

网络计量方法在多学科综合评估中的实证应用

吴茵茵

(华南农业大学图书馆, 广东 广州 510642)

〔摘要〕学科建设始终是高等教育中的重中之重, 而科学客观的评估则是实现学科建设健康发展的重要推动力。本文利用信息计量学中的网络链接分析方法, 对美国大学的计算机科学、生物医学工程和心理学三个学科的网络影响力进行了综合对比分析, 发现网络链接分析方法能够有效地表征不同学科之间的特点, 从而为多学科建设的科学评估提供了新的思路。

〔关键词〕学科建设; 网络链接分析; 信息计量学

〔中图分类号〕G350 〔文献标识码〕A 〔文章编号〕1008-0821(2010)04-0037-04

Webometric Research in the Evaluation of Different Academic Disciplines

Wu Yinyin

(Library, South China Agricultural University, Guangzhou 510642, China)

〔Abstract〕The development of the academic discipline is crucial for the advanced education, and evaluation of discipline in a scientific and objective way is of great help to this progress. The present study applied the webometric analysis to explore the linking patterns of three disciplines in the universities of the United States: computer science, biomedical engineering, and psychology. The results showed that this method can effectively reveal the specific characteristics of different academic disciplines, therefore providing a new way to evaluate the development of multiple academic disciplines objectively.

〔Key words〕discipline development; web linking analysis; webometrics

信息时代发展的今天, 海量的信息正在进行数字化和网络化, 因而提取网络数据进行各个方面的分析为社会发展和科学研究提供了更好的服务。网络信息计量学^[1] (Webometrics) 在这种需要下应运而生, 其中心内容就是充分利用强大的搜索引擎, 建立基于数学统计和科学计算的数学模型来对海量的信息数据进行提取和分析。其中链接分析方法是最为常用和重要的信息计量学方法^[2]。近几年来, 国内外已有学者利用网络信息计量学来研究大学网络影响力^[3-5], 侧重点在于研究大学整体排名和网络链接指数的相关性。研究结果表明总体网络影响因子并不适合对大学网站进行评价。其中一个重要原因在于研究者没有将大学整体影响力和学科单独的影响力分开, 也就是说忽视了学科建设对于大学整体发展的影响力。

作为高等教育发展的重中之重, 学科建设是实现高等院校快速发展的必需和保障, 而对学科建设进行科学客观评估则是实现这一过程的重要推动力。现在评估方法包括两类^[6]: 一类是同行评议的定性评价方法, 而另一类则是

包括文献计量和成本效益方法在内的定量评议方法。目前学科建设评估的主要方法是采取定性定量相结合的方法。但是从各种定量分析方法的具体应用来看, 它们仍然无法摆脱定性分析的影响, 如层次分析法中人们需对问题每一层次各种因素的相对重要性进行主观判断来决定其分值, 这就大大影响了分析方法的客观性和结论的可靠性。而随着数字化信息时代的来临和互联网的大规模应用与发展, 高等院校的影响力评价出现了新的变化, 即引入了网络影响力评价这一指标。例如西班牙教育部高等学术委员会在2009年9月公布了最新“全球大学排行榜”, 其中首次引入了大学网站的网络影响力作为大学排名的依据之一。从这里可以看出, 网络的影响力大小已经成为反映大学排名和实力的一项指标, 值得引起特别的重视。

近年来国外开展了网络计量学在学科建设评估中的应用研究^[7-8]。Thelwall领导的研究小组对连接到美国大学各学院有关化学、心理学和历史学等3个学科的链接进行了归类分析, 并且比较一个机构外部链接数和该机构出版物

被引情况。研究结果表明网络计量学指标与心理学和化学学科的论文发表影响力存在显著性的相关,同时这种相关与该学科的国际学术交流有关。这说明网络计量学用于学科建设评估方面不但是有效的,并且它比其他研究方法在效率以及数据更新方面有着更大的优势。

目前虽然网络计量学在学科建设评估中应用已有所研究,但是它们或者是针对单个学科的研究^[7],或者是缺乏多个学科之间的纵向比较研究^[8]。基于此,本研究选取了心理学、生物医学工程和计算机科学3个学科,以美国在这3个学科排名前十位的高校为研究对象,分析它们的网络链接分析数据的表达特征,探讨网络计量学在高校多学科建设评估方面的可能性。

1 研究方法

本研究选取了美国高校在心理学、计算机科学和生物医学工程3个学科方面的前十位来获取网络链接分析数据,其选取依据是根据美国新闻周刊2010年美国大学研究院排名。

本研究使用的是AltaVista搜索引擎,因为AltaVista比Google更为稳定,并明显优于Excite和Lycos^[9-10]。表1以斯坦福大学为例来说明AltaVista搜索引擎的搜索语法,所

用方法来自于Li^[11]:

表1 AltaVista 搜索语法

项 目	语 法
链接总数	D = LINK: cs. stanford. edu
外部链接总数	E = LINK: cs. stanford. edu AND NOT HOST: stanford. edu
edu 域名的外部链接总数	edu - E = LINK: cs. stanford. edu AND DOMAIN: edu AND NOT HOST: stanford. edu
com 域名的外部链接总数	com - E = LINK: cs. stanford. edu AND DOMAIN: com AND NOT HOST: stanford. edu
net 域名的外部链接总数	net - E = LINK: cs. stanford. edu AND DOMAIN: net AND NOT HOST: stanford. edu
链接效率 (Index efficiency, IE) = 外部链接总数/链接总数 = E/D	

2 数据分析和结果

笔者在2009年11月19日上午9时至11时对选定的美国高校的计算机科学、生物医学工程和心理学学科进行了网络链接数据的测定,具体结果如表2~4所示:

表2 美国高校计算机专业网络分析数据

美国高校 计算机科学前十名	链接总数	外部链接数	edu 域名 外部链接数	com 域名 外部链接数	net 域名 外部链接数
麻省理工学院	90 400	31 200	18 500	17 700	16 800
斯坦福大学	850 000	171 000	96 600	89 300	79 200
加州大学伯克利分校	1 010 000	364 000	128 000	117 000	107 000
卡内基梅隆大学	775 000	139 000	76 800	70 500	65 800
伊利诺伊大学香槟分校	363 000	83 900	45 600	42 500	41 000
康奈尔大学	485 000	92 500	49 100	45 100	41 100
普林斯顿大学	470 000	87 300	45 500	42 000	38 400
华盛顿大学	1 560 000	464 000	130 000	120 000	107 000
佐治亚理工学院	161 000	45 600	23 800	22 600	21 100
德州大学奥斯汀分校	176 000	51 000	30 000	29 000	27 000

表3 美国高校生物医学工程专业网络分析数据

美国高校 生物医学工程前十名	链接总数	外部链接数	edu 域名 外部链接数	com 域名 外部链接数	net 域名 外部链接数
约翰霍普金斯大学	14 800	756	371	363	332
麻省理工学院	5 560 000	1 310 000	578 000	506 000	441 000
加州大学圣地亚哥分校	653	16	7	6	5
杜克大学	21 500	1 550	656	628	591
华盛顿大学圣路易斯分校	12 900	787	472	466	452
凯斯西储大学	45 000	13 300	2 230	1 920	1 620
佐治亚理工学院	17 400	1 060	558	546	516
密歇根大学	17 700	1 270	661	650	622
宾夕法尼亚大学	56 700	14 700	3 520	3 030	2 790
加州大学伯克利分校	1 570	363	148	81	52

表4 美国高校心理学专业网络分析数据

美国高校 心理学前十名	链接总数	外部链接数	edu 域名 外部链接数	com 域名 外部链接数	net 域名 外部链接数
斯坦福大学	80 900	26 700	8 840	6 570	3 620
加州大学伯克利分校	802 000	126 000	39 900	33 200	28 800
哈佛大学	1 440	86	36	36	25
加州大学洛杉矶分校	48 700	10 600	1 690	1 450	1 470
密歇根大学	10 500	392	74	66	64
耶鲁大学	635 000	102 000	28 000	23 900	18 900
伊利诺伊大学香槟分校	27 900	2 280	1 270	1 250	1 130
普林斯顿大学	45 700	9 030	1 530	1 280	1 280
明尼苏达大学双城分校	30 100	2 120	1 060	1 040	988
威斯康星麦迪逊大学	29 800	1 980	1 070	1 050	985

从表2~4中可以看出，链接总数和外部链接总数的数据随着学科的不同而发生变化。Kruskal - Wallis 统计分析结果显示，不同学科之间在链接总数 ($\chi^2 = 11.384$, d.f. = 2, $p = 0.003$) 和外部链接总数 ($\chi^2 = 11.419$, d.f. = 2, $p = 0.003$) 的数据上存在显著性差异。组间分析结果显示计算机科学的链接数据最多，其次是生物医学工程，最少的是心理学。需要指出的是表3中的加州大学圣地亚哥分校和表4中哈佛大学的链接数据与其他高校相比过少，所以没有作为有效数据进行统计分析。

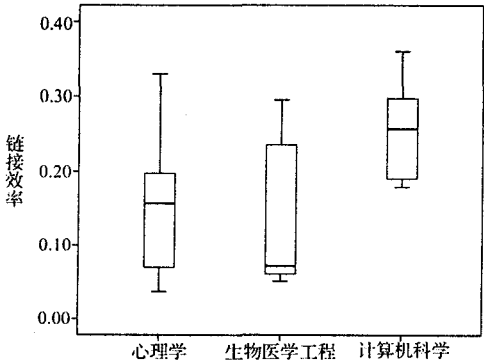


图1 美国高校心理学、生物医学工程、计算机科学专业的链接效率比较

图1中列出了各个学科之间的链接效率，从中可以看出心理学和生物医学工程的链接效率详尽，而计算机科学的链接效率要相对高一些。方差统计分析结果显示，链接效率在各个学科之间存在显著差异 ($F(2,25) = 4.853$, $p = 0.017$)。组间对比分析显示，计算机科学的链接效率要明显高于心理学 ($p = 0.037$) 和生物医学工程 ($p = 0.042$)，而心理学和生物医学工程之间则不存在显著性差异 ($p = 1.000$)。

以前的研究中已经证明链接效率和外部网络影响因子存在线性相关性，为了进一步研究链接效率和学科排名之间的关联，笔者对上述3个学科的大学排名和相对应的链接效率进行线性相关分析，线性拟合结果如图2。统计分析结果显示，计算机科学 ($F(1,8) = 0.040$, $p = 0.846$) 和生物医学工程 ($F(1,7) = 0.779$, $p = 0.407$) 学科的链接效率和大学学科排名之间并不存在关联，而心理学学科中这两者之间的相关性较好 ($F(1,7) = 4.236$, $p = 0.056$)。这些数据进一步说明不同学科之间的网络特征表现是存在差异的。

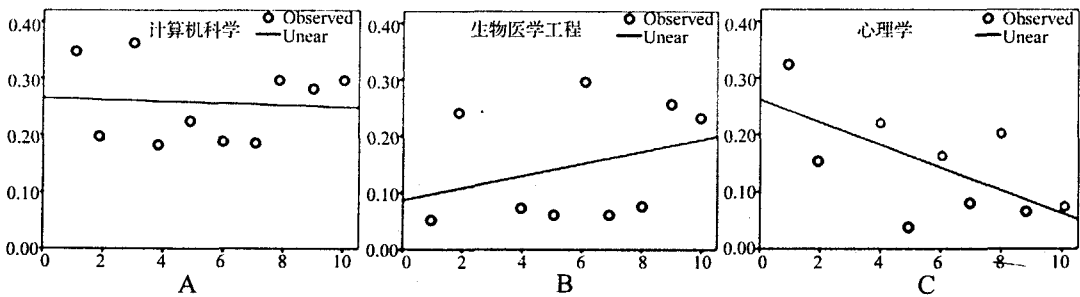


图2 美国高校计算机科学
A. 生物医学工程 B. 心理学 C. 专业链接效率与专业排名的相关分析

3 讨论

网络计量学在各个领域内的应用分析越来越广泛，然

而从学科建设评估的角度对大学学科进行定量评估还甚为少见。本研究第一次应用网络链接分析方法对美国高校的心理学、计算机科学和生物医学工程学科进行了对比分析，

有以下新的发现:

(1) 各种链接类型的链接总数在各个学科之间存在不同程度的差异,这有效地体现了各个学科之间的差异。比如计算机科学的链接总数最多,这是因为这一科学在社会领域、经济领域、工程领域、军事领域等有着极为广泛的应用,社会的需求量大,在信息时代是推动社会发展的强劲动力,因而在各个高校中一直是关注的热点,自然其链接数据包括链接效率就较其他两个学科要高。生物医学工程的历史并不长久,但是近年来随着工程技术的进步和社会对医疗健康需求的不断提高而得到了迅速的发展,日益引起众多高校的关注,加之该学科和生命科学的关系密切,而跨入21世纪以来生命科学明显成为社会关注的重中之重,因此该学科的各项链接指标紧随计算机科学气候甚至有超越之势也就不足为奇了。心理学虽然是一门历史悠久的学科,但是由于其应用的范围相当有限,加之不同层次的人群对其认同度并不相同,所以得到的关注相对另外两个学科来说就明显要少。因此,链接分析的数据可以有有效的反应出各个学科在不同时间的时代特征。

(2) 本研究的结果显示链接分析的数据和学科的大学排名没有直接关联。在以前的网络计量学研究中,网络影响因子和大学整体的排名的线性相关性没有得到一致的证明,但总体上两者并不存在直接关联。笔者曾提出这可能与这种关联忽视了各个大学之间的学科差异性有关,而本研究则表明尽管选取了较好的学科和较为公正的学科大学排名,网络链接数据和大学排名之间还是不存在明显的相关性。其中心理学学科中这两者的线性拟合程度尽管已经很高($p=0.056$),但是还不能由此断定在大样本的情况下这种线性相关性还会持续下去。另外,Thelwall等人的研究发现网络影响因子和科研产出存在正相关关系,即研究实力越强的学科,其网络计量学指标表现就越好,其国际影响力就越高^[8]。更为重要的是,研究发现这种相关性随着学科的不同会发生相应的变化^[12-13]。Aguillo等人^[14]的研究显示网络计量学的各种指标与各主要大学的研究实力存在显著性相关。所以,大学排名可能和网络链接分析的其他指标存在相关性,另一方面网络链接分析指标和大学排名以外的其他指标(科研阐述,投入资金,教职人数等等)的相关性可能更好。这需要进行更深层次的研究。

(3) 笔者之所以选择美国高校而没有选择中国高校,原因在于一方面中国高校还缺乏权威的学科大学排名,其公正性上存在质疑。另一方面在于很多中国高校在这3个学科的网络建设方面还非常粗糙,甚至无法找到高校在该学科网页的单独域名,以至于无法进行网络链接分析。美国高校的研究提示我们中国高校在学科发展的软建设方面仍然任重道远,提升学科的科研实力不仅仅是单纯的科研资金的透出、科研成果的阐述,还要注重其社会乃至

全球所发挥的影响。

参考文献

- [1] Tomas C.A., Ingwersen P. Informetric analyses on the World Wide Web: methodological approaches to "Webometrics" [J]. Journal of Documentation, 1997, 53 (4): 404-426.
- [2] Larson R.R. Bibliometrics of the World Wide Web: an Exploratory Analysis of the Intellectual Structure of Cyberspace [A]. In: Steve Hardin eds. Proceedings of the American Society for Information Science (ASIS) 59th Annual Meeting [C]. Oct21-24, 1996, Baltimore, MD, Medford, NJ: Information Today Inc. for ASIS: 71-78, 1996.
- [3] 邱均平, 安璐. 中文期刊影响因子与网络影响因子和外部链接数的关系研究 [J]. 情报学报, 2003, (4): 398-402.
- [4] 吴茵茵. 网络影响因子实证研究: 基于不同搜索引擎的大学网站影响力分析 [J]. 图书情报工作, 2005, (4): 107-111.
- [5] Ananpour F, Kabiri, P, Otraj Z et al. Webometric analysis of Iranian universities of medical sciences [J]. Scientometrics, 2009, 80: 253-264.
- [6] 姚捷. 高等院校重点学科建设评估研究进展 [J]. 北京中医药, 2008, 27 (5): 396-399.
- [7] Arakaki M, Willett P. Webometric analysis of departments of librarianship and information science: a follow-up study [J]. Journal of Information Science, 2009, 35: 143-152.
- [8] Tang R, Thelwall M. Patterns of national and international Web inlinks to US academic departments: An analysis of disciplinary variation [J]. Scientometrics, 2004, 60 (3): 475-485.
- [9] Thelwall M. The responsiveness of search engine indexes. Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics, 2001, 5 (1).
- [10] Clark S.J, Willett P. Estimate the recall performance of web search engines [J]. Aslib Proceedings, 1997, 49 (7): 184-189.
- [11] Li X., Thelwall M., Wilkinson D., Musgrove P. National and international university departmental web site interlinking. part 1: validation of departmental link analysis, Scientometrics [J], 2005, 64 (2): 151-185.
- [12] Li X, Thelwall M, Wilkinson D, Musgrove P. National and international university departmental web site interlinking. part 2: link patterns, Scientometrics [J], 2005, 64 (2): 187-208.
- [13] Li X, Thelwall M, Musgrove P, Wilkinson D. The relationship between the WIFs or inlinks of computer science departments in UK and their RAE ratings or research productive in 2001. Scientometrics [J], 2003, 57 (2): 239-255.
- [14] Aguillo IF, Granadino B., Ortega JL, and Prieto JA. Scientific research activity and communication measured with cybermetrics indicators, Journal of the American Society for Information Science and Technology [J], 2006, 57 (10): 1296-1302.