

# h 指数应用于微博影响力分析的探索<sup>\*</sup>

——以我国“211 工程”大学图书馆微博为例

周志峰<sup>1 2</sup> 韩静娴<sup>3</sup>

(1. 武汉大学信息管理学院 武汉 430072; 2. 温州大学 温州 325035;  
3. 苏州大学图书馆 苏州 215006)

**摘要** 2005 年美国学者 Hirsch 提出的 h 指数不仅是一种基于引文分析的学术评价方法,更能作为分析微博影响力的新思想和新工具。通过调查我国“211 工程”大学的图书馆在新浪微博平台上开设微博的情况,以其中 25 个认证微博为研究样本,将 h 指数应用于微博影响力的分析。微博的被转发 h 指数和评论 h 指数可以反映微博的传播力和受关注度,从而体现出影响力状况。

**关键词** 信息计量学 微博 h 指数 大学图书馆 微博影响力 新浪微博

中图分类号 G250.2

文献标识码 A

文章编号 1002-1965(2013)04-0063-05

## An Exploration on the Application of H-index in the Analysis of the Impact of Microblogs ——Taking Microblogs of ‘211 Project’ Universities Libraries as Examples

Zhou Zhifeng<sup>1 2</sup> Han Jingxian<sup>3</sup>

(1. School of Information Management, Wuhan University, Wuhan 430072;  
2. Library of Wenzhou University, Wenzhou 325035;  
3. Library of Soochow University, Suzhou 215006)

**Abstract** The h-index proposed by Hirsch in 2005 is not only a method of the academic evaluation based on the citation analysis, but also a new idea and tool to analyze the impact of microblog. This paper surveys 25 microblogs of academic libraries coming from ‘211 project’ universities at www.weibo.com, tries to explore the application of h-index in the analysis of the impact of academic libraries microblogs. The h-index of microblog forwarding and the h-index of microblog comment can reflect the power of communication and degree of concern of microblogs, and then reflect the impact of microblogs.

**Key words** informetrics microblog h-index academic libraries the impact of microblog

### 0 引言

微博(Microblog)将博客与即时通信结合在一起,是一个基于用户关系的 Web2.0 信息分享、传播与获取平台,用户可随时随地通过网页、移动通讯设备、IM 工具等途径更新动态,信息一般限制在 140 字左右,同时其“粉丝”(followers)或其他用户也可通过网页、移动通讯设备等接收、浏览,实现即时信息分享。

全球最早、最著名的微博平台是 Twitter,截至 2012 年 6 月,在全球范围已经拥有 5 亿注册用户<sup>[1]</sup>。2009 年 8 月,新浪网推出微博内测版,成为中国第一

家提供微博服务的门户网站,目前注册用户已经超过 3 亿<sup>[2]</sup>。根据中国互联网络信息中心(CNNIC)第 30 次《中国互联网络发展状况统计报告》,微博已成为用户增长最快的互联网应用模式,截至 2012 年 6 月,微博用户规模已达 2.7 亿,网民使用率达到 50.9%;微博在手机端的增长幅度更加明显,用户数量由 2011 年底的 1.37 亿增至 1.70 亿,增速达到 24.2%<sup>[3]</sup>。

微博自 2006 年出现以来,便受到社会各界的青睐与关注,图书馆也利用微博开展读者服务、舆情监控、学术交流等工作,成为继博客等工具之后的网络新贵工具<sup>[4]</sup>。笔者选择新浪微博平台进行调查,截至 2012

收稿日期:2012-12-31

修回日期:2013-02-05

基金项目:温州市科技计划项目“基于 h 指数的区域科研评价研究”(编号:R20100046)。

作者简介:周志峰(1978-),男,博士研究生,研究方向:信息计量学;韩静娴(1979-),女,馆员,研究方向:图书馆信息服务。

年 10 月 20 日,名称中包含“图书馆”的微博已经超过 1000 个,名称中包含“大学图书馆”和“学院图书馆”的微博共有 421 个,其中认证微博有 88 个。高等院校图书馆利用微博开展的工作主要包括资源推介、活动播报、读者交流、馆情馆讯、服务信息和同行交流等项目<sup>[5]</sup>。

## 1 微博的影响力分析

博客、微博已经成为广泛应用的信息载体和传输媒介,对人们日常工作和生活所产生的影响越来越大<sup>[6]</sup>,因此对其影响力的研究也成为热点。微博影响力的构成涉及多种因素,如粉丝数、评论回复、转发影响力等,分别对应具体的用户行为<sup>[7]</sup>。比较直观的因素是粉丝数,其反映“人气”的旺盛与否,或者说“吸引力”的大小,在一定程度上能折射出微博的影响力,但粉丝数的背后有可能存在“僵尸粉丝”<sup>[8]</sup>。新浪微博数据中心通过博文的篇均转发、篇均评论次数来衡量的传播力和影响力。

一些学者已经将信息计量学的方法应用到对博客、微博等的分析中<sup>[9]</sup>。链接分析是微博影响力研究的重要方法,Jianshu Weng 等基于 PageRank 设计了 Twitter 用户和链接结构的话题相似性影响力排序算法<sup>[10]</sup>。Meeyoung Cha 等选择了用户行为的角度,通过 Twitter 的粉丝、被转发与被@ 状况对微博影响力进行了考察<sup>[11]</sup>。原福永等针对微博的排名机制进行研究,从链接分析和用户行为两个角度对微博影响力进行衡量,建构了微博用户的用户影响力指数模型,通过计算用户被关注度得到微博用户的用户活跃度和微博影响力,进而合成微博用户的用户影响力<sup>[12]</sup>。

## 2 h 指数在微博影响力分析中的应用

在微博影响力的分析中,借鉴和移植信息计量学相关指标有利于扩展研究思路和方法。已有研究者将 h 指数应用于博客、微博影响力的分析中,如 Jos'e De-vezas 等人将 h 指数应用于博客的权威性的评估中<sup>[13]</sup>,利用 h 指数对博客进行排名<sup>[14]</sup>。Romero 等将 h 指数应用到 Twitter 等社会媒体的影响力分析中<sup>[15]</sup>。张晓阳等以科学网为例,研究了科学家博客 h 指数,认为 h 指数是博客影响力评价的有益工具<sup>[16]</sup>。

2.1 h 指数定义与特点 2005 年,美国学者 Hirsch 提出 h 指数,用以评价科学家的学术成就,引起了学术界的关注,拓宽了学术评价的视野。h 指数概念由 Hirsch 提出,并得到其他学者的修正,定义为:将科学家 S 发表的论文按照被引次数从高到低排列,相同被引次数论文具有不同的序号( Rousseau 建议按发表时间倒序排列),当且仅当前 h 篇论文每篇被引次数至

少为 h,同时第 h+1 篇论文的被引次数小于 h+1 时,科学家 S 的 h 指数值为 h<sup>[17-18]</sup>。

通过 h 指数在评价研究中的实证运用,可将其特征概括为以下几点:计算相对简单;h 指数兼顾了学术成果的质量与数量;结果相对稳定、准确;具有数据源敏感性和测评领域敏感性,数据来源和测评领域不一样,获取的结果也不一样;具有测算时间区间敏感性等<sup>[19]</sup>。

h 指数可用在相应信息源的评价分析中<sup>[20]</sup>,最初用于评价科学家的学术成就,之后经过众多研究者的延伸和扩展,被应用到学术期刊评价、高校等学术机构评价、专利评价等方面<sup>[21-23]</sup>。

2.2 h 指数应用于微博数据分析 h 指数是基于引文数据提出的,但其基本思想与方法可被用于其他相似的数据环境中,微博被转发、评论数据集就符合这样的数据环境特征。微博博文与论文有类似之处,微博被转发、评论的数量对应论文被引用频次,这样就可以像分析论文一样分析微博及其被转发/评论的数据了。由此来看,h 指数应用于微博数据分析具有可行性和必要性。

2.2.1 h 指数应用于微博影响力分析的可行性。微博数据在逻辑上由紧密相关、不可分割的两个部分组成:一是微博发表博文的数据,二是微博被转发/评论的数据。如果将这些数据按照被转发/评论的次数从高到低排列,就会形成类似于 h 指数原始定义中论文数与引文数的数据序列。以表 1 为例,将一定时间段内的微博博文按照被转发/评论数由高到低排列,前 7 篇博文的被转发/评论次数均大于或等于 7,从序号 8 开始博文被转发/评论次数皆小于 8,那么可得出此微博的被转发/评论 h 指数为 7。被转发/评论 h 指数为 7 的含义就是:在一定时间段内,某微博起码有 7 篇博文的被转发/评论频次大于等于 7。

表 1 微博数据中被转发/评论 h 指数计算示例

| 序号    | 微博博文  | 被转发/评论数 |
|-------|-------|---------|
| 1     | M1    | 25      |
| 2     | M2    | 19      |
| 3     | M3    | 16      |
| 4     | M4    | 15      |
| 5     | M5    | 12      |
| 6     | M6    | 9       |
| 7     | M7    | 7       |
| 8     | M8    | 5       |
| ..... | ..... | .....   |

2.2.2 h 指数应用于微博数据分析的必要性。通过对微博的数据分析,可测度其在社会化网络中信息传播的影响力,一般利用以下指标:一是绝对数据——粉丝数量、发表博文数、博文被转发/评论总数

等;二是相对数据——篇均被转发/评论数等。这些指标虽然能够在一定程度上揭示微博的影响力状况,但存在明显的不足,比如:粉丝数量、发表博文数、博文被转发/评论总数与开博时间的长短关联太大;篇均被转发/评论数则掩盖了博文集合内部的差别;还会出现个别博文的高被转发和评论,从而提高了整体被转发数和评论数;以上原因都可能造成不能客观地评价微博整体的传播力和影响力状况。

$h$  指数相对于传统学术评价指标的良好特性也在微博数据分析中得到体现。相对于微博评价的常用指标,微博被转发/评论的  $h$  指数增强了不同微博之间的可比性,能够更加客观地反映微博信息传播力与影响力。

### 3 实例研究

3.1 样本来源 本文将“大学图书馆”的选取范围限定在“211 工程”高校中。目前,全国共有 112 所“211 工程”高校,其中华北电力大学、中国石油大学、中国地质大学、中国矿业大学拥有两个校区,实际高校数为 116 所<sup>[25]</sup>。

本文以新浪微博平台作为微博数据的来源,并为了保证微博的权威性和官方代表性,需选取通过实名认证的微博。通过检索,获取了 25 个大学图书馆微博,并对它们相关的数据进行了采集和汇总,具体可见表 2。数据检索时间:2012 年 7 月 16-20 日。

3.2 微博被转发/评论  $h$  指数测算 表 3 以武汉大学图书馆微博为例,描述了微博被转发/评论  $h$  指数是如何计算的。根据  $h$  指数的定义,武汉大学微博被转发  $h$  指数为 14,被评论  $h$  指数为 11,说明了武汉大学图书馆微博中有 14 篇博文的被转发数大于等于 14,有 11 篇博文评论数大于等于 11。

表 2 对 25 所大学图书馆微博被转发/评论  $h$  指数进行计算和汇总,并对各项指标进行了排名。

### 3.3 讨论

3.3.1 数据说明。通过对表 2 中各项指标的比较,我们对涉及微博影响力的因素有以下的认识:

a. 微博的影响力与开设时间有着重要的关联,微博的人气与影响力需要时间积累。20 个图书馆微博开设时间在 2011 年,占样本数的 80%;最早开通微博的是厦门大学图书馆(2009 年 11 月),之后是南京师

表 2 25 所大学图书馆新浪微博数据信息一览表

| 序号 | 图书馆所在大学    | 网址                                  | 开设时间       | 博文数/排名 | 原创博文数/排名 | 博文原创率/排名  | 粉丝数/排名  | 被转发数/排名 | 篇均被转发数/排名 | $h_f$ 指数/排名 | 评论数/排名  | 篇均评论数/排名 | $h_c$ 指数/排名 | $h_m$ 指数/排名 | $h_m$ 指数/排名 |
|----|------------|-------------------------------------|------------|--------|----------|-----------|---------|---------|-----------|-------------|---------|----------|-------------|-------------|-------------|
| 1  | 北京工业大学     | http://weibo.com/bjutlib            | 2011-4-19  | 162/20 | 94/20    | 58.02%/16 | 3141/16 | 317/17  | 1.96/12   | 8/16        | 396/15  | 2.44/6   | 9/11        | 8.5/16      | 8.5/16      |
| 2  | 北京航空航天大学   | http://weibo.com/buaalib            | 2011-1-10  | 728/8  | 690/2    | 94.78%/1  | 4305/12 | 746/14  | 1.02/19   | 9/15        | 1804/5  | 2.48/5   | 18/3        | 13.5/8      | 12.7/9      |
| 3  | 北京化工大学     | http://weibo.com/buctlib            | 2011-6-14  | 430/15 | 278/16   | 64.65%/12 | 1936/19 | 355/16  | 0.83/23   | 5/21        | 284/18  | 0.66/20  | 5/18        | 5/20        | 5/20        |
| 4  | 北京理工大学     | http://weibo.com/u/2710400355       | 2012-4-19  | 60/24  | 16/24    | 26.67%/23 | 396/24  | 55/24   | 0.92/21   | 3/23        | 42/23   | 0.70/19  | 3/23        | 3/23        | 3/23        |
| 5  | 重庆大学       | http://weibo.com/cqulib             | 2011-11-11 | 673/12 | 488/5    | 72.51%/9  | 6668/6  | 2480/4  | 3.68/4    | 20/5        | 2618/3  | 3.89/1   | 20/2        | 20/4        | 20/4        |
| 6  | 东南大学       | http://weibo.com/seulib             | 2011-2-21  | 768/6  | 313/14   | 40.76%/21 | 4314/11 | 1463/11 | 1.90/14   | 15/6        | 582/11  | 0.76/17  | 8/13        | 11.5/11     | 11/11       |
| 7  | 复旦大学       | http://weibo.com/fdulibrary         | 2011-4-20  | 688/11 | 360/9    | 52.33%/18 | 6988/4  | 2447/5  | 3.56/5    | 23/4        | 938/10  | 1.36/12  | 10/10       | 16.5/5      | 15.2/5      |
| 8  | 合肥工业大学     | http://weibo.com/hfutlib            | 2011-6-29  | 708/9  | 526/4    | 74.29%/7  | 1482/20 | 282/19  | 0.40/25   | 5/21        | 354/17  | 0.50/23  | 7/16        | 6/17        | 5.9/17      |
| 9  | 华东师范大学     | http://weibo.com/ecnlib             | 2011-5-3   | 532/14 | 358/10   | 67.29%/11 | 4702/10 | 1703/8  | 3.20/6    | 15/6        | 1399/8  | 2.63/4   | 15/6        | 15/6        | 15/6        |
| 10 | 暨南大学       | http://weibo.com/jnulib             | 2011-3-1   | 1801/2 | 279/15   | 15.49%/24 | 7298/3  | 7528/2  | 4.18/2    | 33/2        | 3153/2  | 1.75/9   | 18/3        | 25.5/2      | 24.4/2      |
| 11 | 南京大学       | http://weibo.com/njulibrary         | 2011-4-19  | 753/7  | 326/12   | 43.29%/19 | 6182/7  | 1499/10 | 1.99/11   | 14/10       | 570/12  | 0.76/18  | 8/13        | 11/12       | 10.6/12     |
| 12 | 南京航空航天大学   | http://weibo.com/nuaalib            | 2011-11-1  | 75/22  | 43/23    | 57.33%/17 | 1314/21 | 72/23   | 0.96/20   | 3/23        | 24/24   | 0.32/25  | 3/23        | 3/23        | 3/23        |
| 13 | 南开大学       | http://weibo.com/nklib              | 2011-5-12  | 220/17 | 157/19   | 71.36%/10 | 877/23  | 192/20  | 0.87/22   | 6/18        | 112/20  | 0.51/21  | 5/18        | 5.5/19      | 5.5/19      |
| 14 | 南京师范大学     | http://weibo.com/njnlib             | 2010-9-6   | 980/3  | 620/3    | 63.27%/13 | 4962/9  | 2362/6  | 2.41/10   | 15/6        | 1799/6  | 1.84/8   | 13/8        | 14/7        | 14/7        |
| 15 | 清华大学       | http://weibo.com/thulib             | 2010-12-13 | 3976/1 | 325/13   | 8.17%/25  | 18314/1 | 16067/1 | 4.04/3    | 35/1        | 11604/1 | 2.92/2   | 25/1        | 30/1        | 29.6/1      |
| 16 | 上海财经大学     | http://weibo.com/ilovesufelib       | 2011-4-29  | 899/5  | 825/1    | 91.77%/2  | 3875/14 | 1051/12 | 1.17/18   | 11/13       | 447/13  | 0.50/22  | 7/16        | 9/14        | 8.8/15      |
| 17 | 上海交通大学     | http://weibo.com/sjtulib            | 2011-6-21  | 211/19 | 179/18   | 84.83%/5  | 1978/18 | 313/18  | 1.48/17   | 7/17        | 184/19  | 0.87/14  | 5/18        | 6/17        | 5.9/17      |
| 18 | 同济大学       | http://weibo.com/tongjiuniv-library | 2011-9-8   | 400/16 | 334/11   | 83.50%/6  | 3773/15 | 772/13  | 1.93/13   | 12/11       | 428/14  | 1.07/13  | 9/11        | 10.5/13     | 10.4/13     |
| 19 | 武汉大学       | http://weibo.com/whulibrary         | 2011-11-13 | 704/10 | 418/6    | 59.38%/15 | 5079/8  | 2027/8  | 2.88/7    | 14/9        | 1064/9  | 1.51/11  | 11/9        | 12.5/10     | 12.4/10     |
| 20 | 西安交通大学     | http://weibo.com/xjtulib            | 2011-10-8  | 87/21  | 53/22    | 60.92%/14 | 2394/17 | 133/22  | 1.53/16   | 6/18        | 72/21   | 0.83/15  | 4/21        | 5/20        | 4.9/21      |
| 21 | 西北工业大学     | http://weibo.com/nwpulib            | 2011-10-21 | 66/23  | 60/21    | 90.91%/3  | 1290/22 | 161/21  | 2.44/9    | 6/18        | 52/22   | 0.79/16  | 4/21        | 5/20        | 4.9/21      |
| 22 | 西南财经大学     | http://weibo.com/u/2152468530       | 2011-6-2   | 217/18 | 195/17   | 89.86%/4  | 3939/13 | 610/15  | 2.81/8    | 10/14       | 357/16  | 1.65/10  | 8/13        | 9/14        | 8.9/14      |
| 23 | 厦门大学       | http://weibo.com/xmulibrary         | 2009-11-17 | 563/13 | 413/7    | 73.36%/8  | 8281/2  | 3191/3  | 5.67/1    | 26/3        | 1581/7  | 2.81/3   | 16/5        | 21/3        | 20.4/3      |
| 24 | 中国科学技术大学   | http://weibo.com/ustclib            | 2012-5-25  | 28/25  | 12/25    | 42.86%/20 | 126/25  | 18/25   | 0.64/24   | 2/25        | 12/25   | 0.43/24  | 2/25        | 2/25        | 2/25        |
| 25 | 中国矿业大学(徐州) | http://weibo.com/cumt20051118       | 2011-4-6   | 970/4  | 375/8    | 38.66%/22 | 6920/5  | 1686/9  | 1.74/15   | 12/11       | 1878/4  | 1.94/7   | 14/7        | 13/9        | 13/8        |

注:  $h_f$  指数为微博的被转发(Forwarding)  $h$  指数;  $h_c$  指数为微博的评论(Comment)  $h$  指数。

范大学图书馆(2010年9月)和清华大学图书馆(2010年12月),它们的大多指标都是处于靠前位置;而中国科学技术大学图书馆和北京理工大学图书馆的微博都是在2012年,所以在各项指标的排名中都处于末尾。

表3 武汉大学图书馆微博博文被转发/评论 $h$ 指数测算

| 序号  | 博文               | 被转发数  | 博文               | 评论数   |
|-----|------------------|-------|------------------|-------|
| 1   | 2012年1月11日08:47  | 174   | 2012年1月11日08:47  | 34    |
| 2   | 2011年12月15日11:09 | 89    | 2011年12月15日11:09 | 28    |
| 3   | 2012年2月27日14:51  | 71    | 2012年6月29日08:14  | 17    |
| 4   | 2012年2月10日15:42  | 36    | 2012年6月15日16:31  | 15    |
| 5   | 2012年6月1日16:10   | 25    | 2012年6月16日21:21  | 14    |
| 6   | 2012年4月20日14:24  | 24    | 2012年5月22日15:24  | 13    |
| 7   | 2012年6月7日11:17   | 22    | 2012年2月24日10:18  | 13    |
| 8   | 2012年3月23日16:39  | 22    | 2012年6月9日21:38   | 11    |
| 9   | 2012年1月30日10:21  | 22    | 2012年6月6日10:20   | 11    |
| 10  | 2012年6月28日09:45  | 19    | 2012年5月9日09:09   | 11    |
| 11  | 2012年6月29日08:14  | 17    | 2012年5月31日10:38  | 11    |
| 12  | 2012年6月11日15:20  | 16    | 2012年3月9日14:14   | 10    |
| 13  | 2012年2月23日09:23  | 15    | 2012年3月23日16:39  | 10    |
| 14  | 2011年12月30日11:03 | 14    | 2012年2月27日14:51  | 10    |
| 15  | 2012年6月29日09:26  | 14    | 2012年1月30日10:21  | 10    |
| 16  | 2012年3月21日11:13  | 14    | 2012年7月8日14:49   | 9     |
| 17  | 2012年6月13日16:28  | 13    | 2012年7月3日15:58   | 9     |
| ... | .....            | ..... | .....            | ..... |

注:由于博文内容过长,表中以发布的时间代替

b. 博文数量反映了微博的信息容量和活跃程度,博文的原创率则体现了微博创新程度;博文的数量和质量是吸引用户关注的内容基础,对微博的影响力产生直接的作用。从博文数来看,清华大学图书馆微博发布的最多,达3976篇,但其原创率排名处于末端,大部分博文是转发而来;而原创博文则是上海财经大学图书馆发布最多,达825篇,占91.77%。在此必须要说明无论是转发来的博文还是原创的博文,与信息质量没有直接关系。

c. 粉丝数量反映了微博的“人气”程度,是微博影响力的“群众”基础,它们在博文的传播中既是信息的受众又是传播的主体。从粉丝数量来看,三甲是清华大学图书馆18314个、厦门大学图书馆8281个、暨南大学图书馆7298个。同时,我们在微博影响力的实际分析中,必须还要关注粉丝的活跃度,辨别微博的“僵尸粉”。

d. 微博的一键转发机制,减小了网络信息传播过程中的“摩擦力”,可以便捷地实现分享;微博的评论模块具有强大的交互功能。博文的被转发/评论次数、篇均被转发/评论的次数反映了微博的传播力状况。清华大学图书馆微博的被转发/评论次数排名皆为第一,原因在于其开博时间长、博文数量多以及粉丝多;厦门大学图书馆微博篇均被转发的次数最多,重庆大学图书馆微博篇均评论的次数最多。

e.  $h_f$ 指数和 $h_c$ 指数最高的都是清华大学图书馆,说明了综合考量微博的信息容量与活跃程度(通过博文数量表现出来)和被关注、传播程度(通过博文的被转发/评论数量表现出来),其影响力最高。

f. 25个大学图书馆微博中, $h_f$ 指数核内博文总数为315篇,占有所有博文的1.8%; $h_c$ 指数核内博文总数为247篇,只占有所有博文的1.5%。这些博文都可被视为具有高影响力。

3.3.2 相关性分析。采用SPSS相关分析检验表2中 $h_f$ 指数、 $h_c$ 指数与博文数、粉丝数、被转发数、评论数、篇均被转发数、篇均评论数之间的相关性,以揭示微博影响力相关指标之间的数量关系。

a.  $h_f$ 指数、 $h_c$ 指数与微博数、粉丝数之间的关系。

表4  $h_f$ 指数、 $h_c$ 指数与微博数、粉丝数之间的相关性

| 相关系数 | $h_f$ 指数 |             | $h_c$ 指数    |
|------|----------|-------------|-------------|
|      | 微博数      | 0.778( ** ) | 0.748( ** ) |
|      | 粉丝数      | 0.886( ** ) | 0.846( ** ) |

注: \*\* 为0.01显著性水准(下同)

微博作为一个社交媒体工具,是依靠其信息内容获取用户关注的,从这个角度来说,博文数量多少是其影响力的基础。因此,其与 $h_f$ 指数和 $h_c$ 指数之间呈现正相关的关系,分别达到了0.778和0.748。

微博被转发/评论依靠的是微博粉丝的行为,微博粉丝越多,博文被转发和评论的可能性就越大,在一定程度上说明微博被关注度高、人气高。因此,粉丝数与 $h_f$ 指数和 $h_c$ 指数之间相关系数达到了0.886和0.846。

b.  $h_f$ 指数、 $h_c$ 指数与被转发/评论数之间的关系。

表5  $h_f$ 指数、 $h_c$ 指数与被转发/评论数之间的相关性

| 相关系数 | $h_f$ 指数 |             | $h_c$ 指数    |
|------|----------|-------------|-------------|
|      | 被转发数     | 0.836( ** ) | ——          |
|      | 评论数      | ——          | 0.782( ** ) |

$h_f$ 指数和 $h_c$ 指数延续了 $h$ 指数的优势,但同时仍然是以被转发/评论数为基础的,因此与被转发/评论数呈现高度的相关,相关系数分别达到了0.836和0.782,能够作为一个独立的评价指标出现,也能更客观地体现出微博的传播力和影响力。

c.  $h_f$ 指数、 $h_c$ 指数与篇均转发/评论数之间的关系。

表6  $h_f$ 指数、 $h_c$ 指数与篇均被转发/评论数之间的相关性

| 相关系数 | $h_f$ 指数 |             | $h_c$ 指数    |
|------|----------|-------------|-------------|
|      | 篇均被转发数   | 0.861( ** ) | ——          |
|      | 篇均评论     | ——          | 0.871( ** ) |

篇均被转发/评论数是衡量微博传播力的重要指标。 $h_f$ 指数、 $h_c$ 指数与篇均被转发/评论数也高度相关,相关系数分别达到0.861和0.871。

3.3.3 微博影响力综合 $h$ 指数探讨。在可公开

获得的微博相关数据集合中,只有被转发/评论数据集合具有获取  $h$  指数的数据环境。 $h_f$  指数可体现出微博信息的传播影响力, $h_c$  指数体现微博粉丝的参与和互动程度,进而体现微博的活跃度。如要形成体现微博影响力的综合性  $h$  指数,可将  $h_f$  指数、 $h_c$  指数按照适当的权重进行加权结合,本文提出计算公式为:

$$h_m = r \cdot h_f + (1 - r) \cdot h_c \quad (1)$$

或者

$$h_m = h_f^r \cdot h_c^{1-r} \quad (2)$$

$h_m$  为反映微博影响力的综合性  $h$  指数; $r$  为调节因子,用于调整  $h_f$  与  $h_c$  之间的权重。

具体到本文实例研究,在获取综合性的  $h$  指数过程中,假设对  $h_f$  和  $h_c$  无偏重,即设定  $r = 0.5$ ,进而可以获得体现微博影响力的  $h_m$  指数的值及其排名,见表2。可以看出,经过处理获取的  $h_m$  已经成为独立的、综合性的指数,并具有测度的功能。

排除经验判断和主观假设,在微博影响力评测的实践中, $r$  值如何的设置,还需要经过深入的研究和分析。

#### 4 小 结

$h$  指数应用到微博影响力评价中,具有诸多优势,可以丰富微博评价指标体系,建议微博平台的数据分析中心可以将其引入到评价分析中。

图书馆可通过  $h_f$  指数和  $h_c$  指数对博文进行选择和分析。通过整理高影响力博文为用户导读,分析用户关注的微博主题,以便有针对性地提供信息服务或改良信息服务。

从数据分析看, $h_f$  指数、 $h_c$  指数、 $h_m$  指数高的微博,一般来说都是开设时间长、博文数量多、粉丝多的微博,被转发数、评论数也相应高。因此,要提高微博的影响力,从根本来说还是要提高内容质量,通过时间的积累和与用户形成良性互动,提升其知名度和美誉度,而不是靠灌水、刷屏和“僵尸粉”去拔高单项指标数据,粉饰出虚假繁荣。

#### 参 考 文 献

- [1] Twitter reaches half a billion accounts [EB/OL]. [2012-10-02]. [http://semioast.com/publications/2012\\_07\\_30\\_Twitter\\_reaches\\_half\\_a\\_billion\\_accounts\\_140m\\_in\\_the\\_US](http://semioast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US).
- [2] 新浪微博注册用户突破3亿 [EB/OL]. [2012-10-02]. [http://news.xinhuanet.com/tech/2012-02/29/c\\_122769084.htm](http://news.xinhuanet.com/tech/2012-02/29/c_122769084.htm)
- [3] 第30次中国互联网络发展状况统计报告 [EB/OL]. [2012-09-25]. <http://www.cnnic.cn/hlwfzyj/hlwxxzbj/hlwtjbg/201207/P020120723477451202474.pdf>
- [4] 李 华,赵文伟. 微博客:图书馆的下一个网络新贵工具[J]. 图书与情报,2009(4):78-82
- [5] 刘钟美,张文彦. 高校图书馆的微博新时代[J]. 图书馆理论与实践,2012(4):77-81
- [6] Louis Yu, et al. What Trends in Chinese Social Media [C]. SNAKDD'2011:2-4
- [7] Shaozhi Ye, S. Felix Wu. Measuring Message Propagation and Social Influence on Twitter.com [C]. SocInfo, 2010: 223-228
- [8] 李 军,等. 微博影响力评价研究[J]. 信息安全,2012(3):10-13,27
- [9] Smith A G. Issues in "blogmetrics": Case Studies Using Blog-Pulse to Observe Trends in Weblogs [A]. In: Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics [C]. Madrid, 2007: 659-671
- [10] Weng J S, et al. TwitterRank: Finding topic-sensitive influential Twitterers [A]. In: Proceedings of the 3rd ACM International Conference on Web search and data mining (WSDM 2010) [C]. New York: ACM, 2010: 261-270
- [11] Cha M Y, et al. Measuring user influence in Twitter: The million follower fallacy [A]. In: Proceedings of International AAAI Conference on Weblogs and Social Media (ICWSN 2010) [C]. Washington, Menlo Park: The AAAI Press, 2010
- [12] 原福永等. 微博用户的影响力指数模型[J]. 现代图书情报技术,2012(6):60-64
- [13] Using the H-index to Estimate Blog Authority [EB/OL]. [2012-10-21]. <http://labs.sapo.pt/publications/devezas2011.pdf>
- [14] FEUP at TREC 2010 Blog Track: Using h-index for blog ranking [EB/OL]. [2012-10-21]. [http://trec.nist.gov/pubs/trec19/papers/labs-sapo-up\\_blog.pdf](http://trec.nist.gov/pubs/trec19/papers/labs-sapo-up_blog.pdf)
- [15] Influence and Passivity in Social Media [EB/OL]. [2012-10-21]. <http://www.hpl.hp.com/research/scl/papers/influence/influence.pdf>
- [16] 张晓阳,李晓亮. 科学家博客  $h$  指数评价及其相关性分析[J]. 图书情报工作,2010(2):66-69
- [17] Hirsch J E. An Index to Quantify an Individual's Scientific Research Output [J]. Proceedings of the National Academy of Sciences of the USA, 2005(46):16569-16572
- [18] Rousseau R. New Developments Related to the Hirsch Index [J]. 科学观察,2006(4):23-25
- [19] 叶 鹰.  $h$  指数和类  $h$  指数的机理分析与实证研究导引[J]. 大学图书馆学报,2007(5):2-5
- [20] 叶 鹰.  $h$  指数与  $h$  型指数研究[M]. 北京:科学出版社,2011:51-52
- [21] Braun T, Glanzel W, Schubert A. A Hirsch-type Index for Journals [J]. The Scientist, 19(22):8-10
- [22] 万锦堃,等. 中国部分重点大学  $h$  指数的探讨[J]. 科学观察,2007,2(3):9-16
- [23] 官建成等. 运用  $h$ -指数评价专利质量与国际比较[J]. 科学学研究,2008(5):932-936
- [24] "211工程"学校名单 [EB/OL]. [2012-10-22]. [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s238/201002/xxgk\\_82762.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s238/201002/xxgk_82762.html)

(责编:刘影梅)