

●段宇锋 邱均平

中美大学网站评价的比较研究*

摘 要 基于揭示网站规模、链接特征和影响力三个方面的 21 项指标,对中美财经和商学院网站作对比研究,发现其中的 15 项指标有明显差异。表明中国财经院校网站信息含量、信息组织的科学性、网站影响力等方面低于美国同类院校网站。通过分析,可以为网站评价指标体系的建立奠定基础。表 4。参考文献 11。

关键词 大学网站评价 链接分析 对比研究 中国 美国

分类号 G250.7

ABSTRACT By analyzing the sizes, link characteristics and impacts, the authors make an analysis of websites of some business schools in China and U. S. A., and find obvious differences in 15 indicators.

The analysis can be used as a basis for the evaluation of websites. 4 tabs. 11 refs.

KEY WORDS Evaluation of university websites. Analysis of links. Comparative study. China. U. S. A.

CLASS NUMBER G250.7

1 研究对象和研究方法

1.1 研究对象的确定

本研究以大学网站为研究对象,对其质量进行评价。

(1)美国院校网站样本的确定

根据 U. S. News 在网上公布的研究型院校排行榜(“American’s Best Graduate Schools 2004”),获取排名在前 50 位的美国商学院^[1],以这些院校主页所在的主机作为总体,从中随机产生 20 个作为样本。抽样结果为排名第 1、2、5、7、10、11、12、16、21、22、23、25、30、32、33、37、42、44、45、46 的商学院网站。

(2)中国院校网站样本的确定

从中国管理科学研究院科学研究所武书连等人 2004 年发布的《中国大学评价》上排名在 B+ 以上的经济学院中选择专业性的财经院校作为研究对象^[2-3]。为了确保样本的可比性,我们必须限制网站的类型。经过逐一筛选,在 29 个 B+ 级以上的学院中有 7 个符合我们的标准(c1-c7),见表 1。

1.2 研究所采用的评价指标

1.2.1 衡量网站规模的指标

(1)网站的字节数(x1)

编码的容量也就意味着网站的信息量。编码的多少一般以字节数来衡量,因此网站字节数是衡量网

站规模最常用的指标。

(2)文件数(x2)

文件是计算机组织信息最基本的单位,是逻辑上具有密切联系的信息集合体。信息的编码方式对数据量的影响极大,但很少会影响到系统中文件的数据量。因此,以文件数衡量网站规模可以消除因数据编码方式的差异而造成的影响。

(3)页面文件数(x3)

页面文件是指以“.html”、“.htm”、“.asp”、“.jsp”、“.php”、“.stm”、“.stml”、“.idc”、“.shtml”、“.htx”、“.txt”、“.text”、“.xsp”、“.xml”、“.ixml”、“.cfm”、“.wml”等后缀结尾的文件。尽管所有类型的文件中都可能包含网络链接,但音频、视频文件、图像文件、压缩文件和数据库中的链接以目前的信息处理技术很难分析出来。并且,绝大部分链接存在于页面文件,从数量上,存在于其他形式文件中的链接基本上可以忽略不计。

1.2.2 衡量网站链接特征的指标

(1)链接总数(x4)

指网站中所存在的总的链接数,它是衡量网站链接数量特征最常用的指标之一,对反映网站信息的丰裕度、信息揭示的深度和组织的完备程度具有一定的参考价值。

* 本文是国家自然科学基金资助项目“网络信息计量学的理论、方法与实证研究”(批准号:70273032)的组成部分和研究成果之一。

表 1 样本各指标数据

样本 编号	网址	排 名	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19	x20	x21
1	www.hbs.edu/	1	197248064	6992	2734	42781	34921	7860	40317	2464	759	30839	4610.65	5648.41	25095.17	4892.43	80051.97	15.65	12.77	2.87	14.75	0.90	11.28
2	www.gsh.stanford.edu/	2	479669591	13690	4566	141785	125286	16499	135573	6212	1118	23908	3383.08	3828.60	29072.65	3538.09	77216.61	31.05	27.44	3.61	29.69	1.36	5.24
3	www.kellogg.northwestern.edu/	5	1453609335	29.962	12296	197594	159953	37641	179884	17710	3192	7437	7356.58	9087.73	38617.71	8080.82	82078.45	16.07	13.01	3.06	14.63	1.44	0.60
4	www.fuqua.duke.edu/	7	337213475	2521	952	18046	14244	3802	16060	1986	561	13628	18886.33	23674.07	88693.71	20997.10	169795.31	18.96	14.96	3.99	16.87	2.09	14.32
6	www.tuck.dartmouth.edu/	10	163090027	3423	1058	21368	16077	5291	19329	2039	464	5801	7632.44	10144.31	30824.05	8437.58	79985.30	20.20	15.20	5.00	18.27	1.93	5.48
7	www.darden.virginia.edu/	11	2806297556	21384	6411	175081	75988	99093	78698	96383	2432	51477	16028.57	36930.80	28319.84	35659.07	29116.11	27.31	11.85	15.46	12.28	15.03	8.03
8	www.stern.nyu.edu/	12	51287093	1845	588	23137	19969	3168	21269	1868	314	2227	2216.67	2568.34	16189.11	2411.35	27455.62	39.35	33.96	5.38	36.17	3.18	3.79
9	www.johnson.cornell.edu/	16	254952338	21035	1412	52356	47760	4596	48481	3875	164	750	4869.59	5338.20	55472.66	5258.81	65794.15	37.08	33.82	3.25	34.33	2.74	0.53
10	www.goizueta.emory.edu/	21	551651027	6511	2195	83401	76678	6723	79022	4379	411	652	6614.44	7194.38	82054.30	6980.98	125976.49	38.00	34.93	3.06	36.00	1.99	0.30
11	www.kennan-flagler.mcc.edu/	22	980666627	46266	3586	42126	27854	14272	33630	8496	1804	1524	23279.37	35207.39	68712.61	29160.47	115426.86	11.75	7.77	3.98	9.38	2.37	0.43
12	www.kelley.iu.edu/	23	714226610	14377	4277	58884	37246	21638	38990	19894	4319	9929	12129.38	19175.93	33007.98	18318.20	35901.61	13.77	8.71	5.06	9.12	4.65	2.32
13	www.mgmt.purdue.edu/	25	1290219105	21566	5074	77095	63410	13685	68569	8526	1650	24881	16735.44	20347.25	94279.80	18816.36	151327.60	15.19	12.50	2.70	13.51	1.68	4.90
14	www.biz.uiowa.edu/mba/	30	852596464	20092	2388	112976	62044	50932	66210	46766	1980	6775	7546.70	13741.80	16739.90	12877.16	18231.12	47.31	25.98	21.33	27.73	19.58	2.84
15	www.olin.wustl.edu/	32	211278775	7591	3506	85176	61274	23902	78743	6433	737	8683	2480.50	3448.10	8839.38	2683.14	32842.96	24.29	17.48	6.82	22.46	1.83	2.48
16	www.smeal.psu.edu/	33	1330241837	39290	5004	144915	121028	23887	137096	7819	788	655	9179.46	10991.19	55688.95	9703.00	170129.41	28.96	24.19	4.77	27.40	1.56	0.13
17	wpcarey.asu.edu/	37	57799322	1008	634	19911	16032	3879	18323	1588	314	279	2902.88	3605.25	14900.57	3154.47	36397.56	31.41	25.29	6.12	28.90	2.50	0.44
18	www.freeman.tulane.edu/	42	921745040	18195	6915	355676	281034	74642	333827	21849	2258	16899	2591.53	3279.83	12348.88	2761.15	42187.06	51.44	40.64	10.79	48.28	3.16	2.44
19	www.chssmith.usmd.edu/	44	537197115	14020	6004	54720	46640	8080	49182	5538	1557	7709	9817.199	11517.95	66484.79	10922.64	97002.01	9.11	7.77	1.35	8.19	0.92	1.28
20	mba.vanderbilt.edu/	45	280095987	6450	1215	11276	8503	2773	8838	2438	816	11720	24840.0	32940.84	101008.29	31692.24	114887.61	9.28	7.00	2.28	7.27	2.01	9.65
c1	www.shufe.edu.cn/	-	787126961	15725	6256	51709	44188	7521	45024	6685	3253	142	15222.24	17813.14	104657.22	17482.39	117745.25	8.27	7.06	1.20	7.197	1.07	0.02
c2	www.dufe.edu.cn/	-	246669796	6615	3286	33982	29179	4803	31609	2373	437	0	7258.84	8453.68	51357.44	7803.78	103948.50	10.34	8.88	1.46	9.62	0.72	0.00
c3	www.swufe.edu.cn/	-	76443333	4103	978	3495	2076	1419	2555	940	132	104	21872.20	36822.41	53871.27	29919.11	81322.70	3.57	2.12	1.45	2.61	0.96	0.11
c4	www.cufe.edu.cn/	-	157715237	6468	2383	19343	17534	1809	17597	1746	179	41	8153.61	8994.82	87183.66	8962.62	90329.46	8.12	7.36	0.76	7.38	0.73	0.02
c5	www.uibe.edu.cn/	-	222460145	6409	2703	28501	24686	3815	25219	3282	334	173	7805.35	9011.59	58311.96	8821.13	67781.88	10.54	9.13	1.41	9.33	1.21	0.06
c6	www.jxufe.edu.cn/	-	130743933	3543	1034	6414	5593	821	5620	794	199	64	20384.15	23376.35	159249.61	23664.05	164664.90	6.20	5.41	0.79	5.44	0.77	0.06
c7	www.cueb.edu.cn/	-	14495908	2032	546	6624	6271	353	6290	334	16	94	2188.39	2311.58	41064.90	2304.60	43400.92	12.13	11.49	0.65	11.52	0.61	0.17

(2) 指向内部的链接数

根据不同的划分标准,这一指标又可区分成指向主机内部的链接数(x_5)和指向主机所在域的链接数(x_7)。这类链接与期刊中的自引在形式上极其相似,但研究证实,两者存在着本质的差异。指向内部的链接在数量上通常占总链接数的60%~80%,有些网站所占比例甚至更高,其作用是将独立的网页组织成具有内在逻辑关系和层次的整体,属于结构性链接。网站中指向内部的链接的数量和所占总链接数的比例能在一定程度上反映网站的性质。专业性的、以提供资源和最终信息为目的的网站,由于往往自成体系,内部信息的揭示深度较大而具有较多指向内部的链接。

(3) 指向外部的链接数

指向外部的链接数也可以区分为指向主机以外网络资源链接的数量(x_6)和指向主机所在域以外网络资源链接的数量(x_8)。互联网正是通过它们将网上各相对独立的主机和系统联接成庞大而复杂的体系。指向外部的链接所占比例高的网站往往带有门户网站性质。

(4) 链接密度($x_{11} \sim x_{15}$)

链接密度是度量网站内链接出现频率的指标,单位为“bytes/link”。该值越大,表示网站中链接出现的频率越低,分布越稀疏;反之,则说明网站中链接出现的频率越高。网站中链接的绝对数量与网站规模有关,链接密度能在一定程度上消除由此对网站评价所造成的负面影响。网站信息的组织是有层次的,位于不同层次的网页在功能上具有一定的差异,这种差异会通过字节数和链接数的变化体现出来,从而表现为链接密度的跳跃式变化。因此,通过链接密度我们可以了解网站信息资源的分布情况。

(5) 页面平均链接数($x_{16} \sim x_{20}$)

页面平均链接数反映的是平均每个网页所含链接的数量,单位为“links/page”。一般而言,该值越大,网站信息的揭示深度就越大,信息的自组织程度越高。这一指标与链接密度不同,它与网站字节数无关,因而克服了多媒体文件对衡量链接出现频率的影响。但是,网站的设计风格却会对该指标造成严重影响。

(6) 所链接的网站数(x_9)

所链接的网站数反映了网站信息的跨度和扩散性。任何网站所包含的信息都是有限的,它不可能也

没有必要囊括所有的内容。一个网站所链接到的资源越多,其开放性就越强;用户通过该网站进一步所能获得的信息就越丰富。有些网站尽管存在许多指向外部的链接,但这些链接集中在为数极少的几个网站,如果单纯依据总链接数和指向外部的链接数进行评价,必然得出错误结论。因此,在考察上述指标的同时,以所链接到的网站数作为辅助指标是十分必要的。

1.2.3 衡量网站影响力的指标

(1) 链接到网站的网页数(x_{10})

链接到网站的网页数是评价网站价值和影响力的重要尺度。但是,依靠目前的技术无法精确地获取这一指标,在多数情况下我们只能以 AltaVista、All-TheWeb、Google 等搜索引擎的检索结果替代。

(2) 网络影响因子(Web-IF)(x_{21})

网站规模对链接到网站的网页数具有一定的影响,为消除由此造成的误差, Peter Ingwersen 在 1998 年提出了 Web-IF,其计算公式为“链接到网站的网页数/网站的网页数”。

1.3 研究方法

(1) 获取样本原始数据。以 MetaProducts 公司开发的 Offline Explorer2. 9. 1280 获取样本原始数据。由于技术上的原因,无法读取样本 5 的数据,因而将其从研究中剔除。

(2) 解析样本数据。利用本研究自主开发的链接分析工具——webStat 解析样本数据,获取指标 $x_1 \sim x_9$ 的值。

(3) 利用搜索引擎“alltheweb”的指令“link. all: 网站地址-site:网站地址”获取指标 x_{10} 。

(4) 根据 $x_1, x_3 \sim x_{10}$ 计算指标 $x_{11} \sim x_{21}$ 的值。

(5) 将样本 1~20 和 $c_1 \sim c_7$ 的各指标值汇总至表 1。

(6) 描述样本各指标数据的分布特征,在此基础上分析两者在各指标上存在的差异。

2 数据分析及结果

2.1 样本各指标数据分布特征的描述

以算术平均数、中位数、标准差、最大值、最小值和四分位数描述样本各指标数据的分布特征,结果见表 2、表 3。

表2 美国院校样本各指标的数据分布特征

指标	数据量	算术平均数	标准差	最小值	下四分位数	中位数	上四分位数	最大值
x1	19	709004494	671485125	51287093	211278775	537197115	980666627	2806297556
x2	19	15590.42	12620.67	1008.00	6450.00	14020.00	21384.00	46266.00
x3	19	3727.11	2901.53	588.00	1215.00	3506.00	5074.00	12296.00
x4	19	90437.05	84811.49	11276.00	23137.00	58884.00	141785.00	355676.00
x5	19	68207.42	66174.63	8503.00	19969.00	47760.00	76678.00	281034.00
x6	19	22229.63	26409.89	2773.00	4596.00	13685.00	23902.00	99093.00
x7	19	76423.21	77557.11	8838.00	21269.00	49182.00	79022.00	333827.00
x8	19	14013.84	22741.47	1588.00	2438.0	6212.00	17710.00	96383.00
x9	19	1349.37	1109.59	164.00	464.00	816.00	1980.00	4319.00
x10	19	11882.79	13111.64	279.00	16899.00	7709.00	1524.00	51477.00
x11	19	9626.36	7102.11	2216.67	3383.08	7546.70	16028.57	24840.01
x12	19	13614.23	11318.60	2568.34	3828.60	10144.31	20347.25	36930.80
x13	19	45597.39	30102.23	8839.38	16739.90	33007.98	68712.63	101008.29
x14	19	12439.21	10477.60	2411.35	3538.09	8437.58	18816.36	35659.07
x15	19	81673.88	48861.94	18231.12	35901.61	79985.30	115426.86	170129.41
x16	19	25.587590	12.767358	9.1139241	15.194126	24.294352	37.079320	51.4354302
x17	19	19.750609	10.658575	6.9983539	11.852753	15.195652	27.438896	40.6412148
x18	19	5.8369803	4.9781133	1.3457695	3.0612394	3.9936975	6.1182965	21.3283082
x19	19	21.853851	11.677442	7.2740741	12.275454	18.269376	29.691852	48.2757773
x20	19	3.7337389	4.9221246	0.9012436	1.5625500	2.0065844	3.1596529	19.5837521
x21	19	4.024874	4.1236023	0.1309	5.4830	2.4766	0.5312	14.3151

表3 中国院校样本各指标的数据分布特征

变量	数据量	算术平均数	标准差	最小值	下四分位数	中位数	上四分位数	最大值
x1	7	233665045	256816949	14495908.00	76443333.00	157715237	246669796	787126961
x2	7	6413.57	4463.25	2032.00	3543.00	6409.00	6615.00	15725.00
x3	7	2455.14	1960.01	546.0000000	978.0000000	2383.00	3286.00	6256.00
x4	7	21438.29	17772.11	3495.00	6414.00	19343.00	33982.00	51709.00
x5	7	18503.86	15269.80	2076.00	5593.00	17534.00	29179.00	44188.00
x6	7	2934.43	2581.89	353.0000000	821.0000000	1809.00	4803.00	7521.00
x7	7	19130.57	15746.78	2555.00	5620.00	17597.00	31609.00	45024.00
x8	7	2307.71	2178.96	334.0000000	794.0000000	1746.00	3282.00	6685.00
x9	7	650.0000000	1155.88	16.0000000	132.0000000	199.0000000	437.0000000	3253.00
x10	7	88.2857143	59.1064415	0	41.0000000	94.0000000	142.0000000	173.0000000
x11	7	11840.68	7406.03	2188.39	7258.84	8153.61	20384.15	21872.20
x12	7	15254.80	11758.78	2311.58	8453.68	9011.59	23376.35	36822.41

变量	数据量	算术平均数	标准差	最小值	下四分位数	中位数	上四分位数	最大值
x13	7	79385.15	41687.05	41064.90	51357.44	58311.96	104657.22	159249.61
x14	7	14079.67	9820.85	2304.60	7803.78	8962.62	23264.05	29919.11
x15	7	95599.09	38859.32	43400.92	67781.88	90329.46	117745.24	164664.90
x16	7	8.4538322	2.9009161	3.5736196	6.2030948	8.2655051	10.5442101	12.1318681
x17	7	7.3501426	2.9942684	2.1226994	5.4090909	7.3579522	9.1328154	11.4853480
x18	7	1.1036896	0.3597983	0.6465201	0.7591272	1.2022059	1.4509202	1.4616555
x19	7	7.5854917	2.9497809	2.6124744	5.4352031	7.3843894	9.6192940	11.5201465
x20	7	0.8683405	0.2175791	0.6117216	0.7221546	0.7678917	1.0685742	1.2142064
x21	7	0.0634718	0.0598408	0	0.0172052	0.0618956	0.1063395	0.1721612

绘制上述指标数据分布的直方图,我们发现其分布不符合正态分布规律,分别以对数正态分布的概率密度函数、指数分布的概率密度函数和威布尔分布的概率密度函数进行拟合,只有拟合出的对数正态分布概率密度函数曲线基本呈单峰、对称状。

2.2 中、美样本各指标的对比分析

为了明确本研究所抽取的中、美院校网站的组织特征是否存在差异,我们将各组参数进行对数转换后进行t-检验,结果如表4。

表4 中、美样本各指标差异性分析的t-检验结果

变量	方法	方差	自由度	t值	pr > t	变量	方法	方差	自由度	t值	pr > t
x1_log	Pooled	Equal	24	2.40	0.0247	x12_log	Pooled	Equal	24	-0.41	0.6863
x1_log	Satterthwaite	Unequal	9.59	2.24	0.0497	x12_log	Satterthwaite	Unequal	10.4	-0.40	0.6959
x2_log	Pooled	Equal	24	1.53	0.1379	x13_log	Pooled	Equal	24	-2.26	0.0328
x2_log	Satterthwaite	Unequal	18	1.92	0.0707	x13_log	Satterthwaite	Unequal	17.1	-2.78	0.0129
x3_log	Pooled	Equal	24	1.03	0.3155	x14_log	Pooled	Equal	24	-0.54	0.5909
x3_log	Satterthwaite	Unequal	11	1.04	0.3212	x14_log	Satterthwaite	Unequal	11	-0.55	0.5915
x4_log	Pooled	Equal	24	3.40	0.0023	x15_log	Pooled	Equal	24	-1.02	0.3191
x4_log	Satterthwaite	Unequal	9.98	3.27	0.0085	x15_log	Satterthwaite	Unequal	17.5	-1.26	0.2248
x5_log	Pooled	Equal	24	3.14	0.0045	x16_log	Pooled	Equal	24	4.68	<.0001
x5_log	Satterthwaite	Unequal	9.26	2.87	0.0178	x16_log	Satterthwaite	Unequal	13.8	5.26	0.0001
x6_log	Pooled	Equal	24	3.90	0.0007	x17_log	Pooled	Equal	24	3.80	0.0009
x6_log	Satterthwaite	Unequal	10.9	3.93	0.0024	x17_log	Satterthwaite	Unequal	11	3.84	0.0028
x7_log	Pooled	Equal	24	3.29	0.0031	x18_log	Pooled	Equal	24	5.58	<.0001
x7_log	Satterthwaite	Unequal	9.55	3.07	0.0125	x18_log	Satterthwaite	Unequal	20.4	7.34	<.0001
x8_log	Pooled	Equal	24	2.97	0.0066	x19_log	Pooled	Equal	24	4.10	0.0004
x8_log	Satterthwaite	Unequal	12.3	3.18	0.0077	x19_log	Satterthwaite	Unequal	12.3	4.38	0.0008
x9_log	Pooled	Equal	24	2.92	0.0075	x20_log	Pooled	Equal	24	3.47	0.0020
x9_log	Satterthwaite	Unequal	7.43	2.25	0.0569	x20_log	Satterthwaite	Unequal	23.8	5.24	<.0001
x10_log	Pooled	Equal	24	6.82	<.0001	x21_log	Pooled	Equal	24	3.92	0.0006
x10_log	Satterthwaite	Unequal	9.31	6.27	0.0001	x21_log	Satterthwaite	Unequal	18.4	6.50	<.0001
x11_log	Pooled	Equal	24	-0.74	0.4672						
x11_log	Satterthwaite	Unequal	10.6	-0.73	0.4803						

上述数据表明,在所考察的 21 项指标中,除 x_2 、 x_3 、 x_{11} 、 x_{12} 、 x_{14} 和 x_{15} 外,其余各指标都存在显著差异($P < 0.05$)。

3 讨论

3.1 中、美网站信息含量和组织特征上的差异

(1) 网站规模的对比研究

数据显示,美国商学院网站平均字节数约为 709 兆,平均每个网站文件数为 15590 个,页面文件数为 3727 个;而国内财经院校网站的平均字节数还不到 234 兆,每个网站平均有 6414 个文件,页面文件数为 2544 个。从上述指标的绝对值来说,分别是美国同类型网站的 $1/3$ 、 $1/2$ 和 $2/3$ 。可见,国内大学网站的信息总量明显少于国外同类型网站。并且,我们从网站中文件总数和页面文件数的差距小于字节数的差距还可以推知,国内大学网站中每个网页所含的信息量比国外同类网站中每个网页所含信息量少。

(2) 链接特征的对比研究

①衡量链接绝对量的指标($x_4 \sim x_8$)。美国商学院网站上述五项指标的平均值分别为 90437、68207、22230、76423、14014,中国财经院校网站上述指标平均值为 21438、18504、2934、19131、2308,在绝对数量上,前者分别是后者的 4.2 倍、3.7 倍、7.6 倍、4 倍、6.1 倍,两者之间的差异具有显著的统计学意义。中、美网站在这些指标上的差异首先表明两者的信息总量具有显著的差距;其次,两者在 x_5 、 x_6 上的差异表明我国大学网站信息揭示的深度和组织的完备程度上不如美国同类网站;第三,两者在 x_7 、 x_8 上的差异则说明我国大学网站的扩散能力较差。

②衡量链接出现频率的指标。总链接数、指向内部的链接数和指向外部的链接数都受网站规模的影响,因此,在对网站链接特征进行描述时必须结合链接密度、页面平均链接数等衡量链接出现频率的相对性指标综合评价。与计量链接绝对数量的各指标相对应,衡量链接密度和页面平均链接数的指标也各有 5 个,分别是 $x_{11} \sim x_{15}$ 和 $x_{16} \sim x_{20}$ 。美国商学院网站的 $x_{16} \sim x_{20}$ 指标值分别为 26、20、6、22、4,中国同类网站上述指标值分别为 8.5、7.5、1、7.6、0.9,两者具有显著差异。其中, x_{16} 、 x_{17} 、 x_{19} 是前者的 $1/3$, x_{18} 和 x_{20} 仅为前者的 $1/6$ 和 $1/4$ 。两者在页面平均链接数上的差距再次表明,国内大学网站平均每个网页所含的信息量和信息揭示与组织的完备程度远不如美国同类型网站。

从理论上说,链接密度和页面平均链接数是从不同角度揭示相同的问题,两者的统计分析结果应该一致。但是,本研究的统计数据却显示, $x_{11} \sim x_{15}$ 这 5 项衡量链接密度的指标,除 x_{13} 外,中、美网站都不具有显著差异。笔者认为,各网站多媒体文件的数量、占总文件数的比例和所占字节数的差异是造成这两组指标数据分析结果不一致的主要原因。

③所链接的网站数。数据显示,平均每个美国商学院网站所链接的网站数是 1349 个,而中国同类网站该指标仅为 650,尚不足前者的一半。这一数据表明中国财经类院校网站中信息的跨度和扩散能力低于美国商学院网站,换言之,从我国财经院校网站能获得的相关信息和线索少于美国商学院网站。

(3) 网站影响力的对比研究

两组研究对象在指标 x_{10} 和 x_{21} 上的差异表明,国内财经院校网站的影响力比美国同类学院网站的影响力小。造成这一差距的原因有两个:一方面美国排名在前 50 位的商学院在国际上确实存在很大的影响力,因而这些学院的网站也受到广泛的关注和利用;另一方面,搜索引擎 AlltheWeb 对中国和美国网络信息资源覆盖率的差异也是应该考虑的因素。

(4) 推论

①本研究的统计数据表明,所考察的中、美网站在链接总数和页面平均链接数上的差异比总文件数和页面文件数更为显著,这表明在中国财经院校网站寻找信息所需要经过的中间环节要多于美国商学院网站,这将增加用户获取信息的难度和成本。

②本研究所考察的 21 项指标,中国财经院校网站与美国商学院网站在其中的 15 项上存在显著差异。这些差异表明中国财经院校网站在信息含量、信息组织的合理程度以及网络影响力等方面都不如美国同类网站。更重要的是,这项研究让我们认识到尽管我国互联网建设近年取得了可喜的成就,但在网络信息资源的质量和建设水平上仍与美国存在较大的差距。

3.2 网站评价指标体系

国内外围绕这一主题开展了许多研究,建立了一些评价标准和方法。譬如,Jim Kapoum 提出以准确性、客观性、权威性、时效性和覆盖面作为评价标准^[4];Judith Edwards 认为应该从可获得性、质量和易用性三个方面对网站进行评价^[5]。国内的研究者也提出了一些与此类似的评价指标,譬如,田菁提出的标准包括内容、学术水平、取用方式和连续性与稳定

性^[6];黄奇和郭晓苗认为应考虑的因素包括内容、设计、可用性和可获得性、安全以及其他评价来源^[7];左艺等提出的标准包括范围、内容、图形和多媒体设计、目的及对象、评论和可使用性。另外,研究者认为可以参考搜索引擎、网站的被访问量以及其他评价工具所提供的信息^[8]。但是,这些研究采用的都是定性和半定量方法,不仅容易受研究者主观意志的影响,而且难以适用于大样本研究。因此,笔者认为网站评价指标体系的建立必须基于网站可量化的特征。

3.3 网站评价研究存在的问题

尽管网站评价研究已经引起包括计算机科学、情报学、管理学等各相关领域的高度关注,但从总体而言仍然处于探索阶段。这主要体现在三个方面,即缺乏成熟的理论指导、评价标准有待于进一步探索和完善的,以及缺乏有效的工具和方法。网站评价的理论和标准尽管目前还不成熟,但是毕竟还可以借鉴传统信息资源评价的研究成果,而研究工具和方法却是无法替代的,这已经成为阻碍网站评价研究进展的关键因素。

我们以网站评价定量研究中最常采用的 Web-IF 这个指标为例进行阐述。1998 年 Peter Ingwersen 提出 Web-IF 后立即引起网络信息计量学领域的高度关注,并很快运用于网站评价^[9]。迄今,它的理论意义已经十分明确,在网站评价中的重要性也达成了共识,但是对其实际应用价值却存在巨大的分歧。造成这一问题的根本原因就在于缺乏有效的研究工具。Web-IF 的计算需要两个参数,即链接到网站的网页数和网站所拥有的网页数,这两个数据的获取目前多数研究利用的都是商业搜索引擎。但是,商业搜索引擎不像 SCI 那样是专门为进行引用分析而设计和开发的,它在网络信息的覆盖率、检索结果的准确性等方面都存在严重缺陷。许多研究表明,以商业搜索引擎为工具测定的 web-IF 不能作为网站评价的依据^[10-11]。

本研究也面临同样的问题,除搜索引擎外,我们所使用的网站原始数据获取工具“Offline Explorer”尽管与其他软件相比性能相对优越,但是处理动态网页存在的缺陷仍值得重视。同样,由于缺乏标准的方法和工具,本研究为了解析网站中的链接,不得不耗费大量时间去设计专用的链接分析工具“WebStat”,而该软件只适用于本研究,如需运用到其他环境和研究必须进行改造。因此,在网络评价甚至整个网络信息计量学研究领域,研究方法的探索和工具的开发是最

紧迫的任务。

4 结论

研究表明:①中国财经院校网站与美国商学院网站相比,在网站规模、链接特征和影响力等方面存在显著差异。②本研究从网站规模、链接特征和影响力三个方面构建的 21 项指标具有内在逻辑关系,并且对网站质量具有稳定的区分度,可以作为网站定量评价指标体系的构成要素。③为了促进网站评价和网络信息计量学的发展,必须突破因研究方法和工具的缺陷而造成的瓶颈。

参考文献

- 1 http://www.usnews.com/usnews/edu/grad/rankings/mba/brief/mbarank_brief.php, 2003. 10. 11.
- 2 搜狐教育. 中国大学本科各学科门类、各专业等级定义. <http://learning.sohu.com/2004/01/14/83/article218538368.shtml>, 2004. 3. 15
- 3 搜狐教育. 2004 中国大学经济学 B+ 级以上学校. <http://learning.sohu.com/2004/01/16/26/article218632659.shtml>, 2004. 3. 15
- 4 Kapoun J. Five Criteria for Evaluating Web Pages. <http://www.library.cornell.edu/okuref/webcrit.html>, 2004. 5. 14.
- 5 Edwards J. The Good, the Bad and the Useless: Evaluating Internet Resources. <http://www.ucl.ac.uk/library/>, 2004. 6. 3
- 6 田菁. 网络信息与网络信息的评价标准. 图书馆工作与研究, 2001(3)
- 7 黄奇, 郭晓苗. Internet 上网站资源的评价. 情报科学, 2000(4)
- 8 左艺. 国际互联网上信息资源优选与评价研究方法初探. 情报学报, 1999(4)
- 9 Peter Ingwersen. The calculation of web impact factors. Journal of Documentation. 1998, 54(2)
- 10 Stephen P. Harter, Charlotte E. Ford. Web-based analyses of E-journal impact: Approaches, Problems, and Issues. Journal of the American Society for Information Science. 2000, 51(13)
- 11 刘雁书, 方平. 利用链接关系评价网络信息的可行性研究. 情报学报, 2002, 21(4)

段宇锋 华东师范大学商学院讲师, 博士。通信地址: 上海市。邮编 200062。

邱均平 武汉大学中国科学评价研究中心教授, 博士生导师。通信地址: 湖北武汉。邮编 430072。

(来稿时间: 2004-08-27)