

doi:10.3772/j.issn.2095-915x.2015.06.004

医学类院校的网络影响力分析

雷罗¹, 陈先来¹, 冯麟^{1,2}

(1. 中南大学信息安全与大数据研究院医药信息系, 长沙 410013;

2. 医学信息研究湖南省普通高等学校重点实验室(中南大学), 长沙 410013)

摘要: 随着信息网络的蓬勃发展, 网络的影响力日趋增强, 成为了社会中不可忽视的一股重要力量。本文主要运用网络链接分析的方法, 对医学类院校进行网络影响力分析, 得出院校的网站的的影响力情况。分析结果显示实际得出的网络影响力的排名与院校排名有一定的联系, 但并非密切相关。

关键词: 网络链接分析, 网络影响力分析, 医学类院校

Network Influential Analysis of Medical Colleges

LEI Luo¹, CHEN XianLai¹, FENG Lin^{1,2}

(1. Institute of information security and big data; 2. Institute of information security and big data, Key Laboratory of Medical Information Research (Central South University) College of Hunan Province, Changsha 410013)

Abstract: In this paper, a method using a network link analysis, medical colleges network impact analysis, the influence of the situation Colleges website. The results obtained show the actual network influential rankings and institutions ranking a definite link, but not closely related.

Key words: Network link analysis, network impact analysis, medical colleges

作者简介: 雷罗, 研究生, 中南大学信息安全与大数据研究院, 医学数据处理方向。冯麟, 研究生, 中南大学信息安全与大数据研究院, 医药信息管理方向。陈先来, 男, 博士生导师, 中南大学信息安全与大数据研究院医药信息系, 医学信息系统方向。

一、引言

随着当前社会信息化、网络化的不断发展,网络影响力已经成为各个学科领域不可避免的研究热点。根据中国互联网络信息中心(CNNIC)发布的《第35次中国互联网络发展状况统计报告》报告显示,截至2014年12月,我国网民规模达6.49亿,全年共计新增网民3117万人;互联网普及率为47.9%,较2013年底提升了2.1个百分点。这表明中国互联网络一直保持着蓬勃的发展,网络影响力的重要性随着网民数量的增加也在持续增加。

从引文分析^[1](citation)的概念出发,McKiernan于1996年首先提出了“situation”这一新术语,用以研究网页之间的引用关系。Ingwersen则受期刊影响因子的概念启发于年提出了网络影响因子^[2](Web Impact Factor,WIF)的概念来评价网站在网上的影响力。Rousseau认为对网页链接关系的研究与对发表文章的引文研究相似,但又不尽相同。他对网址的分布模式和进入网页的链接作了分析,发现在他所研究的343个网址中最高层域名服从洛特卡分布,而且对这些网址的引用也符合洛特卡分布,自引比例约为30%^[3]。而且随着越来越多的学者参与其中,链接分析的研究正逐步深入,其应用也越来越广泛。由美国斯坦福大学开发的Google搜索引擎^[4]在许多学者的相关研究中被应用于链接分析,除了搜索引擎之外,越来越多的站长工具也受到网络影响力分析研究人员的欢迎^[5]。通过站长工具可以准确及时的掌握网站的收录情况、alexa排名、PR值(全称为PageRank(网页级别),用来表现网页等级的一个标准,级别分别是0到10,是Google用于评测一个网页“重要性”的一种方法,PR值越高说明该网页越受欢迎(越重要)。)等,国内有许多优秀的站长工具,站长之家(中国站长站,ChinaZ)就是其中之一,其拥有上百万用户、最专业的行业资讯频道、国内最大的建站源码下载中心、站长聚集的站长社

区、最大的建站素材库、最实用的站长工具,以及最大的中文网站流量统计分析系统“站长统计”(每天为中文网站分析60-80亿PV以上)。本文选取站长之家的站长工具和Google搜索引擎作为分析工具。

二、研究方法

(一) 研究对象和途径

本文研究对象主要是国内排名较前的医学类院校,排名取自医学教育网2015年6月的数据,初步计划取前20名的院校(综合性大学则取其独立的医学院网站)作为研究对象,利用站长工具(在实际调查过程中部分学校由于在站长工具内无法查询到部分数据故对其进行了替换,具体替换名单在后文中会予以说明)分别查询各院校网站的PR值、PR输出值、平均的人均页面浏览量、反向链接数、Alexa排名(全球综合排名、中文排名),并利用Google搜索引擎查询20所学校的总网页数和总链接数。利用查到的相应网站的反向链接数、总网页数、总链接数分别计算20所院校的总网络影响因子和外部网络影响因子。分析上述数据并利用平均分析法进行名次法综合评价20所学校的网络影响力。

(二) 检索方法

中国站长之家的站长工具提供:

(1) PR查询, 可得各院校网站的PR值和PR输出值;

(2) Alexa排名: 可得全球综合排名、中文排名、平均的人均页面浏览量、反向链接数;

Google搜索引擎可查询:

(1) 总网页数: 例: site: www.csu.edu.cn

(2) 总链接数: 例: link: www.csu.edu.cn

三、结果与分析

按照以上实验方法检索得出数据如下表 1 和表 2:

表 1 各院校 PR 值和 PR 输出值

学校	网址	PR 值	PR 输出值
中国医学科学院北京协和医学院	www.pumc.edu.cn	7	6.10
北京大学医学部	www.bjmu.edu.cn	8	2.42
复旦大学上海医学院	shmc.fudan.edu.cn	7	3.13
西安交通大学医学部	www.med.xjtu.edu.cn	5	4.40
中山大学中山医学院	zssom.sysu.edu.cn	5	0.28
四川大学华西临床医学院	hx.scu.edu.cn	4	0.33
中南大学湘雅医学院	www.xysm.csu.edu.cn	5	0.39
首都医科大学	www.ccmu.edu.cn	6	5.25
中国医科大学	www.cmu.edu.cn	6	0.43
南京大学医学部	med.nju.edu.cn	6	0.88
哈尔滨医科大学	www.hrbmu.edu.cn	6	0.43
广西医科大学	gxmu.edu.cn	5	0.43
南方医科大学	www.fimmu.com	6	0.25
山西医科大学	www.sxmu.edu.cn	5	1.21
大连医科大学	www.dlmedu.edu.cn	6	0.88
天津医科大学	www.tjmu.edu.cn	7	0.89
沈阳药科大学	www.syphu.edu.cn	6	5.25
山东大学医学院	www.medicine.sdu.edu.cn	5	4.40
福建中医药大学	www.fjtcu.edu.cn	6	1.17
湖南中医药大学	www.hnctcm.edu.cn	6	1.17

表 2 各院校 Alexa 排名

学校	全球综合排名	中文排名	人均页面浏览	反向链接数	总网页数	总链接数
中国医学科学院北京协和医学院	221,421	29,809	5.5	410	231,000	179,000
北京大学医学部	199,596	18,232	1.0	1,273	158,000	181,000
复旦大学上海医学院	9,538	1,031	5.1	9,421	2,460	1,490
西安交通大学医学部	8,771	867	6.8	2,356	66,600	68,100
中山大学中山医学院	15,067	1,863	3.2	2,663	6,720	4,120
四川大学华西临床医学院	23,024	2,595	2.6	2,361	625,000	539,000
中南大学湘雅医学院	17,992	2,404	2.4	1,821	3,370	3,700
首都医科大学	537,007	20,936	1.2	300	70,700	70,600
中国医科大学	202,109	23,695	5.3	642	51,400,000	55,700,000
南京大学医学部	14,865	1,557	2.5	4,303	2,040,000	1,550,000
哈尔滨医科大学	60,649	5,641	3.5	649	132,000	117,000
广西医科大学	167,990	16,572	3.0	5,992	54,700	58,500
南方医科大学	345,746	22,192	9.0	473	133,000	96,600
山西医科大学	208,148	28,691	5.3	262	148,000	151,000
大连医科大学	127,482	22,696	4.6	375	15,600	17,000
天津医科大学	199,534	36,300	6.0	520	11,500	9,360
沈阳药科大学	231,400	54,863	5.3	399	10,600	17,700
山东大学医学院	17,566	2,537	1.9	2,456	2,150,000	3,050,000
福建中医药大学	1,174,780	19,635	1.7	414	12,300	12,400
湖南中医药大学	712,242	27,412	1.2	272	12,800	13,600

利用查到的相应网站的反向链接数、总网页数、总链接数分别计算 20 所院校的总网络影响因子和外部网络影响因子结果如下表 3：（其中：总网络影响因子（整体 WIF）= 总链接数 / 总网页数；外部网络影响因子（外部 WIF）= 反向链接数 / 总网页数）

表 3 各院校影响因子

学校	总网络影响因子	外部网络影响因子
中国医学科学院北京协和医学院	0.77	0.001775
北京大学医学部	1.15	0.008057
复旦大学上海医学院	0.61	3.829675
西安交通大学医学部	1.02	0.035375
中山大学中山医学院	0.61	0.396280
四川大学华西临床医学院	0.86	0.003778
中南大学湘雅医学院	1.10	0.540356
首都医科大学	1.00	0.004243
中国医科大学	1.08	0.000012
南京大学医学部	0.76	0.002109
哈尔滨医科大学	0.89	0.004917
广西医科大学	1.07	0.109543
南方医科大学	0.73	0.003556
山西医科大学	1.02	0.001770
大连医科大学	1.09	0.024038
天津医科大学	0.81	0.045217
沈阳药科大学	1.67	0.037642
山东大学医学院	1.42	0.001142
福建中医药大学	1.01	0.033659
湖南中医药大学	1.06	0.021250

根据所得结果，对 PR 值、PR 输出值、全球综合排名、中文排名、平均的人均页面浏览量、总网络影响因子和外部网络影响因子七项指标分别进行排序，对各单项排序结果按名次进行计分，第一名计 20 分，第二名计 19 分，依次递减，最

后一名计 1 分，最后将各院校网站所得的七项分数相加得出总分并对总分进行排序得出最后的结果，这就是本次采取的平均名次排序的方法。各院校排名计分情况如下表 4：

表 4 单项排名积分情况与总分排序

学校	PR 值	PR 输出值	全球综合排名	中文排名	人均页面浏览	整体 WIF	外部 WIF	总分
西安交通大学医学部	7	17	20	20	19	12	14	109
复旦大学上海医学院	19	15	19	19	13	1	20	106
山东大学医学院	17	17	16	15	5	19	2	91
中南大学湘雅医学院	7	4	15	16	6	17	19	84
大连医科大学	16	9	12	7	12	16	12	84
北京大学医学部	20	14	9	11	1	18	10	83
沈阳药科大学	16	19	5	1	16	20	5	82
天津医科大学	19	10	10	2	18	6	16	81

南京大学医学部	16	9	18	18	7	4	5	77
哈尔滨医科大学	16	7	13	13	11	8	9	77
广西医科大学	7	7	11	12	9	14	17	77
中国医学科学院北京协和医学院	19	20	6	3	17	5	4	74
中山大学中山医学院	7	2	17	17	10	2	18	73
中国医科大学	16	7	8	6	16	15	1	69
首都医科大学	16	19	3	9	3	9	8	67
福建中医药大学	16	12	1	10	4	10	13	66
山西医科大学	7	13	7	4	16	12	3	62
湖南中医药大学	16	12	2	5	3	13	11	62
南方医科大学	16	1	4	8	20	3	6	58
四川大学华西临床医学院	1	3	14	14	8	7	7	54

原排名前 20 的学校依次为：

中国医学科学院北京协和医学院、北京大学医学部、复旦大学上海医学院、上海交通大学医学院、中山大学中山医学院、四川大学华西临床医学院、中南大学湘雅医学院、首都医科大学、中国医科大学、华中科技大学同济医学院、哈尔滨医科大学、浙江大学医学院、南方医科大学、中国药科大学、吉林大学医学院、天津医科大学、重庆医科大学、山东大学医学院、南京医科大学和河北医科大学。

其中上海交通大学医学院、华中科技大学同济医学院、中国药科大学、吉林大学医学院、重

庆医科大学、南京医科大学和河北医科大学由于缺乏中文排名、平均的人均浏览量等数据被更换为南京大学医学部、西安交通大学医学部、山西医科大学、大连医科大学、沈阳药科大学、福建中医药大学和湖南中医药大学，浙江大学医学院由于没有独立的网址（输入浙江大学医学院网址到站长工具则自动指向浙江大学）被替换为西安交通大学医学院。

以下一系列图依次为 PR 值、PR 输出值、全球综合排名、中文排名、平均的人均页面浏览量、总网络影响因子和外部网络影响因子单项排名得分与总分的关系图：

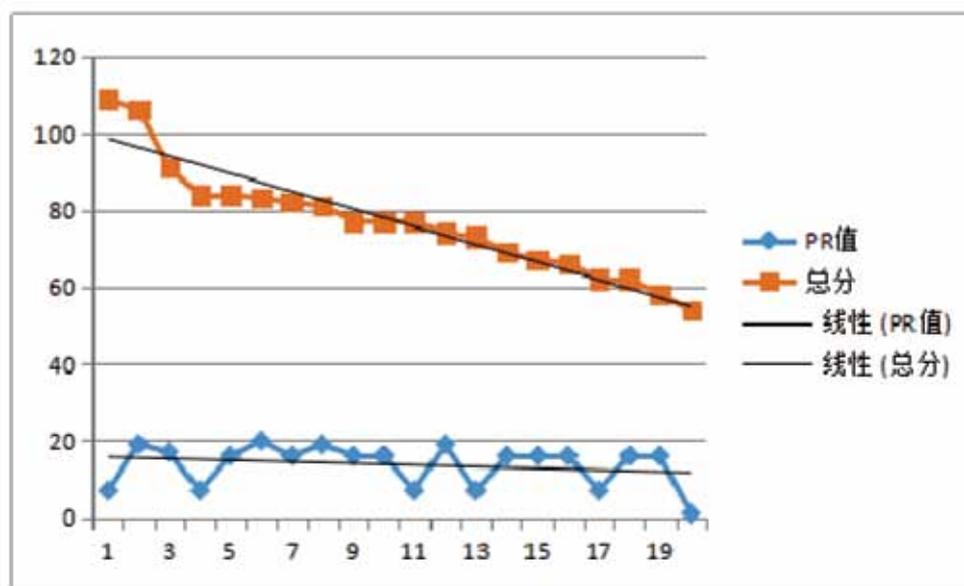


图 1 PR 值单项排名得分与总分的关系

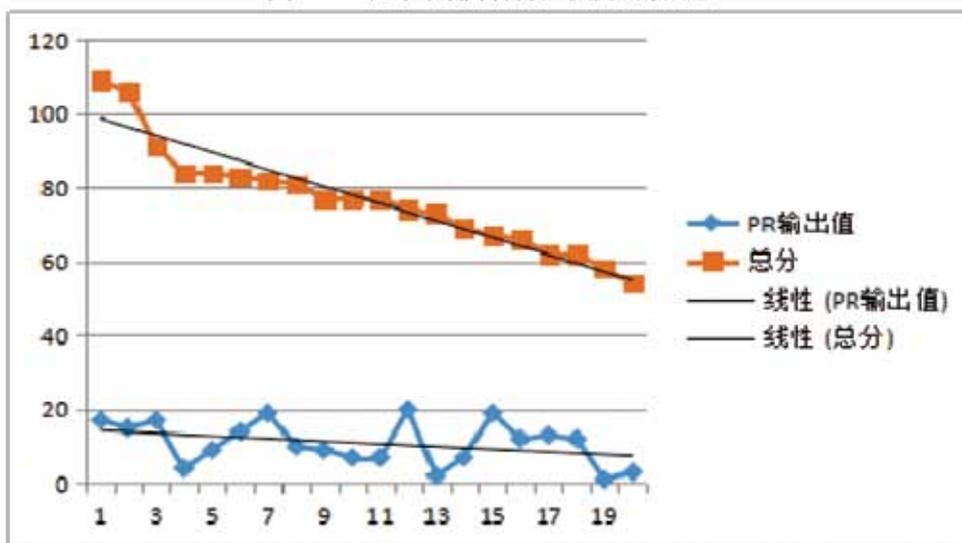


图2 PR输出值单项排名得分与总分的关系

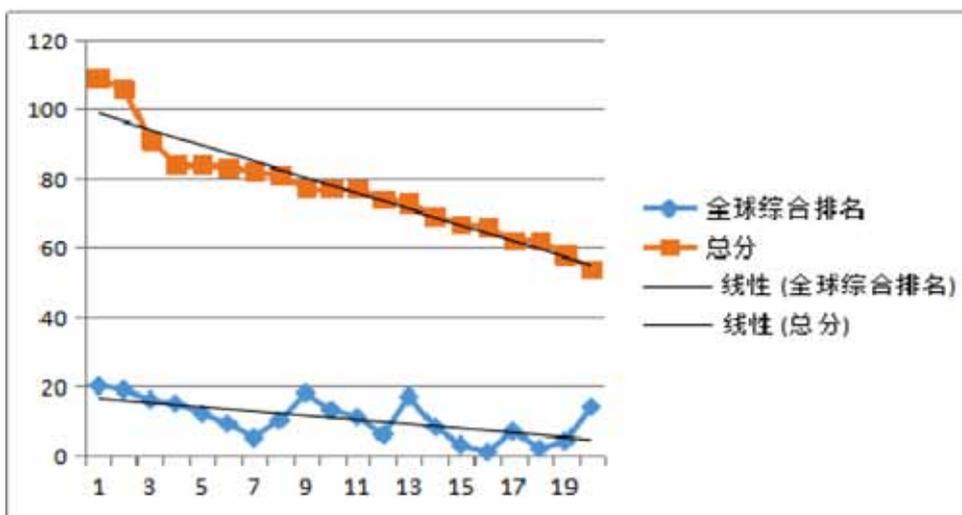


图3 全球综合排名单项排名得分与总分的关系

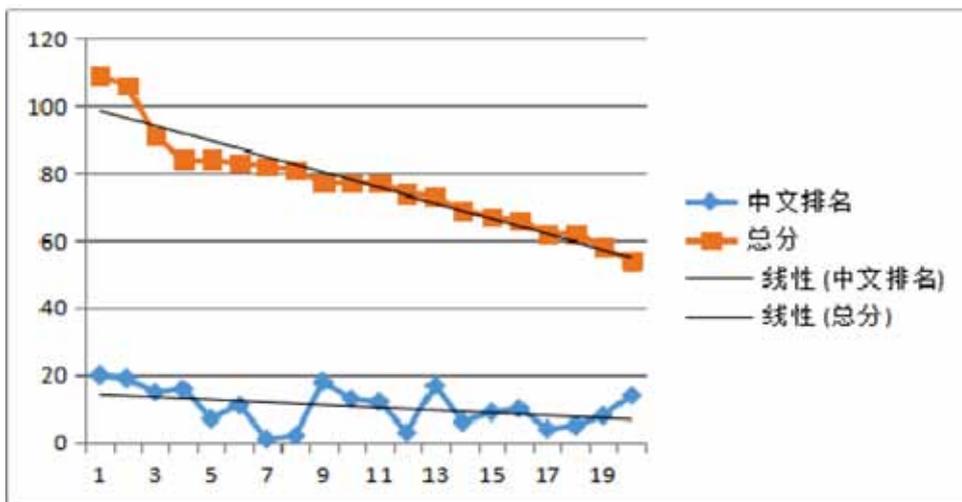


图4 中文排名单项排名得分与总分的关系

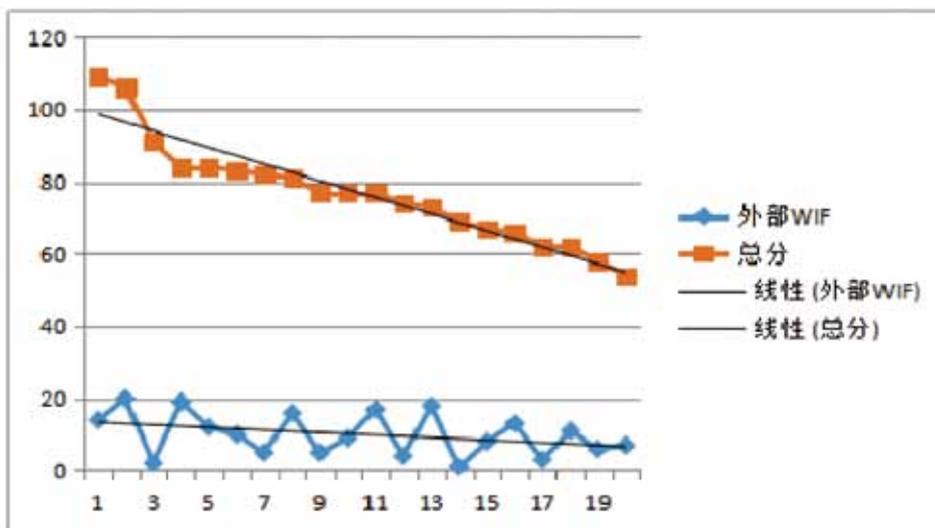


图 5 平均的人均页面浏览量单项排名得分与总分的关系

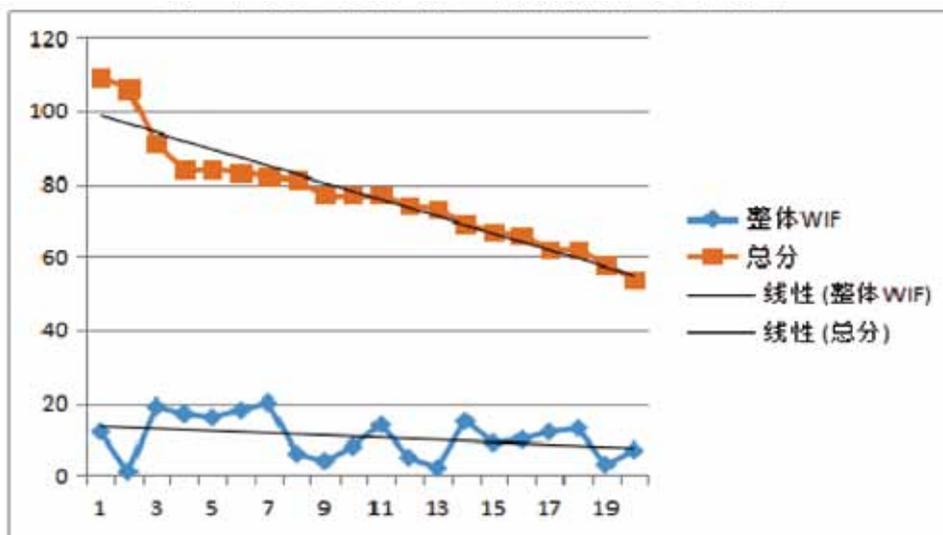


图 6 总网络影响因子单项排名得分与总分的关系

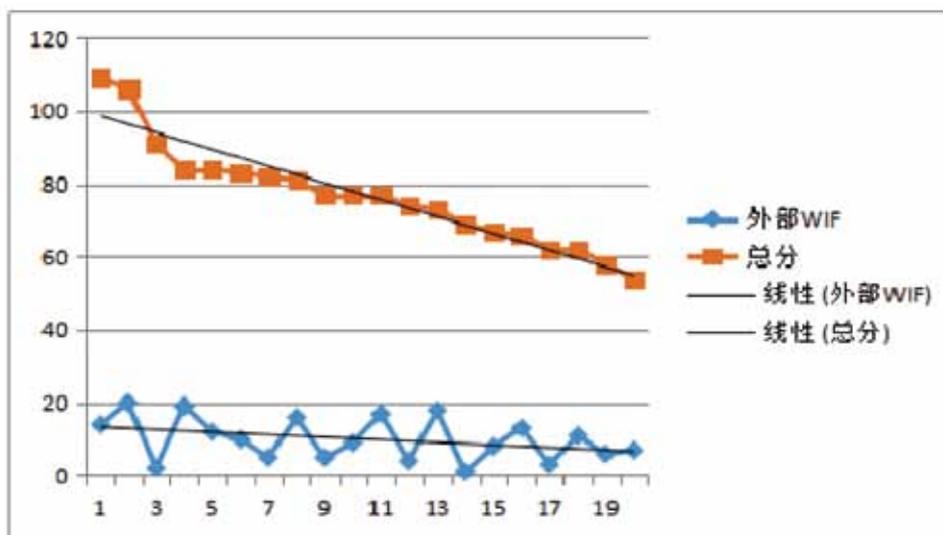


图 7 外部网络影响因子单项排名得分与总分的关系

从最终分析得出的平均排名情况得知,在20所综合实力尚佳且网站建设情况较好的20所医学类院校中,总排名更为靠前的院校网站大多得分较为均衡,且在PR、Alexa排名、网络影响因子三方面有一项或多项得分较高,而且各院校在医学方面的排名和网络影响力排名并没有十分密切的联系。相比之下,与总分变化趋势较为接近的是网站的全球排名和外部网络影响因子。

四、分析与讨论

本文主要运用网络链接分析^[6]的方法,实际得出的网络影响力与国内医学类院校的排名情况有一定的差距,最终网络影响力的排名与院校排名有一定的联系,但并非密切相关。导致这一结果的原因可能有:

(1) 排名取自医学教育网,其排名机制与网络影响力分析机制有差异,且院校排名的分析情况与分析指标及其权重和分析者的主观判断等因素有关,并不存在绝对客观的排名,故任何排名情况都只能作为参考。

(2) 采取的分析工具(站长之家的站长工具、Google搜索引擎)虽然是较为公认的可用的分析工具,但是其分析必然也有一定的局限性,尤其是在中国医学类院校这一小环境之内其适应性不一定能达到最好。

(3) 医学学科特性导致对网络需求的偏低,且我国众多知名医学院已并入综合性大学,其网站成为大学一系列网站的一部分,许多功能在校级或其他网站中得以实现故对医学院网站的影响力估算并不能完全体现整个学院的网络影响力。

(4) 中国的医学院校对于网站的建设的重视程度不一致,网站使用人员的使用意识的强弱有别等原因,均可能导致“名校网无名”的现象。

(5) 虚假链接、临时链接、“友情”链接等

情况的存在仍在一定程度上干扰了网络影响力的结果。

(6) 由于没有经过专家论证等环节,直接采取相等权重的平均名次法来进行排名,忽略了各项指标对网络影响力的贡献大小,可能造成结果的误差。

(7) 链接分析法本身存在的问题:随着研究的深入,人们越来越发现文献之间的引用与网络信息之间的链接存在着显著的差异。譬如,Stephen P.Harter和Charlotte E.Ford^[7]的研究发现,在指向电子期刊的网络链接中,有近半数是结构性自我链接,对于评价文献的学术价值和研究学术交流行为没有意义^[8];具有指向电子期刊的链接的网页,只有不到8%的网页是学术性的;Thelwall通过对英国大学网站中被链接次数最多的100个网页进行分析,其结果也表明网页被链接次数并不能反映其学术价值^[9]。这使得研究者们开始重新反思网络链接分析研究领域最基本的问题。

总的来说,我国的医学类院校的网站建设和使用情况仍不尽人意,全球排名和中文排名普遍偏低,总网络影响因子和外部网络影响因子也都不高,虽然与许多因素有关,但仍表明我国需加大医学类院校的网站建设和宣传工作,培养学生利用网络资源自主学习的能力,这不仅是顺应信息时代发展的要求,也是对学校的建设管理能力和学生自主学习能力——我国大学生如今尤其缺乏的能力的培养。

我国尚处于高等教育的改革阶段^[10],国外高校学生通过网络自主学习是他们大学生活中相当重要的一部分,这不仅提高了学生自主学习的能力,也拓宽了他们的视野,给了他们更广阔的空间去发展。我国也应学习这一点,一方面培养学生自主学习与思考,尽早摆脱“书呆子”、“分数奴隶”和“60分万岁”的情况;另一方面要加

大高校网络建设, 加强网络监督与管理, 为学生提供良好的在线学习资源。当学生有学习的欲望、掌握学习的方法、有充足的知识来源时, 我国高等教育的面貌或许能够得到改善。

参考文献:

- [1] 杨思洛. 引文分析存在的问题及其原因探究 [J]. 中国图书馆学报, 2011(03):108-117.
- [2] 张洋. 网络影响因子研究综述 [J]. 中国图书馆学报, 2010(01):63-79.
- [3] 曹姗姗, 王冲. 基于网页链接与用户反馈的 PageRank 算法改进研究 [J]. 计算机科学, 2014b (12):179-182.
- [4] Langville A N, Meyer C D. Google's PageRank and beyond: The science of search engine rankings[M]. Princeton University Press, 2011.
- [5] 米云. 基于链接分析的网站评价研究 [D]. 大连理工大学, 2013.
- [6] 张洋, 邱均平, 文庭孝. 网络链接分析研究进展 [J]. 图书情报知识, 2005 (6): 3-8.
- [7] Harter S P, Ford C E. Web-based analyses of e-journal impact: approaches, problems, and issues[J]. Journal of the American Society for Information Science, 2000, 51(13): 1159-1176.
- [8] 吴华香, 钟少丹. 链接分析法——网络计量学方法初探 [J]. 情报科学, 2002(01):71-73.
- [9] 邱均平, 陈敬全, 段宇锋. 中国大学网站链接分析及网络影响因子探讨 [J]. 中国软科学, 2003(06):151-155.
- [10] 周远清. 我国高等教育改革与发展的回顾与展望 [J]. 高等教育研究, 2001, 22(1): 1-8.