

# 基于检索特征的科技论文可信等级的评估方法\*

周静<sup>1</sup>, 曾国荪<sup>1</sup>, 曾媛<sup>2</sup>

(1. 同济大学 计算机科学与技术系, 上海 200092; 2. 上海理工大学 管理学院, 上海 200093)

**摘要:** 期刊的影响因子、特征因子、 $h$ -指数等各种评价指标, 为查找高质量的期刊提供了参考依据, 但仍不全面, 且存在不足。为此, 基于传统检索特征, 增加考虑期刊影响力随时间衰减的特性, 突出和扩大对低影响力期刊的区分度, 提出一个增强的论文综合评价公式, 并通过算法形成论文的可信等级。最后, 利用 JCR 统计的数据进行实例分析, 并与期刊的影响因子、特征因子的排序进行比较, 验证其有效性。

**关键词:** 科技论文; 检索特征; 可信等级; 综合评估

**中图分类号:** TP391      **文献标志码:** A      **文章编号:** 1001-3695(2013)03-0820-05

**doi:**10.3969/j.issn.1001-3695.2013.03.046

## Assessment method of trust level for scientific papers based on retrieval features

ZHOU Jing<sup>1</sup>, ZENG Guo-sun<sup>1</sup>, ZENG Yuan<sup>2</sup>

(1. Dept. of Computer Science & Technology, Tongji University, Shanghai 200092, China; 2. School of Management, University of Shanghai for Science & Technology, Shanghai 200093, China)

**Abstract:** There are a variety of evaluations such like journal impact factor, the eigenfactor and  $h$ -index helping to find high-quality journals, but they are still not comprehensive. Based on the traditional retrieval features, this paper considered the journal impact's attenuation characteristics over time, highlighted the degree of distinction between the low influential journals, and proposed an enhanced paper evaluation formula and the algorithm to form the paper's credibility rating. Finally, it used the JCR statistical data to do case analysis and compared with journal eigenfactor to check its effectiveness.

**Key words:** scientific papers; retrieval features; trust level; comprehensive assessment

随着互联网的发展和普及, 网络上的信息越来越多, 形成了一片信息海洋。这虽然促进了知识的传播与共享, 但同时也带来了很多问题, 如大量论文出现在网络上, 使得网络论文数据库不断扩容, 网络论文质量参差不齐, 其检索返回不够精确, 在海量的在线论文中读者很难找到满意和可信的论文。

针对上述的问题, 目前国内外已经有一些研究人员提出一些解决方法。1972年, 美国科技信息研究所所长 Garfield 博士<sup>[1]</sup>在《科学家》(The Scientists) 期刊中叙述了影响因子的产生过程, 他最初提出影响因子的目的是为《现刊目次》(Current Contents) 评估和挑选期刊。目前人们所说的影响因子一般是指从 1975 年开始, 《期刊引证报道》(Journal Citation Reports, JCR) 每年提供上一年度世界范围期刊的引用数据, 给出该数据库收录的每种期刊的影响因子。在这样的背景下, 2007 年, 美国华盛顿大学和加州大学的 Bergstrom 等人<sup>[2]</sup>组成的研究小组发布了一个新的期刊引文评价指标, 即 eigenfactor, 它与影响因子不同, eigenfactor 的基本假设是: 期刊越多地被高影响的期刊所引用则其影响力也越高。2005 年, 美国物理学家 Hirsch<sup>[3]</sup>提出了将论文发表数量与论文被引数量相结合的复合指标, 即  $h$ -指数。自 2005 年提出以来,  $h$ -指数得到了国内外情报学界和科技期刊界的广泛关注, 并由用于科学家个人评价迅速扩展到机构、地区、国家的期刊、基金资助项目、学科研究热点等方

面的科学评价。随着  $h$ -指数被选做科技论文的考核指标越来越普及, 许多学者发现其存在不足, 继而又加以改进。比利时著名科学计量学家 Egghe<sup>[4]</sup>认为: 在评价科学家的科研绩效时, 应充分考虑到高被引论文的贡献, 为此提出了  $g$ -指数。虽然国内的一些研究人员如金壁辉<sup>[5]</sup>等人也提出了相关改进, 但依然没有对科技论文的分级给出可靠和可信的评价指标。因为诸如考虑论文作者贡献的不同的权重引用次数、跨学科的论文比较, 以及如何观察科学家的研究活力的衰退状况等问题都还没有得到很好的解决。为此, 本文基于传统的论文检索评价指标, 对现有的几种科技论文的评价策略加以利弊分析, 并在此基础上予以完善, 提出一个增强的论文综合评价公式, 旨在对科技论文的可信分级提供进一步完善的方法, 对未来网络论文平台的发展有一定的指导意义。

### 1 科技论文的影响因子分析

影响因子 (impact factor, IF) 是美国科学情报研究所 (ISI) 的期刊引证报告中的一项数据, 指的是某一期刊的文章在特定年份或时期被引用的频率, 是衡量学术期刊影响力的一个重要指标。为了叙述准确, 下面给出形式化定义。

**定义 1** 影响因子。假设某期刊  $S$ , 在评估的时间段  $T$  内

收稿日期: 2012-07-24; 修回日期: 2012-09-04      基金项目: 国家“863”计划资助项目(2009AA012201); 国家自然科学基金资助项目(61103068); NSFC-微软亚洲研究院联合资助项目(60970155); 国家教育部网络时代的科技论文快速共享专项研究课题(20110740001); 国家教育部博士点基金资助项目(20090072110035); 上海市优秀学科带头人计划资助项目(10XD1404400)

作者简介: 周静(1989-), 女, 江苏扬州人, 硕士研究生, 主要研究方向为内容可信、信息安全(zj890829@126.com)。

发表的论文总数为  $Y$ ,  $T$  一般为两年,再假设在统计当年  $U$  的被引用总次数  $X$ ,则该期刊的影响因子  $IF_U = X_{(s,T)} / Y_{(s,T)}$ 。

为了清楚理解影响因子的定义,下面举一个例子来说明。

例1 假设某刊2009年的发表论文数量为154篇,2010年内发表论文数量为187篇。在2011年统计评估时,2009年发表论文中被引次数为128次,2010年发表论文中被引用次数为48次。因此,该刊在2009—2010年期间发表的论文总数量为  $154 + 187 = 341$  篇;2009年和2010年发表论文的被引用次数一共为  $48 + 128 = 176$  次。所以,该刊2011年的影响因子为  $IF_{2011} = 176 \div 341 = 0.5161$ 。

根据影响因子  $IF$  的计算规则,对美国《科学引文索引》(Science Citation Index, SCI) 收录的著名刊物进行评估,2011年影响因子前十名的刊物如表1所示。

表1 2011年SCI收录期刊的影响因子排名表<sup>[6]</sup>

rank	abbreviated journal title(S)	ISSN	total cites(X)	IF(T=2)	IF(T=5)
1	CA - Cancer J CLIN	0007 - 9235	9801	94.262	70.216
2	Acta Crystallogr A	0108 - 7673	13944	54.333	24.717
3	New Engl J Med	0028 - 4793	227674	53.484	52.362
4	Rev Mod Phys	0034 - 6861	29868	51.695	48.621
5	Annu Rev Immunol	0732 - 0582	16100	49.271	46.688
6	Nat Rev Mol Cell Blo	1471 - 0072	26837	38.650	41.576
7	Nat Rev Cancer	1474 - 175X	26727	37.178	37.878
8	Nat Genet	1061 - 4036	76301	36.377	32.701
9	Nature	0028 - 0836	511145	36.101	35.241
10	Nat Rev Immunol	1474 - 1733	21080	35.196	33.644

从表1可以看出,论文的影响因子是相对统计值,可以克服规模不同的期刊由于载文量不同所带来的偏差。一般来说,论文的影响因子越大,其学术影响力也越大。但随着研究的深入,影响因子也暴露出诸多缺陷<sup>[7]</sup>,比如容易被人为操纵、统计错误、不能跨学科比较、选源标准问题以及对非英文期刊不公平等。另外,期刊影响因子隐含假设:在剔除论文数量因素后,期刊越多地被引用则其影响力越高。显然,此假设成立需要满足“所有引文重要性等价”的条件。但是每条引文的重要性是否等价是值得讨论的,例如在总被引频次相等的情况下,被《Nature》或《Science》大量引用的论文,其影响力显然会大于只被一些低水平期刊引用的论文。因此,如果要更为准确地反映期刊论文影响,还需考虑每条引文本身的价值。影响因子对于期刊的评价远非完美,理论上和实践上都需要更为准确的期刊评价指标。

## 2 科技论文的特征因子分析

为了弥补影响因子评价论文质量时的不足,2007年,Bergstrom<sup>[2]</sup>提出了论文的特征因子(eigenfactor)概念。Eigenfactor中的词头“eigen-”源自德语,是characteristic(特征、固有、本征的意思),故eigenfactor翻译为特征因子。特征因子具体由特征因子分值和论文影响分值来刻画。

定义2 特征因子分值(eigenfactor score)。假设某期刊

$S$ ,在评估的时间段  $T$  内发表的论文总数为  $Y$ ,  $T$  一般为五年,再假设在统计当年  $U$  的被引用总次数  $X$ ,则该期刊的特征因子分值  $EF_U = X_{(s,T)} / Y_{(s,T)}$ 。

初看上去,期刊  $i$  的特征因子分值  $EF_i$  的定义与影响因子  $IF$  的定义类似,但是特征因子以期刊影响力为权,以更贴近实际的权重网络形式重构了简单的引文网络,其算法较为复杂,详见第4.2节。为了清楚理解特征因子分值的定义,下面举一个例子来说明。

例2 假设读者到图书馆,随机选择了一本期刊的一篇文章,读过这篇文章之后又随机选择了这篇文章的一条参考文献,进入被引期刊,随机地在被引期刊上选择一篇文章;然后再随机地选择这篇文章的一条参考文献,到达另一本期刊,重复进行此过程。该研究者花在某一期刊上的时间与他呆在图书馆的时间的比值,即访问每本期刊的频率,在本质上就是期刊的特征因子值。例如,2007年通过对所有期刊的遍历检索,可得到《Nature》的特征因子分值为1.83870,其意义就是读者在图书馆所花时间的1.8387%里直接访问了《Nature》。通常被一个权威期刊引用的文章,读者会多花些时间认真阅读;另外假如一篇文章被众多期刊引用,在随机选择的过程中,难免会多次被选中。这两种情况都使得读者花在这篇文章所在期刊上的时间较多,也就是特征因子分值较大。

与影响因子  $IF$  比较,期刊特征因子分值的优点<sup>[8,9]</sup>主要有:a)特征因子考虑了期刊论文发表后五年的引用时段,而影响因子只统计了两年的引文时段,后者不能客观地反映期刊论文的引用高峰年份;b)特征因子对期刊引证的统计包括自然科学和社会科学,更为全面、完整;c)特征因子的计算扣除了期刊的自引;d)特征因子的计算基于随机的引文链接,通过特征因子分值可以较为合理地测度科研人员用于阅读不同期刊的时间。

定义3 论文影响分值(article influence score)。假设期刊  $S$  在五年的时间窗口中发表论文数量为  $P_s$ ,在统计内的所有期刊发表论文的总数为  $Q$ ,该期刊所发论文的标准化比值为  $N_s = P_s / Q$ ,论文的特征因子分值为  $EF_U$ ,那么在统计当年  $U$ ,论文影响分值  $AI_U = EF_U / N_s$ 。

论文影响分值旨在测度期刊的相对重要性,论文影响分值的平均值为1.00,大于1.00表明期刊中每篇论文的影响力高于平均水平;小于1.00则表明期刊中每篇论文的影响力低于平均水平。

根据特征因子分值  $EF$  和论文影响分值  $AI$  的法则,对2011年SCI收录的著名期刊重新进行评估,按照特征因子分值的降序排列如表2所示。

由表2可知,强大公众影响力的《Nature》,按影响因子  $IF$  排序,只列在第九名,而按照特征因子分值  $EF$  排序,则排名第一,因此考虑论文质量的评价指标更加让人信服。但是不可否认特征因子也有不足之处<sup>[10]</sup>:a)对于影响力较低的期刊群来说,它们的特征因子分值很低,区分度不大,在我国的SCI期刊中,影响力较低的期刊在特征因子分值的小数点后第4位数才开始出现差别,并且还有多种期刊出现重值现象;b)特征因子计算的数据封闭性比较强,计算准确性目前仍难以检验。

表2 2011年SCI收录期刊的特征因子排名表<sup>[6]</sup>

EF rank	IF rank	Abbreviated journal title(S)	ISSN	total cites	IF (T=5)	EF (T=5)	AI (T=5)
1	9	Nature	0028-0836	511145	35.241	1.74466	19.334
2	151	P Natl A Cad Sci USA	0027-8424	482679	10.591	1.66833	4.959
3	15	Science	0036-8075	469704	31.769	1.46485	16.859
4	238	Phys Rev Lett	0031-9007	335409	7.154	1.24359	3.486
5	427	J Biol Chem	0021-9258	412003	5.498	0.88585	2.191
6	182	J Am Chem Soc	0002-7863	369164	8.979	0.86714	2.753
7	918	Phys Rev B	1098-0121	268704	3.362	0.7833	1.389
8	890	Appl Phys Lett	0003-6951	197445	3.845	0.72217	1.399
9	14	Cell	0092-8674	167587	34.929	0.70472	20.639
10	3	New Engl J Med	0028-4793	227674	52.362	0.69167	21.366

### 3 科技论文的h-指数分析

2005年,美国物理学家 Hirsch<sup>[3]</sup> 提出了一项用于科学家个人科研绩效评价的文献计量学指标,即 h-指数。自提出以来,h-指数得到了国内外情报学界和科技期刊界的广泛关注,并从科学家个人绩效评价,迅速扩展到机构、地区、国家的期刊、基金资助项目、学科研究热点等方面的科学评价<sup>[7]</sup>。

**定义4** h-指数。假设一位科学家发表的 N 篇论文中有 h 篇论文被引次数至少为 h,其他 (N - h) 篇论文中每一篇的被引次数都小于等于 h,那么这位科学家的 h-指数就为 h。它是一项旨在评价科学家个人绩效的指标。

为了清楚理解 h-指数的定义,下面用一个具体算法来说明。设科学家发表的论文为  $p_1, p_2, \dots, p_n$ , n 为发表的论文总数,每篇论文的被引次数分别为  $ref_1, ref_2, \dots, ref_n$ , sort 是降序排序函数,则下面程序段运行后 h 变量中的值就是 h-指数。

```

sort(ref1, ref2, ..., refn);
for i = 1:n;i +
if i > = refi
    h = i - 1;
break;

```

随着对 h-指数研究的深入,目前已经有专家收集了学者、期刊、研究机构、大学、国家共五个层面八组 h-指数,以及论文被引指标数据,实证研究 h-指数与论文总被引数 C 之间的关系。另一方面从 Egghe-Rousseau 模型和 Glanzel-Schubert 模型<sup>[9,11]</sup> 出发,推演出 h-指数与总被引次数 C 之间的幂函数关系模型:  $h = C^{\alpha/(\alpha^2+1)}$ ,该模型与前者的实证结果基本相符。

h-指数作为一项复合指标,优点很明显:a) 尤其适合对科学家个人科研成就作评估;b) 是引文影响力与论文产出力二者相结合的测试指标;c) 是一个稳健的累积指标,单纯论文数量的增长对该指标不产生直接的影响;d) 能测试科学家的持久绩效;e) 不会随着引文数的增加而改变,所以文献类型不受限制;e) 它与其他“有意义”的文献计量学指标有显著相关性。但是缺点也不容忽视:a) 对于刚从事科学研究的年轻学者是不利的;b) 可能造成科学家躺在原有成就上“睡大觉”;c) 无法观察科学家研究活力的衰退情况;d) 与具体学科的交流行为有关,并且不能像论文或引文指标那样进行标准化处理;e) 即

使在相同学科领域也较难找到一个适当的参考标准;f) 不同的人群使用时会有不同的效果。

### 4 论文的综合可信等级评价公式

以上阐述的三种评价指标存在各自的缺点,为了在一定程度上改进这些缺点,本文将提出一个增强的、综合的评价方法,并考虑引文随时间变化的影响,以及低质量论文的区分度。

#### 4.1 引文数量随时间的变化影响

论文被引频次随时间的推移呈现出复杂的走势,从文献计量学的角度来看,论文在发表以后的两至三年之内达到引文的峰值。对于大多数论文而言,发表两至三年以后,这些论文的被引用会有很大的衰减,而对于少数经典性的论文,其被引用会持续很长的时间。此外,快报和简报类文献与综述和评论类文献在引文寿命上也有很大的差别。根据大量现象并观察发现,图1所示是论文被引频数随时间的变化规律<sup>[12,13]</sup>。图1中的参数 t 表示论文发表年限,C 表示论文被引频数,通过回归分析得出其近似的公式为  $C_t = C_\beta e^{-\lambda t - \beta t}$ 。其中:  $C_t$  为经过 t 年之后论文的被引数目;  $C_\beta$  是论文被引达到的峰值;  $\lambda$  为变化强度系数。

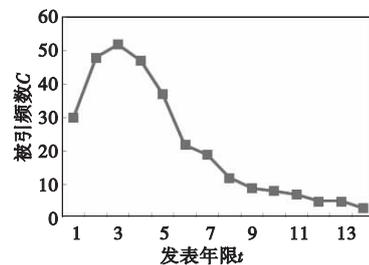


图1 论文被引频数随时间的变化曲线

由此可以类推,一种期刊对读者的影响也是随着年限变化的,越是陈旧的期刊,翻阅它的读者越少,而新出的期刊则成为读者关注的热点。假设期刊发表年份是  $t_0$ ,统计的年份是  $t_n$ ,期刊发表之初的影响力为  $M_0$ ,那么期刊在统计年  $t_n$  的影响力  $M_n$  为

$$M_n = M_0 \times e^{-\lambda(t_n - t_0)} \tag{1}$$

在  $t_0 \sim t_n$  时间段内,不同时间统计期刊的影响力  $M_n$  有着不同的结果,离期刊发表时间  $t_0$  越近的统计,越能够保留先前较高的影响力;反之,若时间间隔越长,一般期刊的影响力衰减得越厉害,直至该期刊无人问津。因此,加入时间因素进行考虑,能够避免一些起点影响力较高的期刊,始终列在排名的前端,而忽视其随时间衰减的状况。

#### 4.2 论文特征因子分值计算方法

在第2章中阐述了特征因子的概念,其最大的特色在于:以期刊影响力为权,以更贴近实际的权重重构了简单的引文网络,下面详细介绍影响权重 w 的计算方法<sup>[14]</sup>。

一个期刊的被引频次越高,它的影响权重越大。因此需要在引文网络中研究影响权重 w,选取 JCR 指标的 7611 篇自然科学和社会科学的引文数据,形成简单的引文网络,再依据该引文网络构建相应的引文矩阵 Z。这个矩阵中的条目  $Z_{ij}$ ,即是至 2011 年,2005—2010 年发表于期刊 i 的文章在 2011 年被引用于期刊 j。建立矩阵 Z 后,除去自引,即设置对角线元素为 0,再标准化 Z,即每列的元素除以对应列的总和,矩阵 Z 经过以上一系列转换后形成的新矩阵即为随机矩阵  $H_{ij}$ 。

但是存在期刊没有引用其他任何期刊的情况,在引文网络中称之为虚悬节点,则在矩阵  $Z$  中该期刊对应的就是全零列,为了计算出影响权重  $w$ ,需要用论文向量  $a$  代替全零列。论文向量  $a$  是一组由  $a_1, a_2, \dots, a_i$  构成的列向量,其定义如下:假设在 5 年的时间窗口中发表于期刊  $i$  的论文数量为  $p_i$ ,在同样 5 年里发表于统计内的所有期刊的论文总数量为  $s$ ,那么论文向量  $a_i = p_i/s$ 。用  $a$  向量取代  $H$  矩阵中这些全零列,产生一个新的矩阵  $H$ 。基于 Google 的 PageRank 算法<sup>[15]</sup>,定义一个新的随机矩阵  $P = \alpha H' + (1 - \alpha)a \cdot e^T$ ,  $\alpha$  为没有虚悬节点的列占总共列的百分比。

要计算出影响权重  $w$ ,即求出新的随机矩阵  $P$  的主特征向量,此影响权重  $w$  可以衡量引文的价值,那么引用次数与影响权重的乘积即为期刊实际提供的引用贡献,即  $H_{ij}w$ ;然后代入定义 2 中的公式,可以求出特征因子分值  $EF = 100 \frac{H_{ij}w}{\sum_i [H_{ij}w]_i}$ 。但是使用这个公式也有缺点,对于排名靠后的期刊区分度不大,相互之间的比较甚至用到了万分位,如特征因子分值为 0.00152 的期刊就有《Nano》《Bryologist》等 11 种<sup>[6]</sup>,故考虑改进公式为

$$EF' = 100 \frac{H_{ij}w}{\sqrt{\sum_i [H_{ij}w]_i}} \quad (2)$$

上述公式缩小分母以增大 EF,并且这种增大 EF 的方式与单纯地扩大 100 倍不一样,对总的引用贡献开方,既缩小了分母,又保留了总引用贡献对计算式的影响。

### 4.3 h-指数的拓展

由第 3 章可知,  $h$ -指数是考察科学家个人科研绩效评价的指标,而本文提出的改进指标是针对期刊的。于是借鉴  $h$ -指数的定义,构造期刊  $h$ -指数的定义如下:假设一期刊发表的  $N$  篇论文中有  $h$  篇论文被引次数至少为  $h$ ,其他  $(N - h)$  篇论文中每一篇的被引次数都小于等于  $h$ ,那么该期刊的  $h$ -指数就为  $h$ 。最终被统计的所有期刊的  $h$ -指数形成这样的一个列向量:  $H = (h_1, h_2, h_3, \dots, h_i)^T$ ,对应地,利用第 4 章介绍的总被引次数与  $h$ -指数的关系模型,可以计算得到  $H$  矩阵为

$$H = \left( \left[ \sum_0^i h_{2j} \right]^{\alpha/(a^2+1)}, \dots, \left[ \sum_0^i H_{ij} \right]^{\alpha/(a^2+1)}, \right. \\ \left. \left[ \sum_0^i H_{ij} \right]^{\alpha/(a^2+1)} \right)^T \quad (3)$$

其中:  $H_{1j}, H_{2j}, \dots$  为第 4.2 节中  $H_{ij}$  中的行向量,  $\sum_0^i H_{ij}$  为第  $i$  本期刊被引总次数。

### 4.4 综合评价公式

通过上面的分析可知,对于期刊的可信质量排名实际就是期刊影响力的排名,本文考虑的影响力受三个因素的影响,即时间、被引用情况、 $h$  值。为了方便研究,将这三个因素用向量  $O = \langle T, E, H \rangle$  表示,其中,  $T$  表示时间因素;  $E$  为特征因子分值即被引用情况因素;  $H$  为期刊特征因子分值。其中  $E, H$  都是列向量,即  $E = \langle EF_1, EF_2, \dots, EF_N \rangle, H = \langle H_1, H_2, \dots, H_n \rangle$ ,  $n$  是矩阵的维数,第  $i$  个期刊的影响力向量为  $O_i = \langle T_i, E_i, H_i \rangle$ 。

利用式(1)~(3)构造综合评价公式,分析如下:在统计有效年份  $t_0 \sim t_n$  中,  $\lambda$  为期刊影响力随时间变化的强度系数,其值越大,影响力衰减的速度越快。对应期刊的  $h$ -指数值作为  $e$  函数指数的分母,更能够体现衰减,因为期刊的  $h$ -指数越

高,说明其质量越高,而一般质量越高的期刊,它的影响力随时间的衰减就越慢;  $h$  作为分母,分母越大,  $\frac{\lambda}{h}$  的值就越小,影响力衰减则越慢。另一方面,综合评价公式主要是基于特征因子的改进,故特征因子分值比时间衰减更重要,为此使用  $(EF')^2$  来提升在综合评价指标中的重要程度。因此,得出增强和综合的评价式如下:

$$\text{paperScore}_i = e^{-\frac{\lambda}{h_i}(t_n - t_0)} \times (EF'_i)^2 \\ \text{paperScore}_{(i,n)} = \begin{cases} 0 & i, j = 0 \\ e^{\frac{-\lambda(t_n - t_0)}{\sum_{k=1}^n h_k}} \times \left( \frac{100 H_{ij} w}{\sqrt{\sum_i [H_{ij} w]_i}} \right)^2 & \text{other} \end{cases}$$

上述公式中,  $\alpha, \lambda$  为参数,后面将会讨论其取值情况;变量  $i$  表示期刊的代号,在统计内的每个期刊都有各自对应的代号  $i$ 。如果  $i$  的值改变,整个引文网络就会跟着改变,那么引文矩阵  $H_{ij}$ 、影响权重  $w$  也随之改变。变量  $n$  表示统计当年距期刊发表的年数,当  $n = 0$  时,  $\text{paperscore} = (EF')^2$ ; 当  $n = +\infty$  时,  $\text{paperScore} = 0$ 。

改进后的论文综合评价式  $\text{paperScore}$  有以下的优点:

a) 考虑了期刊影响率随时间变化的状况,能够体现出期刊的衰减情况。这也符合现实生活中读者查阅期刊的习惯,越是新出版的期刊,越是得到读者的青睐。PaperScore 使用期刊  $h$ -指数,一方面区分了不同质量期刊随时间的衰减强度,质量越高的期刊,影响力越持久;另一方面以不同于特征因子分值的角度对综合式加以补充。

b)  $EF'$  不同于  $EF$ ,通过缩小了公式中的分母而扩大了特征因子分值,便于低指标期刊之间的比较。在增强的综合式中,使用  $(EF')^2$  来提升在综合评价指标中的重要程度,既突显特征因子的优点,又增加了区分度。

然而单凭  $\text{paperScore}$  数值,读者很难对期刊的质量高低有直观的认识,因此本文根据计算出的期刊  $\text{paperScore}$  值划分可信等级。本文规定:  $\text{paperScore}$  值在区间  $[1, +\infty)$  的期刊为一级期刊;在区间  $[0.1, 1)$  为二级期刊;在区间  $[0.01, 0.1)$  的为三级期刊;在区间  $[0.001, 0.01)$  的为四级期刊;在区间  $(-\infty, 0.001)$  的为五级期刊。通过以下算法可以实现对期刊的可信分级。

算法 1 论文可信等级划分算法 PaperTrustGrade()

输入: float paperScore。

输出: grade。

int paperTrustGrade(float paperScore)

```

{
    if (paperScore >= 1)
        grade = 1;
    else if (paperScore >= 0.1 && paperScore < 1)
        grade = 2;
    else if (paperScore >= 0.01 && paperScore < 0.1)
        grade = 3;
    else if (paperScore >= 0.001 && PaperScore < 0.01)
        grade = 4;
    else if (paperScore < 0.001)
        grade = 5;
    return grade;
}
    
```

### 4.5 论文等级评估实例分析

#### 4.5.1 参数的确定

在上述 paperScore 公式中有两个参数  $\alpha$  和  $\lambda$ , 为了保证计算的有效性, 需要确定参数的取值。其中  $\alpha$  为  $h$ -指数中的洛特卡系数, 通过以下实验确定: a) 取一组已知  $h$  值与被引频数的期刊数据留作验证; b) 根据公式  $h = C^{\alpha/(\alpha^2+1)}$ , 以  $\alpha$  为自变量在  $x$  轴上,  $h$  为因变量在  $y$  轴上,  $C$  取一常数作图, 如图 2 所示; c) 用已知数据验证, 如当  $C$  取 20 时, 对应检验数据中  $h = 4$ , 而在图中当取  $\alpha = 2$  时, 可以保证  $h$  与  $C$  这样的关系。

参数  $\lambda$  为期刊影响力随时间变化的强度系数, 其值越大, 影响力衰减的速度越快。图 3 为  $\lambda$  ( $x$  轴) 取不同值时, 某一期刊的 paperScore ( $y$  轴) 的变化情况。根据现实期刊影响力变化的特点, 影响力有一定的衰减且不至于过快, 取  $\lambda = 2$  能比较好地反映影响力的衰减过程。

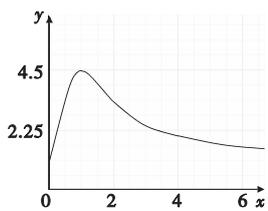


图2  $h$ 值随 $\alpha$ 的变化曲线

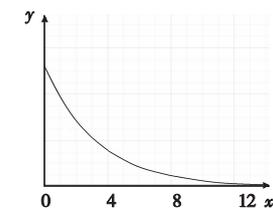


图3 paperScore随 $\lambda$ 的变化曲线

#### 4.5.2 各种评价指标的比较

为探究综合评价 paperScore 公式的合理性与有效性, 采用 2011 年 JCR 的期刊统计数据, 针对不同等级批次的期刊, 计算出其 paperScore 值, 并且与对应的特征因子分值 EF 比较。该实验分四组进行: 第一组是按照 EF 排名靠前的 10 种期刊; 第二组是 EF 排名中间偏前的 10 种期刊; 第三组是 EF 排名中间偏后的 10 种期刊; 第四组是用 EF 排名靠后的 10 种期刊, 它们分别代表不同的质量等级。实验结果如图 4~7 所示。

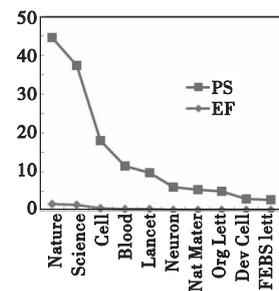


图4 一级论文的PS及EF排名比较

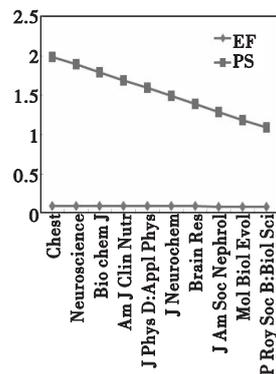


图5 二级论文的PS及EF排名比较

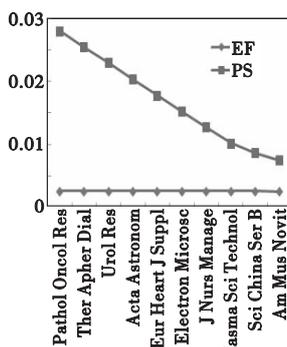


图6 三级论文的PS及EF排名比较

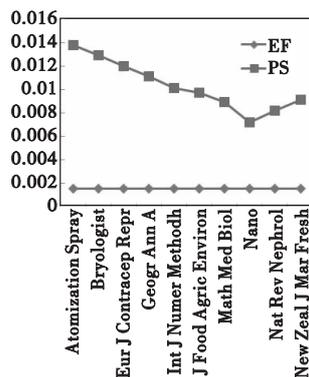


图7 四级论文的PS及EF排名比较

由图 4 所示, 特征因子分值 EF 高的期刊, 对应的 paperScore 也高, 并且后者较前者的数据有更好的可视性。由图 5~7 所示, 对于排名不是很靠前的期刊, EF 已经无法对其排名, 然而 paperScore 依然可以对其排名。因此, 增强型的综合公式 paperScore 的优点显而易见。

### 5 结束语

本文通过分析现存几种科技论文的评价指标, 即影响因子、特征因子、 $h$ -指数, 详细介绍它们的定义、计算方法以及优缺点。并基于这些检索特征的不足, 加以改进, 得到增强型的综合评价指标, 该指标不仅能够体现出期刊的衰减情况, 还能扩大比较分值, 使得低影响力的期刊之间具有区分度。最后, 用实例评估综合公式的可行性与准确性, 同时发现, 使用综合公式比单纯使用一种评价指标得出的排名结果要更加符合读者的要求。但是尽管如此, 还是存在一些缺点, 如跨学科的科技论文无法比较, 论文作者不同贡献的权重引用次数无法计算等。但是, 评价一篇论文质量好坏, 归根到底还是要看其内容的价值, 下一步研究工作将从内容上着手, 研究科技论文的内容可信。

#### 参考文献:

- [1] GARFIELD E. Forms for literature citations [J]. Science, 1954, 120 (3129): 1038-1041.
- [2] BERGSTROM C T. Eigenfactor: measuring the value and prestige of scholarly journals [J]. College & Research Libraries News, 2007, 68(5): 314-316.
- [3] HIRSCH J E. An index to quantify an individual's scientific research output [J]. National Academy of Science of the United States of America, 2005, 102(46): 16569-16572.
- [4] EGGHE L. How to improve the  $h$ -index? [J]. The Scientist, 2006, 20(3): 14.
- [5] 金碧辉. 科学家为自己设计了一项评价指标:  $h$  指数 [J]. 科学观察, 2006, 1(1): 8-9.
- [6] 2011-JCR-SCI-IF [EB/OL]. <http://wenku.baidu.com/view/>.
- [7] 王梅英, 刘雪立, 王璞.  $h$ -指数及其扩展指标在期刊评价中的应用 [J]. 情报杂志, 2012, 31(3): 60-64.
- [8] FRANCESCHET M. Ten good reasons to use the Eigenfactor™ metrics [J]. Information Processing and Management, 2010, 46(5): 555-558.
- [9] 赵星, 高晓强.  $h$ - $C$  幂律关系模型推演及参数分析 [J]. 情报学报, 2011, 30(1): 83-86.
- [10] 张琳. CSSCI 教育学期刊特征因子与影响因子比较分析 [J]. 情报杂志, 2011, 30(7): 34-37.
- [11] 赵星. 期刊引文评价新指标 eigenfactor 的特性研究 [J]. 情报理论与实践, 2009, 32(8): 53-56.
- [12] 张莉, 张凤莲. 缩短发表时滞提高论文的时效性 [J]. 编辑学报, 2003, 15(5): 331-332.
- [13] 库耘, 邹腊年. 论文的发表时间及期刊的论文发表周期辨析 [J]. 中国科技期刊研究, 2007, 18(1): 153-155.
- [14] PageRank [EB/OL]. <http://en.wikipedia.org/wiki/PageRank>.
- [15] Method, Information method [EB/OL]. <http://www.eigenfactor.org>.
- [16] 任胜利. 特征因子 (eigenfactor): 基于引证网络分析期刊和论文的重要性 [J]. 中国科技期刊研究, 2009, 20(3): 415-418.