似度 指数	NA
Zhang等 等人（2019 年）	会议 评审人 分配: ACM	标题和 摘要的 revi$_{ewers}$ ， 先前 发表的 论文	标题和 提案的摘要 提案	Hiepar	召回率, 准确率, 规范化 折扣后 累积 gain	1. 比较 所提出的模型、潜在狄利克雷分配 Word2Vec、卷积神经网络、BERT 双向长短期记忆网络和长短期记忆网络 2. Hiepar 的比较 与词频-逆文档频率、LSI、潜在狄利克雷分配、词移距离、BBA

作者	**Case** 实施 **on**	评审人 简介	提案 简介	方法	性能 指标使用	比较
Kalmukov （2020 年）	会议 评审人 分配: CmpSysTech	摘要 reviewers' 先前 发表的 论文	摘要 of 提案	向量空间模型，TF-IDF	偏好 审稿人	不同模型的比较 TF-IDF加权模型
Yang 等人 等人（2020 年）	会议 评审人 分配: WWW, SIGIR、CIKM	标题, 关键词和 摘要 reviewers' 先前 发表的 论文	摘要 and 关键词 of 提案	LDA	计算 时间, 目标 函数 值, 匹配 度	NA
Pradhan 等人 等人（2020 年）	会议 评审人 分配: CBIM'16	专业知识 领域、共同 作者身份 评审人, 全文 审稿人的 先前 发表的 论文	标题 提案	LDA	质量	比较 所提方法与 EasyChair 和 TPMS 系统
Tan 等人 （2021 年）	期刊 评审人 分配: arXiv，GitHub	标题和 摘要 $revi_{ewers}$ ' 先前 发表的 论文	提案标题和 摘要 提案	语言 模型, 潜在狄利克雷分配、词, and 基于语义的 基础 迭代式 模型 （WSIM）	准确率, 精确率, 召回率, Macro-F1 分数，均值 平均值 精确率, 二元 偏好, 规范化 折扣 累积 gain	潜在狄利克雷分配的比较, SLM、LDA-SLM, Word2Vec、LDA-分支 定界、时间感知, 和基于主题的关键词 余弦相似度
Hoang 等人 等人（2021 年）	期刊 评审人 分配: DBLP	全文的 出版物, 数量 出版物, 数量 引文，h- 评审人指数 审稿人	全文 of 提案	LDA, Doc2Vec	规范化 折扣后的 累积的 收益, 折现 累积 gain	Protasiewicz 的研究 （2014年）、Hoang 等人 （2019年）以及提出的 模型比较
Patil & Mahalle （2021 年）	会议 评审人 分配: Interspeech, NIPS'19, AAAI'14	姓名, 所属机构, 出版物, 以及网络 信息 审稿人	标题, 摘要, 关键词 of 提案	LDA	NA	NA

附录 C：基于优化的方法

作者	**Case** 实现	数学 模型	算法	性能 度量指标使用情况	比较
Hartvigsen 等人 1999年	会议 评审人 分配	最大 权重- 容量限制 运输 问题	NA	计算 time	NA
Guerv´os & Valdivieso (2004年)	会议 评审人 分配: PPSN2002	GAP（整数线性规划）	贪心算法 算法 & 进化 算法	NA	NA
Cook 等（2005年）	项目提案 评审人 分配: 模拟数据	集合覆盖 整数 规划	贪心 算法	计算性 时间，解 质量	1. 4项比较 不同规模的样本数据集 , 2. 集合 覆盖整数 规划和贪心 算法
Hettich 与 Pazzani（2006）	项目提案 评审人 分配：美国国家科学基金会	NA	爬山法 算法	总和 残差项 权重	比较 备选数据集 审稿人和替代 方法
Janak 等人（2005年）	研究项目 提案评审人 分配：美国国家科学基金会	GAP（整数线性规划/混合整数规划）不适用（使用	CPLEX）	目标 函数值	1. 比较 推荐模型和 人工分配, 2. 四个数据集的比较 不同规模的样本数据集 不同规模
Sun 等人 2007年	研发项目 提案评审人 分配：国家自然科学基金委员会	GAP（整数线性规划）	单纯形法 算法	专业水平, 评审人工作量 审稿人	NA
Li 等人（2007年）	项目提案 评审人 分配: 天津NSF	GAP（整数线性规划）	遗传 算法	平均适应度值，最佳 适应度值	标准 遗传算法, 基于遗传算法的 关于.....机制 正反馈与 自适应变异通过 机制
Goldsmith & Sloan（2007年）	会议 评审人 分配	网络流	NA	NA	NA
Li 等（2008年）	项目提案 评审人 分配: 天津NSF	复杂的 离散 优化 问题	蚁群 优化	平均适应度值，最佳 适应度值	比较蚂 蚁优化 基于遗传算法的 基于.....机制 正反馈与 通过自适应变异 信息素
Karimzadehgan & Zhai（2009年）	期刊审稿人 分配: 模拟数据	GAP（整数线性规划）	整数线性规划算法, 贪心 算法	平均值 置信度	整数线性规划 算法与贪心 算法
Conry 等人（2009年）	会议 评审人 分配: ICDM'07	GAP（整数线性规划）	NA	评估 评分	比较 所提出的模型与 Taylor 的（2008年）模型

作者	**Case** 实现	数学 模型	算法	性能 使用的度量指标	比较
Xu等人（2010年）	研究项目 提案评审人 分配：国家自然科学基金委员会	GAP（整数线性规划）	贪心 随机化 自适应搜索过程， 遗传 算法	平均值 专业知识， 评审人工作量 审稿人	NA
Huang 等人（2010年）	期刊审稿人 分配： 模拟数据	最小成本 流量 问题	NA	近似 稳定性， 满意度	NA
Tang 等人（2010年）	会议 评审人 分配： KDD'08 ，KDD'09 和 ICDM'09	最小值 凸成本 流问题	所提出的 算法	精确率	比较 所提出的模型与 贪心算法
Tang 等人（2012年）	会议 评审人 分配： KDD'08 ，09 和 ICDM'09	最小值 凸成本 流问题	所提出的 算法	匹配 度，负载 方差， 专业知识 专业能力方差，CPU time	权威平衡比较 基础网络 流，贪心基线
Chen 等人（2012年）	项目提案 评审人 分配： 模拟数据	GAP（整数线性规划）	遗传 算法	评审人的偏好 审稿人， 时间的 评审人的流动 审稿人更换 会议， 计算 time	小样本 数据集与大型 示例数据集
Karimzadehgan & Zhai (2012)	会议 评审人 分配：ACM 数字图书馆	GAP（整数线性规划）	贪心 算法	覆盖率， 置信度， 平均值 置信度	贪心算法的比较 算法与分支定界 算法
Xue等人(2012年)	会议 评审人 分配： KDD'08 ，09，ICDM'09	GAP	匈牙利 算法	NA	NA
Charlin 等人（2012年）	会议 评审人 分配： NIPS'08 ，10	GAP（整数线性规划）	整数线性规划算法	NA	NA
Long等人（2013年）	会议 评审人 分配： KDD'06 ，10	MaxTC-PRA	贪心 用于 MaxTC-PRA	分配 公平性	分配比较 由不同类型的 利益冲突
Tayal等人（2014年）	研发项目 提案评审人 分配： 模拟数据	成对- 受约束 问题	随机化 算法， 标准 分配 算法	匹配 度	比较 所提方法和 reviewers' preferences
Wang 等人（2013年）	期刊审稿人 分配： 商学院中的 中国	GAP(MIP)	两阶段 随机- 有偏贪心 算法	总成本， 计算 time	CPLEX 比较 与两阶段 随机偏置贪心 算法

作者	**Case** 实现	数学 模型	算法	性能 度量指标使用情况	比较
Li & Watanabe（2013年）	研究项目 提案评审人 分配：国家自然科学基金委员会	GAP（整数线性规划）	匈牙利 算法	Time 消耗， 指派 在 约束条件	比较 所提出的模型与 匈牙利算法
Li 等人（2013年）	研究项目 提案评审人 分配：A 大学	GAP（整数线性规划）	自适应 并行遗传 算法	平均值 计算 时间，最佳 最优 自适应值， 平均值 最优 自适应值	自适应比较 并行遗传 算法，使用两个 处理器核心，自适应并行 遗传算法，使用 四个处理器核心、一种随机 搜索算法，以及一种 遗传算法
Charlin & Zemel（2013年）	会议 评审人 分配： NIPS'10 ，ICML'12	GAP（整数线性规划）	NA	NA	NA
Cechlarova 等人等人（2014年）	科学项目 提案评审人 分配： 斯洛伐克研究与发展 局	网络流	NA	总 工作负载 评审人	比较： 所提方法与 微调模型
Daş 与 Gökçen (2014年)	研发项目 提案评审人 分配：资助机构	模糊线性规划	NA	匹配 度	各种可行性程度的比较 可行性程度
Silva 等人（2014年）	研发项目 提案评审人 分配：国家自然科学基金委员会	GAP（整数线性规划）	匈牙利 算法	准确率， 召回率， 精确率，F-分数	比较 向量空间模型、ATM、LSI、无尺度 网络和整数 规划（使用 TF-IDF)
Kolasa & Krol（2011年）	会议 评审人 分配： KNS'09 ，KES'09	GAP	蚁群 优化与 遗传 算法， 遗传 算法与 禁忌搜索 算法	计算 时间， 资格标准 ，质量 解的	1. 比较遗传 算法、禁忌搜索， 蚁群优化， 蚁群优化 和遗传算法，以及 遗传算法和禁忌 搜索 2.KNS、KES 和 合成数据集 比较
Kou 等人（2015a、2015b）	会议 评审人 分配： SIGKDD'08 ，09， ICDM'08 ，09， SDM'08 ，09， CIKM'08 ，'09	WGRAP	阶段加深 贪心 算法	响应时间， 最优比 ， 优越比 比率	1. 分支定界， 暴力搜索和 整数线性规划比较， 2. 阶段比较 加深贪心 算法，贪心 算法，阶段加深 贪心算法与 随机改进， 最佳审稿人组 贪心、整数线性规划和稳定 匹配

作者	**Case** 实现	数学 模型	算法	性能 度量指标使用	比较
Liu 等人 （2016 年）	研发项目 提案评审人 分配：国家自然科学基金委员会	容量约束 运输 问题	单纯形法 算法	平衡性和 合理性， 分配 时间， 初始化 时间、效应 避免 利益冲突	人工分配的比较 分配、决策 支持系统以及 基于互联网的科学 信息系统
Yue 等人 （2017 年）	科学项目 提案评审人 分配：国家自然科学基金委员会	NP-C 问题	随机化 算法	适应度， 研究 强度， 学术 协会	NA
Ogunleye 等人 （2017 年）	会议 评审人 分配：NIPS	GAP（整数线性规划）	NA	NA	NA
Jin 等人（2017年）	期刊审稿人 分配： 万方， ArnetMiner	GAP（整数线性规划）	整数线性规划算法	NA	NA
Jin 等人 (2018b)	期刊审稿人 分配： 万方， ArnetMiner	GAP（整数线性规划）	整数线性规划算法	NA	NA
Li & Hou (2018)	会议 评审人 分配：IEEE INFOCOM'15,16	GAP（整数线性规划）	整数线性规划 求解器	NA	NA
Lian 等人 (2018)	会议 评审人 分配： PrefLib	GAP(MIP)	网络流 算法	平均效用， 胡佛指数， 基尼指数	1. 比较 调和型有序加权平均、线性有序加权平均和几何型 OWA 2. 比较 递减有序加权平均向量和递增有序加权平均 向量 3.比较 功利主义和平等主义
Guo 等人 （2018）	会议 评审人 分配： SIGIR'07, NIPS'06	无K环 分配	无K环 分配 算法， 整数 规划， 贪心 算法	总 适用性 得分	1. 两个数据集的比较： NIPS'06 和 SIGIR'07 2. 不同循环长度 比较 3. 比较 带有 k-循环 自由约束的分配与 忽略 k- 的分配 无环约束
Yeşilçimen & Yıldırım (2019年)	研发项目 提案评审人 分配： 随机样本	GAP(MIP)	OBH算法	计算 时间，上 界	1. 比较 Cook 等人（2005年）的数学模型 Cook 等人（2005年）以及 文中提出的模型 本研究 2. 线性规划的比较 与半

作者	**Case** 实现	数学 模型	算法	性能 度量指标的使用	比较
					确定性规划 松弛
Mirzaei 等人 (2019年)	会议 评审人 分配：AGM SIGIR、PubMed	不受约束 多方面 审稿团队 分配 基于潜在 研究 领域 （UMARTA- 潜在研究领域）， 受约束的 多方面 审稿团队 分配 使用潜在 研究 领域 （CMARTA- LRA)	阶段加深 贪心 算法, 贪心 正向- 选择，阶段 加深 贪心 算法- 随机 改进	覆盖率, 平均值 置信度	1. 比较 UMARTA-LRA 与 CFLA 2. 比较 CMARTA-LRA、CFLA, 和整数线性规划 3. 比较 UMARTA-KL 和 UMARTA-LRA 4. 整数线性规划- 概率潜在语义分析、CFLA-PLSA、 WCGRA-PLSA、 CMARTA-LRA-ILP、 CMARTA-LRA
Stelmakh 等人 （2019 年）	会议 评审人 分配: CVPR'17、18, MIDL	GAP（整数线 性规划）	PeerReview4All 算法	公平性，相似度之和 , 相似度和的均值 相似度, 计算 Time	1. 比较 PeerReview4All 与 TPMS（Charlin 和 Zemel，2013年），ILPR （Garg 等人，2010年），硬 算法 2. 比较 PeerReview4All 与 TPMS（Charlin 和 Zemel，2013年），FAIRIR （Kobren 等人，2019年）， FAIRFLOW（Kobren 等人 等人，2019年）
Kobren 等人 （2019 年）	会议 评审人 分配: MIDL、CVPR'17	GAP（整数线 性规划）	FAIRIR（公平 通过匹配 迭代 松弛） FAIRFLOW （最小成本流 启发式方法 用于构建 公平匹配）	计算 时间，论文 得分，数量 论文数量 分配给一位 评审人, 目标 函数得分,	TPMS 的比较, FAIRIR、FAIRFLOW, 和 PR4A 算法
Yang 等人 （2020 年）	会议 评审人 分配: WWW、SIGIR、 和 CIKM	GAP（整数线 性规划）	MBBPSO	计算 时间、适应度 函数值, 匹配 度	粒子群优化算法、 二进制粒子群优化算法、MBBPSO 和 遗传算法
Jecmen 等人 （2020 年）	会议 评审人 分配: ICLR'18	成对- 受约束 问题	随机化 算法	相似度之和, 分配 概率	1. 成对约束 成对受约束 与分区约束 问题 2. 随机化算法 与标准分配 算法比较

作者	**Case** 实现	数学 模型	算法	性能 所用度量指标	比较
					3. 四个 样本数据集 大小不同
Xu 等人（2020年）会议	评审人 分配： ICLR'17, 18	防策略性- 基于分区 分配	分而治之的 Rank 分配 算法	NA	NA
Pradhan 等人 （2020 年）	会议 评审人 分配： ICBIM'16	均衡 多任务 分配	加权- 矩阵 分解- 基于贪心 算法	质量	比较 EasyChair、TPMS 和 所提方法
Kat（2021年）	研究项目 提案评审人 分配： TÜBİTAK	GAP（整数线 性规划）	NA	得分总和， 评分均值 水平，比率 的标准 偏差为 评分 百分比至 均值	比较 所提出的模型与两个 整数规划 具有不同 目标函数
Boehmer 等人 （2021 年）	会议 评审人 分配： ICLR'18	加权 z- 无环同行评审 同行评审	贪心 算法	代理比例 代理人	整数线性规划与 拟议启发式方法的比较
Jecmen 等人 （2021 年）	会议 评审人 分配： ICLR'18, SIGIR'07, PrefLib3，Bid1	两阶段 论文 分配 问题	随机划分	预言机最优 解的 相似度， 平均值 分配 值	1. 比较 不同的数据集 2. 比较 不同的分配负荷 参数
Payan & Zick （2022 年）	会议 评审人 分配： MIDL、CVPR， CVPR'18	公平分配 问题	贪心 评审人轮 询算法	功利主义 社会福利， 纳什社会 福利， EF1违反的数量 ， 最小值 论文得分， 运行时间，嫉妒总和 的 所有论文 对的基尼 系数 所有论文得分的 得分	GRRR 的比较， FairIR 和 FairFlow 算法（Kobren 等人 等人，2019年），TMPS （Charlin 和 Zemel， 2013年），以及 PeerReview4All （Stelmakh 等人，2019年）
Dhull 等人 （2022 年）	会议 评审人 分配： ICLR'18	策略真实 通过分区 分配	随机 划分， 着色 算法， 断环 算法， 多重划分 算法	总计 相似度， 评审得分	随机划分的比较 ，循环 中断算法， 着色算法， 多重划分算法， 以及最优的非 策略激励相容分配
Leyton-Brown 等人（2022年）	会议 评审人	GAP(MIP)	NA	置信度 评审人， 百分比	1. 所提出的算法与 所提出的算法与

作者	**Case** 实现	数学 模型	算法	性能 度量指标使用情况	比较
	分配： AAAI'21			减少 软约束 违反次数， 墙钟时间， 数量 迭代次数， 数量 论文	匹配算法的 AAAI'20 2. 所提出方法与 所提出的方法和 所提出的方法 不含软约束条件 以及单阶段版本 的 AAAI'21 3. 运行时间与 会议规模 比较

参考文献

Alon, N., Fischer, F., Procaccia, A., & Tennenholtz, M.（2011年）。我们之和：从选择者中进行策略无关的选择。载于《第13届理性与知识理论方面会议论文集》，101–110。Alghamdi, R., & Alfalqi, K.（2015年）。文本挖掘中的主题建模综述。《国际高级计算机科学与应用杂志》，6(1)，147–153。https://doi.org/10.14569/ijacsa.2015.060121 Anjum, O., Gong, H., Bhat, S., Hwu, W.M., & Xiong. J.（2019年）。PaRe：一种利用共同主题空间进行论文—审稿人匹配的方法。载于《2019年自然语言处理实证方法会议暨第9届自然语言处理国际联合会议论文集（EMNLP-IJCNLP）》，518–528，中国香港。计算语言学协会。10.18653/v1/D19-1049 Andrade-Navarro, M., Palidwor, G. A. & Perez-Iratxeta, C.（2015年）。Peer2ref：一种利用作者消歧来查找同行评审人的网络工具。《生物数据挖掘》，5–14。doi:10.1186/1756-0381-5-14 Basu, C., Hirsh, H., Cohen, W. W., & Nevill-Manning, C.（2001年）。技术论文推荐：结合多种信息来源的研究。《人工智能研究杂志》，14，241 262 10 1613 739 – 。https://doi.org/ . /jair. Beltagy, I., Lo, K., & Cohan, A.（2019年）。SciBERT：面向科学文本的预训练语言模型。arXiv。https://doi.org/10.48550/arXiv.1903.10676 Benferhat, S., & Lang, J.（2001年）。会议论文分配。《国际智能系统杂志》，16(10)，1183–1192。https://doi.org/10.1002/int.1055 Biswas, H. K., & Hasan, M. M.（2007年）。利用出版物和领域知识构建研究者档案：自动审稿人分配的应用。载于《国际信息与通信技术会议论文集》，82–86。https://doi.org/10.1109/ICICT.2007.375347 Blei, D. M., Ng, A. Y., & Jordan, M. I.（2003年）。潜在狄利克雷分配。《机器学习研究杂志》，3(4–5)，993–1022。https://doi.org/10.1016/b978-0-12-411519-4.00006-9 Boehmer, N., Bredereck, R., & Nichterlein, A.（2021年）。打击串通圈困难重重，但并非不可能。arXiv预印本，arXiv:2112.08444。https://doi.org/10.48550/arXiv.2112.08444 Bojanowski P., Grave E., Joulin A., & Mikolov T.（2017年）。利用子词信息丰富词向量。《计算语言学协会会刊》，5，135-146。Caldera, C., Eggeling, E., Schr, M., & Fellner, D. W.（2014年）。PRIMA——迈向自动化评审—论文匹配度计算。载于《CONTENT 2014：第六届创意内容技术国际会议论文集》，70–75。

Cechlárová, K., Fleiner, T., & Potpinková, E.（2014年）。为研究资助申请分配评估者：斯洛伐克研究与发展局案例。《科学计量学》，99(2)，495–506。https://doi.org/10.1007/s11192-013-1187-1 Charlin, L., Zemel, R., & Boutilier, C.（2012年）。论文匹配优化框架。载于《第27届人工智能不确定性会议论文集》，86–95。Charlin, L., & Zemel, R. S.（2013年）。多伦多论文匹配系统：自动论文—审稿人分配系统。ICML同行评审与出版模型研讨会（PEER），28。https://www.cs.toronto.edu/~zemel/documents/tpms.pdf Chen, R. C., Shang, P. H., & Chen, M. C.（2012年）。项目评审人分配问题的两阶段方法。《先进材料研究》，452–453，369–373。https://doi.org/10.4028/www.scientific.net/AMR.452-453.369 Chughtai, G.R., Lee, J., Shahzadi, M., Kabir, A., & Hassan, M.A.（2019年）。一种基于本体的主题特定文章推荐模型，用于推荐最匹配的审稿人。《科学计量学》，122，249-265。Conry, D., Koren, Y., & Ramakrishnan, N.（2009年）。会议论文分配问题的推荐系统。载于《第3届ACM推荐系统会议论文集》，357–360。https://doi.org/10.1145/1639714.1639787 Cook, W. D., Golany, B., Kress, M., Penn, M., & Raviv, T.（2005年）。将提案最优分配给评审人，以促进有效排名。《管理科学》，51(4)，655–661。https://doi.org/10.1287/mnsc.1040.0290 Daş, G. S., & Göçken, T.（2014年）。一种用于审稿人分配问题的模糊方法。《计算机与工业工程》，72(1)，50–57。https://doi.org/10.1016/j.cie.2014.02.014 Daud, A., Li, J., Zhou, L., & Muhammad, F.（2010年）。通过广义时间主题建模进行时序专家发现。《基于知识的系统》，23(6)，615–625。https://doi.org/10.1016/j.knosys.2010.04.008 Davoodi, E., Kianmehr, K., & Afsharchi, M.（2013年）。基于语义社交网络的专家推荐系统。《应用智能》，39(1)，1–13。https://doi.org/10.1007/s10489-012-0389-1 Deerwester S., Harshman, R., Susan T., George W., & Thomas K.（1990年）。基于潜在语义分析的索引。《美国信息科学学会会刊》，41(6)，391–407。Devlin, J., Chang, M. W., Lee, K., & Toutanova, K.（2018年）。BERT：用于语言理解的深度双向变换器预训练。arXiv预印本。arXiv:1810.04805v2。1-16。Dhull, K., Jecmen, S., Kothari, P., & Shah, N. B.（2022年）。通过分区实现同行评估的策略无关化：评估者专业知识方面的代价。arXiv预印本，arXiv:2201.10631。https://doi.org/10.48550/arXiv.2201.10631 Duan, Z., Tan, S., Zhao, S., Wang, Q., Chen, J., & Zhang, Y.（2019年）。基于句对建模的审稿人分配。《神经计算》，366，97–108。https://doi.org/10.1016/j.neucom.2019.06.074 Dumais, S. T., & Nielsen, J.（1992年）。将投稿稿件自动分配给审稿人。《SIGIR论坛》（ACM信息检索特别兴趣小组），233–244。Ekinci, E., & Omurca, S. İ.（2020年）。NET-LDA：一种基于文档语义相似度的新型主题建模方法。《土耳其电气工程与计算机科学杂志》，28(4)，文章30。https://doi.org/10.3906/elk-1912-62 Farahat A., & Chen, F.（2006年）。利用主成分分析改进概率潜在语义分析。载于《第11届欧洲计算语言学协会分会会议论文集（EACL）》，意大利特伦托，第105–112页。

Ferilli, S., Di Mauro, N., Basile, T. M. A., Esposito, F., & Biba, M. (2006). 用于审稿人分配的自动主题识别。《计算机科学讲义》（包括《人工智能讲义》和《生物信息学讲义》子系列），4031 LNAI（6月），721–730。https://doi.org/10.1007/11779568_78Fiez, T., Shah, N. B., & Ratliff, L. (2020). 一种用于优化同行评审中论文竞标的超级* 算法。载于第36届人工智能不确定性会议论文集，124，600–608。Garg, N., Kavitha, T., Kumar, A., Mehlhorn, K., & Mestre, J. (2010). 将论文分配给评审人。《Algorithmica》，58(1)，119–136。https://doi.org/10.1007/s00453-009-9386-0Goldsmith, J., & Sloan, R. H. (2007年5月). 人工智能会议论文分配问题。AAAI研讨会——技术报告，53–57。Green, S. M., & Callaham, M. L. (2011). 基于质量和可靠性的期刊同行评审人分层系统的实现。《急诊医学年鉴》，57(2)，149–152。Guo, L., Wu, J., Chang, W., Wu J., & Li, J. (2018). 会议评审系统中的无K环分配。载于国际计算、网络与通信会议（ICNC）论文集，第542–547页。doi: 10.1109/ICCNC.2018.8390343。Hartvigsen, D., Wei, J. C., & Czuchlewski, R. (1999). 会议论文—审稿人分配问题。《决策科学》，30(3)，865–876。https://doi.org/10.1111/j.1540-5915.1999.tb00910.xHettich, S., & Pazzani, M. J. (2006). 提案审稿人挖掘：美国国家科学基金会的经验教训。载于2006年ACM SIGKDD知识发现与数据挖掘国际会议论文集，862–871。Hoang, D. T., Nguyen, N. T., Collins, B., & Hwang, D. (2021). 用于解决审稿人分配问题的决策支持系统。《控制论与系统》，52(5)，379–397。https://doi.org/10.1080/01969722.2020.1871227Hofmann, T. (1999). 概率潜在语义索引。载于SIGIR-99论文集，35–44。Howard, J., & Ruder, S. (2018). 用于文本分类的通用语言模型微调。ACL，计算语言学协会。http:// arxiv.org/abs/1801.06146。Huang, M., Liu, B., & Hong, L. (2010). 关于将论文分配给审稿人。第2届数据库技术与应用国际研讨会。https://doi.org/10.1109/DBTA.2010.5659009Janak, S. L., Taylor, M. S., Floudas, C. A., Burka, M. K., & Mountziaris, T. J. (2005). 一种用于美国国家科学基金会评审小组分配问题的新型有效整数优化方法：一种多资源偏好约束广义分配问题。载于美国化学工程师学会年会论文集，609，6066。Jecmen, S., Zhang, H., Liu, R., Fang, F., Conitzer, V., & Shah, N. B. (2021). 两阶段论文评审中的近最优审稿人分组与会议实验设计。arXiv预印本 arXiv:2108.06371。https://arxiv.org/abs/2108.06371Jecmen, S., Zhang, H., Liu, R., Shah, N. B., Conitzer, V., & Fang, F. (2020). 通过随机审稿人分配缓解同行评审中的操纵行为。arXiv预印本 arXiv:2006.16437。http://arxiv.org/abs/2006.16437Jin, J., Geng, Q., Mou, H., & Chen, C. (2018a). 用于审稿人推荐的作者—主题—话题模型。《信息科学杂志》，45(4)，554–570。https://doi.org/10.1177/0165551518806116Jin, J., Geng, Q., Zhao, Q., & Zhang, L. (2017). 将研究兴趣趋势整合到审稿人分配中。载于第26届万维网国际会议伴随会议论文集，澳大利亚珀斯，第1233–1241页。http://dx.doi.org/10.1145/3041021.3053053。

Jin, J., Niu, B., Ji, P., & Geng, Q. （2018b）. 考虑研究兴趣的审稿人分配整数线性规划模型。《运筹学年鉴》，291(1–2), 409–433。https://doi.org/10.1007/s10479-018-2919-7 Kale, A., Kharat, R., Bodkhe, S., Apte, P., & Dhonde, H. （2015年）. 面向会议管理系统的公平论文审稿人自动分配。载于《计算、通信、控制与自动化国际会议论文集》，ICCUBEA 2015, 408–411。https://doi.org/10.1109/ICCUBEA.2015.85 Kalmukov, Y. （2020年）. 基于向量空间文本分析模型的审稿人自动分配。载于《ACM 国际会议论文集系列》，229–235。https://doi.org/10.1145/3407982.3408026 Kalmukov, Y., & Rachev, B. （2010年）. 论文审稿人自动分配的现有方法与算法比较分析。《信息技术与控制》，20-31。Kalyan, K. S., Rajasekharan, A., & Sangeetha, S. （2021年）. AMMUS：自然语言处理中基于 Transformer 的预训练模型调查。arXiv，https://doi.org/10.48550/arXiv.2108.05542 Karimzadehgan, M., & Zhai, C. （2009年）. 面向委员会审稿分配的受约束多方面专业知识匹配。载于第18届 ACM 信息与知识管理会议论文集，1697–1700。Karimzadehgan, M., & Zhai, C. （2012年）. 用于受约束的多方面委员会评审分配的整数线性规划。《信息处理与管理》，48(4), 725–740。https://doi.org/10.1016/j.ipm.2011.09.004 Karimzadehgan, M., Zhai, C. X., & Belford, G. （2008年）. 面向评审分配的多方面专业知识匹配。载于信息与知识管理国际会议论文集，1113–1122。https://doi.org/10.1145/1458082.1458230 Kat, B. （2021年）. 评审委员分配问题的算法与决策支持系统：TUBITAK 案例。《加齐大学工程与建筑学院学报》，36(1), 69–87。https://doi.org/10.17341/gazimmfd.631071 Kim, J., & Lee, E. （2018年）. 了解开发者的审稿专长：一种基于潜在狄利克雷分配的审稿人推荐方法。《Symmetry》，10(4), 5–7。https://doi.org/10.3390/sym10040114 Kobren, A., Saha, B. and McCallum, A. （2019年）. 带有局部公平性约束的论文匹配。载于第25届 ACM SIGKDD 知识发现与数据挖掘国际会议论文集，KDD '19，第1247–1257页，纽约，美国。ACM。Kolasa, T., & Krol, D. （2011年）. 论文—审稿人分配问题算法调查。《IETE 技术评论》（电子与电信工程师学会，印度），28(2), 123–134。https://doi.org/10.4103/0256-4602.78092 Kou, N. M., Hou, L. U., Mamoulis, N., & Gong, Z. （2015a）. 基于加权覆盖的审稿人分配。载于 ACM SIGMOD 数据管理国际会议论文集，2031–2046。https://doi.org/10.1145/2723372.2723727 Kou, N. M., Leong Hou, U., Mamoulis, N., Li, Y., Li, Y., & Gong, Z. （2015b）. 基于主题的审稿人分配系统。《VLDB基金会论文集》，8(12), 1852–1855。https://doi.org/10.14778/2824032.2824084[数据集] Kreutz, C. K., & Schnekel, R. （2020年）. RevASIDE 数据集：MOL'17、BTW'17、ECIR'17。Zenodo。https://doi.org/10.5281/zenodo.4071874 Le, Q., & Mikolov. T. （2014年）. 句子和文档的分布式表示。载于机器学习国际会议论文集，1188–96。中国北京：JMLR.org。

Leyton-Brown, K., Mausam, Nandwani, Y., Zarkoob, H., Cameron, C., Newman, N., Raghu, D.（2022年）。大型会议中的论文与审稿人匹配。arXiv预印本，arXiv:2202.12273。https://arxiv.org/abs/2202.12273 Li, B., & Hou, Y. T.（2016年）。IEEE INFOCOM新型自动化审稿分配系统。IEEE Network，30(5),18-2。Li, H., & Chen, N.（2011年）。基于知识集理论的同行评审专家遴选方法研究。载于第二届人工智能、管理科学与电子商务国际会议论文集，2，2248–2251。ttps://doi.org/10.1109/AIMSEC.2011.6011167 Li, J., Peng, J., & Wei, Y.（2013年）。用于专家分配问题的自适应并行遗传算法。载于第六届计算智能与设计国际研讨会论文集，1，23–26。https://doi.org/10.1109/ISCID.2013.13 Li, N. N., Zhang, J. N., Gu, J. H., & Liu, B. Y.（2007年）。使用改进遗传算法解决专家分配问题。载于第六届机器学习与控制论国际会议论文集，2（8月），934–937。https://doi.org/10.1109/ICMLC.2007.4370276 Li, N. N., Zhao, Z., Gu, J. H., & Liu, B. Y.（2008年，7月）。用于专家分配问题的蚁群优化算法。载于第七届机器学习与控制论国际会议论文集，2（7月），660–664。https://doi.org/10.1109/ICMLC.2008.4620487 Li, X., & Watanabe, T.（2013年）。基于审稿人匹配度的论文自动分配给审稿人。Procedia Computer Science，22，633–642。https://doi.org/10.1016/j.procs.2013.09.144 Lian, J. W., Mattei, N., Noble, R., & Walsh, T.（2018年）。会议论文分配问题：使用有序加权平均分配不可分物品。载于AAAI人工智能会议论文集，32(1)。https://doi.org/10.1609/aaai.v32i1.11484 Littman, M. L.（2021年）。串通圈威胁计算机科学研究的诚信。ACM通信，64(6):43–44。Liu, D., Xu, W., Du, W., & Wang, F.（2015年）。如何为同行评审选择合适的专家：大数据语境下的一种智能推荐方法。Data Science Journal，14，1–11。https://doi.org/10.5334/dsj-2015-016 Liu, O., Wang, J., Ma, J., & Sun, Y.（2016年）。一种用于研发项目选择中审稿人分配的智能决策支持方法。Computers in Industry，76，1–10。https://doi.org/10.1016/j.compind.2015.11.001 Liu, P., & Dew, P.（2004年）。利用语义网技术改进学术界内的专业知识匹配。载于I-KNOW论文集，奥地利格拉茨，第70–378页。Liu, X., Suel, T., & Memon, N.（2014年）。论文—审稿人分配类别的稳健模型及主题描述符。载于RecSys2014论文集，25–32。Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. & Stoyanov, V.（2019年）。Roberta：一种稳健优化的BERT预训练方法。arXiv预印本，arXiv:1907.11692。Long, C., Wong, R. C. W., Peng, Y., & Ye, L.（2013年）。关于良好且公平的论文—审稿人分配。载于IEEE数据挖掘国际会议论文集，1145–1150。https://doi.org/10.1109/ICDM.2013.13 Mai, X., Ding, G., & Wang, J.（2012年）。感知权威性的专家搜索：面向国家自然科学基金委员会的算法与系统。载于IET会议论文集，585–588。https://doi.org/10.1049/cp.2012.1047 Mauro, N. D., Basile, T. M. A., & Ferilli, S.（2005年）。GRAPE：科学会议管理系统的专家审稿分配组件。《应用人工智能创新》《计算机科学讲义》，3533，Springer，柏林，海德堡。https://doi.org/10.1007/11504894

Medakene, A. N., Bouanane, K., & Eddoud, M. A.（2019年）。论文审稿人分配问题中匹配度计算的新方法。载于《理论与应用计算机科学国际会议》会议录。
https://doi.org/10.1109/ICTAACS48474.2019.8988127 Merelo-Guervós, J. J., & Castillo-Valdivieso, P.（2004年）。使用贪心／进化组合算法进行会议论文分配。《计算机科学讲义》（包括《人工智能讲义》和《生物信息学讲义》子系列），3242，602–611。
https://doi.org/10.1007/978-3-540-30217-9_61 Mikolov, T., Chen, K., Corrado, G., & Dean, J.（2013年）。向量空间中词表示的高效估计。arXiv预印本 arXiv:1301.3781。
https://doi.org/10.48550/arXiv.1301.3781 Mimno, D., & Mccallum, A.（2007年）。用于论文与评审人匹配的专长建模。载于*KDD'07*会议录，500-509。10.1.1.87. 5674.Pdf。500–509。Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J.（2022年）。基于深度学习的文本分类：全面综述。《ACM计算调查》，54(3)，1-40。Mirzaei, M., Sander, J., & Stroulia, E.（2019年）。基于潜在研究领域的多方面评审团队分配。《信息处理与管理》，56(3)，858–878。
https://doi.org/10.1016/j.ipm.2019.01.007 Miller, G. A.（1995年）。Wordnet：英语词汇数据库。《ACM通信》，38(11):3941。Mittal, K., Jain, A., & Vaisla, K. S.（2019年）。理解审稿人分配问题及其问题与挑战。载于第四届物联网：智能创新与应用国际会议论文集，1–6。
https://doi.org/10.1109/IoT-SIU.2019.8777727 Müngen, A. A., & Kaya, M.（2018年）。从语境中提取学术文章的摘要和关键词。《社会网络分析与挖掘》，8(1)，0。
https://doi.org/10.1007/s13278-018-0524-z Nguyen, J., Sánchez-Hernández, G., Agell, N., Rovira, X., & Angulo, C.（2018年）。使用有序加权平均进行会议审稿分配的决策支持工具。《模式识别快报》，105，114–120。https://doi.org/10.1016/j.patrec.2017.09.020 Ogunleye, O., Ifebanjo T., & Abıodun T., & Adebiyi, A.（2017年）。使用Word2vec的论文审稿人分配系统框架提案。载于CUCEN2017会议录，211-218。Papagelis, M., Plexousakis, D., & Nikolaou, P. N.（2005年）。CONFIOUS：管理科学会议的电子投稿与评审流程。《计算机科学讲义》（包括《人工智能讲义》和《生物信息学讲义》子系列），3806 LNCS，711–720。
https://doi.org/10.1007/11581062_78[数据集] Patil, A. H.（2021年）。Kaggle数据集：NIPS2019。
https://www.kaggle.com/abolihpatil/nips2019-paper-for-reviewer-assignment-ahp-pnm Patil, A. H., & Mahalle, P. N.（2021年）。出版流程中的公平且准确的评审：一种基于学习的主动方法，将评审人分配给手稿。《土耳其计算机与数学教育杂志》，12(4)。
https://doi.org/10.17762/turcomat.v12i4.455 Payan, J., & Zick, Y.（2021年）。我将拥有排序！优化排序以实现公平的审稿人分配。arXiv预印本，arXiv:2108.02126。
https://doi.org/10.48550/arXiv.2108.02126 Peng, H., Hu, H., Wang, K., & Wang, X.（2017年）。时间感知与基于主题的审稿人分配。《高级应用数据库系统》，DASFAA《计算机科学讲义》，10179，145 157，10 1007 978 3 319 55705 2 11 – Springer，Cham。https://doi.org/ . / - - - - _ Pennington, J., Socher, R., & Manning, C.（2014年）。GloVe：用于词表示的全局向量。载于自然语言处理中的经验方法会议（EMNLP）会议录，1532- 1543。

Pradhan, D. K., Chakraborty, J., Choudhary, P., & Nandi, S.（2020年）。一种基于利益冲突的会议论文分配系统自动化贪心方法。《计量学报》，14(2), 101022。https://doi.org/10.1016/j.joi.2020.101022 Protasiewicz, J.（2014年1月）。一种用于遴选评审人的支持系统。载于 IEEE系统、人类与控制论国际会议论文集，3062–3065。https://doi.org/10.1109/smc.2014.6974397 Protasiewicz, J., Pedrycz, W., Kozłowski, M., Dadas, S., Stanisławek, T., Kopacz, A., & Gałęzewska, M.（2016年）。一种用于评审问题的评审人与专家推荐系统。《知识基础系统》，106, 164–178。https://doi.org/10.1016/j.knosys.2016.05.041 Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. & Liu, P. J.（2020年）。探索统一文本到文本转换器的迁移学习极限。《机器学习研究杂志》，21(140), 1–67。Rodriguez, M. A., & Bollen, J.（2008年）。一种确定同行评审人的算法。载于信息与知识管理国际会议论文集，319–328。https://doi.org/10.1145/1458082.1458127 Romanov, V., & Khusainova A.（2019年）。英语和俄语形态嵌入的评估。第3届自然语言处理向量空间表示评估研讨会论文集，明尼阿波利斯，美国，77–81。Rosen-Zvi, M., Griffiths, T., Steyvers, M., & Smyth, P.（2004年）。面向作者和文档的作者-主题模型。载于第20届人工智能不确定性会议（UAI）论文集，加拿大班夫，第487–494页。Sanh, V., Debut, L., Chaumond, J., & Wolf, T.（2019年）。DistilBERT，一种BERT的蒸馏版本：更小、更快、更便宜、更轻量。arXiv预印本 arXiv:1910.01108。Shah, N. B.（2022年）。同行评审中的挑战、实验与计算解决方案概述（扩展版）。https://www.cs.cmu.edu/~nihars/preprints/SurveyPeerReview.pdf Shon, H.S., Han, S. H., Kim, K. A., Cha, E. J., & Ryu, K. H.（2016年）。面向国家科研管理机构的基于大数据的提案评审人推荐系统。《信息科学杂志》，1-12。Silva, T., Jian, M., & Chen, Y.（2014年）。研发项目选择的流程分析方法。《ACM管理信息系统汇刊》，5(4)。https://doi.org/10.1145/2629436 Snyder, H.（2019年）。作为研究方法论的文献综述：概述与指南。《商业研究杂志》，104（8月），333–339。https://doi.org/10.1016/j.jbusres.2019.07.039 Sun, Y. H., Jian, M., Fan, Z. P., & Wang, J.（2007年）。一种用于审稿人分配的混合知识与模型方法。载于夏威夷系统科学国际年会论文集，1–10。https://doi.org/10.1109/HICSS.2007.17 Stelmakh, I, Shah, N. B., & Singh A.（2019年）。PeerReview4All：公平且准确的同行评审审稿人分配。arXiv预印本，arXiv:1806.06237。https://arxiv.org/abs/1806.06237 Tan, S., Duan, Z., Zhao, S., Chen, J., & Zhang, Y.（2021年）。基于词和语义特征的改进审稿人分配。《信息检索杂志》，24(3), 175–204。https://doi.org/10.1007/s10791-021-09390-8 Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z.（2008年）。ArnetMiner：学术社交网络的提取与挖掘。载于第14届ACM SIGKDD知识发现与数据挖掘国际会议论文集，拉斯维加斯，内华达州，2008年8月24–27日，第990–998页。Tang, W., Tang, J., Lei, T., Tan, C., Gao, B., & Li, T.（2012年）。论多种约束下专业知识匹配的优化。《神经计算》，76(1), 71–83。https://doi.org/10.1016/j.neucom.2011.04.039

Tang, W., Tang, J., & Tan, C.（2010年）。通过基于约束的优化实现专业知识匹配。载于国际网络智能会议论文集，1, 34–41。https://doi.org/10.1109/WI- IAT.2010.133 Tayal, D. K., Saxena, P. C., Sharma, A., Khanna, G., & Gupta, S.（2014年）。一种利用二型模糊集和模糊函数解决审稿人分配问题的新方法。《应用智能》，40(1), 54–73。https://doi.org/10.1007/s10489-013-0445-5 Taylor, C. J.（2008年）。会议论文分配给审稿人的最优方法。技术报告 MS-CIS-08-30，宾夕法尼亚大学。Thurner S., & Hanel. R（2011年）。在理性科学家的世界中进行同行评审：走向平均值的选择。《欧洲物理学杂志B》。Tian, Q., Ma, J., Liang, J., Kwok, R. C. W., & Liu, O.（2005年）。一种用于有效研发项目选择的组织决策支持系统。《决策支持系统》，39(3), 403–413。https://doi.org/10.1016/j.dss.2003.08.005 Tian, Q. Ma, J., & Liu O.（2002年）。一种用于研发项目选择的混合知识与模型系统。《专家系统与应用》，265-271。Tran, H. D., Cabanac, G., & Hubert, G.（2017年）。面向会议程序委员会的专家推荐。载于第11届国际信息科学研究挑战会议（RCIS）论文集，第221-232页。doi: 10.1109/RCIS.2017.7956540。Wang, C., Han, J., Jia, Y., Tang, J., Zhang, D., Yu, Y., & Guo, J.（2010年）。从研究出版物网络中挖掘导师—学生关系。载于KDD论文集，203–212。Wang, F., Chen, B., & Miao, Z.（2008年）。审稿人分配问题调查。《计算机科学讲义》（含《人工智能讲义》和《生物信息学讲义》子系列），5027 LNAI(60704048), 718–727。https://doi.org/10.1007/978-3-540-69052-8_75 Wang, F., Shi, N., & Chen, B.（2010年）。审稿人分配问题的全面调查。《国际信息技术与决策》，9(4), 645–668。https://doi.org/10.1142/S0219622010003993 Wang, F., Zhou, S., & Shi, N.（2013年）。组间审稿人分配问题。《计算机与运筹研究》，40(5), 1351–1362。https://doi.org/10.1016/j.cor.2012.08.005 Xu, Y., & Du, Y.（2014年）。一种用于审稿人推荐的三层网络模型。载于第6届国际商业智能与金融工程会议论文集，552–556。https://doi.org/10.1109/BIFE.2013.115 Xu, Y. H., Guo, X. T., Xu, L., Chen, Y., & Zhuang, Y. Y.（2012年）。用于审稿人推荐的研究分析。载于国际管理科学与工程会议论文集，213–217。https://doi.org/10.1109/ICMSE.2012.6414185 Xu, Y., Ma, J., Sun, Y., Hao, G., Xu, W., & Zhao, D.（2010年）。一种将审稿人分配给提案的决策支持方法。《专家系统与应用》，37(10), 6948–6956。https://doi.org/10.1016/j.eswa.2010.03.027 Xu, Y., Zhao, H., Shi, X., & Shah, N. B.（2019年）。论会议同行评审中的策略证明。载于IJCAI 人工智能国际联合会议论文集，616–622。https://doi.org/10.24963/ijcai.2019/87 Xu, Y., Zhao, H., Shi, X., Zhang, J., & Shah, N. B.（2020年）。论会议同行评审中的策略证明。arXiv预印本，arXiv:1806.06266。https://doi.org/10.48550/arXiv.1806.06266 Xu, Y. H., & Zuo, X. I.（2016年）。一种基于潜在狄利克雷分配模型的文本挖掘方法，用于为研究项目选择提案推荐评审人。载于第13届国际服务系统与服务管理会议（ICSSSM）论文集，1-5。doi: 10.1109/ICSSSM.2016.7538568。

Xue, N., Hao, J. X., Jia, S. L., & Wang, Q. (2012). 基于区间模糊本体的同行评审指派方法。载于 IEEE国际电子商务工程会议论文集，55–60。https://doi.org/10.1109/ICEBE.2012.19 Yan, S. (2017). 利用基于学术网络的利益冲突进行论文审稿人指派。《国际知识工程期刊》，3(2)，65–73。https://doi.org/10.18178/ijke.2017.3.2.089 Yang, C., Liu, T., Yi, W., Chen, X., & Niu, B. (2020). 通过语义建模识别专业知识：一种用于审稿人分配问题的改进型 BBPSO算法。《应用软计算期刊》，94，106483。https://doi.org/10.1016/j.asoc.2020.106483 Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R. & Le, Q. V. (2019). XLNet：用于语言理解的通用自回归预训练。《神经信息处理系统进展》，32，5753–5763。Yarowsky, D., & Florian, R. (1990). 减轻会议主席的负担：迈向数字化论文路由助手。《科学》，220–230。Yeşilçimen, A., & Yıldırım, E. A. (2019). 用于审稿人分配问题的多项式规模公式及基于优化的启发式方法。《欧洲运筹学杂志》，276(2)，436–450。https://doi.org/10.1016/j.ejor.2019.01.035 Yin, H., Cui, B., Lu, H., & Zhao, L. (2016). 用于评审指派的专家团队发现。载于人工智能技术与应用会议（TAAI）论文集，1–8。https://doi.org/10.1109/taai.2016.7932314 Yong, Y., Yao, Z., & Zhao, Y. (2021). 基于知识图谱和规则匹配的审稿人推荐框架。载于 IEEE国际信息通信与软件工程会议（ICICSE）论文集，199–203。http://dx.doi.org/10.1109/ICICSE52190.2021.9404099 Yue, M., Tian, K., & Ma, T. (2017). 一种用于科研项目评审的准确、公正的专家指派方法。《数据与信息科学学报》，2(4)，65–80。https://doi.org/10.1515/jdis- 2017-0020 Zhai, C., & Lafferty, J. (2001). 语言模型平滑方法在专用信息检索中的应用研究。SIGIR论坛（ACM信息检索特别兴趣小组），334–342。https://doi.org/10.1145/383952.384019 Zhang, Q., Zhang, M., Chen, T., Sun, Z., Ma, Y., & Yu, B. (2019). 卷积神经网络加速的最新进展。《神经计算》323，37-51。Zhang, D., Zhao, S., Duan, Z., Chen, J., Zhang, Y., & Tang, J. (2019). 一种用于论文审稿人推荐的层次化透明表示多标签分类方法。arXiv。https://doi.org/10.48550/arXiv.1912.08976 Zhao, S., Zhang, D., Duan, Z., Chen, J., Zhang, Y. ping, & Tang, J. (2018). 一种新颖的论文审稿人推荐分类方法。《科学计量学》，115(3)，1293–1313。https://doi.org/10.1007/s11192-018-2726-6