

系统科学

基于 Rough 集理论的图书馆网站 综合评价方法研究

王建龙

(河海大学商学院, 常州 213022)

摘要 图书馆网站是现代化图书馆自我展示的窗口, 如何科学、合理地评价图书馆网站以提高其建设质量, 显得尤为重要。通过对图书馆网站的分析, 建立了其评价指标体系, 利用灰色统计的方法来确定模糊测评矩阵, 并利用 Rough 集理论来确定相应评价指标的权重, 然后运用模糊方法对网站进行了综合评价, 从而为提高图书馆网站建设质量提供了有益的理论参考。

关键词 图书馆网站 评价指标 模糊综合评价 灰色统计 粗糙集

中图法分类号 N945.16; **文献标志码** A

随着网络技术的飞速发展, 越来越多的图书馆都在 Internet 上建立了自己的网站, 为读者提高网络化的信息服务。网站作为图书馆新的形象标志, 把图书馆馆藏特色、信息服务、研究力量、参考咨询等众多方面的综合能力在网络上进行全方位的展示^[1]。因此网站是现代化图书馆自我展示的窗口, 是图书馆传统业务在时间和空间上的延伸与拓展^[2]。如果说图书馆发展最大的变化是以网络为中心而不再以图书馆建筑为中心, 以利用网络提供有序的数字化信息为中心, 而不再以提供传统印刷型资源为中心, 则这一切都取决于图书馆网站建设的成功与否。但由于各种原因, 图书馆的网站在建设水平上良莠不齐, 客观上需要对网站做出评价。因此, 对网站进行科学、合理的评价就显得尤为重要。网站评价能够很好地促进图书馆改进网站的

内容及服务功能, 促进网站的发展, 以及指导如何提高网站质量。

目前对图书馆网站的评价大都是利用模糊综合评判的方法, 而利用此方法, 指标权重分配选取是否合理将直接影响到评判结果。从目前的研究情况来看, 人们对图书馆网站建设的认识处于一种灰色的系统状态, 因而用灰色系统理论来确定图书馆网站的模糊测评矩阵具有一定的适用性和科学性。而 Rough 集的最大特点是不需要提供求解问题时所需处理的数据集合之外的任何先验知识, 仅对实测数据本身进行分类处理即可发掘隐含知识和揭示潜在的数据内部规律; 并且能在保留关键数据信息的前提下对数据进行化简并求出知识的最小表达, 故可以考虑利用 Rough 集理论来确定各评价指标的权重。本文研究的内容就是采用一种集成的方法对图书馆网站模型进行分析评价, 为图书馆提高其建设质量提供了一种可行的方法。

2010年4月7日收到

作者简介: 王建龙(1988—), 男, 江苏兴化人, 研究方向: 供应链管理。E-mail pangqh@ hhuc. edu. cn。

1 图书馆网站评价指标体系的建立

由于图书馆网站的建设是一项系统工程,故评价时需要对众多因素综合考虑,而评价时往往有众多因素无法定量分析,因此怎样把一些定性的因素和定量的因素统一在同一个评价尺度上,是评价研究的关键。故在建立评价指标时一般应遵循以下的原则:①简洁、完备性原则。指标体系应能全面反映图书馆网站建设质量的情况,但指标数目尽可能的少,各指标之间不应有强相关性。②客观、可

考核原则。指标设定过程尽量不受主观因素的影响,数据来源要真实可靠以确保结果的真实性和可考核性。③可扩充原则。指标应根据图书馆网站不同阶段的要求对指标体系进行修改、增加和删除,并可根据具体情况将质量指标进一步具体化。④导向性原则。指标体系的建立和评价产生的效果能有利于端正图书馆网站建设的指导思想,提高服务水平和服务质量,并易于和同类部门网站相比较。

基于以上原则考虑,并结合相关文献^[1-8],我们设计了如下的指标体系,见图1。

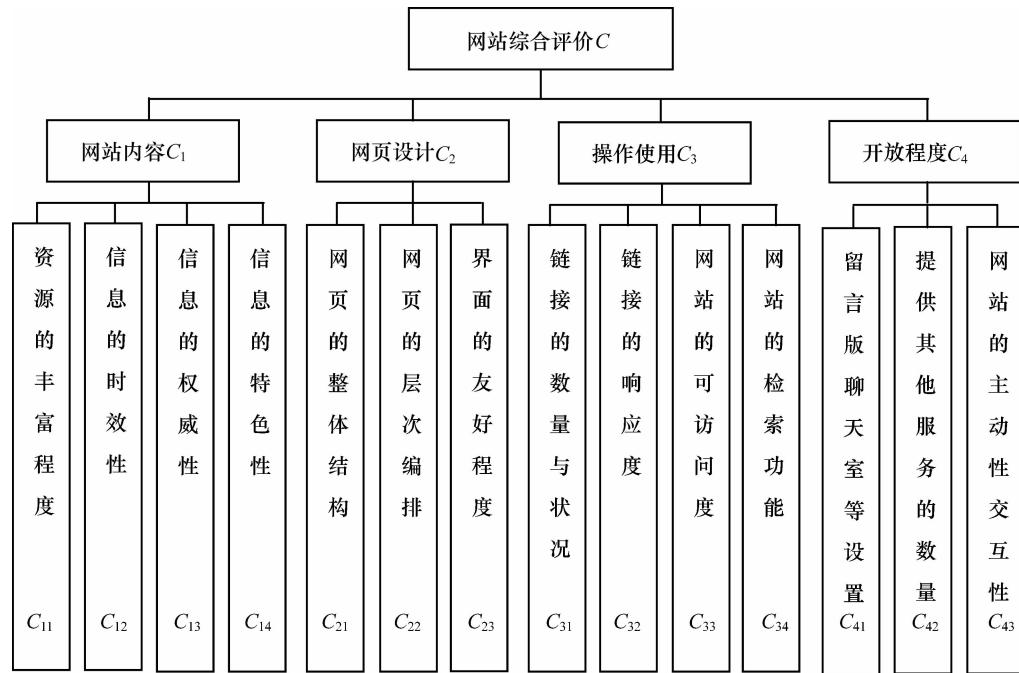


图1 网站综合评价指标体系

2 综合评价步骤

2.1 建立评价因素(即评价指标)集

如图1所示,分两个层次,建立4个因素集合。其中第一层次为: $C = \{C_1, C_2, C_3, C_4\}$; 第二层次为: $C_1 = \{c_{11}, c_{12}, c_{13}, c_{14}\}$, $C_2 = \{c_{21}, c_{22}, c_{23}\}$, $C_3 = \{c_{31}, c_{32}, c_{33}, c_{34}\}$, $C_4 = \{c_{41}, c_{42}, c_{43}\}$ 。

2.2 确定评价等级集合 V

这是一个表示评价目标优劣程度的集合。用 $V = (v_1, v_2, \dots, v_m)$ 表示,其中 v_i 表示第 i 等评价级别, m 表示评价等级数。这里采用四级评判方法,即 $m = 4$, $V = (v_1, v_2, v_3, v_4) = (\text{优}, \text{良}, \text{中}, \text{差})$ 。为了更加清楚的表示出评价等级,我们采用 Delphi 方法,对上述等级进行评分,记为 $D = (d_1, d_2, \dots, d_m)$, 即 v_i 等级的分值为 d_i , 并假定 $D = (d_1, d_2, d_3, d_4) = (10, 7, 5, 2)$ 。

2.3 计算灰色模糊评判矩阵 $R^{[9,10]}$

设有 r 位专家参加评价。我们把第 s 位专家对第 i 个指标的评价量样本记为 l_{si} , 将 r 位专家对所评价的图书馆网站的评价数据记为样本矩阵

$$\begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{r1} & l_{r2} & \cdots & l_{rn} \end{bmatrix}, \text{用灰色统计法由确定的各评价}$$

标准函数(灰数的白化函数),求出 l_{si} 属于第 j 类评价标准的权 $f_j(l_{si})$, 据此求出评判矩阵的灰色统计

数 n_{ij} 和总灰色统计数 n_i , 其中: $n_{ij} = \sum_{s=1}^r f_j(l_{si})$,

$n_i = \sum_{j=1}^m n_{ij}$ 。综合 r 位专家对第 i 个评价因素主张第

j 种评价标准的灰色权值为 $r_{ij} = \frac{n_{ij}}{n_i}$, 则由 r_{ij} 构成的

矩阵为 $R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{bmatrix}$ 。为了方便统计计

算,这里选择的是常用的三种白化函数。对于上类形态灰数 $\otimes \in [l_1, +\infty)$ 、中类形态灰数 $\otimes \in [0, l_1, 2l_1]$ 与下类形态灰数 $\otimes \in (0, l_1, l_2)$, 其白化函数分别定义如下:

$$(1) f_1(l_{si}) = \begin{cases} l_{si}/l_1, & l_{si} \in [0, l_1] \\ 1, & l_{si} \in [l_1, +\infty) \\ 0, & l_{si} \in (-\infty, 0); \end{cases}$$

$$(2) f_2(l_{si}) = \begin{cases} l_{si}/l_1, & l_{si} \in [0, l_1] \\ 2 - l_{si}/l_1, & l_{si} \in [l_1, 2l_1] \\ 0, & l_{si} \notin (0, 2l_1]; \end{cases}$$

$$(3) f_3(l_{si}) = \begin{cases} 1, & l_{si} \in [0, l_1] \\ \frac{l_2 - l_{si}}{l_2 - l_1}, & l_{si} \in [l_1, l_2] \\ 0, & l_{si} \notin (0, l_2]. \end{cases}$$

2.4 确定指标的权重集合 W

用 $W = (w_1, w_2, \dots, w_n)$ 表示且 $\sum_{i=1}^n w_i = 1$, 其

中 w_i 表示第 i 项指标的权重。在这里,用粗糙集理论来确定各个指标的权重。下面为利用粗糙集计算指标权重的具体步骤^[11-13]:

2.4.1 首先建立关系数据模型

这里,将每个评价指标视为条件属性,则条件属性集合为 $C = \{c_1, c_2, \dots, c_n\}$; 将各个专家根据每个评价指标打分后的最后综合得分 y 视为决策属性,则决策属性集合为 $D = \{y\}$ 。第 k 个待评价网站的各个指标值和最后综合得分视为某个知识系统的一条信息,则可定义 $u_k = (c_{1k}, c_{2k}, \dots, c_{nk}, y_k)$, 从而论域为 $U = (u_1, u_2, \dots, u_m)$, 则由 u_k 构成的二维信息表就是关于要评价的关系数据模型。

2.4.2 从最低一层指标开始

建立其对父指标的知识表达系统(KRS)。

2.4.3 对 KRS 进行数值化处理

数值化并删去重复行。

2.4.4 计算

知识 R_D 对知识 R_C 的依赖程度为 $\gamma_{R_C}(R_D)$, 即计算评价指标集合 C 对决策属性指标 y 的依赖程度。

2.4.5 计算依赖程度

对每个评价指标 c_i , 计算知识 R_D 对知识 $R_{C-\{c_i\}}$ 的依赖程度 $\gamma_{R_{C-\{c_i\}}}(R_D)$, $i = 1, 2, \dots, n$

2.4.6 计算第 i 种评价指标的重要性

$$\sigma_D(c_i) = \gamma_{R_C}(R_D) - \gamma_{R_{C-\{c_i\}}}(R_D); \\ i = 1, 2, \dots, n.$$

2.4.7 归一化处理

可计算出第 i 种评价指标的重要性。

$$w_i = \frac{\sigma_D(c_i)}{\sum_{i=1}^n \sigma_D(c_i)}, i = 1, 2, \dots, n.$$

2.5 计算模糊评价矩阵

$$B = W \cdot R.$$

2.6 综合评价

有时根据最大隶属度原则并不能真实反映该图书馆网站处于什么样的水平。这里利用前面确定的相应等级分值来计算该图书馆网站所处的等级与分值,即 $F = B \cdot D^T$ 。通过此分值,也可将该图

图书馆网站与其他同类网站相比较。

从上可以看出,这种方法简单易行,有很大的灵活性,具有较强的可信度,弥补了其它方法的不足。

3 案例分析

站的建设水平进行评价。表1为各个图书馆网站指标数据和专家总体评分数值。为了简明说明问题,表3为相关专家对表1中第一个图书馆网站各指标的评分值。需要说明的是,图书馆网站指标数据的计算方法可参见相关计算方法,这里给出的数据已经过相应处理。

这里聘请7位相关专家对18个相关图书馆网

表1 指标值及专家评分表

论域 企业	指标值							专家总体评分				
	c_1	c_2	c_3	c_5	E_1	E_2	E_3	E_4	E_5	E_6	E_7	平均分 f
1	0.278 9	1.739 0	0.569 2	0.817 0	7	8	5	9	6	7	8	7.142 9
2	0.770 4	0.351 1	0.997 5	0.976 4	6	7	7	8	9	7	7	7.285 7
3	1.007 8	0.643 3	0.757 0	1.021 9	7	8	7	8	9	8	7	7.714 3
4	1.369 5	0.989 6	0.812 3	1.200 7	8	9	7	9	7	8	9	8.142 9
5	1.082 4	1.597 4	1.117 4	0.960 6	9	8	8	9	10	7	9	8.571 4
6	1.491 0	0.679 6	1.746 5	1.023 4	8	8	7	8	9	8	9	8.142 9
7	0.346 8	1.023 1	0.963 4	0.564 2	7	9	6	5	8	8	8	7.285 7
8	0.632 5	0.874 5	0.221 7	1.320 1	6	7	7	6	8	7	8	7.285 7
9	1.412 1	0.742 3	0.657 8	0.854 1	9	9	8	7	6	8	7	7.714 3
10	1.541 4	1.321 1	0.679 8	1.021 3	9	7	6	9	8	7	8	7.714 3
11	0.758 8	0.413 4	0.756 3	1.042 3	8	9	7	7	8	7	7	7.571 4
12	0.312 4	0.869 5	0.675 2	0.845 2	7	5	8	7	6	7	7	6.714 3
13	1.231 4	0.765 3	1.342 6	1.023 1	8	7	7	8	8	7	9	7.714 3
14	0.865 4	0.421 5	1.233 2	0.634 2	5	6	5	7	6	8	7	6.285 7
15	0.561 8	0.756 2	0.557 4	1.554 1	5	4	6	5	7	6	5	5.428 6
16	0.428 5	0.521 4	0.631 2	0.742 9	3	5	4	6	6	7	6	5.285 7
17	0.642 3	1.023 1	0.854 6	1.246 3	6	7	7	8	7	6	5	6.571 4
18	0.875 8	0.426 9	1.314 5	0.756 4	7	8	7	6	8	7	6	7.285 7

将表1中各属性值进行离散化处理:

网站内容:(1—>1.5;2—1~1.5;3—0.5~1;4—<0.5);网站设计:(1—>1.5;2—1~1.5;3—0.5~1;4—<0.5)。

操作使用:(1—>1.5;2—1~1.5;3—0.5~1;4—<0.5);开放程度:(1—>1.5;2—1~1.5;3—0.5~1;4—<0.5)。

平均得分:(1—<7.5;2—7.5~8;3—8~8.5;4—>8.5)。

根据以上划分,可得到知识表达系统简化表,见表2。

表 2 知识表达系统简化表表

论域	c_1	c_2	c_3	c_4	f	论域	c_1	c_2	c_3	c_4	f
1	4	1	3	3	1	10	1	2	3	2	2
2	3	4	3	3	1	11	3	1	3	2	2
3	2	3	3	2	2	12	4	3	3	3	1
4	2	3	3	2	3	13	2	3	2	2	2
5	2	1	2	3	4	14	3	4	2	3	1
6	2	3	1	2	3	15	3	3	3	1	1
7	4	2	3	3	1	16	4	3	3	3	1
8	3	3	4	2	1	17	3	2	3	2	1
9	2	3	3	3	2	18	3	4	2	3	1

表 3 某企业的详细评价表

	c_1	c_2	c_3	c_4
E_1	4	8	6	7
E_2	6	9	7	9
E_3	5	8	6	8
E_4	5	8	7	7
E_5	3	8	7	8
E_6	4	9	6	7
E_7	3	7	6	8

由表 2 可以得到:

$$\frac{U}{ind(c_2, c_3, c_4)} = \{(1), (2), (3, 4), (5), (6), (7), (8), (9, 12, 16), (10, 17), (11), (13), (14, 18), (15)\};$$

$$\frac{U}{ind(c_1, c_3, c_4)} = \{(1, 7, 12, 16), (2), (3, 4), (5), (6), (8), (9), (10), (11, 17), (13), (14, 18), (15)\};$$

$$\frac{U}{ind(c_1, c_2, c_4)} = \{(1), (2, 14, 18), (3, 4, 6, 13), (5), (7), (8), (9), (10), (11), (12, 16), (15), (17)\};$$

$$\frac{U}{ind(c_1, c_2, c_3)} = \{(1), (2), (3, 4, 9), (5), (6), (7), (8), (9), (10), (11), (12, 16), (13), (14, 18), (15), (17)\};$$

$$\frac{U}{C} = \{(1), (2), (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (13), (14), (15), (16), (17), (18)\};$$

$$\frac{U}{f} = \{(1, 2, 7, 8, 12, 14, 15, 16, 17, 18), (3, 9,$$

$10, 11, 13), (4, 6), (5)\}$ 。

故: $pos(c_2, c_3, c_4)(f) = \{1, 2, 5, 6, 7, 8, 11, 13, 14, 15, 18\} = 11$

$pos(c_1, c_3, c_4)(f) = \{1, 2, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18\} = 14;$

$pos(c_1, c_2, c_4)(f) = \{1, 2, 5, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18\} = 14;$

$pos(c_1, c_2, c_3)(f) = \{1, 2, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18\} = 15;$

$posC(f) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18\} = 18;$

故可以求得: $\gamma_C(d) = 18/18 = 1$, $\gamma_{|c_2, c_3, c_4|}(f) = \frac{11}{18} = 0.6111$, $\gamma_{|c_1, c_3, c_4|}(f) = \frac{14}{18} = 0.7778$,

$\gamma_{|c_1, c_2, c_4|}(f) = \frac{14}{18} = 0.7778$, $\gamma_{|c_1, c_2, c_3, c_5|}(f) = \frac{15}{18} = 0.8333$,

故可以进一步求得: $\gamma_C(d) - \gamma_{|c_2, c_3, c_4|}(f) = 0.3889$, $\gamma_C(d) - \gamma_{|c_1, c_3, c_4|}(f) = 0.2222$,

$\gamma_C(d) - \gamma_{|c_1, c_2, c_4|}(f) = 0.2222$, $\gamma_C(d) - \gamma_{|c_1, c_2, c_3|}(f) = 0.1667$ 。

经归一化处理, 即可求得相关指标权重为: $W = (w_1, w_2, w_3, w_4) = (0.3567, 0.2339, 0.2339, 0.1755)$ 。

根据我们前面的评价等级集合, 相应的灰数及其白化函数如图 2 所示。

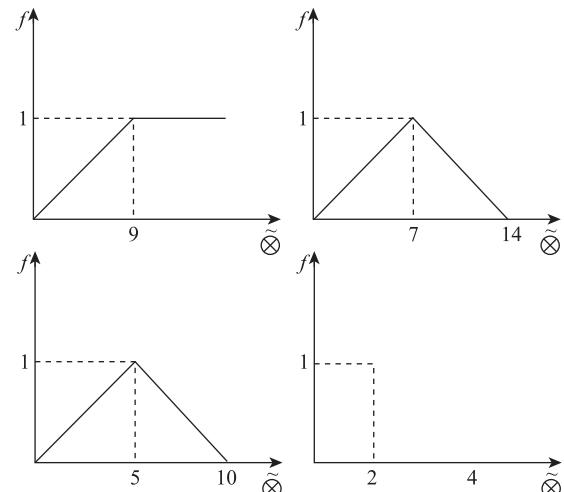


图 2 白化函数

据图 2 与表 3, 可求出评判矩阵的灰色统计数 n_{ij} 和总灰色统计数 n_i , 并进而得到灰色模糊测评矩阵为:

$$R = \begin{bmatrix} 0.2344 & 0.3014 & 0.3938 & 0.0703 \\ 0.4282 & 0.3960 & 0.1758 & 0 \\ 0.3043 & 0.3913 & 0.3043 & 0 \\ 0.3875 & 0.4059 & 0.2066 & 0 \end{bmatrix},$$

故模糊评价矩阵为 $B = W \cdot R = (0.3230, 0.3629, 0.2891, 0.0251)$, 则综合评价结果为 $Z = B \cdot D^T = 7.2654$ 。按照设定的评价等级, 该图书馆网站建设的质量水平属于良好状态。此结果与原先得到的平均数相比, 可以看出分值有所提高。

4 结语

网站是图书馆对外交流的窗口, 是图书馆新的形象标志。由于现实世界中亦此亦彼的中间过渡现象大量存在, 与以往对图书馆网站评价的研究相比, 本文利用一种综合评价方法研究的结果更具有现实意义和可操作性, 使对图书馆构成要素的整体质量的评价更加切实可行, 为客观、科学、合理的评价图书馆网站建设质量提供了有效而实用的方法。但需要指出的是, 不同的图书馆由于处于再网络化建设发展的不同阶段, 因而评价指标也需要区别对待。

参 考 文 献

- 王尊新. 高校图书馆网站评价指标体系研究. 现代图书情报技术, 2005;3:60—62
- 伍晓星. 图书馆网站评价指标浅析. 现代图书情报技术, 2004;4:50—53
- 陆志强, 贾志宏. 如何评价图书馆网站. 图书馆理论与实践, 2002;(4):31—32
- 朱兰芳. 谈谈图书馆的网站建设. 现代情报, 2004;1:146—148
- 王晓丽. 基于专业类网站评价体系研究. 情报杂志, 2002;10:31—33
- 刑中芳. 图书馆综合评价体系研究. 华南理工大学学报(社会科学版), 1999;1(1):141—144
- 陈雅 郑建明. 网站评价指标体系研究. 中国图书馆学报, 2002;28(5):57—60
- 王道林. 网站评价的数学模型. 泰山学院学报, 2004;26(3):6—9
- 徐维祥, 张全寿. 一种基于灰色理论和模糊数学的综合集成算法. 系统工程理论与实践, 2001;4:114—119
- 徐维祥, 张全寿. 信息系统项目评价 DHGF 集成法. 计算机工程与应用, 2000;5:60—62
- 张喜成, 汪江洪. 粗糙集综合评价法在公交服务水平评价中的应用. 统计与决策, 2006;2:46—48
- 刘佳, 魏彩巧, 王西彬. 粗糙集综合评价法在绿色制造评价中的应用研究. 重庆环境科学, 2003;25(12):64—67
- 钟波, 肖智, 周家启. 组合预测中基于粗糙集理论的权重确定方法. 重庆大学学报, 2002;25(7):127—130

Research on Comprehensive Evaluation Method for Library Website Based on Rough Set Theory

WANG Jian-long

(Business School, Hohai University, Changzhou 213022, P. R. China.)

[Abstract] Library website is the window for the modern library to show its service. So it is important to evaluate the library website to improve its construction quality. The factors of library website are first analyse, then established evaluation index systems. It uses grey statistic method to establish fuzzy evaluation matrix and obtains the weight of every index by rough set theory, then appraises library website in terms of fuzzy comprehensive evaluation. It gives a beneficial academic reference to promote the construction quality of library website.

[Key words] library Website evaluation indicators fuzzy comprehensive evaluation grey statistic rough set