



# 体育院校图书馆网站链接特征的计量分布

司虎克, 马洪亮, 周敦敏, 彭忱

**摘要:** 运用链接分析法, 以全国体育院校图书馆网站为样本, 考察了6个主要衡量网站链接特征的指标。结果表明, 全国体育院校图书馆网站的链接特征各项指标中存在较大差异, 数据量都不符合正态分布; 网站中指向内部的链接在数量和所占比例高于指向外部的链接; 通过R型聚类分析, 将各指标特征比较一致体育院校图书馆网站大致分为4类, 指出体育院校图书馆网站的建设水平差异较大, 仅有少数网站的布局结构比较合理, 内容比较完善, 大部分体育院校图书馆的网站建设还需进一步加强。

**关键词:** 体育院校; 图书馆; 网站; 链接; 计量

中图分类号: G80-05 文献标识码: A 文章编号: 1006-1207(2009)02-0033-06

## Measurement Distribution of the Characteristics of Website Hyperlinks of the PE College Libraries

SI Hu-ke, MA Hong-liang, Zhou Dunmin, et al.

(Shanghai University of Sport, Shanghai 200438, China)

**Abstract:** By the method of hyperlink analysis and taking the PE college library websites in China as the samples, the author examines 6 indices measuring the characteristics of website hyperlinks. The result shows that there exists great difference in the different indices. The measurements of these indices do not follow the lognormal distribution. The number and proportion of the website's self-links are much larger than those of the outgoing-links. Through R-type cluster analysis, PE college library websites, whose characteristics of the different indices are quite similar, can be divided into 4 categories. The article points out that there is a great difference in the construction of PE college library websites. Only a few websites have appropriate layout and perfect contents. Most of the PE college library websites need to be improved.

**Key words:** PE college; library; website; hyperlink; measurement

## 1 前言

加强体育信息工作, 组织信息化建设重点工程, 积极推进体育院校图书馆的信息化建设, 是在网络环境下体育信息资源建设与保障的一个重要环节。以网络数据为代表的现代信息环境为体育院校图书馆信息建设的发展提供了良好的机遇, 同时也冲击着传统的体育院校图书馆信息资源建设的工作模式, 这就使我们必须进一步审视体育院校图书馆信息资源建设工作, 重新定位体育院校图书馆信息资源建设工作的职能, 探索网络环境下体育院校图书馆信息资源建设的新思路。

由于图书馆网站建设是网络环境下体育信息资源建设中的一个重要环节, 这就要求对我国各体育院校图书馆网站建设特点作深入研究, 然而从全国各个体育院校图书馆网站规模、数量与质量现状看, 还存在着诸多问题和不足。因此, 对我国各个体育院校图书馆网站展开实证研究, 有助于我们进一步认识网络环境下图书馆网站建设的特征, 更好地把握体育信息化建设的规律, 因而具有重要的实践意义。

## 2 研究现状

近年来我国一些学者分别从现代信息技术在体育中的应用与发展、网络环境下的信息资源管理、体育信息管理的理论构

建及应用系统开发研究、体育信息网络的建设与发展前景、全国高等体育教育系统文献信息资源共享模式研究、奥运信息门户的设计与实现、体育信息整合暨区域间信息共享的研究、我国体育管理信息化的研究、北京奥运会对我国体育信息化建设的影响等不同角度和层面, 对我国体育信息化建设过程中的热点和难点问题进行了有益地探索与研究。<sup>[1-14]</sup>同时还分析了我国的体育信息网络化建设与国外的主要差距。

经过对文献的整理和归纳发现, 关于我国各体育院校图书馆信息资源建设的研究仍存不足, 值得我们对我国体育院校图书馆网络信息资源建设中的一个重要环节给予关注, 即各体育院校图书馆体育网站建设状况不如人意, 发展不平衡, 网站互相链接不合理, 还缺乏对体育院校网站的系统研究及其相应评价, 反映出我国各体育院校网站建设仍处于初步阶段, 其建设与发展及其辐射影响力还很有限。这正是本文研究的逻辑起点。

## 3 研究方法

### 3.1 对象

以全国除港澳台外的14所体育院校图书馆主页所在的网站作为研究总体, 由于西安体育学院和成都体育学院图书馆的主页通过宽带利用因特网浏览器无法浏览, 故从中选取其

收稿日期: 2009-01-20

基金项目: 2007年上海市教委社会科学研究项目(07ZS122)

作者简介: 司虎克(1955-), 男, 教授, 博士生导师, 主要研究方向: 网络环境下的体育信息化建设。

作者单位: 上海体育学院图书馆, 上海 200438



余的 12 所体育院校图书馆网站作为样本。它们分别是：北京体育大学、上海体育学院、武汉体育学院、首都体育学院、天津体育学院、南京体育学院、吉林体育学院、广州体育学院、沈阳体育学院、山东体育学院、哈尔滨体育学院、河北体育学院 12 所体育院校的图书馆网站。

### 3.2 文献综述法

通过对复旦大学图书馆 ISI 外文文献检索、上海体院图书馆、中国期刊网、清华同方数据库、维普中文科技期刊网相关检索，确定与本文有关的参考文献，并对文献进行分析与梳理，了解与把握国内外研究现状，确立自己的研究目标以及研究重点。

### 3.3 链接分析法

参照与运用链接解析工具 WebStat，统计相关各指标数据。本研究利用 WebStat 获取网站的总链接数、指向内部的链接数、指向外部的链接数和被链接网站数，以及网站的字节数和页面文件数，进而计算各类链接密度和页面平均各类链接数。<sup>[15-16]</sup>

#### 3.3.1 衡量网站链接特征的指标

网络链接是网站内部信息组织以及与外部联系的方式，网站链接则是网站链接属性的总和。体育院校图书馆网站具有为读者提供体育信息资源服务的功能，通过对链接各指标的分析，探讨体育院校图书馆网站的性质、信息量、信息的组织原则及网站各指标的分布规律，对于促进体育院校图书馆网站的体育信息资源有着重要的意义。

#### 3.3.2 总链接数 (T-NUM)

一般来讲，体育网站的链接总数越多，网络信息的组织体系就是越完备，信息揭示程度越高，而网站链接总数是网站内部所含有的网络链接总量，它是衡量网站链接数量特征最主要的指标之一。

#### 3.3.3 自链接数 (S-NUM)

网站的自链接数是指网站中的自我链接的数量。一般说来，自链接数在总链接数中所占比例高的网站，往往其内部的体育信息揭示较为充分，能较好的反映出网站自身具有较完备的体系。像图书馆这一类的资源性网站这一指标值较

大，占总链接数的比例也较高。

#### 3.3.4 出链接数 (OUT-NUM)

出链接数是指网站中指向主机所在域以外资源的链接数量，通过建立该指标并据此考察图书馆网站，可以判断图书馆网站所建立的网络信息资源的性质是否具有重要的参考价值。

#### 3.3.5 被链接网站数 (L-NUM)

被链接网站数是指网站的出链所指向的链宿的数量，据此可以考察与评价图书馆网站的辐射力和该网站与相关网站联系的密切程度。网站指向外部的链接数越多，其开放性就越强，用户通过该网站能进一步获得的相关信息越丰富。

#### 3.3.6 链接密度 (Link Density-LD)

链接密度是度量链接出现频率的指标，单位为 bytes/link，它消除了网站规模对链接数的影响。该值越大，表示网站中链接出现的频率越低，分布越稀疏，反之亦然。该指标可以反映图书馆网站所含各类体育信息量和对信息揭示的深度所具有的重要价值。

#### 3.3.7 页面平均链接数 (Links/page, LP)

页面平均链接数也是度量链接出现频率的指标，单位为 links/page。与链接密度不同，它揭示的是平均每个网页中链接的数量，因此，该指标与网站字节数无关，但会受网站设计风格的影响。

### 3.5 数理统计法

运用统计分析软件 SPSS11.0 处理体育院校图书馆网站各相关指标数据的分布特征及其差异性。

## 4 结果与分析

### 4.1 体育院校图书馆网站链接特征各指标的计量分布特征

利用 SPSS 统计分析软件描述了样本各指标数据量的集中和离散趋势，以及分布特征。同时，以频数分布和正态分布的概率密度函数拟合各指标统计量的分布情况，通过偏度与峰度检验明确个指标的分布趋势。

#### 4.1.1 体育院校图书馆网站链接数的分布特征

链接数是衡量网站链接特征的常用指标，见表 1 与表 2。本研究表明，体育院校图书馆网站的平均总链接数为

表 1 体育院校图书馆网站各指标统计量  
Table 1 Indices Measurement of the PE College Library Websites

样本编号	总链接数	自链接数	出链数	被链接网站数	网站总字节数	页面数	总链接密度	自链接密度	出链接密度	页面平均		
										总链接数	自链接数	出链接数
1	1681	1398	283	28	1664932	105	991	1191	5883	16	13	3
2	1087	508	579	256	69178118	61	63618	136070	119479	18	9	9
3	2930	2426	496	367	93719841	120	31986	38628	189028	24	20	4
4	597	445	152	33	37864361	89	63382	85012	249108	7	5	2
5	522	212	311	231	15696290	68	30046	74109	50535	8	3	5
6	362	235	128	24	28664375	10	79140	122184	224642	35	23	12
7	295	211	84	16	34083590	27	115616	161687	405757	11	8	3
8	583	306	277	175	80979215	47	138901	264811	292133	13	7	6
9	611	175	436	347	2699895	103	4419	15428	6192	6	2	4
10	687	521	165	19	15549799	34	22648	29823	94127	20	15	5
11	842	566	276	236	5109607	51	6067	9031	18486	17	11	6
12	270	222	49	14	2296127	19	8498	10362	47245	14	11	3



表2 体育院校图书馆网站各指标统计量偏度与峰度检验一览表

Table II Indices Measurement Skewness and Kurtosis of the PE College Library Websites

	总链接数	自链接数	出链数	被链接网站数	网站总字节数	页面数	总链接密度	自链接密度	出链接密度	页面平均		
										总链接数	自链接数	出链数
有效值	12	12	12	12	12	12	12	12	12	12	12	12
偏度	2.176	2.321	0.547	0.435	0.918	0.250	0.930	1.209	0.753	1.097	0.566	1.474
偏度标准误差	0.637	0.637	0.637	0.637	0.637	0.637	0.637	0.637	0.637	0.637	0.637	0.637
峰度	5.019	5.342	-0.576	-1.544	-0.485	-1.217	-0.129	1.214	-0.351	1.519	-0.424	2.038
峰度标准误差	1.232	1.232	1.232	1.232	1.232	1.232	1.232	1.232	1.232	1.232	1.232	1.232

872, 自链接数为602, 占总数的69%, 出链接数为270, 占31%; 可以看出, 指向内部的链接占总链接数的大部分。其中, 北京体育大学、上海体育学院、武汉体育学院及哈尔滨体育学院图书馆网站总链接数明显多于其它体育院校图书馆网站; 自链数、总字节数、页面数等指标也存在较大的差异。说明全国体育院校图书馆网站建设水平参差不齐, 存在较大差距。

为了解体育院校图书馆网站各指标数据的分布是否符合正态分布, 对各指标统计量进行了偏度与峰度检验。显示, 总链接数、自链接数、页面平均出链接数和自链接密度均呈比正态分布更陡峭的正偏尖顶峰分布; 而出链接数、被链接网站数、字节数、页面数、总链接密度、出链接密度、页面平均总链接数、和页面平均自链接数则呈比正态分布更平坦的正偏平顶峰分布。总体来讲各项指标都不呈现正态分布特征。

4.1.2 体育院校图书馆网站各指标数据的概率分布

对所收集的数据进行概率分布统计, 如图1至图12所示。可以看出, 总链接数、出链接数、页面数、页面平均总链接数、页面平均自链接数和页面平均出链接数等数据的分布特征成交比较明显的单峰对称分布, 其余的指标数据峰状对称分布不明显。

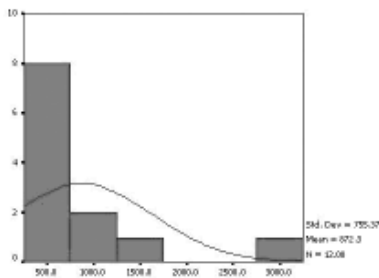


图1 体育院校图书馆网站总链接数 (T-NUM) 的数据分布特征  
Figure 1 Data Distribution Characteristics of T-NUM of the PE College Library Websites

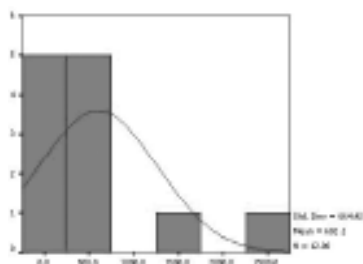


图2 体育院校图书馆网站自链接数 (S-NUM) 的数据分布特征  
Figure 2 Data Distribution Characteristics of S-NUM of the PE College Library Websites

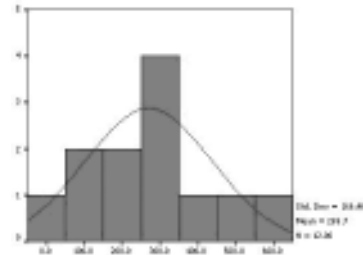


图3 体育院校图书馆网站出链接数 (OUT-NUM) 的数据分布特征  
Figure 3 Data Distribution Characteristics of OUT-NUM of the PE College Library Websites

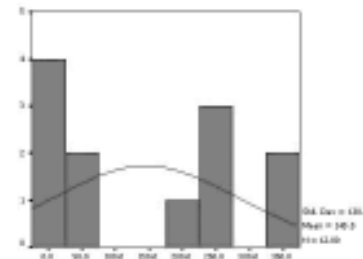


图4 体育院校图书馆网站被链接网站数 (L-NUM) 的数据分布特征  
Figure 4 Data Distribution Characteristics of L-NUM of the PE College Library Websites

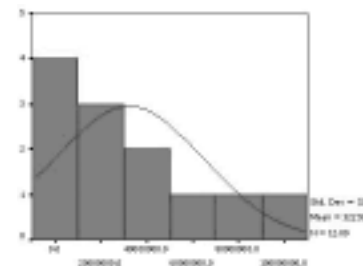


图5 体育院校图书馆网站总字节数的数据分布特征  
Figure 5 Data Distribution Characteristics of the Total Number of Bytes of the PE College Library Websites

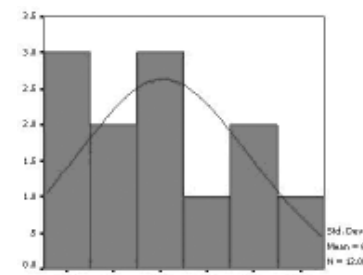


图6 体育院校图书馆网站页面数的数据分布特征  
Figure 6 Data Distribution Characteristics of the Web Pages of the PE College Library Websites

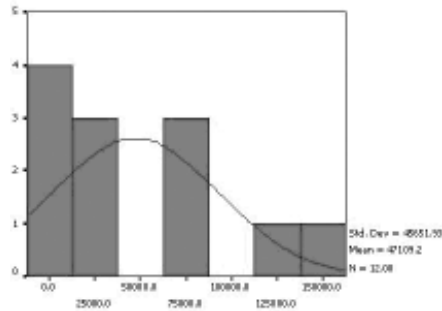


图7 体育院校图书馆网站总链接密度 (LD-T) 的数据分布特征  
Figure 7 Data Distribution Characteristics of LD-T of the PE College Library Websites

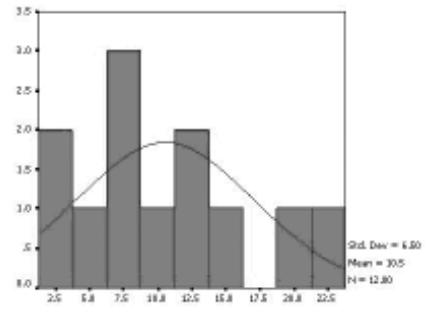


图11 体育院校图书馆网站页面平均自链接数 (LP-S) 的数据分布特征

Figure 11 Data Distribution Characteristics of LP-S of the PE College Library Websites

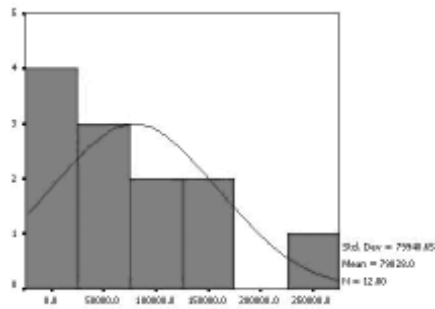


图8 体育院校图书馆网站自链接密度 (LD-S) 的数据分布特征  
Figure 8 Data Distribution Characteristics of LD-S of the PE College Library Websites

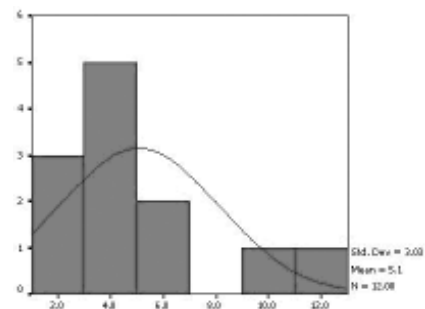


图12 体育院校图书馆网站页面平均出链接数 (LP-OUT) 的数据分布特征

Figure 12 Data Distribution Characteristics of LP-OUT of the PE College Library Websites

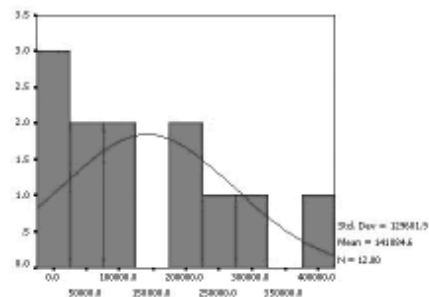


图9 体育院校图书馆网站出链接密度 (LD-OUT) 的数据分布特征  
Figure 9 Data Distribution Characteristics of LD-OUT of the PE College Library Websites

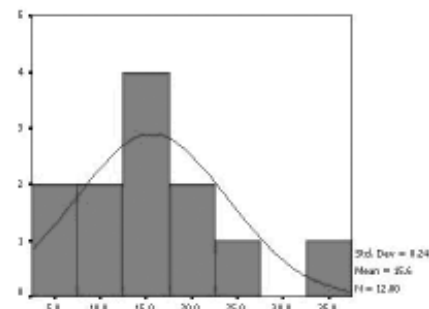


图10 体育院校图书馆网站页面平均链接数 (LP-T) 的数据分布特征

Figure 10 Data Distribution Characteristics of LP-T of the PE College Library Websites

#### 4.1.3 体育院校图书馆网站链接密度的分布特征

与总链接数不同, 链接密度消除了网站规模和页面文件大小所造成的影响。由表3 体育院校图书馆网站各指标统计量的数据分布特征中可知体育院校图书馆网站的链接密度平均为 47109bytes/link, 而 Almind 的研究仅为 635bytes/link (1997), 邱均平等人对 101 个美国商学院和医学院网站中抽取 40 个作为研究对象得出的链接平均密度为 6129bytes/link (2005)<sup>[7]</sup>。

表3 体育院校图书馆网站各指标统计量的数据分布特征  
Table III Data Distribution Characteristics of the Indices Measurement of the PE College Library Websites

指标	数据量	算术平均数	中位数	最大值	最小值
总链接数	12	872	604	2930	270
自链接数	12	602	376	2426	175
出链接数	12	270	277	579	49
被链接网站数	12	146	104	367	13
字节数	12	32292179	22180333	93719841	1664932
页面数	12	61	56	120	10
总链接密度	12	47109	31016	138901	991
自链接密度	12	79028	56369	264811	1191
出链接密度	12	141885	106803	405757	5883
页面平均总链接数	12	16	15	35	6
页面平均自链接数	12	11	10	23	2
页面平均出链接数	12	5	4	12	2



出现如此巨大差距的原因可能是：第一网站类型的不同。图书馆网站主要提供有利于研究的信息资源，为揭示信息的深度所含的链接密度就大。而综合性大学的网站往往提供综合的多方面的信息，信息深度不高的信息相对较多，链接密度相对就小。第二统计方法的差异。在计算网站字节数时除页面文件外，还包含了图像文件、压缩文件、用户自定义文件和其他文件等，页面文件只占总文件数的19.32%，其余类型的文件占80.68%，这对链接密度的计算会直接造成巨大影响。第三网页制作技术的进步对链接密度也有一定的影响。

4.1.4 体育院校图书馆网站页面平均链接数的分布特征

页面平均链接数与链接密度相比具有更大的优越性，它与网站字节数无关，因而消除了多媒体文件、压缩文件等非页面文件的影响。由表4可知：体育院校图书馆网站页面平均链接数为16 links/page，页面平均自链数为11 links/page，页面平均出链数为5 links/page。1997年Almind和Ingwersen的研究表明，平均每个网页有9.09个链接。比较这两项研究可以看出，随着时间的推移，网页平均链接数是递增的。邱均平

等人研究表明，美国商学院网站的页面平均链接数为25.59 links/page，医学院网站的页面平均链接数为21.85 links/page (2005)；说明我国体育院校图书馆网站的建设水平与国外综合大学网站相比，水平还较低。

4.1.5 体育院校图书馆网站被链接网站数的分布特征

被链接网站数从不同角度揭示了同一个问题，即网站的辐射力。同时，网站之间的链接关系和数量还是网站评价的重要依据，当然，被链接次数多的网站并不一定就具有很高的学术性和质量。由表3可知，体育院校图书馆网站平均链接到146个其他网站，但是最大值为367，而最小值仅为13，也说明体育院校图书馆网站建设的水平相差很大。

4.1.6 体育院校图书馆网站各指标显著性比较

多独立样本非参数检验分析样本数据是推断样本来自的多个独立总体是否存在显著差异。SPSS多独立样本非参数检验一般推断多个独立总体的均值或中位数是否存在显著差异。通过对2007年1月至5月体育院校图书馆网站各指标数据进行多独立样本非参数检验，其结果见表4。

表4 多独立样本K-W检验结果
Table IV Multi-Independent Samples of K-W Test

Table with 6 columns: Chi-Square, df, Asymp. Sig., 总链接数T-NUM, 自链接数S-NUM, 出链数OUT-NUM, 被链接网站数L-NUM, 网站总字节数, 页面数

由表4多独立样本K-W检验结果可以看出，总链接数、自链接数、出链数、等指标的P均小于显著水平0.05，因此可以认为体育院校图书馆网站各指标存在显著差异。

4.2 体育院校图书馆网站各指标统计量Q型聚类分析

随着社会进入信息时代，信息量越来越大，传播速度快，而传统图书馆信息吞吐能力、传播速度等方面已不能满足读者要求。如何使体育院校图书馆在当今激烈的网络环境下立足、长久并可持续发展，如何搞好体育信息资源建设并使其效益最大化是体育院校图书馆网站建设过程中亟待解决的问题。

层次聚类分析中的Q型聚类可以使具有共同特点的样本聚齐在一起，以便对不同类的样本进行分析。由体育院校图书馆网站各指标统计量Q型聚类分析样本归属情况可知，北京体育大学图书馆网站属于第一类；上海体育学院、天津体育学院、广州体育学院沈阳体育学院及哈尔滨体育学院图书馆网站属于第二类；武汉体育学院图书馆网站属于第三类；首都体育学院、南京体育学院、吉林体育学院、山东体育学院及河北体育学院图书馆网站属于第四类。

从图13体育院校图书馆网站各指标统计量Q型聚类分析树形图中可以看出，各类之间的距离在25的坐标内。北京体育大学图书馆网站的综合水平位居体育院校前列；上海体育学院、天津体育学院、广州体育学院沈阳体育学院及哈尔滨体育学院图书馆网站的特征较一致，网络信息的组织体系比较完备，信息揭示程度高，网站内部体系比较完备；武汉体育学院图书馆网站由于其自身的链接特征自属一类；首都体育学院、南京体育学院、吉林体育学院、山东体育学院及河北体育学院图书馆网站的特征较一致，这几个网站的总链接

数、自链接数、出链数以及被链接网站数目近似，差异不大，其网站内部体系还需进一步完善。这说明我国体育院校图书馆网站在信息资源建设的道路上已经迈出了坚实的步伐，但是图书馆网站建设水平依然参差不齐，各体育院校间差距依然很大，仅有少数网站的布局结构比较合理，内容比较完善，大部分体育院校图书馆的网站建设还需进一步加强。

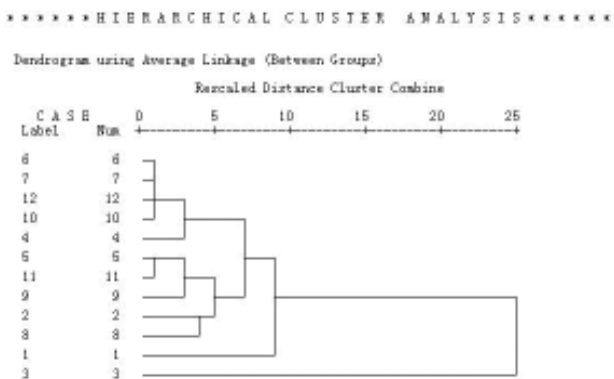


图13 体育院校图书馆网站各指标统计量Q型聚类分析树形图
Figure 13 Q-Type Cluster Analysis Tree of the Indices Measurement of the PE College Library Websites

加强体育院校图书馆网站的信息资源建设是网络环境下体育院校图书馆发展的必由之路，这是因为社会信息化的大趋势必然要求并推动作为社会信息化重要组成部分的图书馆加快网站及其信息资源建设。因此，必须充分利用先进的信息技术，加快图书馆的网络建设，完善网站结构，迅速提高其信息处理和传播的能力，缩小差距，以便满足信息时代



发展的迫切需要,为体育院校乃至整个国家体育事业的信息资源建设做出应有的贡献。

## 5 结论

5.1 我国各体育院校图书馆网站的建设水平差异较大,其中网站中指向内部的链接在数量和出现频率上高于指向外部的链接,以及网站链接的各项指标并不呈现正态分布特征。

5.2 在我国各体育院校图书馆网站中,仅有少数网站的布局结构比较合理,内容比较完善,而大部分体育院校图书馆网站建设还需进一步加强。

5.3 我国各体育院校图书馆网站链接特征可以分为4种类型,即北京体育大学图书馆网站网络信息的组织体系比较完备,信息揭示程度高,网站内部体系比较完备属于第一类;其次是上海体育学院、天津体育学院、广州体育学院沈阳体育学院及哈尔滨体育学院图书馆网站的指标特征近似,差异不大,处于中间水平属于第二类;武汉体育学院图书馆网站属于第三类;而首都体育学院、南京体育学院、吉林体育学院、山东体育学院及河北体育学院图书馆网站则属于第四类,其网站内部体系还需进一步加强与完善。

## 参考文献

- [1] 廖慧平, 尧旭华, 等. 我国体育信息网络化发展现状分析[J]. 山东体育学院学报, 2005, 21(2).
- [2] 马铁. 现代信息技术在体育中的应用与发展[J]. 体育科技信息, 1997(3): 56.
- [3] 董小英. 网络环境下的信息资源管理[D]. 北京大学博士学位论文, 1997: 12.
- [4] 董伦红. 体育信息管理的理论构建及应用系统开发研究[D]. 北京体育大学博士学位论文, 2002: 4.
- [5] 张江南. 体育信息网络的建设与发展前景[J]. 情报学报, 2003, 21(1): 86-88.
- [6] 洪雁. 奥运信息门户的设计与实现[M]. 武汉大学硕士学位论文, 2005: 4.
- [7] 胡彪, 饶坚, 等. 体育信息整合暨区域间信息共享的研究[J]. 武汉体育学院学报, 2006, (2): 5-8.
- [8] 尹博. 我国体育管理信息化的研究[J]. 北京体育大学学报, 2006, 29(5): 611-613.
- [9] 邱均平, 黄晓斌, 段宇峰, 等. 网络数据分析[M]. 北京: 北京大学出版社, 2004.
- [10] 刘刚, 娄策群. 政府网站评价研究与应用分析[J]. 国书情报知识, 2006(1): 104-109.
- [11] 段宇峰, 刘伟. 电子政务信息资源网络影响力评价指标体系研究[J]. 情报资料工作, 2006(1).
- [12] 文红斌, 马飞. 北京奥运会对我国体育信息化建设的影响[J]. 商场现代化, 2006, 469(6): 213.
- [13] 周兰君, 张天健. 国内外体育信息网络化发展现状及对策研究[J]. 广州体育学院学报, 2003, 23(6).
- [14] 廖慧平, 刘兴杰. 我国体育信息网络化建设的指导思想、目标、思路与原则[J]. 武汉体育学院学报, 2005, 39(1).
- [15] 段宇峰. 网络链接分析与网站评价研究[M]. 北京: 北京图书馆出版社, 2005: 157-158.
- [16] 邱均平, 陈敬全, 段宇峰. 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, 6: 151-155.
- [17] 邱均平, 段宇峰. 网络信息计量学研究(III)——大学网站链接特征指标的计量分布规律[J]. 情报学报, 2005, 24(4).

(责任编辑: 陈建萍)