

●沙勇忠 欧阳霞

网络信息计量学研究方法的三维框架^{*}

摘 要 网络信息计量学研究方法在网络信息计量学学科体系中具有基础地位。网络信息计量学方法主要来源于3个学科领域,即信息计量学、文献计量学和科学计量学方法;数字统计学方法以及网络技术和软件技术。提出由理论研究方法、数据收集方法和结果展现方法构成的网络信息计量学研究方法的三维框架,以促进网络信息计量学科的学科范式的形成。图1。表1。参考文献15。

关键词 计量学 网络信息计量学 方法 方法体系 三维框架

分类号 G250.73

ABSTRACT The authors discuss some aspects of webometrics, including its roles, origins and methods. They propose that theoretical research method, data gathering method and result representation method constitute a three-dimensional framework of the webometric research methodology.

1 fig. 1 tab. 15 refs.

KEY WORDS Webometrics. Methodology. Methodological system. Three-dimensional framework.

CLASS NUMBER G250.73

网络信息计量学(Webometrics)是随着网络的发展和网络信息资源的开发利用而兴起的一门新兴学科,目的是通过对网上信息的计量研究,揭示其新的特征、数量关系和内在规律,为网上信息的有序化组织、网络信息资源的优化配置及网络管理的科学化提供理论支持和定量依据。

1 方法在网络信息计量学中的基础地位

网络信息计量学的学科性质和特点决定了它对方法和工具有强烈的依赖和需求。方法在网络信息计量学学科体系中具有基础性的地位。

网络信息计量学的内容体系是由它的理论、方法和应用3个部分组成,理论是基础,方法是手段,应用是目的,三者相辅相成^[1]。作为手段的方法是完整的网络信息计量学学科体系不可缺少的组成部分。

方法是网络信息计量学理论和应用之间的桥梁。方法在理论和应用之间的联系和沟通,促进了网络信息计量学学科体系的形成与进化。

网络信息计量学方法与理论和技术有密切的联系,很多研究方法本身就是一种理论或网络技术。一方面,一些网络信息计量学研究方法本身就是理论,如Situation, WIF, Co-Situation, Web-coupling等,既是网络信息计量学独特的研究方法,又是网络信息计量学最基本的学科概念和理论要素;另一方面,网络信息计量

学研究方法具有显著的技术特性,需要依附于技术或以技术的形式展现才能实现其功效,如网络搜索引擎分析、链接分析、Web挖掘分析等,既是网络信息计量学重要的数据收集和信息分析方法,又是典型的网络信息技术。撇开技术批判的意义,海德格尔所揭示的信息技术作为一种“展现方式”促进了世界的技术化展现,在网络信息计量学中则是一种内在的学科规定或学科特性。

方法及其应用研究一直在网络信息计量学研究中占据核心地位。M. A. Boudourides 和 G. P. Antypas 指出,自从 Abraham 等人开始网络信息计量学研究以来,方法问题一直是该学科的重要问题^[2]。网络信息计量学的产生被认为是将信息计量学的方法应用到WWW网上信息计量的结果^[3], Moses A. Boudourides 等人更是直截了当地认为,网络信息计量学本身就是一种面向WWW网络的方法^[4]。方法及其应用研究一直是网络信息计量学研究的重点,网络信息计量学是随着其方法的发展而不断发展的。

2 网络信息计量学的方法来源

从学科起源上考察,网络计量学是由网络技术、信息计量学、数学统计学等相互交叉渗透而形成的。

2.1 信息计量学、文献计量学和科学计量学方法

根据 Bjorneborn 和 Ingwersen 的观点,信息计量

^{*} 本文系兰州大学“985工程”项目“网络信息计量学”的研究成果之一。

学、文献计量学和科学计量学(简称“三计学”)是网络信息计量学的母体学科,网络信息计量学的研究内容被文献计量学和信息计量学全部涵盖,与科学计量学有一定程度的交叉^[5]。“三计学”作为一组具有紧密联系的学科,其研究对象、内容、方法许多方面存在相似甚至重合之处,并向着信息计量学的统一学科方向发展。网络信息计量学被认为是“三计学”和技术计量学在新的信息网络时代经过革命性改造的结果^[6]。“三计学”的理论方法融入到网络信息计量学中,网络信息计量学的理论方法也必将围绕“三计学”而发展^[7]。

另外,“三计学”的经典定律在网络环境中得到了新的应用,成为网络信息计量学新的研究思路和方法。例如,Jakob Nielsen发现网页点击率和网页链接率都符合Zipf定律^[8]。R. Rousseau认为,他所研究的343个网址中最高层域名服从Lotka分布且对这些网址的引用也符合Lotka分布^[9]。

“三计学”中的许多概念和方法也被引入到网络信息计量学。1990年开始,信息计量学方法被应用到电子交流领域的研究^[10],之后,T. C. Almind和P. Ingwersen指出,将“三计学”的方法应用于网络就是网络信息计量学。“三计学”中广泛应用的书目分析法、内容分析法、关联数据分析法、数学模型法、关键词统计分析法等,在网络信息计量学研究中正得到广泛应用。网络信息计量学中的链接分析法也是对“三计学”中的引文分析法进行改造的结果。

2.2 数学统计学方法

数学统计学方法是网络信息计量学进行定量研究的重要方法来源和基础。网络信息计量学研究广泛采用“指标—模型—应用”的研究路径^[11],而模型构建方法的思路和算法来源于数学的模型和思想。网络数据的统计和分析处理方法,如相关分析、聚类分析、回归分析、统计检验等大都源于数学统计学。网络信息计量学的一些特殊方法,如链接分析法、图论、路径分析法、网络地图、Pagerank等定量研究方法是统计学和数学算法为基础的,实现网络信息计量学研究的软件工具,也是以数学统计学为基础来实现其功能的。

2.3 网络技术和软件技术

网络技术和软件技术是网络信息计量学定量研究的技术手段。一方面,网络信息计量学的研究对象是基于网络技术的网络信息,而网络信息的传播、展现、控制等均依赖于特定的网络技术,因此对网络信

息的研究方法一定程度上来源于网络技术和软件技术。另一方面,网络信息计量学数据收集方法、理论研究方法和结果展现方法的实现都依赖于一定的软件工具。从某种意义上讲,网络信息计量学方法的推进在很大程度上依赖于网络技术和软件技术的发展。网络信息计量学方法的技术来源突出体现了其学科性质和方法特点。

3 网络信息计量学方法的三维框架

按照网络信息计量学的学科追求和当前已经呈现的研究图景,理论研究(含应用研究)、数据收集和结果展现是网络信息计量学的3个主要工作领域和学科知识范畴。因此,我们提出由理论研究方法、数据收集方法和结果展现方法所构成的网络信息计量学研究方法的三维框架(图1)。其中,理论研究方法是三维框架中的核心方法和中心维度,数据收集方法和结果展现方法是理论研究方法实现的必要补充,并具有不可替代的特定功能。需要指出的是,这3个维度没有绝对的界限,各种具体方法在3个维度上的“位置”是相对而言的。

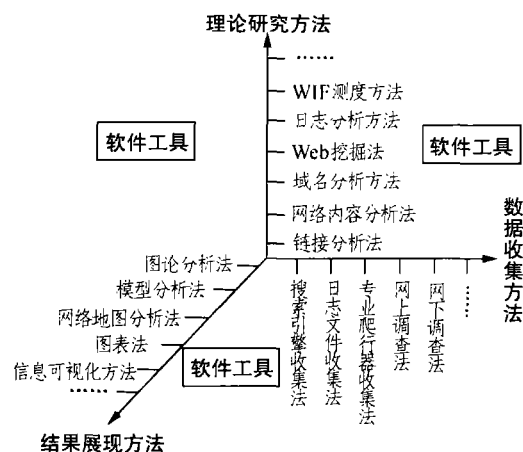


图1 网络信息计量学研究方法的三维框架

3.1 理论研究方法

理论研究方法主要包括网络内容分析法、网络结构分析法和网络利用分析法。理论研究方法在很大程度上体现了网络信息计量学的内容框架,是网络信息计量学作为一门独立学科的前提和基础,体现了学科特点。理论研究方法在方法体系中占据核心地位。

3.2 数据收集方法

网络信息计量学方法问题总是与作为实证研究基础的网络数据收集相关的^[12]。网络数据收集分析

不仅是网络信息计量学的研究内容,而且是网络信息计量学的主要研究方法^[13]。数据收集方法是网络信息计量学的基础性方法。

根据现有的数据收集技术和软件工具,数据收集方法可分为搜索引擎收集法、专业爬行器收集法、日志文件收集法和网上网下调查法。数据收集会影响研究结果的有效性,而网络信息计量学研究数据的可靠与否又取决于采集工具。数据收集方法的提高,十分依赖于软件技术的发展和相关软件功能的完善。

3.3 结果展现方法

它主要包括网络地图分析法、模型分析法、图论分析法、路径分析法、图表法、信息可视化方法等。网络环境中,结果展现方法有两方面的功能。一方面,将杂乱无章的网络结构,用可视化的图、表和模型等展现出来,同时又促进了网络知识发现的研究。例如,Watts等人用图论的方法展现了网络的“小世界”拓扑(Small World Topology)结构^[14]。Paul Terpstra利用网络地图方法展现了任意选择的1600个.com、.edu、.net网站的密集分布图^[15]。另一方面,无论是网络信息计量学的理论研究还是实证研究,用结果展现方法,能形象、简洁地反映网络信息计量学的研究结果,不仅增强研究结果的表现力,也便于人们接受和理解。从发展趋势来看,网络结构和信息的可视化研究,是信息科学和网络信息计量学研究的前沿领域和难点,其发展必将推进结果展现方法的不断更新和发展。

3.4 软件工具

网络信息计量学研究方法的三维框架中,软件工具是一个贯穿并渗透于整个方法体系的不可缺少的要素。无论是理论研究方法、数据收集方法还是结果展现方法的实现和推进,都需要相应的软件工具的支持,这正是网络信息计量学技术性特点的体现。图1对软件工具的这种角色作了形象表达。

4. 与三维框架方法相对应的软件工具体系

软件工具的辅助和支持是网络信息计量学研究的前提和基础。根据网络信息计量学研究对软件支持的需求以及现阶段国内外网络信息计量分析软件研发的现状,这里提出与网络信息计量学三维框架方法体系相对应的软件工具体系(见表1),并列举一些具有代表性的软件。网络信息计量学软件开发将不断向以下方向迈进,并日益成为软件产业的一支重要力量:面向网络,内嵌于网络技术和网络环境;在功能上,将进一步细分,实现多功能集成;向自动化、智能

化信息处理方向发展;日益规范化标准化;多版本软件开发成为一种趋势。随着软件技术含量越来越高,网络信息计量学软件开发主体将集中在专业的软件开发公司,软件的界面将越来越友好,软件的操作性能将向着便于用户使用的方向发展。

表1 与方法体系相对应的网络信息计量学软件工具体系

数据收集方法	搜索引擎	Altavista
		Alltheweb
	网上调查软件	NetRaker—— 在线调查方案
	专业爬行器	Connetivity Server (CS2)
理论研究方法	服务器软件	Microsoft Internet Information Services 5.0
	网络内容分析软件	基于词典的内容分析软件
		CATPAC
		HAMLET
		开发环境软件
		DIMAP
		WebCAT
		注释辅助软件
		Atlas-ti
	网络结构分析软件	链接分析软件
		Xenu's Link Sleuth
		Checkweb
		TreeDec Tool
网站利用分析软件	网站结构软件	Mapping the web in-fome
		WebSAT
	评估分析软件	Max——智能浏览代理方案
		WebVIP
	用户分析软件	VisVIPtool
		WebTrends Log Analyzer
结果展现方法	日志分析软件	Web-IAM
		SPSS, Excel
	通用软件	Webtracer、Internet Cartographer

本文提出的网络信息计量学方法的三维框架体系,是基于当前可获得的学术资源和网络信息计量学的学科追求所做的一种理论概括和尝试。目的在于通过清晰可见的结构,展现网络信息计量学领域的方法积累图景,并为相关学科方法的有效移植和新方法的创造提供一种具有导向意义的框架。(转第41页)

结合各项认证制度在 E-learning 的计划中,相信可以收到很好的效果。

参考文献

- 1 ASTD'S Source for E-learning. <http://WWW.learningcircuits.org/glossary.html#e>(2005-04-02 查询)
- 2 上海市科教智力开发研究所. 美国教育部教育技术白皮书. 2001
- 3 http://www.cisco.com/en/US/netsol/ns460/networking_solutions_packages_list.html(2005-04-10 查询)
- 4 何克抗. E-learning 与高校教学的深化改革. 中国电化教育, 2002(2)
- 5 Clark Aldrich. Simulations and the Future of Learning: An Innovative (and Perhaps Revolutionary) Approach to e-

(接第 32 页)按照科学评价的通行规则,其解释能力和生命力还有待在实践中进一步验证。

从发展前景看,网络信息计量学的方法积累和方法创新将沿着三维框架的 3 个维度不断拓展、细化、综合和集成,吸收网络技术、数据库技术和计算机技术的最新成果,不断研发支持方法功能实现的软件工具,并围绕链接分析法、搜索引擎分析法、Web 挖掘分析法这三类网络信息计量学的核心方法进行有效的方法建设,强化方法的实证研究和效能检验。随着方法三维框架的不断充实和丰富,网络信息计量学的方法体系将日臻完善并体现出自己的特色,从而促进网络信息计量学新的学科范式的形成。

参考文献

- 1 邱均平. 信息计量学(一):信息计量学的兴起和发展. 情报理论与实践, 2000(1)
- 2 Boudourides, M. A., Antypas, G. P. A Simulation of the Structure of the World-Wide Web. Sociological Research Online, 7(1)
- 3 Almind, T. C., Ingwersen P. Informetric analyses on the world wide web: methodological approaches to "webometrics". Journal of Documentation, 1997, 53(4)
- 4 Boudourides, M. A., Sigrist, B., Alevizos, P. D. Webometrics and the Self-organization of the European Information Society. 1999. <http://hyperion.math.upatras.gr/webometrics>
- 5 Bjorneborn, L., Ingwersen, P. Towards a basic framework of webometrics. Journal of the American Society for Science and Technology, special issue on webometrics, 2004, 55(14)
- 6 蒋国华. 迎接科学计量学应用的新时代——第二届科研

Learning. Pfeiffer(September 5, 2003)

- 6 Fred McCrea, R. Keith Gay&Rusty Bacon. Riding the Big Wave: A white paper on the B2B e-learning industry. 2000
- 7 廖肇弘. 企业 E-learning 市场现状与发展趋势. <http://www.online-edu.org/newarticle/articles/18/308.html> (2005-03-20 查询)

邓仲华 武汉大学信息管理学院副教授, 发表论文 20 余篇, 著作 4 部. 通信地址: 武汉大学. 邮编 430072.

陈远 武汉大学信息管理学院副教授, 发表论文 40 余篇. 通信地址同上。

郭梁 武汉大学信息管理学院硕士研究生, 研究方向: 管理信息系统. 通信地址同上。

(来稿时间: 2005-06-14)

绩效定量评价国际学术会议暨第六次全国科学计量学与情报学年会. 科学学研究, 2001(6)

- 7 Thelwall, M., Vaughan, L. Lennart Bjorneborn. Webometrics. Annual Review of Information Science and Technology (ARIST) 39, 2005
- 8 Nielsen, J. Zipf Curves and Website Popularity. 1997 <http://www.useit.com/alertbox/zipf.html>
- 9 Rousseau, R. Sitations: an exploratory study. Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics, 1997(1). <http://www.cindoc.csic.es/cybermetrics/articles/vlilpl.pdf>
- 10 Paisley, W. The future of bibliometrics. Scholarly communication and bibliometrics. Sage, 1990
- 11 徐久龄, 刘春茂, 刘亚轩. 网络计量学的研究. 见: 张力治主编. 情报学进展(第 3 卷). 北京: 航空工业出版社, 1999
- 12 Bjorneborn, L., Ingwersen, P. Perspectives of webometrics. Scientometrics, 2001, 50(1)
- 13 邱均平等. 网络数据分析. 北京: 北京大学出版社, 2004
- 14 J. Watts. Small worlds: the dynamics of networks between order and randomness, Princeton University Press, Princeton, N. J. 1999
- 15 Terpstra, P. Mapping Cyberspace with GIS. The 18th Annual ESRI User Conference. San Diego, California, 1998. 3

沙勇忠 兰州大学管理学院副教授、博士、副院长. 通信地址: 兰州市天水南路 222 号. 邮编 730000.

欧阳霞 兰州大学管理学院情报学专业硕士研究生. 通信地址同上。(来稿时间: 2005-09-13)