

学术期刊“红点指标”的定义与应用*

马 峥 俞征鹿

中国科学技术信息研究所,100038,北京

摘 要 为了促进学术期刊加强稿源组织建设,引导期刊编辑出版单位关注和追逐学科发展,基于关键词统计,设计了学术期刊“红点指标”,用于测度期刊论文与其所在学科研究热点的重合度。“红点指标”的意义在于从内容层面对期刊的质量和影响潜力进行预先评估。这有别于通常采用的引文分析方法等后评估方法,为推进期刊建设工作和期刊评价工作开创了一个新的思路和新的角度。“红点指标”将应用于学术期刊调整定位方向、跟踪学科热点、优化稿源结构、凝练期刊特色、发现优质作者、挖掘竞争情报等方面。

关键词 学术期刊;红点指标;引证报告;关键词;热点

Definition and application of an indicator for academic journals: Red Point Factor (RPF)//MA Zheng, YU Zhenglu

Abstract To help academic periodicals attract more manuscripts, guide periodical editors and publishers to focus on and pursue the development of disciplines, based on key words statistics, this paper designs the Red Point Factor (RPF) for academic periodicals, which is used to measure the coincidence degree of journal papers and their subject research hotspots. The significance of the RPF is to evaluate the quality and potential impact of periodicals from the content level. This is different from the commonly used citation analysis methods, such as the post evaluation method. RPF will open a new way of thinking and new perspective in promoting the periodical construction and the journal evaluation. RPF will be applied to adjust orientation of academic periodicals, track subject hot spots, optimize the structure of manuscripts, compact periodical features, find quality authors, and excavate competitive intelligence.

Keywords academic journal; red point factor; CJCR; keyword; hotspot

Authors' address Institute of Scientific and Technical Information of China (ISTIC), 100038, Beijing, China

DOI:10.16811/j.cnki.1001-4314.2018.01.031

学术期刊作为科学研究活动的重要组成部分,具备报道和传播各个学科领域科学发现和技术创新成果的功能。通过对一个学科中重要的,具有代表性的学术期刊内容的归纳,可以观察到研究活动的热点分布情况。通常来说,一个学科领域中有较多的科研工作者集中开展研究工作,并且有较多的科研产出成果的研究方向,是该学科的研究热点方向或主题。对于一个学术期刊,其所刊载的内容和同领域的研究热点方

向或主题重合的程度,可以作为一个判断学术期刊定位是否合理、观察学术论文的传播效果、评估期刊影响力的拓展空间的参考指标和分析工具,进而为加强学术内容组织和稿源建设工作提供数据支持^[1]。本研究设计“红点指标”,用于测度期刊论文与其所在学科研究热点的重合度。

1 “红点指标”的定义

在红点指标的设计中,包含以下几个概念:

1) 关键词:学术期刊上刊载论文的关键词是根据论文所论述的问题、思路、方法、对象、创新点、应用等方面,选择与研究工作和论文内容密切的少量词语,用于辅助读者、文献管理者、信息资源组织建设者等开展相关工作,更准确高效地检索和组织论文文献资源。一般情况下,关键词会优先选用主题词,即从自然语言中优选出来并经过规范化的、表达各种概念的语词^[2]。在主题词之外的自由词作为补充,也可用于标引关键词,选用的自由词须具备概念明确、语义专指、词形简练、实用性强的特点。通常认为,关键词能反映一篇论文的研究方向、研究主题。关键词是文献计量学领域广泛使用的文献特征信息,关键词的聚类研究、共词研究等方法,都是发现文献分布规律、关联强度的可靠手段,大数据基础上能够实现对知识机构和领域发展的客观判断。通过关键词的分析可以实现学科热点的发现和监测^[3]。

2) 高频关键词集合:在特定的时间窗口内,在同一个学科类别中,根据全部期刊所发表的论文的关键词统计,按照出现频次从多到少排序,能列入出现频次前1%的关键词,定义为特定时间窗口内该学科的高频关键词。高频关键词集合所组成的研究方向和主题的集合表征了特定时段,一个学科中的研究热点。在实际应用中,特定的评价时间窗口可以根据评价和研究工作的目的进行调整。科研热点转移的速度与学科发展进步的速度有关,但是一般评价时间窗口的宽度不宜超过1年。从提高可操作性的角度设计,以论文出版年作为评价时间窗口统计高频关键词较为适当。在科学计量学应用中,常采用通过高频关键词统计方法对学科领域发展进行分析^[4],或对某一个具体主题或研究方向的科研现状进行描述^[5]。

* 国家自然科学基金项目(15BTQ059)

3)红点指标 RPF(Red Point Factor): 一篇论文可能标引多个关键词,若其中有任一个或一个以上的关键词属于高频关键词集合,就认定该论文的关键词与高频关键词相重合。在评价时间窗口内,被评价期刊发表的关键词与其所在学科同期高频关键词重合的论文,在该期刊同期发表全部论文中所占的比例,定义为该期刊的红点指标。其定义式为

$$F_{RP} = \frac{N_{\text{fkw}}}{N_{\text{tot}}},$$

式中 F_{RP} 为期刊的红点指标, N_{tot} 为论文总数, N_{fkw} 为与高频关键词重合的论文数(关键词 \in {高频关键词集合})。

红点指标数值最高为 100%,代表该期刊评价时间窗口内所有发表的论文的关键词都能与同期同学科高频关键词相重合;数值最低为 0,代表该期刊评价时间窗口内发表的所有论文的关键词与同期同学科高频关键词都不重合。

学术期刊所发表的文献类型可能非常丰富,并不是所有类型的文献都用于报道学术发现和技术创新成果。为了更加准确清晰地反映研究活动的热点分布和变化,计算红点指标的论文范围限定于学术论文,包括论著、研究述评、一般论文和工业工程设计等类型^[6]。

2 中国科技核心期刊“红点指标”的分布

中国科技核心期刊(中国科技论文统计源期刊)是中国科技论文与引文数据库(CSTPCD)的来源期刊。CSTPCD 始建于 1987 年,由中国科学技术信息研究所研发。其建设目的是通过对中国科技人员在国内外发表论文的数量和被引用情况进行统计,对我国科研活动的产出规模、效率、质量和影响力进行分析和监测。CSTPCD 收录中国科技核心期刊重点选择各个领域最重要的、最具影响力的、能覆盖本领域学术发现和技术创新成果的期刊。2016 年度,CSTPCD 收录的中国科技核心期刊包括 2 008 种自然科学领域期刊,大约占我国科技期刊总数的 1/3,覆盖了自然科学综合、理学、农学、医学、工程技术、管理等 6 大部类 112 个学科类别。同时 CSTPCD 还收录 395 种社会科学领域核心期刊,覆盖社会科学综合、人文艺术、政治经济、传播教育、统计等 5 大部类 40 个学科类别^[7]。因此利用 CSTPCD 作为数据来源,可以较全面地观察了解各个学科类别期刊论文的研究方向和研究主题的分布,可以较准确地统计发现各个学科类别的研究热点。

若一篇论文同时著录了相同概念的中文关键词和英文关键词,在统计时分别计为 2 个不同关键词。根据 CSTPCD2016 统计,全国中国科技核心期刊共发表

了论文 56 万 4 857 篇,著录使用了 149 万 3 974 个关键词,出现 414 万 8 377 次。每篇论文平均使用 7.3 个关键词,超过 70% 的论文设置了 5 ~ 10 个关键词。在每个学科类别中分别计算 2016 年度高频关键词集合,并以此为基础计算各学科类别中每种期刊 2016 年度红点指标。

统计发现,CSTPCD 收录的全部 152 个学科类别(包括自然科学领域和社会科学领域)共 2 396 种期刊中,红点指标在 [75%, 100%] 区间的期刊有 213 种,分布在 73 个学科。其中红点指标达到 100% 的期刊共有 4 种,分布在 4 个学科。红点指标在 [50%, 75%) 的期刊有 1 019 种,分布在 124 个学科。红点指标在 [25%, 50%) 的期刊有 1 062 种,分布在 147 个学科。红点指标在 [0, 25%) 的期刊有 313 种,分布在 109 个学科,其中红点指标数值为 0 的期刊共有 29 种,分布在 13 个学科。其中部分期刊属于交叉领域,分别同时归入 2 个或 3 个学科,并在多学科内计算相应的红点指标。

3 “红点指标”的应用

“红点指标”的意义在于从内容层面对期刊的质量和影响潜力进行预先评估。这有别于通常采用的引文分析方法等后评估方法,为推进期刊建设和期刊评价工作开创了一个新的思路和新的角度。对于期刊编辑出版单位来说,可以从以下几个方面应用“红点指标”辅助开展相关工作:

1) 准确了解学科发展动向,调整期刊定位和方向。学术期刊的核心任务是以论文的形式报道和传播学术研究活动和技术开发活动的成果。对于学术期刊的编辑出版单位来说,只有准确了解学科的发展动向,才能精准把握、及时调整学术期刊的定位和发展方向,力求顺应学科发展态势,提高学术出版传播的效率和价值。通过各个学科“红点指标”计算过程中所提取的高频关键词集合的构成,可以了解一个年度时间截面上,学科研究活动的分布状态。若是持续监测同一学科的高频关键词集合随时间变化的情况,就可以准确了解该学科的发展动向。

2) 辅助编委会对学科热点布局做出判断,提高追踪和把握热点能力。学术期刊的编委会成员,包括主编或执行主编、副主编和编委,对期刊所在学科的热点是最具有发言权的,他们来自科研活动的第一线,具备广阔的学术视野、丰富的信息来源、敏锐的学术观察能力,并能基于此对当前和未来的学科热点形成准确的判断。然而这样的判断往往存在较大的主观性,容易形成较大的分歧,而且并不是所有编委会成员都能全

身心投入到期刊建设工作中^[8],因此编辑出版单位通过对“红点指标”的分析,可以梳理和归纳期刊所在学科热点分布状态,作为客观数据支持,辅助用于提高自身对学科热点的追踪和把握能力。

3) 优化稿件的构成结构,提高约稿组稿针对性。学术期刊所刊登的文献具备多种形式,研究内容上也具备多种选择。如果按照内容与学科内主流研究活动的关注点之间的关系,大体可以分为:①前沿探索性研究;②热点关注性研究;③非主流冷门性研究;④跨学科交叉性研究;⑤跟踪模仿性研究。每一种期刊根据自身的定位,可能会发表不同类型的研究成果,而且各种类型成果论文之间的数量结构有所侧重。稿件的构成结构是否合理取决于期刊的定位与实际情况是否一致。通过对“红点指标”构成的分析,可以测度一个学术期刊中,发表的“热点关注性研究”成果的数量和比例,用于帮助期刊编辑出版单位根据自身发展目的,认识实际稿源构成与目标的差距,从而调整优化稿件的构成结构设计,并进一步为有针对性地开展约稿和组稿工作提供参考。

4) 凝练期刊的学科特色,增加对优质稿源的竞争力。学术期刊对优质稿源的竞争力来源于期刊在学科中所具备的特点。在一些学科领域,期刊同质化现象非常明显,期刊之间特点不突出,削弱了期刊对学术资源的竞争力,不利于形成促进学术出版可持续发展的良性竞争环境,也不利于学术期刊扩大影响,形成品牌效应。通过“红点指标”的计算,可以发现比较期刊在学科中报道和关注的研究主题与方向。若某一种期刊能跟踪到某一个主题方向上较多的热点关注性论文,则表明期刊在这个方向上有潜力汇聚相对较多优质稿源。通过对一种期刊具备挖掘优质稿源的潜力学科的梳理,可以凝练出期刊的学科特色。将期刊各方面的资源集中到特色建设中,可以较有效地增加期刊对优质稿源的竞争力。

5) 发现优质作者群,识别学科内的核心研究团队。通常来说,各个学科的核心研究团队所从事的研究工作,主导着各学科的研究活动的主流方向,代表着学科发展进步的关键学术节点。这些核心研究团队是学术期刊追求的优质作者群体,它们的研究活动往往是学科内具有较高学术水平的研究探索,它们的论文往往能产生较广泛和长远的学术影响。通过“红点指标”的计算,可以发现那些长期有能力围绕学科热点研究方向和主题开展研究活动并发表论文的作者群。这个作者群所在的研究团队,能够发现、跟踪学科发展

方向,甚至能引领学科的发展方向。与高被引论文和高被引作者团队的统计数据相结合分析,可以发现和识别各个学科内的核心研究团队。以作者与关键词作为变量的关联分析与耦合分析可以进一步挖掘科研人员之间、科研团队之间的协作网络^[9]。

6) 挖掘竞争情报。学术期刊所处的环境,是一个激烈竞争的环境^[10]。在期刊之间、期刊编辑出版单位之间、期刊主办单位之间、期刊主管单位之间,都存在着相互竞争的关系。通过对竞争对手与自己各项学术指标的对比分析,可以准确发现各自长处和不足,及时采取对应措施,在竞争中取得先机。通过对对标期刊或期刊群的“红点指标”的数值构成的分析,可以了解对手把握学科热点的能力、在热点之外所关注的前沿或冷门领域的分布情况、核心作者群与热点研究领域的联系等方面的信息,与其他指标分析结果及相关信息进行汇总,挖掘竞争情报,促进期刊在激烈竞争中不断发展进步。

4 参考文献

- [1] 汪艺桦. 跟踪报道热点课题保持期刊与学科同步发展[J]. 中国科技期刊研究, 1997, 8(增刊1): 91
- [2] 文献主题标引规则:GB/T 3860—2009[S]. 北京: 中国标准出版社, 2009
- [3] 潘玮, 牟冬梅, 李茵, 等. 关键词共现方法识别领域研究热点过程中的数据清洗方法[J]. 图书情报工作, 2017, 61(7): 111
- [4] 齐书宇, 胡万山. 2005—2014年国内科技管理研究重要文献高频关键词共现分析[J]. 中国科技论坛, 2016(3): 12
- [5] 夏火松, 龙瑾, 李芳, 等. 基于高频关键词的离群点监测与异类知识研究: 从文献分析视角[J]. 情报杂志, 2017, 36(5): 181
- [6] 马峥. 科技期刊刊载文献的类型分布和传播规律: 以中国科技核心期刊为例[J]. 编辑学报, 2017, 29(2): 108
- [7] 潘云涛, 马峥. 2017年版中国科技期刊引证报告(核心版)自然科学卷[M]. 北京: 科学技术文献出版社, 2017
- [8] 陈翔. 学术期刊编委会履行职责中常见问题分析[J]. 编辑学报, 2007, 19(6): 453
- [9] 牟冬梅, 琚沅红, 郑晓月, 等. 基于学科内容的科研人员隐性合作关系研究[J]. 情报理论与实践, 2017(7): 17
- [10] 马峥, 潘云涛, 武夷山. 基于引文分析的科技期刊竞争压力评价及学科间比较研究[J]. 情报学报, 2013, 32(10): 1026

(2017-12-04 收稿;2017-12-28 修回)