

应用模糊分析层次法可靠评测 Web 资源质量

朱 焱

(西南交通大学信息科学技术学院 成都 610031)

摘 要 建立了一个 Web 资源质量评测系统。该系统将 AHP 算法与模糊逻辑相结合,以解决 Web 资源质量评测中模糊和不确定问题。另外,模糊 AHP 算法中集成了评测置信度和风险指标,以增强评测精度和提高 Web 资源质量评测的可靠性。

关键词 Web 资源质量评测,模糊 AHP 方法,置信区间,风险系数

中图法分类号 TP393 **文献标识码** A

Reliably Evaluating Web Source Quality with the Fuzzy AHP

ZHU Yan

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)

Abstract The success of Web-based information systems greatly depends on Web quality. To assess and screen high quality Web information, a system for evaluating Web source quality was constructed. By combining AHP approach and fuzzy logic, the fuzzy and uncertainty in Web quality assessment were treated. The confidence level of evaluators and the risk index of the assessment were integrated in Fuzzy AHP for enhancing the precision of the measurement and improving the reliability of the Web source quality evaluation.

Keywords Web source quality evaluation, Fuzzy AHP method, Confidence interval, Risk index

1 商务智能与 Web 资源质量挑战

以互联网为基础的商务智能的最终目标是为相关组织制定合理的决策提供快速适时的帮助。PricewaterhouseCoopers(普华永道)早在 1998 年出版的调查报告中就列举了如下数据:446 家快速发展企业的 CEO 表示,Web 信息已经成为他们企业信息的主要来源。其中,82%的高技术企业使用 Web 信息作为“竞争者信息的来源”,73%的企业利用 Web 作为“统计或数据资源”。68%的企业使用 Web 信息“达到了销售上的领先地位”。

Web 资源作为新兴的信息(数据)源既提供了巨大好处也带来了不可忽视的风险。Meta Group 的一项调查显示,41%的数据仓库项目无法成功,主要原因是糟糕的数据源质量导致了错误的决策,这种现象被称为“垃圾进,垃圾出”。不同于传统的数据源(数据文件,数据库等),基于 Web 的信息源具有动态性、自主性、用户类型多样等突出特点,这些特点增加了 Web 资源质量问题的严峻性。

因此,要保证基于 Web 的各类应用,特别是商务智能应用的成功,关键是要评测并遴选出高质量的 Web 资源,这既是重要的技术问题又是应用系统成功的前提。

1.1 Web 资源质量管理面临的四大挑战

(1)Web 资源的高动态性。Web 信息内容更新频繁;网站和网页的结构与版面经常修改以满足使用者的各种需求;大量新的信息和网页被添加,而一些已存在的信息和网页又会

从互联网上消失。要利用基于 Web 的信息辅助市场分析和决策制定,资源的稳定性是质量的一个重要方面。

(2)Web 信息的高自主性。Web 资源没有一个全球性的质量管理机构,也没有统一的质量检查标准,从而无法像传统的出版物(如书籍、报纸、杂志等)和传统的数据资源(如数据库等)那样被仔细地评价、检查和校正。结果是,Web 信息中包含着许多错误的、不完整的、不一致的数据或者模糊的事实,甚至由于不良的表达(例如:缺少单位或时间标记),正确的数据也会变得难以使用。不同 Web 资源之间的数据集成也变得十分困难。

(3)Web 环境的模糊性。某些质量特性很难用精确的数值来度量。例如,可运行性指出某个网站处于活跃状态,主页面与其所有的链接页面间的超链是通达的,用户点击超链后,页面的读取时间在用户接受的范围之内。然而,对点击超链到页面出现所耗时间的容忍度是因人而异的,因此,不同的人评定的 Web 资源可运行性质量值有一定的差异,用唯一“确定的”数值度量是不符合实际,也是不灵活的。

(4)Web 信息应用的多样性。Web 资源的用户群体数量和类型急剧增加,各类用户群体在技术背景、Web 信息的使用目的和使用方式上都有很大区别^[1]。例如:使用 Web 股票证券信息的用户对信息在可信度和时效性方面的质量要求很高。零售企业利用 Web 收集某一地区的人口统计信息,以便有针对性地制定营销方案,这些企业则对信息的可信度、完整性和集成度有很高的要求。因此,Web 信息应用的多样性对

Web 资源的质量提出了多方面的要求。

1.2 Web 资源质量评测研究分析与本文的贡献

Web 资源质量评测是崭新的研究领域,当前还处在起步阶段,没有成熟的解决方法。引入其他领域的技术解决 Web 质量问题是目前行之有效的措施,例如将运筹学中多属性决策方法(MADM)应用于 Web 资源质量评测。

Web 资源质量度量是一个复杂的问题,各种因素相互影响,模糊性很明显,然而基于精确的“0/1”逻辑的计算机无法如人工评测那样较好地判断和处理不确定性问题,针对这类问题,模糊数学与模糊逻辑是公认的解决方案。

Web 资源质量评测环境复杂,评测者对质量指标的理解和对 Web 资源质量的要求有较大差别,这些差别直接影响评测性能。换言之,如果评测者有丰富的 Web 资源质量知识,能很好地理解 Web 资源质量评测模型、评测指标和评测过程,而且评测模型稳定,满足实际要求,则评测结果的可信度就高,决策失败的风险就小。

经典的信息资源质量管理工作包括麻省理工的 Total Data Quality Management(TDQM)项目^[2],Naumann 等人的质量驱动的数据整合系统^[3],评价物品拍卖网站性能的 WebQual 系统^[4],以及国内研究者应用模糊综合评价法对课堂教学质量进行定量分析^[5]。这些研究工作中的大部分针对传统的数据资源,而没有考虑 Web 信息源这种特殊资源对质量管理的挑战。

近年,一些学者认识到 Web 资源质量的特殊性,针对其质量问题进行了研究。主要成果有 Olsina 等的网页质量评价模型^[6,7]和笔者的 Web 资源质量初级评测系统^[8,9]。然而这些工作的重点放在建立 Web 资源评测体系,设计评测准则,研究评测算法,没有考虑 Web 资源评测过程中的模糊性,以及评测者的主观赋值引入的评测可信度和风险问题。

本文研究了经典 MADM 方法中的层次分析(AHP)算法^[10],并将模糊逻辑与 AHP 相结合,对 Web 资源质量展开模糊评测。在模糊 AHP(FAHP)算法的应用过程中,考虑了评测者的置信程度和风险影响,设置了可信度和风险度阈值,以改进评测的性能。本文的主要贡献为:

- 研究模糊逻辑(Fuzzy Logic)与多属性决策支持方法的有机结合,解决 Web 资源质量量化与评测的不确定性。
- 研究将评测者的信任度和评测失败的风险量化,集成到质量评测过程中,优化评测判断,提高评测的可信度。

2 Web 资源质量评测系统

Web 资源质量评测系统由 6 部分组成:

(1) 设定评测目标:根据实际应用的质量要求设定质量评测总体目标。

(2) 建立评测模型:通过分析 Web 资源/信息的静态与动态质量特征,建立多维多层次的 Web 资源质量评测模型,设计综合的质量评测准则。

在建模过程中,将 Web 资源视为特殊的软件产品,它具有功能性、使用性、可维护性等传统软件产品的共性,以及传统软件产品不具备的动态性、自主性等特点。为度量 Web 资源的特点,笔者提出了 Web 资源/信息质量模型,WebQM^[8]。该模型在 Web 源质量 (WebSQ), Web 信息质量(WebIQ)和应用相关质量(WebAQ)3 个主维和 12 个子维上量化 Web 资

源/信息质量。通过综合考察 Web 资源评测目标和待评测对象,建立单目标多层次的 Web 资源/信息质量评测模型,如图 1 所示。

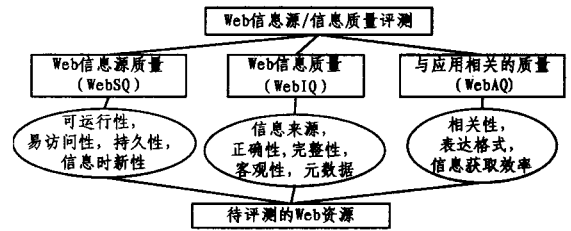


图 1 Web 资源/信息质量评测模型-WebQM

(3) 确定待评测的 Web 资源:根据应用域预选一组相关的 Web 资源。预选工具可以利用信息检索引擎,如 Google 等;

(4) 质量评测:应用算法根据质量评测准则评测待选的 Web 资源,得到资源在每项准则下的质量评测分量值。质量分量值可以采用不同方法获得,FAHP 方法中采用对评测参数两两比较的方法。该方法首先采用自然语言词汇定义 Web 资源质量,应用模糊集理论将自然语言定义的模糊系统量化,并且研究设计结合模糊逻辑的评测算法使计算机自动完成评测集成运算,得到更符合实际的结果。

(5) 基于置信度和风险系数的质量评测值集成:在 FAHP 算法中,采用置信度和风险系数过滤、集成各项质量值,得到每个 Web 资源的质量综合评测值。

3 结合置信度和风险系数的模糊层次分析算法

模糊层次分析(FAHP)方法的两个关键是:准则的权重值通过成对的重要性比较获得;待评测的 Web 资源依据高一层的准则成对地进行相对性能比较。然后根据准则的权重和 Web 资源的质量分量进行性能值综合,最后按综合性能进行排序,确定高质量的 Web 资源。

设在 Web 资源质量评测空间中有 M 个待评测对象, N 个质量准则,应用结合了置信度阈值和风险系数的 FAHP 进行评测的步骤如下:

(1) 根据 Saaty 建立的比较量级确定性偏好的度量词汇级别如表 1 所列,以及确定三角型模糊隶属成员函数如图 2 所示。

表 1 质量性能语言度量及其模糊取值

语言度量定义	模糊取值
同等重要	$\tilde{1}(1,1,3)$
适度重要	$\tilde{3}(1,3,5)$
中度重要	$\tilde{5}(3,5,7)$
高度重要	$\tilde{7}(5,7,9)$
极其重要	$\tilde{9}(7,9,9)$
	$\tilde{1/9}(0.11,0.11,0.14)$
	$\tilde{1/2}(0.25,0.5,0.5), \tilde{1/4}(0.17,$
	$\tilde{2}(2,2,4), \tilde{4}(2,4,6),$
中间值	$0.25,0.5), \tilde{1/6}(0.13,0.17,0.$
	$\tilde{6}(4,6,8), \tilde{8}(6,8,8)$
	$\tilde{25}),$
	$\tilde{1/8}(0.13,0.13,0.17)$

(2) 建立质量准则的重要性成对比较矩阵 \tilde{W} :

$$\tilde{W} = [\tilde{w}_{ij}]_{N \times N}, \text{ 其中: } i=1, \dots, N, j=1, \dots, N; \quad (1)$$

其中, \tilde{w}_{ij} 是第 i 个准则与第 j 个准则的重要性模糊比较值;

\tilde{w}_{ij} 为第 j 个准则与第 i 准则的重要性模糊比较值, 为 \tilde{w}_{ij} 值的倒数。如果某个准则与自身相比较, 则值为“同等重要”, 即 1。

(3) 建立每个待评测 Web 资源在各个准则下的性能比较矩阵 \tilde{X} :

$$\tilde{X}_k = [\tilde{x}_{ij}]_k, \text{ 其中: } i=1, \dots, M, j=1, \dots, M, k=1, \dots, N \quad (2)$$

\tilde{X}_k 是第 k 个准则下待评测项性能比较矩阵, \tilde{x}_{ij} 是第 i 个与第 j 个 Web 资源在第 k 个准则下的模糊比较值, \tilde{x}_{ji} 是第 j 个与第 i 个 Web 资源在第 k 个准则下的模糊比较值, 是 \tilde{x}_{ij} 值的倒数。如果某个准则与自身相比较, 则值为“同等重要”。

(4) 结合评测者置信度 α 确定更准确的模糊取值范围。

统计学中的置信度反映了评测者对评测的把握程度。如果置信度高, 则评测过程可信度高, 得到的结果更准确。置信度的设定必须从评测者的技术背景和对评测模型的理解出发。

本文在 FAHP 中结合评测者的置信度, 以提高评测的精度。以 \tilde{X}_k 为例, 用置信度阈值 α 水平截取性能比较矩阵中的每个模糊值 \tilde{x}_{ij} (图 3 中, $\alpha=0.5$), 获得区间值 $(x_{ij}^{\alpha}, x_{ij}^{\alpha})$, \tilde{X}_k 转换为 \tilde{X}_k^{α} 。

区间的左右边界值可按下列公式计算。

$$x_{ij}^{\alpha} = L_{ij} + \alpha(M_{ij} - L_{ij}), x_{ij}^{\alpha} = R_{ij} - \alpha(R_{ij} - M_{ij}) \quad (3)$$

其中 L_{ij}, M_{ij}, R_{ij} 分别是 \tilde{x}_{ij} 模糊值的 3 个分量。 $0 < \alpha \leq 1$, α 越大, 可信度越高。

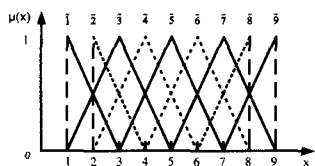


图 2 三角形模糊隶属成员函数

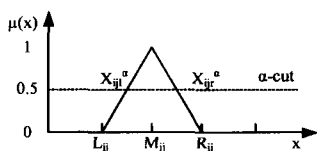


图 3 置信度阈值对模糊评测值进行筛选

(5) 质量评测失败的风险指标 β 取决于评测模型的稳定性和与应用要求的符合程度。本文的 WebQM 模型通过经典的多准则决策支持方法验证是稳定和符合实际的。因此, 使用该模型对 Web 资源质量评测可以设置一个较低的风险系数 β 。例如, 将 β 引入性能比较矩阵 \tilde{X}_k 得到 $\tilde{X}_{k,\beta}$, 新的分量计算公式如下:

$$x_{ij,\beta}^{\alpha} = \beta x_{ij}^{\alpha} + (1-\beta)x_{ij}^{\alpha} \quad (4)$$

其中, $x_{ij,\beta}^{\alpha}$ 代表第 i 个 Web 资源与第 j 个资源在第 k 个质量评测准则下性能比较值, 该比较值结合了可信度 α 和风险指数 β 。 $0 < \beta \leq 1$, β 越大, 风险越高。

(6) 采用计算本征向量的方法分别对 \tilde{W} 和 \tilde{X}_k^{α} 规范化, 得到评测准则归一化的权值向量 \tilde{V} 和所有待评测项在所有准则下的归一化性能值矩阵 \tilde{Y} 。 \tilde{V} 中的值即是各准则的权重, \tilde{Y} 中的值是待评测的 Web 资源在各个准则下的质量分量值。获得 \tilde{V}, \tilde{Y} 常用的计算方法是:

(a) 输入矩阵的列的各单元除以该列的总和得到规范化的值;

(b) 每一行的总和构成新的一列;

(c) 列的各单元除以该列的总和得到规范化的值。

(7) 计算待评测 Web 资源综合的质量值 R , 并排序。

$$r_i = \sum_k \tilde{v}_k \times \tilde{y}_{ik} \quad i=1, \dots, M, k=1, \dots, N \quad (5)$$

4 Web 资源质量评测实例

本文选择 4 个真实的 Web 资源(分别以 A, B, C, D 标识)进行评测。为了在有限的篇幅里清晰地展示评测过程, 从 WebQM 3 个维的 12 个质量评测准则中, 每维各选取 2 个准则进行分析, 被选择出来的质量准则包括可运行性、可访问性、正确性、完整性、相关性和表达格式。

根据表 1 的规则对准则进行成对比较, 例如可运行性比可访问性适度重要, 值为 3, 而可访问性比较可运行性, 得 1/3。同理得到质量准则模糊权值比较矩阵 \tilde{W} , 如表 2 所列。

表 2 质量准则模糊权值比较矩阵 \tilde{W}

	可运行性	可访问性	正确性	完整性	相关性	表达格式
可运行性	1	3	1/2	2	1/5	2
可访问性	1/3	1	1/5	1/2	1/9	1/2
正确性	2	5	1	4	1/3	4
完整性	1/2	2	1/4	1	1/7	1
相关性	5	9	3	7	1	7
表达格式	1/2	2	1/4	1	1/7	1

同理, 将待评测的 Web 资源在 6 个质量评测准则下进行分项质量评测, 产生 6 个质量性能比较矩阵。其中, 相关性准则下 4 个待评测资源的质量分量比较矩阵如表 3 所列。对建立的 7 个 FAHP 比较矩阵分别进行一致性检验。如果各矩阵一致性满足 Saaty 的定义 (≤ 0.1), 矩阵中质量分项值是合理的。如果不满足, 比较矩阵需要重新建立。

表 3 4 个待评测资源的相关性准则质量分量比较矩阵

相关性比较	A	B	C	D
A	1	1/6	2	1/4
B	6	1	8	2
C	1/2	1/8	1	1/6
D	4	1/2	6	1

针对权值矩阵和 6 个质量性能比较矩阵, 应用结合了 α, β 阈值的 FAHP 算法步骤(4)~(7)进行质量值集成, 计算出 4 个待评测 Web 资源质量综合分值为: A: 0.2095, B: 0.3797, C: 0.201, D: 0.2183。由此, 4 项 Web 资源质量由高到低排序为: B, D, A, C。其中, B 是质量最佳的 Web 资源。

结束语 本文应用模糊逻辑与多属性决策相结合的方法解决 Web 资源质量模糊评测的问题, 研究中应用 α -cut 方法确定置信区间, 应用风险系数 β 对 Web 资源的评测值进行过滤筛选, 在评测过程中集成评测者的信任度和评测风险指标, 从而提高评测结果的准确性和可信度。

下一步的工作将研究改进模糊评测过程中一致性的判定问题, 以及确定敏感的质量指标, 研究这些指标对质量序列的影响。

参考文献

- [1] 中国互联网信息中心. 第二十一次中国互联网络发展状况统计报告 [OL]. <http://www.cnnic.cn/html/Dir/2008/01/17/4966.htm>

(下转第 267 页)

用不同的网络结构做预测。

组合模型中聚类数设定为 5, BP 网络训练参数与 ES-BP 中各训练参数相同。网络输出为 0 时表示工人不患病, 为其它数值时表示工人患病年龄的预测值。

4.3 组合模型建模与单独 BP 建模分析比较

对 300 组工人矽肺队列数据分别用 2 种组合模型对工人患病年龄做预测, 从预测精度和模型复杂度角度分别对两种组合模型做评估, 将组合建模预测与单独使用 BP 建模预测做比较。

图 3 给出采用 ES-BP 组合模型与单独使用 BP 网络对 50 组预测样本数据中工人患病年龄的预测情况。由图中看出 ES-BP 组合模型预测值比 BP 预测更接近工人实际患病年龄。

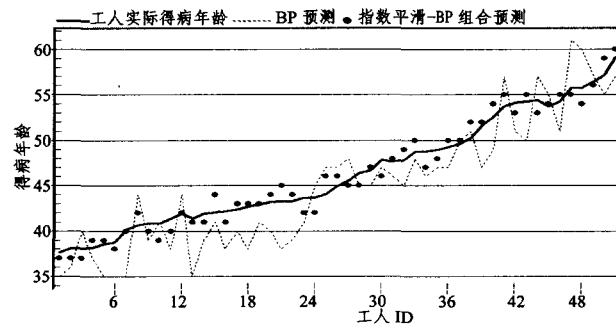


图 3 ES-BP 预测与 BP 预测

表 1 将 FCM-BP 组合建模与 BP 单独建模训练开销做比较。由表中数据分析得知, 由于 FCM-BP 组合模型做预测时先对数据进行类别判断, 使每一小类数据集具有较大相似性, 因此对每小类做网络训练时数据集数目相对减少, BP 网络训练次数、运行时间和模型复杂度都减小很多。

表 1 FCM-BP 组合建模执行效率分析

预测模型	BP	FCM 聚类-BP 组合	
训练次数	5000	类 1	350
		类 2	400
		类 3	370
		类 4	450
		类 5	475
运行时间/s	75.35	45.26	

表 2 显示了模型评价的结果。分别计算组合建模与 BP 单独建模预测的平均相对误差 (MRE)、均方误差 (MSE)、平

均绝对误差 (MAE) 和残差平方和 (SSE), 表明两种组合建模预测的误差都低于 BP 模型预测的误差。

表 2 组合建模预测与 BP 预测误差分析

患病年龄预测	BP	指数平滑-BP	FCM 聚类-BP
MRE	16.1732%	12.3684%	13.4763%
MSE	0.1576	0.1274	0.1463
MAE	0.1453	0.1358	0.1164
SSE	0.1475	0.1263	0.1183

结束语 本文基于粉尘接触量、吸烟量、接尘时间等影响矽肺的各致病因子, 研究了分别采用 ES-BP 和 FCM-BP 组合模型的矽肺预测方法, 并和单一采用 BP 建模预测进行了对比。经实验比较分析知: ES-BP 组合模型能结合矽肺时间队列数据的非线性特性, 综合各种致病因素, 从时间段上对工人患病年龄做出比 BP 更准确的预测; FCM-BP 组合模型对矽肺数据进行聚类, 显著降低了网络训练的复杂度和训练时间。

由此可知, 影响因子较多且非线性联系显著的时间序列数据, 更适合应用组合模型进行预测, 且两种组合模型预测精度均高于单独使用 BP 建模的预测精度, 有较好的应用价值。

参考文献

- [1] 陈卫红, 张小康, 王海椒, 等. 锡矿作业工人粉尘接触和队列死因分析[J]. 环境与职业医学, 2007(24): 9-12
- [2] Scalia A P C, Barreto S M, Siqueira A L, et al. Continued Exposure to Silica After Diagnosis of Silicosis in Brazilian Gold Miners [J]. American Journal of Industrial Medicine, 2006, 49: 811-818
- [3] 甘健胜, 陈国龙. 线性组合预测模型及其应用[J]. 计算机科学, 2006, 33(9): 191-194
- [4] Chen W, Eva H, Chen J Q. Risk of Silicosis in Cohorts of Chinese Tin and Tungsten Miners, and Pottery Workers (I): An Epidemiological Study [J]. American Journal of Industrial Medicine, 2005, 48: 1-9
- [5] 何书元. 应用时间序列分析[M]. 北京: 北京大学出版社, 2004
- [6] 高新波. 模糊聚类分析及其应用[M]. 西安: 西安电子科技大学出版社, 2004
- [7] 邓赵红, 王士同. 鲁棒性的模糊聚类神经网络[J]. 软件学报, 2005, 16(8): 1415-1422
- [8] 杨奎河, 王宝树, 赵玲玲. 基于神经网络的预测模型中输入变量的选择[J]. 计算机科学, 2003, 30(8): 139-140, 143

(上接第 223 页)

- [2] Wang R. A Product Perspective on Total Data Quality Management[J]. Communications of the ACM, 1998, 41(2): 58-65
- [3] Naumann F. From Databases to Information Systems-Information Quality Makes the Difference// Proc. of the International Conference on Information Quality (IQ). USA, 2001
- [4] Barnes S, Vidgen R. Assessing the Quality of Auction Web Sites//Proc. of the 34th Annual Hawaii International Conference on System Sciences. USA, 2001
- [5] 洪月华. 基于模糊综合评价的课堂教学质量数据挖掘[J]. 计算机科学, 2008, 35(2): 154-156, 170
- [6] Olsina L, Rossi G. Measuring Web Application Quality with Web-

- QEM[J]. IEEE MultiMedia, 2002, 9(4): 20-29
- [7] Mich L, Franch M, Gaio L. Evaluating and Designing Web Site Quality[J]. IEEE MultiMedia, 2003, 10(1): 34-43
- [8] 朱焱, 唐慧佳, 马永强. 基于 ISO/IEC9126 的 Web 资源质量评测系统[J]. 西南交通大学学报, 2008, 43(2): 253-257
- [9] Zhu Y. Group Assessment of Web Source/Information Quality Based on WebQM and Fuzzy Logic//G. Wang, et al., eds. Lecture Notes in Artificial Intelligence, vol. 5009. Germany: Springer Verlag, 2008: 660
- [10] Saaty T. The Analytic Hierarchy Process. McGraw-Hill, Inc., 1980