

# 我国省级以上公共图书馆网站的链接分析

王宏鑫

(信阳师范学院图书馆, 信阳 464000)

【摘要】评价公共图书馆的网站建设质量,可以运用链接分析方法。通过 google 搜索各省级以上公共图书馆网址,使用“高级搜索”获得计量指标并进行分析,从而给出各馆网站建设质量按不同指标的排序和综合排序,以及各指标的优选排序。

【关键词】公共图书馆; 网站; 链接分析

【中图分类号】G 25

## Analysis on Web Sites Links of Public Library over the Provincial Grade in Our Country

WANG Hongxin

【Abstract】We can evaluate public library web site using links analysis method. With google, we can search web site addresses and use “advanced search” to get and analysis measuring index, then the public library web site rank and index rank can be presented out.

【Keywords】public library; web sites; links analysis

随着互联网的日益普及,越来越多的读者习惯利用网络来查询自己所需的信息,公共图书馆作为地区文献信息中心,收藏有大量的书刊、报纸和声像资料,是互联网中最基本、最重要的信息源,它有条件提供丰富多彩的网络信息服务。因此,如何建设好公共图书馆的网站,如何评价公共图书馆的网站建设质量,并提供读者满意的网络信息服务,成为值得公共图书馆网站建设者们深思的一个问题。本文拟就我国省级以上公共图书馆网站进行链接分析。

### 1 研究方案的设计

#### 1.1 研究对象的选择

为了研究的方便,我们选择省级以上综合性公共图书馆,不包括台、港、澳的图书馆,另外,由于黑龙江省图书馆、海南省图书馆、青海省图书馆、宁夏回族自治区图书馆、新疆维吾尔自治区图书馆尚无独立网站开通,或搜索不到其网址,因此也被排除掉。

#### 1.2 研究指标的选择<sup>[1,2]</sup>

(1) 网页数:一般是指运用搜索引擎搜索到的某网站内的网页数。

(2) 总链接数:一般是指运用搜索引擎搜索到

的与某网站存在链接的网页总数。

(3) 内链接数:一般是指运用搜索引擎针对某网站范围内搜索得到的与该网站存在链接的网页数。

(4) 外链接数:一般是指运用搜索引擎针对某网站范围外搜索得到的与该网站存在链接的网页数。

(5) 链接测度偏差:理论上讲,应该存在等式:总链接数=内链接数+外链接数,但实际搜索中存在偏差。链接测度偏差= {总链接数- [内链接数+外链接数]} / 总链接数。但这对本文来讲并没有评价分析的意义,因此也不被使用。

(6) 网络影响因子:

网络影响因子= 总链接数/网页数

外部网络影响因子= 外链接数/网页数

#### 1.3 研究工具的选择

##### 1.3.1 商业性搜索引擎

可以从 google, Alltheweb 等具有高级检索功能的搜索引擎中挑选。其中 google 功能强大,使用方便,除以上网络信息计量学指标外,google 还可给出其他具有评价意义的数据,如“类似网页”用于搜索与某网页相关的网页,其数值可以反映网站的个

性化程度;“链接”用于搜索与该网页存在链接的网页,在本文中是与某网站主页存在链接的网页,其数值反映其他网站对其重视的程度。

### 1.3.2 数据处理工具

根据需要和使用方便,我们选 Excel 即可。

## 2 数据来源和搜集方法

### 2.1 数据搜集方法

首先,通过 google 搜索各省级以上公共图书馆网址,然后,使用“高级搜索”获得计量指标。

下面以国家图书馆网址 www.nlc.gov.cn 为例来说明搜索过程。在“高级搜索”的检索框里按要求

填入 www.nlc.gov.cn,点击检索按钮,获得相关数据。或者使用下列检索式进行检索。

表 1 检索式列表

检索项目	检索式
“链接”数	link:www.nlc.gov.cn
类似网页数	related:www.nlc.gov.cn
内链接数	“www.nlc.gov.cn”site:www.nlc.gov.cn
外链接数	“www.nlc.gov.cn”-site:www.nlc.gov.cn
网页数	site:www.nlc.gov.cn
总链接数	“www.nlc.gov.cn”

### 2.2 网站链接数据

按以上方法所获数据列入表 2。

表 2 省级以上公共图书馆网站链接数据统计表

图书馆名称	网 址	总链接数	内链接数	外链接数	网页数	网络影响因子	外部网影响因子	类似网页数	链接数
国家图书馆	www.nlc.gov.cn/	5 930	4 260	1 680	4 260	1.392	0.394	28	867
首都图书馆	www.clcn.cn.net	1 940	1 640	121	1 680	1.155	0.072	31	66
天津图书馆	www.tjl.tj.cn/	3 070	3 250	104	3 250	0.945	0.043	31	38
河北省图书馆	www.helib.net	333	272	50	273	1.22	0.183	30	34
山西省图书馆	www.lib.sx.cn/	9	7	2	7	0.992	0.022	0	0
内蒙古图书馆	www.nmglib.com.cn/	43	19	24	19	2.263	1.263	31	38
辽宁省图书馆	www.lnlib.com/	4 550	3 320	164	3 370	1.35	0.049	31	29
吉林省图书馆	www.jlplib.com.cn/	706	671	23	680	1.038	0.034	31	22
上海市图书馆	www.library.sh.cn/	9 030	8 790	325	8 880	1.017	0.037	31	1 350
南京图书馆	www.jslib.org.cn/	1 770	1 690	31	1 740	1.017	0.019	29	14
浙江图书馆	www.zjlib.net.cn/	1 080	889	105	907	1.191	0.116	31	180
安徽省图书馆	www.ahlib.com/	31	26	5	26	1.192	0.192	0	7
福建省图书馆	www.fjlib.fz.fj.cn/	71	31	40	31	2.29	1.29	30	43
江西省图书馆	www.jxlib.gov.cn/	23	6	17	6	3.833	2.833	30	20
山东省图书馆	www.sdlib.com/	1 540	859	54	865	1.78	0.062	31	19
河南省图书馆	www.henanlib.gov.cn/	163	125	35	127	1.284	0.276	30	43
湖北省图书馆	www.library.hb.cn/	865	831	14	838	1.032	0.017	31	13
湖南图书馆	www.library.hn.cn/	6 490	8 460	90	8 690	0.747	0.01	29	52
广东省立中山图书馆	www.zslib.com.cn/	901	648	198	655	1.376	0.302	31	94
广西壮族自治区图书馆	www.gxlib.org.cn/	175	113	24	113	1.549	0.212	30	90
重庆图书馆	www.cqlib.cn/	327	322	5	335	0.976	0.015	0	13
四川省图书馆	www.sclib.org/	400	347	53	347	1.153	0.153	31	48
贵州省图书馆	www.lib.guz.cei.gov.cn	3	0	3	0			0	0
云南省图书馆	www.ynlib.cn/	117	113	3	114	1.026	0.026	0	20
陕西省图书馆	www.sxlib.org.cn/	566	481	54	485	1.167	0.111	31	51
甘肃省图书馆	www.gslib.org/	14	11	3	11	1.273	0.273	0	15

注:搜索时间为 2004-07-19,20 30- 22 30

### 3 数据分析

#### 3.1 网站链接数据说明

在以上的统计表中,“总链接数”是 google 搜索到的包含某网站 URL 的网页数,事实上这些网页与该网站存在链接,它反映了网站被链接的多少;“内链接数”是 google 搜索到的包含某网站 URL 的网页数中,来自该网站的网页数,它反映了网站内部结构的完备性;“外链接数”是 google 搜索到的包含某网站 URL 的网页数中,来自该网站外的网页数,它更好地反映了网站的质量;“网页数”是 google 搜索到的某网站包含的网页数,它在一定程度上反映了网站的大小;“网络影响因子”=“总链接数”/“网

页数”,它反映了网站网页被链接的总的平均水平;“外部网络影响因子”=“外链接数”/“网页数”,它是网站网页被外部链接的平均水平,被认为更好地反映了网站的水平;“类似网页数”是 google 搜索到的与某网站主页类似和相关的网页数,它反映了该网站主页的个性化程度;“链接数”是 google 搜索到的与某网站主页有链接的网页数。

对一个网站来讲,除“类似网页数”指标越小越好外,其他指标都是越大越好。表 3 是按不同指标的排序结果。

从表3结果可以看出,按不同指标的排序结果差别比较大,很难说按哪一种指标排序更好。因此,需要一种综合指标来排序。下面我们给出关联度排序。

表 3 按不同指标的排序结果列表

排序	总链接数	内链接数	外链接数	网页数	网络影响因子	外部网络影响因子	类似网页数	链接数
1	上海市图书馆	上海市图书馆	国家图书馆	上海市图书馆	江西省图书馆	江西省图书馆	甘肃省图书馆	上海市图书馆
2	湖南图书馆	湖南图书馆	上海市图书馆	湖南图书馆	福建省图书馆	福建省图书馆	安徽省图书馆	国家图书馆
3	国家图书馆	国家图书馆	广东省中山图书馆	国家图书馆	内蒙古图书馆	内蒙古图书馆	云南省图书馆	浙江图书馆
4	辽宁省图书馆	辽宁省图书馆	辽宁省图书馆	辽宁省图书馆	山东省图书馆	国家图书馆	山西省图书馆	广东省中山图书馆
5	天津图书馆	天津图书馆	首都图书馆	天津图书馆	广西壮族区图书馆	广东省立中山图书馆	重庆图书馆	广西壮族区图书馆
6	首都图书馆	南京图书馆	浙江图书馆	南京图书馆	国家图书馆	河南省图书馆	贵州省图书馆	首都图书馆
7	南京图书馆	首都图书馆	天津图书馆	首都图书馆	广东省中山图书馆	甘肃省图书馆	国家图书馆	湖南图书馆
8	山东省图书馆	浙江图书馆	湖南图书馆	浙江图书馆	辽宁省图书馆	广西壮族区图书馆	南京图书馆	陕西省图书馆
9	浙江图书馆	山东省图书馆	山东省图书馆	山东省图书馆	河南省图书馆	安徽省图书馆	湖南图书馆	四川省图书馆
10	广东省中山图书馆	湖北省图书馆	陕西省图书馆	湖北省图书馆	甘肃省图书馆	河北省图书馆	江西省图书馆	福建省图书馆
11	湖北省图书馆	吉林省图书馆	四川省图书馆	吉林省图书馆	河北省图书馆	四川省图书馆	福建省图书馆	河南省图书馆
12	吉林省图书馆	广东省中山图书馆	河北省图书馆	广东省中山图书馆	安徽省图书馆	浙江图书馆	河南省图书馆	内蒙古图书馆
13	陕西省图书馆	陕西省图书馆	福建省图书馆	陕西省图书馆	浙江图书馆	陕西省图书馆	广西壮族区图书馆	天津图书馆
14	四川省图书馆	四川省图书馆	河南省图书馆	四川省图书馆	陕西省图书馆	首都图书馆	河北省图书馆	河北省图书馆
15	河北省图书馆	重庆图书馆	南京图书馆	重庆图书馆	首都图书馆	山东省图书馆	内蒙古图书馆	辽宁省图书馆
16	重庆图书馆	河北省图书馆	广西壮族区图书馆	河北省图书馆	四川省图书馆	辽宁省图书馆	广东省中山图书馆	吉林省图书馆
17	广西壮族区图书馆	河南省图书馆	内蒙古图书馆	河南省图书馆	吉林省图书馆	天津图书馆	四川省图书馆	云南省图书馆
18	河南省图书馆	广西壮族区图书馆	吉林省图书馆	云南省图书馆	湖北省图书馆	上海市图书馆	浙江图书馆	江西省图书馆
19	云南省图书馆	云南省图书馆	江西省图书馆	广西壮族区图书馆	云南省图书馆	吉林省图书馆	陕西省图书馆	山东省图书馆
20	福建省图书馆	福建省图书馆	湖北省图书馆	福建省图书馆	上海市图书馆	云南省图书馆	首都图书馆	甘肃省图书馆
21	内蒙古图书馆	安徽省图书馆	重庆图书馆	安徽省图书馆	南京图书馆	山西省图书馆	山东省图书馆	南京图书馆
22	安徽省图书馆	内蒙古图书馆	安徽省图书馆	内蒙古图书馆	山西省图书馆	南京图书馆	辽宁省图书馆	重庆图书馆
23	江西省图书馆	甘肃省图书馆	云南省图书馆	甘肃省图书馆	重庆图书馆	湖北省图书馆	天津图书馆	湖北省图书馆
24	甘肃省图书馆	山西省图书馆	甘肃省图书馆	山西省图书馆	天津图书馆	重庆图书馆	上海市图书馆	安徽省图书馆
25	山西省图书馆	江西省图书馆	贵州省图书馆	江西省图书馆	湖南图书馆	湖南图书馆	吉林省图书馆	山西省图书馆
26	贵州省图书馆	贵州省图书馆	山西省图书馆	贵州省图书馆	贵州省图书馆	贵州省图书馆	湖北省图书馆	贵州省图书馆

3.2 网络链接关联度排序

由于对一个网站来讲,上述各指标,除“类似网页数”指标是越小越好外,其他指标都是越大越好。

因此我们可以在表2中插入一个参考行,其中与“类似网页”列对应的数值取所在列的最小值,其他数值取对应列的最大值。然后,由点关联度公式:

$$V(X_0(k), X_i(k)) = \frac{\min_k (X_0(k) - X_i(k)) + \max_k (X_0(k) - X_i(k))}{\max_k (X_0(k) - X_i(k)) + 4513.5} = \frac{4513.5}{\max_k (X_0(k) - X_i(k)) + 4513.5}$$

取  $\alpha = 0.5$ , 及关联度公式:

$$V_{0i} = V(X_0, X_i) = \frac{1}{n} \sum_{i=1}^n V(X_0(k), X_i(k))$$

计算得关联度数值记入上表中,并按关联度大小排序得下表。

按关联度排序,由于综合了各指标的优缺点与特

性,因而,排序结果更客观可靠。

对于表4中的数据,我们还可以选用关联度数据列为参考列,同样利用上述关联度公式计算其他指标数据列与该参考列的关联度(姑且称其为“指标关联度”)来看哪一指标更好,并排序。结果见表5。

表4 网络链接关联度排序表

排序	图书馆名称	外部网络影响因子	总网络影响因子	总链接数	内链接数	外链接数	网页数	类似网页数	链接数	关联度
0	参考行	2.833	3.883	9030	8790	1680	8880	0	1350	1
1	上海市图书馆	0.037	1.017	9030	8790	325	8880	31	1350	0.9700000
2	湖南图书馆	0.010	0.747	6490	8460	90	8690	29	52	0.8801250
3	国家图书馆	0.394	1.392	5930	4260	1680	4260	28	867	0.8107500
4	辽宁省图书馆	0.049	1.350	4550	3320	164	3370	31	29	0.7397500
5	天津图书馆	0.043	0.945	3070	3250	104	3250	31	38	0.7290000
6	首都图书馆	0.072	1.155	1940	1640	121	1680	31	66	0.7092500
7	南京图书馆	0.019	1.017	1770	1690	31	1740	29	14	0.7068750
8	浙江图书馆	0.116	1.191	1080	889	105	907	31	180	0.7017500
9	山东省图书馆	0.062	1.780	1540	859	54	865	31	19	0.6996250
10	广东省立中山图书馆	0.302	1.376	901	648	198	655	31	94	0.6992500
11	湖北省图书馆	0.017	1.032	865	831	14	838	31	13	0.6963750
12	吉林省图书馆	0.034	1.038	706	671	23	680	31	22	0.6950000
13	陕西省图书馆	0.111	1.167	566	481	54	485	31	51	0.6941250
14	四川省图书馆	0.153	1.153	400	347	53	347	31	48	0.6923750
15	重庆图书馆	0.015	0.976	327	322	5	335	0	13	0.6918750
16	河北省图书馆	0.183	1.220	333	272	50	273	30	34	0.6915000
17	广西壮族自治区图书馆	0.212	1.549	175	113	24	113	30	90	0.6906250
18	河南省图书馆	0.276	1.284	163	125	35	127	30	43	0.6900000
19	云南省图书馆	0.026	1.026	117	113	3	114	0	20	0.6896250
20	福建省图书馆	1.290	2.290	71	31	40	31	30	43	0.6892500
21	内蒙古图书馆	1.263	2.263	43	19	24	19	31	38	0.6887500
22	甘肃省图书馆	0.273	1.273	14	11	3	11	0	15	0.6887500
23	安徽省图书馆	0.192	1.192	31	26	5	26	0	7	0.6887500
24	山西省图书馆	0.022	0.992	9	7	2	7	0	0	0.6886250
25	江西省图书馆	2.833	3.833	23	6	17	6	30	20	0.6882500
26	贵州省图书馆			3	0	3	0	0	0	0.5846667

表5 各指标的关联排序表

指标	关联度	外部网络影响因子	总网络影响因子	类似网页数	链接数	外链接数	内链接数	网页数	总链接数
指标关联度	1	0.99987	0.99984	0.99515	0.97798	0.95450	0.81630	0.81491	0.80520

(下转第44页)

其优越的一面,但也有不少局限。一方面是方法本身存在的缺陷,另一方面就是所采用的计算方法,如公因子的提取方法,可能并不是非常适合证券投资基金的绩效评价。对以上两方面的缺陷还需要做大量深入地研究和分析,这也是以后的研究方向。

#### 参考文献

- 1 郑晓辉,肖 慧. 中国证卷投资基金绩效来源分析. 经济科学,2002(5)
- 2 徐尚友,郑垂勇. 因子分析法在投资项目财务综合评价中的应用. 技术经济与管理研究,2003(1)

- 3 刘 芳等. 对我国投资基金业绩进行评价的实证研究. 运筹与管理,2003(6)
- 4 胡良平. 现代统计学与 SA S 应用. 北京:军事医学科学出版社,2000
- 5 王 聪. 证券投资基金绩效评估模型分析. 经济研究,2001(9)
- 6 沈维涛,黄兴李. 我国证券投资基金业绩实证研究与评价. 经济研究,2001(9)

#### 作者简介

龙子泉,武汉大学商学院管理科学与工程系教授。  
赵红艳,武汉大学商学院管理科学与工程系硕士研究生。

(上接第 28 页)

解网络信息的分布情况,掌握网络信息交流规律;可以分析站点的集中度和联系程度,确定核心信息来源;可以优化网络站点的结构,进行合理配置资源;有助于网络信息资源的开发利用;有利于研究网上的信息交流和学科知识结构特点,掌握学科发展状况和发展趋势。总的来说,研究网络环境下引用行为具有十分重要的意义,网络的引用分析评价问题还有待进一步探讨。

#### 参考文献

- 1 邱均平. 文献计量学. 北京:科技文献出版社,1988
- 2 思萌(黄晓斌). 引文分析的作用与局限性及其改进. 图书馆建设,1992(6):17-20
- 3 邓宏炎等. 网络参考文献初探. 武汉大学学报(人文社会

科学版),2000(1):130-134

- 4 Blaise Cronin. Bibliometrics and beyond: some thoughts on web-based citation analysis. Journal of Information Science,2001(1):1-7
- 5 Julian Warner. A critical review of the application of citation studies to the research assessment exercises. Journal of Information Science,2000(6):453-460
- 6 Lennart Björneborn, Peter Ingwersen. Perspective of webometrics. Scientometrics,2001(1):65-81

#### 作者简介

黄晓斌,中山大学信息管理系副教授、情报学博士。国际图联教育与培训部通讯委员。曾在美国伯克利加州大学信息管理与系统学院、佛罗里达州立大学信息学院作访问研究员。先后发表论文 90 多篇,出版著作(含参加编写)5 部。主要研究方向为数字图书馆技术、网络信息资源开发利用。

(上接第 32 页)

表 5 的结果中“指标关联度”表明了以“关联度”为参照,其他各网络信息计量指标的优劣度,反映了外部网络影响因子等指标评价网站的可行性,以及指标的可信程度,实证性的给出了各指标的优劣排序。

#### 4 小结

综上所述,我们通过 google 搜索引擎,获得我国省级以上公共图书馆网站的网络链接数据,并进行了分析。介绍了利用 google 获得各项网络链接分析指标的方法,将各馆网站按不同指标进行了排序,

以及综合排序,同时,也给出了各网络信息计量指标在网络链接分析中的地位的实证分析及排序。相信这些结果对网络信息计量研究有一定意义。

#### 参考文献

- 1 杨 涛. 网络信息计量学实证研究:对国内 20 个大学网站的分析. 图书情报工作,2003,(9):61-66
- 2 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. 情报学报,2003,22(4):398-402

#### 作者简介

王宏鑫,男,信阳师范学院图书馆研究馆员。