

• 信息计量与科学评价 •

多指标核心作者识别方法研究

张子婷 郑彦宁* 袁 芳

(中国科学技术信息研究所, 北京 100038)

摘 要: [目的/意义] 将无偏指标引入评价体系对某一研究领域的核心作者进行准确识别。本文构建了一个研究领域规范化的引文影响力指标, 提出了综合多指标的核心作者识别方法。[方法/过程] 本研究基于 Web of Science 数据库, 以某一研究领域的学术论文为数据基础, 利用普赖斯定律确定核心作者候选人, 综合发文量、被引频次、 h 指数以及研究领域规范化后的引文影响力 4 个指标, 利用综合指数法和熵值法确定每位学者的最后得分, 将高于 100 分的学者确定为该研究领域内的核心作者。[结果/结论] 最后, 利用上述方法在 DNA 计算领域进行了应用分析, 按照综合得分 $k > 100$ 的标准遴选出 DNA 计算领域的 32 位核心作者。

关键词: 核心作者; 识别方法; 引文影响力; 综合指数法; 熵值法

DOI: 10.3969/j.issn.1008-0821.2020.07.015

(中图分类号) G250.252 (文献标识码) A (文章编号) 1008-0821 (2020) 07-0144-08

Research on Multiple-Indicators Method on Identifying the Core Authors

Zhang Ziting Zheng Yanning* Yuan Fang

(Institute of Scientific and Technical Information of China, Beijing 100038, China)

Abstract [Purpose/Significance] An unbiased index was introduced into the evaluation system to accurately identify the core authors in a research field. This paper constructed a research field normalized citation impact index, and proposed a multiple-indicators method on identifying the core author. [Method/Process] Based on the Web of Science Core database this research, used articles of a research field as data basis. This paper used Price law to determine core author candidates. Based on the statistic data of quantity, cited frequency of papers, h -index and research field normalized citation impact, the core authors were evaluated by comprehensive index method. And scholars whose scores were higher than 100 were determined as the core authors in this field. [Result/Conclusion] Finally, the above method was used to identify the core authors in the field of DNA computing, and 32 core authors in this field were selected according to the criteria of comprehensive score $k > 100$.

Key words: core author; identifying method; citation impact; comprehensive index method; entropy method

作者是科学研究工作的主体, 其科研能力直接决定着科研产出的数量和质量^[1]。核心作者是指那些在各学科领域造诣较高、学术活动较频繁、发表论文较多且影响较大的作者。他们在学科中发挥着导向作用, 不断将学科研究推向新的纵深^[2]。对于领域内的研究人员, 关注高水平学者的科研成

果是快速把握前沿研究内容的重要方法之一。对于领域内的期刊, 也可面向核心作者组约优秀稿件, 提高期刊的学术水平。另外, 对在某一研究方向上潜心科研多年的学者们的研究成果进行分析, 可以发现该研究领域在某一研究方向上的主题演变过程。在这些过程中, 准确识别研究领域内的核心作

收稿日期: 2020-04-03

基金项目: 国家重点研发计划课题“全球创新主体创新感知系统”(项目编号: 2019YFA0707203); 博士后基金第 61 批之颠覆性技术的识别方法研究(项目编号: 2017M610970)。

作者简介: 张子婷(1996-), 女, 硕士研究生, 研究方向: 产业竞争情报。袁芳(1989-), 女, 博士后, 研究方向: 产业竞争情报。

通讯作者: 郑彦宁(1965-), 男, 研究员, 博士生导师, 研究方向: 竞争情报、情报学理论方法。

者显得尤为重要。需要特别指出的是,本文研究的核心作者是识别某一研究领域内有一定学术造诣及研究影响力的作者,并非确定某篇文章的多个作者中贡献较大的作者。

1 相关研究

目前已有一些利用综合指标确定核心学者的研究,主要应用于期刊层面。陈涛^[3]从普赖斯定律、 h 指数和总被引频次3个方面来确定出《档案学研究》的核心作者。赵基明等^[4]从作者单篇论文被引情况和作者全部论文被引次数两个角度测定《中国图书馆学报》的核心作者。钟文娟^[5]利用普赖斯定律确定核心作者候选人并结合作者发文量和被引量两个指标运用综合指数法测定出《图书馆建设》的核心作者。胡臻等^[1]运用普赖斯定律、综合指数法和期刊 h 指数等方法对《四川图书馆学报》2006-2015年核心与扩展核心作者进行分析。李智毅等^[6]综合运用发文量、第一作者发文量、综合贡献度、高引作者、高被引论文等多重方法,系统分析了文献作者的影响力及其合作关系,从中识别出一批核心学者群体。许晶晶^[7]利用CNKI 2004-2009年6年间建筑类图书所涉及的作者及其发文量、 h 指数、主题连续研究数,根据指标的重要性确定权重,最终遴选出Top100作为核心作者。此外,还有一部分研究着重对某一学科领域的核心学者进行挖掘,并基于此识别研究主题。Chien T W等^[8]探讨作者的研究领域及个人影响因素与作者姓名排序之间的关系,利用基尼指数进行评价。Wang Q等^[9]利用Insight软件进行作者网络分析,进一步解释了核心作者的贡献和趋势。支凤稳等^[10]利用主观提取法确定国外馆藏数字资源语义化领域的核心作者列表和作者合著网络。徐健等^[11]、江文华等^[12]通过普赖斯定律分别识别情报学领域和图书馆学领域的核心作者,在此基础上进行社群划分并分别识别各社群研究主题。田文灿等^[13]遴选出Altmetrics领域的23名核心作者,从微观层面上对这些核心作者的研究主题和方向进行演变分析。

研究指标选取方面,主要依赖发文量、总被引次数、篇均被引次数、 h 指数等,这些指标均未考虑文章发表年份、文献类型、学科领域等影响。例

如,在某一研究领域内有两篇分别发表于2009年与2019年的文章,其被引频次均为10,但其引文影响力并不能简单等同。由于受到出版年的影响,出现两篇文章出现被引频次相同但引文影响力明显不相同的情况。目前已有的无偏指标主要有基于爱思唯尔Scopus数据库的学科规范化的论文影响力(Field-Weighted Citation Impact, FWCI)和基于汤森路透Web of Science数据库的学科规范化的论文影响力(Category Normalized Citation Impact, CNCI)。这两种指标较可以排除出版年、学科领域与文献类型的影响^[14],常用于考察机构、国家、个人等的论文影响力。在应用层面,主要是为期刊刊发和约稿提供支持。目前极少有研究对某一研究领域的核心作者进行识别。

2 研究方法

2.1 研究思路

本文研究思路见图1。本文依据Web of Science数据库,以某一研究领域得到的学术论文为数据基础。对作者进行数据清洗后,利用普赖斯公式确定核心作者候选人,然后构建评价体系,最后确定权重、计算得分、遴选出研究领域内的核心作者。

在指标选取时,我们发现,被引频次虽可以衡量某篇论文的学术影响力,但是对于两篇被引频次相同的论文来说,发表年份、文献类型以及所处的学科领域不同都会导致其学术影响力并不等同。故本文除使用常用的发文量、被引频次、 h 指数这3个指标外,还构建了一个研究领域规范化后的引文影响力指标(Research Field Normalized Citation Impact, RFNCI),一定程度上排除出版年、文献类型、学科领域对被引频次产生的影响,具体指标定义见2.3。

在评价方法方面,本文采取综合指数法的无量纲化规则对4个指标进行无量纲化,利用熵值法确定各指标权重。得到权重后赋予每位学者一个综合评分,所得评分大于100分则表示其水平高于领域内的平均水平,确定其为核心作者。

2.2 确定核心作者候选人

为保证分析结果的准确性,本文采用DDA(Derwent Data Analyzer)软件消歧和人工校对相结合的方式对作者信息进行清洗^[11]。首先将姓名相

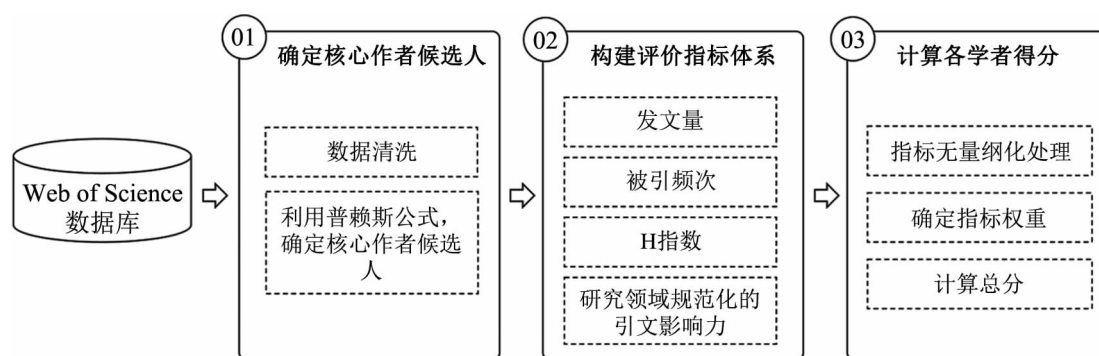


图 1 研究思路

同且作者一级单位相同的作者判定为同一作者，其次人工逐个分析姓名相同而单位不同的情况，查阅补充资料对这些作者进行处理。数据清洗完毕后，利用普赖斯公式 ($M=0.749\sqrt{N_{\max}}$) 计算出核心作者候选人的最少发文量。其中 M 为最小论文发表数， N_{\max} 为所数据集中最高产作者的发文量，发表论文数在 M 篇以上的作者即为核心作者候选人。

2.3 构建评价体系

本文选取发文量、被引频次、 H 指数以及研究领域规范化后的引文影响力 4 个指标。各指标的定义及优劣势如下：

1) 发文量：指在某一特定研究领域中，某一作者的全部发文数量，记为 D_i 。其中 i 为作者编号，下同。发文量体现学者在该领域的科研产出数量，但无法衡量科研成果的质量。

2) 被引频次：某作者的全部发文的被引你频次总数，记为 C_i 。被引频次体现科研成果质量，反应论文的被使用和受重视的程度^[15]，但受发文量及高被引论文影响。

3) h 指数：一位作者的 h 指数等于当且仅当具有 h 篇引文次数至少为 h 的论文，同时其余论文的引文数都小于 h 时的 h 值，记为 h_i ^[15]。 h 指数结合数量和质量两个方面综合分析学者的个人学术成就，用以评价学者的持续绩效。但是 h 指数仍存在以下局限：一是对低被引论文和高被引论文缺乏灵敏度；二是缺乏区分度；三是 h 值只升不降，无法反映学者的持续表现^[3]。

(4) 研究领域规范化的引文影响力 (Research Field Normalized Citation Impact, RFNCI)：该研究领域中的一篇文献的研究领域规范化的引文影响力 (RFNCI) 是通过其实际被引次数除以同文献类型、

同出版年、同研究领域文献的期望被引次数获得的。其中，同文献类型、同出版年、同研究领域文献的期望被引次数 (记为 e) 为该研究领域某一年所有的研究论文 (Articles) 总被引次数除以论文总数。

$$e_{r,t,d} = \frac{\sum_r \sum_t \sum_d c}{\sum_r \sum_t \sum_d p} \quad (1)$$

$$RFNCI = \frac{c}{e_{r,t,d}} \quad (2)$$

其中， e 为期望被引率， c 为总被引频次， p 为论文总数， r 为研究领域， t 为出版年， d 为文献类型。

对于一位作者在该研究领域的一组论文，RFNCI 的值为每篇论文的 RFNCI 的平均值。

$$RFNCI_i = \frac{\sum_{a=1}^{D_i} RFNCI}{D_i} \quad (3)$$

RFNCI 是一个相对无偏的影响力指标，排除了出版年、学科领域与文献类型的影响。在某一研究领域中，如果 RFNCI 的值等于 1，说明该 (组) 论文的被引表现与全球水平相当，如果 RFNCI 大于 1 表明该 (组) 论文的被引表现高于全球水平，否则低于全球平均水平。

2.4 指标无量纲化处理及权重确定

多指标系统的每一种评价方法，都面临着两个问题：一是按何种规则将每个指标无量纲化；二是以何种方式区分各指标的相对重要性^[16]。本文采取综合指数法的无量纲化规则对 4 个指标进行无量纲化，利用熵值法确定各指标权重。

2.4.1 综合指数法

综合指数评价法是运用数学方法处理多重相关

指标而得到的一个综合性指标的方法,用以较全面地反映事物总体变动方向和变动程度。综合指数法可以将多个不能同度量现象的数值,分别变换为能同度量数值。由于各评价指标的量纲、数量级及指标优劣的取向存在较大差异,在对这些指标进行合成之前必须进行标准化处理。采用统计学方法对指标进行处理,以消除多指标综合评价中计量单位差异和指标数量级、相对数形式差别^[17-18]。

设 n 为评价学者个数, p 为指标个数, x_{ij} 为第 i 个对象第 j 项指标值,分别求 x_{ij} 的均值 \bar{x}_j , 得:

$$\bar{x}_j = \frac{1}{n} \sum x_{ij} \quad (4)$$

按照下式将 x_{ij} 无量纲化:

$$x'_{ij} = \frac{x_{ij}}{\bar{x}_j} \times 100 \quad (5)$$

并称 x'_{ij} 为 x_{ij} 的折算指数,即无量纲化后第 i 个学者第 j 个指标的数值。

再利用熵值法求得各指标权重 ω_j , 进行加权计算,便可得到每位学者的综合得分。

$$k_i = \sum_{j=1}^p \omega_j x'_{ij} \quad (6)$$

2.4.2 熵值法

熵是系统不确定性的度量,可以用于度量已知数据所包含的有效信息量和确定权重^[19]。熵值法是一种客观赋权法,其根据各项指标观测值所提供的信息的大小来确定指标权重。在计算指标权重时,若某个指标中的各个数值之间变化不大,则该指标在综合分析中的贡献小,即权小,反之则权大^[20]。由于该方法计算结果可信度较大,自适应功能强,故本文选取熵值法作为确定权重的方法。其主要步骤如下:

计算第 i 个学者第 j 项指标值的比重:

$$Y_{ij} = \frac{x'_{ij}}{\sum_{i=0}^n x'_{ij}} \quad (7)$$

指标信息熵的计算:

$$e_j = -k \sum_{i=1}^m (Y_{ij} \times \ln Y_{ij}), \text{ 令 } k = \frac{1}{\ln m}, \text{ 有 } 0 \leq e_j \leq 1 \quad (8)$$

信息熵冗余度的计算:

$$d_j = 1 - e_j \quad (9)$$

指标的权重:

$$\omega_j = \frac{d_j}{\sum_{j=1}^p d_j} \quad (10)$$

3 应用分析

3.1 数据获取与清洗

3.1.1 数据来源

1994 年,美国加利福尼亚法学的 Adleman 博士在《科学》期刊上首次发表了关于 DNA 分子生物计算方法的开创性文章^[21]。此后,越来越多的学者在 DNA 计算领域进行研究与探索,不断将研究推向纵深。本文以科睿唯安的 SCI-E 数据库中 1994 年至今 DNA 计算领域的学术论文数据源,检索式为 TS=((“DNA Comput*” OR “DNA Based Comput”) OR (“Nucleic Acid” or DNA or “Nucleotide*” or “Oligonucleotide*”) AND (“Molecular Comput*” OR “Molecular Based Comput*” OR “Molecular Scale Comput*” OR “Biological Comput*” OR “Biomolecular Comput*” OR “Biocomput*”)) OR (“Data Structure” OR “Hamilton* Path Problem*” OR “Hamilton* Circuit Problem*” OR “Hamilton* Path” OR “Hamilton* Circle*” OR “SAT Problem*” OR “NP-complete Problem*” OR “NP-hard Problem*” OR (“0-1 Program* Problem*”)) AND (“Nucleic Acid” or DNA or “Nucleotide*” or “Oligonucleotide*”))) , 文献类型为 Article (排除 Proceedings Paper), 检索日期为 2019 年 11 月 8 日,得到 1 268 篇论文。

3.1.2 核心作者候选人确定

首先利用 DDA 和人工消歧方式对人名进行消歧,发现最高产作者的发文量为 46 篇。根据普赖斯公式,本文研究领域的 M 值为 5.080,即在 DNA 计算领域发文 5 篇及以上的学者为该领域的就核心作者候选人,共有 103 位。统计各项指标数值时发现有位作者的 H 指数为 0,这对后续熵值法权重计算会造成一定影响,考虑到其 $RFNCI$ 值为 0.042,远低于平均值,故从候选人集中剔除。最终得到核心作者候选人 102 位。

3.2 核心作者遴选

3.2.1 各指标值获取及无量纲化处理

1) 指标数值获取。每位作者的发文量及总被引频次都可直接获得, H 指数及 $RFNCI$ 值需要进

行一定的计算。其中，每年的同文献类型、同出版年、同研究领域文献的期望被引次数（记为 e ）如表 1 所示。由此计算每位学者每篇文献的 $RFNCI$ 值，进而计算每位学者的 $RFNCI$ 值。

表 1 1994–2019 年 DNA 计算领域学术论文的期望被引次数

年份	发文量	总被引频次	e
1994	3	2 293	764.33
1995	1	655	655.00
1996	11	229	20.82
1997	18	1 151	63.94
1998	18	452	25.11
1999	20	806	40.30
2000	20	2 829	141.45
2001	17	803	47.24
2002	46	1 360	29.57
2003	38	3 203	84.29
2004	44	1 885	42.84
2005	53	1 004	18.94
2006	42	810	19.29
2007	52	946	18.19
2008	56	2 340	41.79
2009	58	1 085	18.71
2010	70	1 590	22.71
2011	60	1 008	16.80
2012	78	1 618	20.74
2013	70	1 529	21.84
2014	82	1 371	16.72
2015	82	923	11.26
2016	68	615	9.04
2017	95	901	9.48
2018	88	307	3.49
2019	78	60	0.77

2) 无量纲化处理。利用式 (5) 对每位学者的 4 项指标数据进行无量纲化处理， x_{ij} 的均值 \bar{x}_j 如表 2 所示。

表 2 各指标数据的均值

文章数量 (\bar{x}_1)	总被引频次 (\bar{x}_2)	h 指数 (\bar{x}_3)	$RFNCI$ (\bar{x}_4)
9.03	212.29	4.80	1.05

3.2.2 指标权重计算

利用熵值法进行指标权重计算，具体步骤见 2.3.2。由式 (7) 计算每位学者各项指标的比重，由式 (8) 得到各指标信息熵 e_j 的值分别为 0.96、0.84、0.89 和 0.98；由式 (9) 得到信息熵冗余度分别为 0.04、0.16、0.11 和 0.02；最终由式 (10) 得到各指标权重分别为 0.11、0.50、0.33 和 0.06，如表 3 所示。

表 3 熵值法确定各指标权重

指标信息熵	信息熵冗余度	指标权重
e_1 0.96	d_1 0.04	ω_1 0.11
e_2 0.84	d_2 0.16	ω_2 0.50
e_3 0.89	d_3 0.11	ω_3 0.33
e_4 0.98	d_4 0.02	ω_4 0.06

3.2.3 核心作者确定

对各指标数据进行无量纲化处理以及确定权重之后，根据式 (6) 计算每位学者的综合得分，综合得分越高，代表候选人的综合学术水平越高，其对该研究领域的贡献也越大。笔者按照综合得分 $k > 100$ 的标准遴选出 DNA 计算领域内 32 位核心作者，具体名单见表 4。

由表 4 可以看出，共有 32 位学者进入 DNA 计算领域的核心作者行列。其中我国有 7 位学者，主要来自中国科学院、北京大学、中南政法大学、国立高雄科技大学、四川大学等机构。值得注意的是，102 位核心作者候选人中半数学者来自中国，但是进入核心学者行列的只有 7 名。这一现象表明，在 DNA 计算领域尽管我国学者发文较多，但在基础科研产出质量与影响力方面仍存在明显差距。综合得分第一的是 Paun G，其于 2000 年在期刊 JOURNAL OF COMPUTER AND SYSTEM SCIENCES 上发表的文章 “Computing with Membranes” 被引 1 207 次，包括 3 篇领域中的高被引论文，激发了中国、西班牙、罗马、意大利以及英国等 25 个国家的科研人员的研究思路。Paun G 的文章均发表于 1996–2002 年，属于 DNA 计算领域早期研究学者，奠定了该领域研究的理论基础。综合得分第二和第三的 Shapiro Ehud 教授和 Adar Rivka，均

表 4 DNA 计算领域核心作者的 4 项指标值及其综合得分

序号	作 者	所属机构	发文 年份	文章 数量	总被引 频次	<i>h</i> 指数	规范化 的引文 影响力	综合 得分
1	Paun G	罗马尼亚科学院	1996–2002	10	1 608	8	2.32	473.55
2	Shapiro Ehud	魏茨曼科学研究所	2001–2012	9	1 312	9	3.57	442.80
3	Adar Rivka	魏茨曼科学研究所	2001–2011	6	1 209	6	4.45	438.75
4	Benenson Yaakov	瑞士联邦理工学院	2001–2014	7	1 221	7	4.02	430.62
5	Winfree Erik	加州理工学院	1998–2019	9	946	7	4.12	371.46
6	樊春海	中国科学院	2005–2017	6	474	6	5.67	303.79
7	Stojanovic Milan N	哥伦比亚大学	2003–2019	6	742	5	3.44	295.93
8	Remacle Francoise	希伯来大学	2010–2015	5	508	5	4.79	282.00
9	Corn R M	威斯康星大学	1997–2001	8	832	8	1.80	271.92
10	Smith L M	威斯康星大学	1997–2005	10	828	10	1.62	270.47
11	Willner Itamar	希伯来大学	2005–2015	7	564	7	3.84	270.31
12	Stefanovic Darko	新墨西哥大学	2003–2019	11	738	6	2.10	260.39
13	Condon A E	威斯康星大学	1997–2002	6	730	6	1.72	240.52
14	Liu Q H	威斯康星大学	1997–2011	6	655	6	1.65	220.68
15	Keinan Ehud	利普斯研究所	2001–2012	5	510	4	2.61	212.74
16	Rozenberg Grzegorz	莱顿大学	1998–2007	12	554	10	1.34	199.41
17	Phillips Andrew	微软研究院	2009–2019	9	220	5	3.17	168.28
18	汪尔康	中国科学院	2011–2019	12	328	8	1.99	164.19
19	许 进	北京大学	2002–2018	46	333	11	0.35	159.17
20	董绍俊	中国科学院	2011–2019	9	249	6	2.03	140.67
21	陈 旭	中南政法大学	2002–2019	6	285	5	1.88	139.44
22	Kolpashchikov Dmitry M	佛罗里达州立大学系统	2005–2017	8	211	7	2.08	133.14
23	Kari Lila	滑铁卢大学	1996–2017	20	304	9	0.59	125.66
24	Chang Weng-Long	国立高雄科技大学	2003–2015	16	273	10	0.83	122.22
25	Chen Jianchi	四川大学	2012–2017	5	190	5	1.98	119.13
26	Hagiya Masami	东京大学	1999–2016	9	332	6	0.57	114.49
27	Deiters Alexander	宾夕法尼亚州立高等教育系统	2012–2017	5	179	5	1.85	112.42
28	Reif John H	杜克大学	1999–2019	12	189	7	1.40	111.56
29	Cardelli Luca	微软研究院	2009–2018	7	186	5	1.60	108.56
30	Song Tianqi	杜克大学	2016–2019	5	52	4	2.66	106.72
31	Ellington Andrew D	得克萨斯大学奥斯汀分校	1996–2018	8	217	6	1.22	106.40
32	Katz Evgeny	克拉克森大学	2008–2019	12	185	7	1.26	106.30

来自以色列的魏茨曼科学研究所 EHUD SHAPIRO 实验室，该实验室主要从事分子计算机、基因编辑、细胞谱系 3 个方面的研究，其中分子计算机项目的长期愿景是创建配备医疗知识的纳米生物计算

机，用于诊断疾病并自主治疗疾病。综合得分排名第六的学者是来自中国科学院的樊春海院士，樊春海院士在生物传感器、DNA 纳米技术与 DNA 计算以及生物光子学领域做出了系统出色的研究工作。

2017年,他作为中国代表参加了第23届DNA计算与分子编程国际会议,并发表演讲。这是领域内最重要的国际学术会议。2018年,他与另外两位教授作为共同主席举办了在中国举办的第24届DNA计算与分子编程国际会议。发文量最多的是来自北京大学的许进教授,发文46篇,主要从事生物计算数据编码与模型构建的理论和研究方法研究,先后5次任国际生物计算机大会主席,为生物计算机的发展做出了突出贡献。另外,32位核心作者在DNA计算领域共计发文238篇,而每位作者发文数单纯相加为312,说明32位核心作者之间存在一定的合作关系。

为了验证该方法的可行性,笔者利用检索式在Web of Science核心合集数据库中检索DNA计算领域的综述文章,选取了5篇综述类文章,将本文选取的核心作者名单与这5篇综述文章的参考文献作者进行匹配。这5篇综述类文章中,文献[22](2009,被引频次363)、文献[23](2019,被引频次34)、文献[24](2019,被引频次34)为领域内的高被引论文,另外选取了分别发表于2006年的文献[25](被引频次109)和2014年的文献[26](被引频次101)进行补充。经过匹配发现,32位核心作者中有26位作者的文献至少被这5篇文章中的一篇引用,一定程度上证明了本文研究方法的可行性及结果的准确性。

4 结 论

本文首先根据普赖斯定律确定核心作者候选人集合,然后构建基于某一特定研究领域的无偏评价指标 $RFNCI$,综合发文量、总被引次数和 H 指数等多个指标,利用综合指数法和熵值法对学者的学术贡献进行全方位评价。无偏评价指标在一定程度上减少了出版年、文献类型和研究领域对被引次数产生的影响,从更客观角度评价论文的影响力。本文在指标选取、权重确定均较为客观、科学。之后,利用本文提出的方法在DNA计算领域进行应用分析,从102位核心作者候选人中遴选出32位核心作者,包括7位国内学者。综合得分位列前3的作者分别是Paun G、Shapiro Ehud和Adar Rivka;樊春海是国内学者中得分最高的学者;许进是发文最多的学者。结合各位作者在领域内所获得的成

就,可以从侧面印证本文研究方法的可行性。

当然,本文仍存在一定的局限性。第一,对于作者的消歧目前的技术几乎无法做到完全精确,本文也不例外;第二,虽引进一定程度上消除出版年、文献类型、研究领域的影响的相对无偏的指标,但评价指标的设置方面仍未做到面面俱到。另外,检索式的制定过程虽力求完善但难免会出现部分文章未涵盖在内的情况。未来将朝着以下几个方向开展研究:1)核心作者对研究主题演化如何产生影响;2)如何在研究主题演化路径识别中体现核心作者的作用。

参 考 文 献

- [1] 胡臻,龙兴跃.基于普赖斯定律与综合指数法的核心作者和扩展核心作者分析——以《四川图书馆学报》为例[J].四川图书馆学报,2016,(3):74-76.
- [2] 苏志芳,张建中,胡惠芳.基于模糊综合评判的中文社科图书“核心作者”决策研究[J].图书情报工作,2010,54(1):42-45,41.
- [3] 陈涛.基于CSSCI的《档案学研究》h指数和h衍生指数研究及核心作者的测定[J].甘肃科技纵横,2010,39(5):100-102.
- [4] 赵基明,舒明全.基于CSSCI的《中国图书馆学报》h指数及核心作者测定[J].中国图书馆学报,2008,34(2):98-102.
- [5] 钟文娟.基于普赖斯定律与综合指数法的核心作者测评——以《图书馆建设》为例[J].科技管理研究,2012,32(2):57-60.
- [6] 李智毅,杨晓春,吴广印,等.基于文献计量的军民融合文献核心作者分析[J].情报工程,2018,4(5):105-115.
- [7] 许晶晶.基于引文分析的核心作者研究——以建筑类图书为例[J].图书馆,2015,(5):89-92.
- [8] Chien T W, Chow J C, Chang Y, et al. Applying Gini Coefficient to Evaluate the Author Research Domains Associated with the Ordering of Author Names A Bibliometric Study [J]. Medicine, 2018, 97(39):11.
- [9] Wang Q, Li R R. Research Status of Shale Gas: A Review [J]. Renewable & Sustainable Energy Reviews, 2017, (74) 715-720.
- [10] 支凤稳,郑彦宁,杜薇薇.国外馆藏数字资源语义化研究现状分析[J].现代情报,2018,38(12):126-32.
- [11] 徐健,毛进,叶光辉,等.基于核心作者研究兴趣相似性网络的社群隶属研究——以国内情报学领域为例[J].图书情报工作,2018,62(12):57-64.

- [12] 江文华, 徐健, 李纲, 等. 基于研究兴趣相似性网络的我国图书馆学研究社群分析 [J]. 现代情报, 2019, 39 (9): 21–27.
- [13] 田文灿, 胡志刚, 王贤文. 科学计量学视角下的 Altmetrics 发展历程分析 [J]. 图书情报知识, 2019, (2): 4–11.
- [14] 吴伟, 姜天悦, 余敏杰. 我国高水平大学基础研究与世界一流水平的群体性差距——基于学科规范化的引文影响力分析 [J]. 现代教育管理, 2017, (4): 18–23.
- [15] 崔建强, 刘文娟. 运动人体科学领域学者学术影响力分析 [J]. 沈阳体育学院学报, 2011, 30 (6): 27–31.
- [16] 秦寿康. 综合评价原理与应用 [M]. 北京: 电子工业出版社, 2003.
- [17] 韩晓明, 王金国, 石照耀. 基于主成分分析和熵值法的高校科技创新能力评价 [J]. 河海大学学报: 哲学社会科学版, 2015, 17 (2):
- [18] 陈辉, 林超辉, 夏承鹏, 等. 基于 PCA 和综合指数法的高水平理工科高校科技成果转化绩效评价体系构建 [J]. 科技管理研究, 2019, (22): 48–54.
- [19] 陆添超, 康凯. 熵值法和层次分析法在权重确定中的应用

- [J]. 电脑编程技巧与维护, 2009, (22): 19–20, 53.
- [20] 戴西超, 张庆春. 综合评价中权重系数确定方法的比较研究 [J]. 煤炭经济研究, 2003, (11): 37.
- [21] Adleman L M. Molecular Computation of Solutions to Combinatorial Problems [J]. Science, 1994, 266 (5187): 1021–1024.
- [22] Pepke S, Wold B, Mortazavi A. Computation for ChIP-seq and RNA-seq Studies [J]. Nature Methods, 2009, 6 (11): S22–S32.
- [23] Mergny J-L, Sen D. DNA Quadruple Helices in Nanotechnology [J]. Chemical Reviews, 2019, 119 (10): 6290–6325.
- [24] Simmel F C, Yurke B, Singh H R. Principles and Applications of Nucleic Acid Strand Displacement Reactions [J]. Chemical Reviews, 2019, 119 (10): 6326–6369.
- [25] Ezziene Z. DNA computing: applications and challenges [J]. Nanotechnology, 2006, 17 (2): R27–R39.
- [26] Stojanovic M N, Stefanovic D, Rudchenko S. Exercises in Molecular Computing [J]. Accounts of Chemical Research, 2014, 47 (6): 1845–1852.

(责任编辑: 陈 媛)

(上接第 143 页)

- [29] Priem J, Taraborelli D, Groth P, et al. Altmetrics: A Manifesto [EB]. <http://Altmetrics.org/manifesto/>, 2015.
- [30] Priem J, Groth P, Taraborelli D. The Altmetrics Collection [J]. Plos One, 2012, 7 (11): e48753.
- [31] Patel A, Chavda J. Measuring Research Impact: Bibliometrics, Social Media, Altmetrics, and the BJGP [J]. British Journal of General Practice, 2016, 66 (642): E59–E61.
- [32] Park H, Park H W. Research Evaluation of Asian Countries Using Altmetrics: Comparing South Korea, Japan, Taiwan, Singapore, and China [J]. Scientometrics, 2018, 117 (2): 771–788.
- [33] Rodriguez C T, de Araujo R F. Altmetrics and Citation Indicators Applied to Scientific Production in Science Open: Descriptive Analysis for Brazil, Spain and Portugal [J]. Bibliotecas-Anales De Investigacion, 2018, 14 (2): 124–137.
- [34] Sugimoto C R, Work S, Lariviere V. Scholarly Use of Social Media and Altmetrics: A Review of the Literature [J]. Journal of The Association for Information Science and Technology, 2017, 68 (9): 2037–2062.
- [35] Saeed-Ul Hassan, Imran M, Gillani U. Measuring Social Media Activity of Scientific Literature: An Exhaustive Comparison of Scopus

- and Novel Altmetrics Big Data [J]. Scientometrics, 2017, 113 (2): 1037–1057.
- [36] Thelwall M, Haustein S, Larivière V, et al. Do Altmetrics Work? Twitter and Ten Other Social Web Services [J]. Plos One, 2013, 8 (5): e64841.
- [37] Thelwall M, Wilson P. Mendeley Readership Altmetrics for Medical Articles: An Analysis of 45 Fields [J]. Journal of The Association for Information Science and Technology, 2016, 67 (8): 1962–1972.
- [38] Sud P, Thelwall M. Evaluating Altmetrics [J]. Scientometrics, 2014, 98 (2): 1131–1143.
- [39] Shu F, Lou W, Haustein S. Can Twitter Increase the Visibility of Chinese Publications? [J]. Scientometrics, 2018, 116 (1): 505–519.
- [40] Xia F, Su X, Wang W, et al. Bibliographic Analysis of Nature Based on Twitter and Facebook Altmetrics Data [J]. Plos One, 2016, 11 (12): e0165997.
- [41] Zoller D, Doerfel S, Jaschke R. Posted, Visited, Exported: Altmetrics in the Social Tagging System Bib Sonomy [J]. Journal of Informetrics, 2016, 10 (3): 732–749.

(责任编辑: 陈 媛)