

基于多指标的学术会议动态评价研究*

张洋 叶月

(中山大学资讯管理学院, 广州 510006)

摘要: 为更好地评价学术会议等级, 本文提出基于多指标的学术会议动态评价研究方法, 收集人工智能领域2007—2014年的学术会议数据, 共235个会议64 929篇论文的信息。通过7个指标与会议等级进行相关性分析, 并应用Topsis模型对会议进行排名, 得到每个会议的动态排名与综合分析结果。

关键词: 会议评价; 同侪声望; 学术影响; Topsis模型

中图分类号: G350

DOI: 10.3772/j.issn.1673-2286.2018.07.005

知识经济时代, 科学技术对社会发展起到前所未有的推动作用, 科技发展全球化趋势日益明显、科技发展内涵不断丰富、科技工作范围不断扩大, 共同促进了科学评价的发展。学术会议作为学术交流的重要形式, 在推广先进技术、传播科研成果、促进学科发展、活跃学术思想、发现和培养科学人才等方面起着重要的作用^[1]。大量的学术会议和会议出版物受到越来越高的重视, 产生了会议排名的需求, 对计量学领域提出了新要求。

由于学术会议的评价标准缺乏正式明确的指标, 且对学术会议的评价多以定性研究为主, 这可能会导致评价结果的失真。因此, 建立定量的、客观的学术会议评价体系, 使其不仅能够独立反映科技机构和科研人员的科学生产率, 并能够与期刊评价共同完善科学评价体系。本文提出基于多指标的学术会议动态评价方法, 与多指标评价期刊方法类似, 可以减少单一指标对会议评价的客观影响。同时为评价会议提供新思路, 减少定性评价对结果的主观影响。

1 文献综述

1.1 会议评价现状

1.1.1 国内会议评价

国内比较著名的计算机科学类会议排名, 多指由中

国计算机学会 (CCF) 发表的《中国计算机学会推荐国际学术会议和期刊目录》, 自2010年至今共发布4版, 第1版、第2版主要依靠专家定性评价, 第3版开始引入会议录取率和引用情况, 2014年对第4版目录更新, 分3个阶段完成: 提议受理阶段, 领域责任专家 (或专家组) 审议和推荐阶段, 以及终审专家组投票表决阶段。将计算机领域会议分为十类, 共A、B、C三个等级, A类指国际上极少数的顶级会议, 鼓励我国学者突破; B类指国际上知名和非常重要的会议, 有重要的学术影响, 鼓励国内同行投稿; C类指国际学术界认可的重要会议。

1.1.2 国外会议评价

国外关于会议评价的排名通常有3种: ①CORE Computer Science Conference Rankings, 由澳大利亚计算研究与教育学会 (CORE) 于2005年发起, 几乎每年都会更新排名, 目前最新排名为2017年, 会议分为A+、A、B、C四个等级, 主要用会议录取率、ISI科学引文数据网数据排名计算得来。②Qualis排名, 由巴西教育部公布, 将H-index作为会议的表现措施。基于H-index百分比, 会议被分为A1 (最好)、A2、B1、B2、B3、B4、B5 (最差) 七个等级。③MSAR排名, 由微软学术会议现场评价得出, 指标类似于H-index, 计算每个会议出席作者的出版物数量和出版物引用分布。本排名并未分等级, 是按照指标进行整体排序。

*本研究得到国家社会科学基金项目“新型网络环境下学术期刊影响力的计量分析与评价研究” (编号: 14BTQ067) 资助。

1.2 会议评价指标现状

评价方法通常采用文献计量学指标,从简单的引文计数到复杂的指数,如H-index^[2]或G指数^[3]。尽管这些指标在近几年被广泛使用,但也有很多学者认为这存在一定问题。例如,Laloe^[4]认为H-index倾向于较大的研究团体而非学者个人评价。Souto等^[5]通过开发半自动工具来评估计算机科学会议,该工具考虑了会议分类的各种标准,这些标准主要由会议出版商、赞助商、计划委员会等组成。纵观该方法发现,在较少的会议评价中效果比较明显,在大量学术会议评价中,这些信息很难获得。

Martins等^[6]提出会议录取率的概念,并认为会议录取率可以作为评价会议的重要指标,主要考虑了会议录取论文量对引文数的影响,从而反作用于会议整体的评价,此外又提出几个新的用于评价国际学术会议的指标:会议影响因子(CIF)、会议引文影响(CCI)、组合会议因子(CCF)和会议因子(C-Factors)。但在后续的研究中发现,会议录取率并不能很好地传递一个会议的质量,Martins等发现会议论文录取率为10%~15%的论文被引量比录取率为15%~20%的论文被引量低。这可能是因为会议筛选过程较为严格,导致一些可能提高会议影响力的论文被剔除。故本文没有采用会议录取率这一指标。其他评价指标现状见表1。

表1 部分会议评价指标现状

指标来源	指标介绍	不足
Martins等 ^[6]	会议影响因子、会议引文影响、会议规模与会议周期	不具备客观性,会议规模、会议周期很难与会议等级呈线性关系,高质量会议规模一般较大,而低质量的会议有时为提高知名度也会刻意提高自身规模
Kungas等 ^[7]	平均提交量、平均录取量、平均录取率、发表论文数、引用量与平均引用量	录取率和文献计量学指标区分高层次会议和其他等级会议的效果明显,但区分中层会议和低层会议的作用不大
Elmacioglu等 ^[8]	项目委员会成员的数量、项目委员会的平均出版量、项目委员会合作作者的平均数量、项目委员会的亲密周期性与项目委员的中间性	最后仅采用社会网络分析法(SNA)得出质量越高的会议,论文作者通常与其他作者的亲密值越高,并没有对会议进行评价
Nelakuditi等 ^[9]	同侪声望(Peer reputation, PR),利用学校的排名和每篇论文第一作者的所属机构确定会议的质量。其用数组(如1/3, 20,表示1/3的文章来自于前20的大学)表示一个会议的质量	PR仅用所属机构来评价一个学者是不全面的,而且忽视了会议地点的重要性。另外,只考虑第一作者的所属机构排名,忽视了其他作者,且PR只考虑美国大学的排名,但是研究表明当考虑除美国以外的大学时,值会有很大的变动
Loizides等 ^[10]	改进PR值、影响因子与H-index	仅包含56个会议,数据集偏小,偶然性很大;K-means聚类方法并不能对会议进行排名,不适合评价研究

综合考量以上指标存在的问题,本文在Loizides等的研究基础上进行改进,选用人工智能领域2007—2014年共8年的数据,采用PR等四类共7个指标进行会议评价,通过Topsis模型融合各个指标,得到每个会议的历年动态排名,并对同一会议历年的接近程度值取平均,得到综合排名,并解读该排名产生的原因。

2 数据来源

在文献数据库研究中,计算机科学是一个分支学科,研究人员通常倾向于在有声望的会议上发表他们的作品,而不是在主要的数据库期刊上发表。解释这一

现象的原因之一是鉴于该领域的迅速发展,需要缩短传播周期^[11]。Eckmann等^[12]发现,在计算机视觉领域中,具有先导性的期刊论文被引用次数超过没有先导性的期刊论文,先导性论文指在扩展的期刊版本之前发表的会议论文。这一事实很好地反映了计算机科学中常见的出版模式,首先在会议记录中发表论文的最初版本,随后是扩展的期刊版本。故本文所用数据选取计算机领域的研究热点——人工智能,由CCF提供的《中国计算机学会推荐国际学术会议和期刊目录》(2014版)的人工智能领域39个会议,共收集8年数据,得到235个会议信息。

本研究截取了2007—2014年的会议,数据获取分

为以下两个步骤。

(1) 在Scopus数据库上人工导出2007—2014年人工智能方面的会议论文数据,包括作者、标题、年份、施引文献、归属机构、会议名等信息,共235个会议64 929篇论文信息,根据前文统计方法,得出A类会议47个,B

类会议69个,C类会议119个。

(2) 基于导出的标题和作者信息,从Scopus数据库中获取每位作者的机构来源和H-index,共获得190 983位作者机构来源的对应关系。论文数量分布示例如表2。

表2 会议论文数量分布(示例)

单位:篇

会议名称	2007	2008	2009	2010	2011	2012	2013	2014
AAAI	577	582	350	472	553	619	599	715
CVPR	542	743	521	462	439	467	473	541
ICCV	392	-	542	-	644	-	457	-
ICML	170	161	162	161	157	245	286	321
IJCAI	528	-	362	-	502	-	542	-
NIPS	207	-	730	293	302	375	361	417
ACL	134	239	399	268	399	219	332	290

3 研究方法

3.1 指标获取

3.1.1 PR值

PR (1/3) 与PR (1/4),即表示1/3或1/4的作者来自于前多少名的大学,在Nelakuditi等^[9]论文中计算的是第一作者的值,在Loizides等^[10]方法中则计算了论文所有作者,本文结合以上两位学者的方法,分别计算出每一个会议的第一作者和所有作者的PR (1/4) 与PR (1/5) 平均值(由于会议规模比较庞大,在计算PR值时1/4与1/3区分不明显,故采用1/4和1/5的PR平均值)。Nelakuditi等的研究中采用了美国大学排名,原因是评价的会议所收到的论文大部分都来自美国的大学。Loizides在大学排名选择上做了改进,选择莱顿世界大学排名,在一定程度上解决了来自美国大学之外的学者投稿问题。莱顿世界大学排名并没有通过对指标赋权重得出综合排名,而是仅给出500所大学单项指标的排名。

除莱顿世界大学排名之外,比较著名的还有QS世界大学排名(QS World University Rankings, QS)、美国权威的《美国新闻与世界报道》(U.S.News & World

Report)。各大学排名指标分布见表3。

CCF提供的会议分级应用指标主要基于定性评价与定量评价两方面,定性评价主要是专家审议,定量评价主要是论文被引量与会议录取率,在比较三个大学排名中,发现《美国新闻与世界报道》排名在声誉指标和文献计量指标中占比最高达到80%,与本文研究最相符,而QS为70%,莱顿为单一指标排名,且不涉及声誉指标。将以上三所大学排名得到的PR平均值与会议等级进行相关性分析,得到相关系数r(见表4),其中,PR15F代表论文第一作者的PR (1/5) 平均值,PR14F代表论文第一作者的PR (1/4) 平均值,PR15A与PR14A分别代表所有作者的PR (1/5) 与PR (1/4) 的平均值,通过计算得知,《美国新闻与世界报道》大学排名在人工智能领域的PR值效果最好,故本文在最终分析时选择了《美国新闻与世界报道》世界大学排名。

3.1.2 H-index

H-index是Hirsch^[2]在2005年提出来的,自提出以来得到广泛应用,但该指数的缺点是不能完全代表作者作品的动态信息,而是强调了作者在同行中的认可程度。H-index指如果作者的N篇论文中有H篇论文被至少引用了H次,那么剩下的N-H篇论文引用次数均不超

表3 大学排名指标

名称	一级指标	二级指标	权重
QS世界大学排名	学术领域的同行评价	-	40.0%
	全球雇主评价	-	10.0%
	单位教职论文引用数	-	20.0%
	教师/学生比例	-	20.0%
	国际学生比例	-	5.0%
	国际教师比例	-	5.0%
莱顿世界大学排名	科研影响力	P (科研论文总量)	-
		MCS (平均引文次数)	-
		MNCS (标准化的平均引用次数)	-
		PP10% (被引率最高的前10%的论文)	-
	科研协作力	PPcollab (与其他机构合作的论文比例)	-
		PP int collab (跨国合作的论文比例)	-
		PP ui collab (与企业合作的论文比例)	-
		MGCD (与其他机构合作的平均距离)	-
《美国新闻与世界报道》大学排名	学术声誉指标	全球学术声誉	12.5%
		地区学术声誉	12.5%
		论文总数	12.5%
	文献计量指标	归一化引用影响指数	10.0%
		总被引次数	10.0%
		高被引论文数	12.5%
		高被引论文比例	10.0%
	学校层面指标	国际合作	10.0%
		授予博士学位数	5.0%
		教师平均授予博士学位数	5.0%

表4 相关系数r值

大学排名	PR15F	PR14F	PR15A	PR14A
QS	-0.36	-0.39	-0.37	-0.40
莱顿	-0.41	-0.42	-0.40	-0.42
美国新闻与世界报道	-0.41	-0.44	-0.43	-0.45

过H篇。本文计算了2007—2014年每个会议的所有作者的H-index平均值,在Loizides等^[10]的论文中提到获取作者H-index时,会出现作者姓氏与其他论文作者相同的情况,需要手动检查出版物是否应该包含在H-index的计算中。本文在获取H-index时通过论文名称去查找

作者信息,有效避免了重名问题。

3.1.3 PImpact

评价会议的质量,研究会议发表的论文对科学界的影响至关重要,论文的被引量一直作为科研评价的一个重要指标。Meho等^[13]在其论文中指出谷歌学术在论文覆盖面上优于Scopus和WOS数据库。本文分别在谷歌学术、WOS和Scopus三个数据库上获得论文的被引量,发现在Scopus上获得的人工智能领域的被引量与谷歌学术上的被引量相差不大,但覆盖面要优于WOS。由此可知,在每一个会议中,论文的引用量是随着时间的推移不断增长的,Loizides的论文提出2~5年的时间窗口能更好地评价会议质量。Harzing^[14]在研究中指出会议论文不同于期刊论文,由于其传播速度快、传播效果明显,一般在发表3年后达到被引峰值,所以本研究选择时间窗口为3年的论文,截取会议数据最终到2014年,获得论文被引量截至2017年。

对于评价会议来说,由于会议的规模不同,不同会议的被引量或相同会议不同年份的被引量不能直接相互比较,除非对引文数量进行归一化处理,以纠正会议规模产生的误差。正如Waltman^[15]在专注于医学研究的工作中指出,即使是完全相关的研究领域也需要对规模进行规范化处理,这些领域因为规模的不同会有不同的引用实践。本文采用如下两个步骤计算会议的影响力。第一步:计算每一篇会议论文的影响力分数,见公式(1)。

$$P_i = \frac{P_j}{\sum_{j=1}^n P_j} \quad (1)$$

P_i 是每一篇论文的影响力分数,其中 p_j 是本研究中每一篇论文截至2017年11月15日的被引量, $\sum_{j=1}^n P_j$ 是每年A、B、C三类会议所有论文被引量的和, n 是每年所有会议发表的论文数。

第二步:计算每一个会议的影响力分数,见公式(2)。

$$PImpact = \frac{\sum_{i=1}^m P_i}{m} \quad (2)$$

PImpact指每一个会议的影响力分数, $\sum_{i=1}^m P_i$ 是每年每个会议的所有论文的影响力分数和, m 是每一个会

议的论文数,通过上述公式,可以得到每一个会议每年的归一化影响力分数。

3.1.4 零被引率

Ye等^[16]在评价学者学术影响力时提出“学者迹”的概念,指在评价学者时,不能只关注该学者的高被引论文和H-index,应对作者发表的所有论文进行综合考虑,既要考虑高被引,又要考虑低被引甚至是零被引论文,将一个学者的学术成果被引量用二维矩阵表示出来,通过计算各位学者的矩阵迹对学者进行排名。如果一些学者的矩阵迹中出现零值或小于零的值,那么在一定程度上说明该学者已经在走下坡路或者只是昙花一现,对学术界的影响力已经处于下滑阶段,对于大学尤其是“双一流”大学来说,学者迹是一个考量学者学术影响力的指标之一。

本文采用Ye等的零被引论文计算方法,很容易判断,质量越高的会议应该有越低的零被引率(Cite0),零被引率可由会议中零被引的论文数量除以所有论文数量计算得出。通过计算每一个会议的零被引率,以及会议排名的相关系数,得到系数绝对值大于0.5。

3.2 会议评价方法

3.1详细地介绍了本研究方法需要用到的指标,并且对每个指标的获取与处理做了说明,本节将介绍指标在会议评价研究中的具体应用。

3.2.1 数据预处理

将3.1中介绍的每个指标值整理如下,得到数据表5。表中Conference是会议名,Year是会议召开时间,

表5 学术会议的7个值(示例)

Conference	Year	Papers	Label	Level	PR15F	PR14F	PR15A	PR14A	H-index	PImpact	Cite0
AISTATS	2007	83	C	1	89.81	130.05	115.66	183.96	22.20	0.50	0.13
GECCO	2008	464	C	1	417.59	496.36	375.34	455.10	12.40	0.38	0.22
ACL	2010	262	A	3	156.63	216.38	164.28	227.77	12.90	1.32	0.05
PPSN	2010	127	B	2	332.24	407.16	320.55	404.38	13.40	0.54	0.10
AAAI	2010	456	A	3	225.70	284.43	213.85	274.70	15.80	0.53	0.22
EMNLP	2014	223	B	2	143.10	217.71	137.20	215.90	11.00	1.92	0.15

Papers是每年发表的论文数,Label是会议在CCF提供的《中国计算机学会推荐国际学术会议和期刊目录》中的等级,Level是将会议等级转换为对应的数字,3表示A级、2表示B级、1表示C级,为方便后续的数据处理,其后的列分别代表了3.1中提到的7个指标,指标名称做了简化处理,其中“F”代表第一作者,“A”代表所有作者。

3.2.2 相关分析

通过对每个指标与会议等级的相关性进行分析,得到两个指标之间相关系数的大小,进而得出指标在评价学术会议时的重要性排序。

由于本文的数据是非等距数据,且应用于学术会议评级,所以采用Spearman等级相关系数来衡量两

组数据的相关性。将表5中的数据与会议的等级进行Spearman相关性分析。

计算得出每个指标与会议等级相关性的正负,其中正相关的有PImpact和H-index,负相关的有Cite0、PR15A、PR14A、PR14F和PR15F。

PImpact指标在评价会议时与会议等级有很强的正相关性,可以理解为,质量越高的会议在学术界的影响力越大,所发表的论文得到的被引量越高,在同年的其他会议中脱颖而出,而质量低的会议平均影响力得分会低很多。本文所提出的零被引率指标经过相关性研究,发现其与会议等级关联比较强,说明这个指标可以应用到会议评价中,质量高的会议被认为存在更少的零被引论文,与分析结果一致。PR值对应的4个指标与会议等级呈现负相关,说明会议等级越高,其PR值越低。

H-index在评价会议时不如其他6个指标更具有可比性。

3.2.3 Topsis模型

本节采用Topsis模型对会议进行排名,具体步骤如下。

(1) 建立原始数据矩阵;

(2) 将原始数据中的7个指标与会议等级进行同趋势化处理,除PImpact、H-index这2个指标外,其余指标采用倒数法进行处理;

(3) 将(2)得到的数据进行无量纲化处理;

(4) 正负理想方案如下:

正理想方案A+= (0.265 306 571、0.333 458 745、0.255 589 752、0.307 194 146、0.103 943 245、0.253 609 901、0.427 646 153);

负理想方案A-=(0.013 594 623、0.017 749 833、0.014 780 225、0.020 088 078、0.027 667 289、0.002 408 146、0.003 634 992);

(5) 计算评价对象各指标值与正负理想方案的欧

式距离,并计算与正理想方案的接近程度,得到235个会议的排名。

其中,美国人工智能协会(American Association for Artificial Intelligence, AAAI)会议在排名上与CCF提供的排名不同,国际计算机视觉识别会议(IEEE Conference on Computer Vision and Pattern Recognition, CVPR)的论文被引量与作者质量要优于其他会议,这可能是由于多数计算机领域的杰出学者在选择上可能更倾向于容易被收录的CVPR会议。同时,研究发现同一会议的不同年份也有区别,其排名分先后顺序,本文进一步讨论同一会议的不同年份排名趋势,并对会议进行重新排名。

3.2.4 会议排名动态分析

从上文分析可知,每一个会议每年的排名不是固定不变的,是随着时间不断变化的,本节将对每一个会议,每一年的排名进行分析,部分示例见图1。

在A类会议中, CVPR排名比较靠前,从2007年

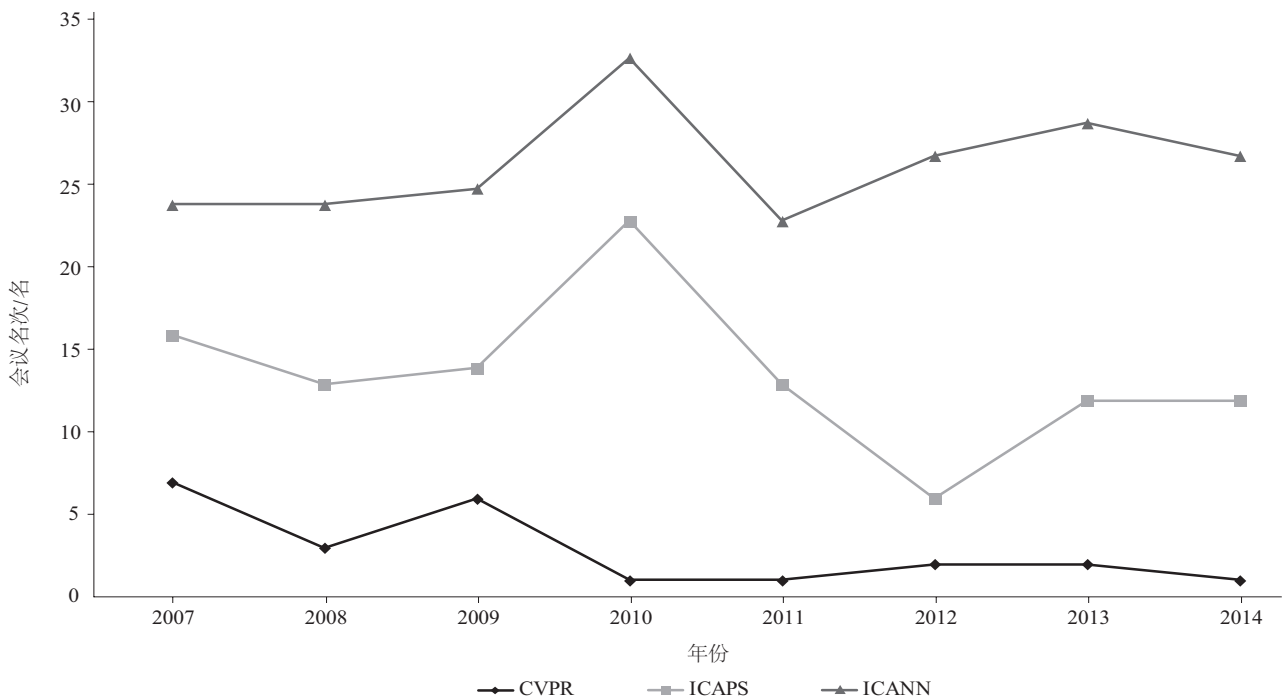


图1 会议历年排名(示例)

的第7名到近几年的第1名或第2名,正面反映了CVPR在学者中的影响,其被引量高于同类会议,且参会学者多来自于排名靠前的大学。CVPR在2010年发表了题为“ImageNet: A Large-Scale Hierarchical Image

Database”的论文,此篇文章被认为开创了采用机器学习进行图片识别比赛的先河。此后,每年的CVPR会议上都会刊载关于ImageNet的文章,吸引大量的优秀学者参加,很大程度上增加了该会议论文的被引量。

ICAPS (International Conference on Automated Planning and Scheduling) 是B类会议中排名幅度波动比较大的一个会议, ICAPS是关于智能规划和调度技术理论与应用研究成果的主要会议。该会议由之前的国际和欧洲规划会议合并而成, 近几年对该领域的研究越来越少, 导致该会议被重视程度越来越低, 由图1可以看出, 自2012年以来, 该会议的排名一直在下降。

在C类会议中, ICANN (International Conference on Artificial Neural Network) 是关于神经网络的会议, 但该领域属于小众领域, 所以在排名方面不如其他会议, 而且该领域的大部分学者对会议的重视程度不如其他领域的学者, 所以该会议排名较低。

由上述分析可知, 对每个会议的每年排名情况进行分析是必要的, 随着时间的推移, 会议更新速度与期刊相比迅速很多, 排名情况也不尽相同, 对每年排名分析将有助于学者更好地了解各个会议的情况。

3.2.5 会议综合分析

对每个会议的历年接近程度值取平均值, 将CCF提供的会议排名作为对照组, 与其进行对比分析, 得到每一个会议的排名情况(见表6)。

在表6中, 会议排名结果是按照平均程度接近值从

表6 会议排名结果

等级	本研究排名	CCF排名	重合率
A类	CVPR, COLT, ICML, NIPS, AISTATS, ECCV, ICCV	AAAI, CVPR, ICCV, ICML, IJCAL, NIPS, ACL	57.14%
B类	UAI, ILP, NAACL, ALT, EMNLP, BMVC, ICAR, FGR, IJCAI, ICAPS, ICRA, AAMAS	COLT, EMNLP, ECAI, ECCV, ICRA, ICAPS, ICCBR, COLING, KR, UAI, AAMAS, PPSN	41.67%
C类	ACL, KR, AAAI, PPSN, ACML, ACCV, IROS, ICB, ICPR, COLING, GECCO, IJCNN, ICDAR, ICCBR, ICANN, ICONIP, PRICAI, ICTAI, ECAI, KSEM	ACCV, CONLL, GECCO, ICTAI, ALT, ICANN, FGR, ICDAR, ILP, KSEM, ICONIP, ICPR, ICB, IJCNN, PRICAI, NAACL, BMVC, IROS, AISTATS, ACML	65.00%

小到大排列的, 重合率指在每类会议中, 本研究的会议排名结果与CCF提供的排名结果重合的比例。由表可知, 本研究对其中一些会议来说并不占优势, 如AAAI和ACL会议, 在CCF提供的会议名单中这两个会议皆是A类会议, 但采用本研究方法得到的会议结果, 将其分为C类会议。本文认为CCF提供的会议列表中, 有些会议举办历史悠久, 其在学界的地位举足轻重, 但随着时间的推移, 其影响或地位受到一些新兴会议的冲击, 导致引用量或排名有所下降。同时了解到CCF提供的每类会议目录是不分先后顺序的, 而本研究发现, 同类会议也存在差距, 同为A类会议也是分先后顺序的。

4 结论与讨论

(1) 本研究采用多个指标对会议进行定量动态评价, 与传统定性评价不同的是, 评价结果不受主观因素影响, 根据客观指标得出会议等级。由表6可以看出, 本研究得出的排名与CCF提供的排名重合率均在41.67%以上, 说明此方法在一定程度上适用于会议评价研究。需要补充的是, 本文考虑的会议都是质量较高的会议, A、B、C三类会议区分度不明显, 因此对其排名更加复杂, 在此条件下, 本研究方法能得到较好的排名结果, 说明其可以很容易地扩展到任何质量的会议。本文发现会议历年排名是浮动变化的, 且部分会议综合排名与CCF存在差异, 产生差异的原因与会议本身和指标选取都有关系, 下一阶段的研究会引入替代计量指标对会议进行评价。

(2) 本文共讨论了7个指标在评价方法中的应用, 其中PImpact与会议相关系数的绝对值最大, 并且与会议呈正相关, 说明其在会议评价中所占比重较大, 其次是Cite0。PR值选择第一作者或所有作者对结果影响不大。H-index指标在评价会议时的相关性不如PR指标, 本文认为这是由于H-index没有排除时间因素对学者的影响, H-index随着时间的推移是越来越大的, 后续研究在评价其他会议时将考虑时间因素对H-index的影响, 如谷歌学术提出的H5指标。后续会加入更多的指标到本研究方法中, 提高会议评价的准确率。

(3) Loizides等改进了Nelakuditi等提出PR指标计算方法, 认为计算论文所有作者PR值的效果要优于计算论文第一作者的PR值, 本文分别计算了PR(1/5)和PR(1/4)的第一作者和所有作者的值, 发现两者差距并不是很大, 在表4中四个指标与会议的相关系数都处

于0.4左右,可以作为评价指标采用。本文认为Loizides的数据经过筛选,是造成该问题的主要原因,为更快地计算PR值,可以只采用第一作者的值。此外,在收集作者机构来源时发现,人工智能领域会议的作者有一部分来源于公司、企业或研究所,不仅局限于大学,那么未来的工作会考虑机构排名对PR指数的影响,而不局限于大学排名。

参考文献

- [1] 冯长根. 重视学术交流作用 提高学术交流质量 [J]. 学会, 2003 (4): 6-8.
- [2] HIRSCH J E. An index to quantify an individual's scientific research output [J]. Proceedings of the National Academy of Sciences, 2005, 102 (46): 16569-16572.
- [3] EGGHE L. Theory and practice of the g-index [J]. Scientometrics, 2006, 69 (1): 131-152.
- [4] LALOË F, MOSSERI R. Bibliometric evaluation of individual researchers: not even right...not even wrong! [J]. Europhysics News, 2009, 40 (5): 26-29.
- [5] SOUTO M A M, WARPECHOWSKI M, OLIVEIRA J P M D. An ontological approach for the quality assessment of computer science conferences [M] //Advances in Conceptual Modeling—Foundations and Applications. Springer Berlin Heidelberg, 2007: 202-212.
- [6] MARTINS W S, GONCALVES M A, LAENDER A H F, et al. Assessing the quality of scientific conferences based on bibliographic citations [J]. Scientometrics, 2010, 83 (1): 133-155.
- [7] KUNGAS P, KARUS S, VAKULENKO S, et al. Reverse-engineering conference rankings: what does it take to make a reputable conference? [J]. Scientometrics, 2013, 96 (2): 651-665.
- [8] ELMACIOGLU E, LEE D. Oracle, where shall I submit my papers? [J]. Communications of the Acm, 2008, 52 (2): 115-118.
- [9] NELAKUDITI S, GRAY C, CHOUDHURY R R. Snap judgement of publication quality: How to convince a dean that you are a good researcher [M]. ACM, 2011.
- [10] LOIZIDES O S, KOUTSAKIS P. On evaluating the quality of a computer science/computer engineering conference [J]. Journal of Informetrics, 2017, 11 (2): 541-552.
- [11] SAKR S, ALOMARI M. A decade of database conferences: a look inside the program committees [J]. Scientometrics, 2012, 91 (1): 173-184.
- [12] ECKMANN M, ROCHA A, WAINER J. Relationship between high-quality journals and conferences in computer vision [J]. Scientometrics, 2012, 90 (2): 617-630.
- [13] MEHO L I, YANG K. Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar [J]. Journal of the American Society for Information Science and Technology, 2007, 58 (13): 2105-2125.
- [14] HARZING A W. A longitudinal study of google scholar coverage between 2012 and 2013 [J]. Scientometrics, 2014, 98 (1): 565-575.
- [15] WALTMAN L. A review of the literature on citation impact indicators [J]. Journal of Informetrics, 2016, 10 (2): 365-391.
- [16] Ye F Y, Bornmann L, Leydesdorff L. H-based 13-type multivariate vectors: multidimensional indicators of publication and citation scores [J]. COLLNET Journal of Scientometrics and Information Management, 2017, 11 (1): 153-171.

作者简介

张洋, 男, 1975年生, 博士, 教授, 博士生导师, 研究方向: 网络信息计量学、科研评价、期刊评价。
叶月, 男, 1991年生, 硕士研究生, 研究方向: 网络信息计量学、会议评价, E-mail: yy0455@126.com。

Research for Dynamic Evaluation of Academic Conferences Based on Multiple Indicators

ZHANG Yang YE Yue

(Sun Yat-sen University School of Information Management, Guangzhou 510006, China)

Abstract: In order to better evaluate the level of international academic conference, this paper proposes a research method of dynamic evaluation of academic conferences based on multiple indicators. Data of international academic conference is collected in the field of artificial intelligence from 2007 to 2014, which is including paper information of 64 929 papers in 235 conferences. The correlation analysis was carried out through 7 indicators and meeting grades, and the ranking of meetings was conducted by using Topsis, and the dynamic ranking and comprehensive analysis results of each meeting were obtained.

Keywords: Conference Evaluation; Peer Reputation; Academic Impact; Topsis Model

(收稿日期: 2018-06-15)