

文章编号: 1001-9081(2014) S1-0287-04

# 基于改进层次分析的数据质量综合评估

严浩<sup>1,2\*</sup>, 袁杭萍<sup>1</sup>, 刁兴春<sup>2</sup>, 周星<sup>2</sup>

(1. 解放军理工大学 指挥信息系统学院 南京 210007; 2. 总参第六十三研究所 南京 210007)

(\* 通信作者电子邮箱 yanhao0114@yeah.net)

**摘要:** 针对数据质量综合评估, 采用二次变权的思想对传统层次分析(AHP)法进行改造, 通过使用局部状态指数变权算法对影响因子的实测值进行激励或者惩罚, 再将影响因子常数向量与激励或者惩罚过后的影响因子实测值变权向量进行合成并二次变权, 从而得到局部变权向量, 并将该局部变权向量作为影响因子的最终权重向量。通过二次变权, 既反映了专家对影响因子相互间重要程度的经验值, 又体现了最终决策者对影响因子实测值不同状态所想体现的激励或惩罚的调节措施, 使得评估的结果更加贴近实际, 可信度更高。

**关键词:** 改进层次分析; 数据质量; 综合评估; 指数; 变权

**中图分类号:** TP391 **文献标志码:** A

## Comprehensive data quality assessment based on improved analytic hierarchy process

YAN Hao<sup>1,2\*</sup>, QIU Hangping<sup>1</sup>, DIAO Xingchun<sup>2</sup>, ZHOU Xing<sup>2</sup>

(1. College of Command Information Systems, PLA University of Science and Technology, Nanjing Jiangsu 210007, China;

2. The 63rd Research Institute of PLA General Staff Headquarters, Nanjing Jiangsu 210007, China)

**Abstract:** In order to assess data quality comprehensively, traditional Analytic Hierarchy Process (AHP) is modified using the concept of twice weight variation. The measured value of impact factor was stimulated or punished by the algorithm of local state exponential weight variation. Then the measured value of impact factor which had been stimulated or punished was integrated with the constant weight vector of impact factor. Local variable weight vector was generated through the second weight variation. And the local variable weight vector was determined as the final weight vector of impact factor. By the twice weight variation, the expert's preference on the importance of impact factor and the simulation and punishment on different state of the measured value of impact factor employed by the decision maker are put into consideration at the same time, and the result of assessment is more realistic and believable.

**Key words:** improved Analytic Hierarchy Process (AHP); data quality; comprehensive assessment; exponent; weight variation

## 0 引言

数据是信息系统的“血液”, 缺少数据的信息系统产生不了重大效益, 而数据质量差更会使信息系统沦为观赏系统。针对数据, 开展数据质量综合评估, 是提高信息系统数据质量的基础和必要前提<sup>[1]</sup>。通过对数据质量进行客观、有效、合理的评价, 有助于了解并掌握数据在质量方面的能力与不足, 明确数据的使用价值和水平, 从而为后续采取相应的处理方法提供科学依据。

本文在以正确性、准确性、一致性、完整性为质量维度建立数据质量评估指标体系并预先已知相关指标值和权重的基础上, 着重对数据质量综合评估方法进行了研究, 将突出对指标实测值“好”、“差”等不同状态的激励或惩罚的思想引入到数据质量综合评估工作中, 将传统层次分析法中固定的权重设置改进为基于指数的变权设置, 体现评估人员对不同指标状态的偏好程度, 从而使评估更加趋于科学、合理。

## 1 数据质量评估指标体系

目前对数据质量形成的共识是: 数据质量是指数据满足

用户需要或适用于特定应用的程度。它不仅仅考虑到与质量有关的各个方面, 而且强调数据满足用户的需求, 是数据产品满足指标、状况和要求能力的特征总和。

评估是指评估主体在根据确定的目标对评估客体进行系统分析的基础上, 按照一定的评估标准对评估客体相关属性做出判断的专业活动<sup>[2]</sup>。相应地, 数据质量评估则是以数据为评估客体的特定评估活动。数据质量评估是数据质量管理中的关键过程和基础部分, 根据数据质量评估的结果, 对数据进行修改和完善, 进一步分析数据质量问题产生的原因, 提出保障数据质量的改进措施。所以说要对数据质量实施有效的管理和控制, 数据质量评估将是关键的一个环节。

数据质量评估涉及评估指标体系和评估方法等两个方面的研究。

在数据质量评估指标体系方面, 国内外学者利用经验、实践和理论的方法建立了许多不同的评估指标体系<sup>[3-6]</sup>, 文献[7]提出了基于网络层次分析法的风险评估指标体系构建方法, 文献[8]提出了基于信息熵理论的评估指标体系构建方法。但是目前还没有形成一个能被广泛认同的评估指标体

收稿日期: 2013-10-17; 修回日期: 2013-12-14。 基金项目: 中国博士后科学基金特别资助项目(201003797)。

作者简介: 严浩(1977-), 男, 江苏南京人, 高级工程师, 博士研究生, 主要研究方向: 软件工程、数据库技术; 袁杭萍(1965-), 女, 浙江杭州人, 教授, 博士, 主要研究方向: 指控理论与技术; 刁兴春(1964-), 男, 江苏泰州人, 研究员, 硕士, 主要研究方向: 软件工程、数据库技术; 周星(1988-), 男, 四川广安人, 博士研究生, 主要研究方向: 信息系统、数据处理。

系,因为不同应用领域的不同数据质量需要不同的评估指标体系<sup>[9]</sup>。

在开展数据质量综合评估方法研究前,以正确性、准确性、一致性、完整性为质量维度建立数据质量评估指标体系,如图 1 所示,并以此作为数据质量综合评估研究的指标依据。

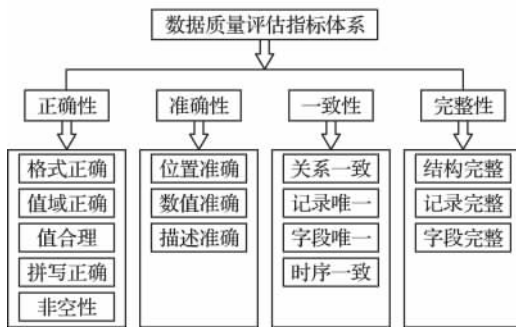


图 1 数据质量评估指标体系

## 2 改进层次分析法的计算模型

层次分析法<sup>[10-11]</sup>是常用的一种综合评估方法,可以实现定性问题向定量的转化,从而使复杂问题系统化、层次化。该方法已被广泛应用于金融、规划、企业管理和决策分析等领域,并取得了较好的成绩与效果。该方法的基本思想是根据待解决问题的性质和要达到的最终目标,将待解决问题有关的元素分解为不同的影响因子,然后依据影响因子之间的相互隶属及关联关系,形成影响因子层次间自上而下的逐层支配关系,最终形成一个有序的金字塔形的阶梯型树状结构。然后,依据一定的规则,业内专家或评委凭借其职业素养和经验,应用两两比较的方法,打分评定出该阶梯型树状结构中同一层次中各影响因子之间的相对重要性,形成判断矩阵,并对结果的准确性进行一致性检验评判。使用该方法时主要涉及两类数据,即影响因子的权重和影响因子的实测值。而权重确定工作,基本只对影响因子的权重进行赋值,最常用的就是直接对影响因子的权重进行常数型赋权重,而不考虑各影响因子的实测值所呈现的状态,也就是不考虑由于影响因子实测值的不同所应体现的不同。这种“一次确定”定“终身”的影响因子权重处理方法显然有它的局限性。由于常数只能反映影响因子之间的相对重要程度,无法对影响因子实测值进行判定和调节,而如果最终的决策者对影响因子的实测值有某种喜好的话,那么常规的层次分析法就起不到任何调节作用。所以,如果在使用层次分析法的时候,在考虑影响因子权重的同时,如果能将决策者对实测值的喜好也考虑一并进行调节和局部变权,实测值“好”的影响因子给予激励,实测值“差”的影响因子给予相应的惩罚,那无疑将大大增加最终评估结果的科学性和合理性。

局部变权方法如下:

$$M^0 = (M_1^0, M_2^0, \dots, M_n^0)$$

$M^0$  表示影响因子常数向量,则影响因子实测向量、影响因子实测变权向量和局部变权函数定义如下:

定义 1  $n$  阶局部变权向量。给定映射  $M: [0, 1]^n \rightarrow (0, 1)^n$  称向量  $M(X) = (M_1(X), M_2(X), \dots, M_n(X))$  为一个  $n$

阶局部变权向量,如果满足条件:  $\sum_{i=1}^n M_i(X) = 1$ 。

考虑到最终的决策者对于影响因子实测值的喜好有三种可能:第一种是只对高于某一取值的影响因子实测值进行奖励;第二种是只对低于某一取值的影响因子实测值进行惩罚;第三种是对高于某一取值的影响因子实测值进行奖励的同时,对低于某一取值的影响因子实测值进行惩罚。

对于第一种可能:对  $\forall i \in \{1, 2, \dots, n\}$  均存在  $\beta_i \in (0, 1)$ , 使得  $M_i(X)$  关于  $x_i$  在  $[\beta_i, 1]$  上单调增,体现奖励的思想。

对于第二种可能:对  $\forall i \in \{1, 2, \dots, n\}$  均存在  $\alpha_i \in (0, 1)$ , 使得  $M_i(X)$  关于  $x_i$  在  $[0, \alpha_i]$  上单调减,体现惩罚的思想。

对于第三种可能:对  $\forall i \in \{1, 2, \dots, n\}$  均存在  $\beta_i \in (0, 1)$ , 且  $\alpha_i \leq \beta_i$ , 使得  $M_i(X)$  关于  $x_i$  在  $[0, \alpha_i]$  上单调减,而在  $[\beta_i, 1]$  上单调增,体现奖励与惩罚并举的思想。

由于指数函数具有:当底大于 1 时,增长快,下降慢;底小于 1 时,增长慢,下降快的性质,所以选用指数函数构造局部变权函数。针对之前描述的最终的决策者对于影响因子实测值的喜好有三种可能,分别构造相应的局部状态指数变权函数。

第一种:只对高于某一取值的影响因子实测值进行奖励。其局部状态指数变权函数的形式如下所示:

$$S_i(x) = \begin{cases} \gamma, & 0 \leq x_i < \beta_i \\ b^{x_i} \times m, & \beta_i \leq x_i \leq 1 \end{cases}$$

其中:  $b > 1$ ,  $m$  为调节因子,以使  $b \times m = 1$ ;  $\beta_i$  是对影响因子  $i$  进行奖励的标准值,满足  $\beta_i \in (0, 1)$ ;  $\gamma \in (0, 1)$  是不进行惩罚和奖励时的局部状态值。

由变权函数的性质可知:在 0 到  $\beta_i$  范围内不变,在  $\beta_i \leq x_i \leq 1$  范围内单调递增。

第二种是只对低于某一取值的影响因子实测值进行惩罚。其局部状态指数变权函数的形式如下所示:

$$S_i(x) = \begin{cases} a^{x_i}, & 0 \leq x_i < \alpha_i \\ \gamma, & \alpha_i \leq x_i \leq 1 \end{cases}$$

其中  $a < 1$ ;  $\alpha_i$  是对影响因子  $i$  进行惩罚的标准值,满足  $\alpha_i \in (0, 1)$ ;  $\gamma \in (0, 1)$  是不进行惩罚和奖励时的局部状态值。

由变权函数的性质可知:在  $0 \leq x_i < \alpha_i$  范围内单调递减,在  $\alpha_i \leq x_i \leq 1$  范围内不变。

第三种是对高于某一取值的影响因子实测值进行奖励的同时,对低于某一取值的影响因子实测值进行惩罚。其局部状态指数变权函数的形式如下所示:

$$S_i(x) = \begin{cases} a^{x_i}, & 0 \leq x_i < \alpha_i \\ \gamma, & \alpha_i \leq x_i < \beta_i \\ b^{x_i} \times m, & \beta_i \leq x_i \leq 1 \end{cases}$$

其中:  $a < 1$ ,  $b > 1$ ;  $m$  为调节因子,以使  $b \times m = 1$ ;  $\alpha_i, \beta_i$  分别为对指标  $i$  进行惩罚和奖励的标准值,满足  $\alpha_i, \beta_i \in (0, 1)$ , 且  $\alpha_i \leq \beta_i$ ;  $\gamma \in (0, 1)$  是不进行惩罚和奖励时的局部状态值。

由变权函数的性质可知:在  $0 \leq x_i < \alpha_i$  范围内,单调递减;在  $\alpha_i$  到  $\beta_i$  范围内,不变;在  $\beta_i \leq x_i < 1$  范围内单调递增。

## 3 改进层次分析法数据质量综合评估算例分析

对数据质量开展综合评价,首先要对数据质量评估指标体系中的每项维度进行分析计算;然后再进行综合评价。维度分析计算时,以正确性和准确性两个维度为例进行评估值

计算。

### 3.1 正确性

正确性指标由格式正确、值域正确、值合理、拼写正确、非空性等五个分指标构成。首先采用层次分析法确定这五个分指标的相对权重。

假设请专家对分指标的相对权重进行判断,得到以下的权重表格,如表 1 所示。

表 1 专家对于正确性分指标相对权重判断

分指标	格式正确	值域正确	值合理	拼写正确	非空性
格式正确	1	1/9	1/7	1	1/5
值域正确	9	1	5	9	3
值合理	7	1/5	1	7	5
拼写正确	1	1/9	1/7	1	1/7
非空性	5	1/3	1/5	5	1

按照层次分析法的步骤求解相对权重。按照专家的判断建立判断矩阵  $A_1$ :

$$A_1 = \begin{bmatrix} 1 & 1/9 & 1/7 & 1 & 1/5 \\ 9 & 1 & 5 & 9 & 3 \\ 7 & 1/5 & 1 & 7 & 5 \\ 1 & 1/9 & 1/7 & 1 & 1/7 \\ 5 & 1/3 & 1/5 & 5 & 1 \end{bmatrix}$$

采用求和法近似计算该矩阵的最大特征值对应的单位特征向量,即分指标对应的权重向量,将矩阵  $A_1$  按列归一化得矩阵  $B_1$ :

$$B_1 = \begin{bmatrix} 0.0435 & 0.0633 & 0.0220 & 0.0435 & 0.0213 \\ 0.3913 & 0.5696 & 0.7709 & 0.3913 & 0.3191 \\ 0.3043 & 0.1139 & 0.1542 & 0.3043 & 0.5319 \\ 0.0435 & 0.0633 & 0.0220 & 0.0435 & 0.0213 \\ 0.2174 & 0.1899 & 0.0308 & 0.2174 & 0.1064 \end{bmatrix}$$

再按行求和得矩阵  $C_1$ :

$$C_1 = [0.1936 \quad 2.4423 \quad 1.4087 \quad 0.1936 \quad 0.7619]^T$$

经再次归一化得  $W_1$ :

$$W_1 = [0.0387 \quad 0.4885 \quad 0.2817 \quad 0.0387 \quad 0.1524]^T$$

一致性检验:

$$C.I. = 0.1022 \quad R.I. = 1.12 \quad C.R. = 0.0912 < 0.1$$

### 3.2 准确性

准确性指标由位置准确、数值准确、描述准确等三个分指标构成。首先采用层次分析法确定这三个分指标的相对权重。

假设分别请两位专家对分指标的相对权重进行判断,得到权重表格如表 2、3 所示。

表 2 专家 A 对于准确性分指标相对权重判断

分指标	位置准确	数值准确	描述准确
位置准确	1	1/5	9
数值准确	5	1	7
描述准确	1/9	1/7	1

表 3 专家 B 对于准确性分指标相对权重判断

分指标	位置准确	数值准确	描述准确
位置准确	1	1/3	7
数值准确	3	1	5
描述准确	1/7	1/5	1

按照层次分析法的步骤求解相对权重。按照专家 A、专

家 B 的判断分别建立判断矩阵。

由专家 A 的判断建立判断矩阵  $A_2$ :

$$A_2 = \begin{bmatrix} 1 & 1/5 & 9 \\ 5 & 1 & 7 \\ 1/9 & 1/7 & 1 \end{bmatrix}$$

采用求和法近似计算该矩阵的最大特征值对应的单位特征向量,即分指标对应的权重向量,将矩阵  $A_2$  按列归一化得矩阵  $B_2$ :

$$B_2 = \begin{bmatrix} 0.1636 & 0.1489 \\ 0.8182 & 0.7447 \\ 0.0182 & 0.1064 \end{bmatrix}$$

再按行求和得矩阵  $C_2$ :

$$C_2 = [0.8420 \quad 1.9746 \quad 0.1834]^T$$

经再次归一化得  $W_2$ :

$$W_2 = [0.2807 \quad 0.6582 \quad 0.0611]^T$$

由专家 B 的判断建立判断矩阵  $A_3$ :

$$A_3 = \begin{bmatrix} 1 & 1/3 & 7 \\ 3 & 1 & 5 \\ 1/7 & 1/5 & 1 \end{bmatrix}$$

采用求和法近似计算该矩阵的最大特征值对应的单位特征向量,即分指标对应的权重向量,将矩阵  $A_3$  按列归一化得矩阵  $B_3$ :

$$B_3 = \begin{bmatrix} 0.2414 & 0.2174 & 0.5385 \\ 0.7241 & 0.6522 & 0.3846 \\ 0.0345 & 0.1304 & 0.0769 \end{bmatrix}$$

再按行求和得矩阵  $C_3$ :

$$C_3 = [0.9972 \quad 1.7609 \quad 0.2418]^T$$

经再次归一化得  $W_3$ :

$$W_3 = [0.3324 \quad 0.5870 \quad 0.0806]^T$$

由于分指标个数为 3 个,判断矩阵很容易做到完全一致,故对于准确性指标不需要再进行一致性检验的相关计算。

对两位专家的判断进行平均得到分指标的相对权重。

$$W = \begin{bmatrix} W_2 \\ W_3 \end{bmatrix} = \frac{W_2 + W_3}{2} = \begin{bmatrix} 0.30655 \\ 0.62260 \\ 0.07085 \end{bmatrix}$$

在完成了分指标的权重计算后,分指标权重与分指标值的乘积再求和,就能得到该项指标的评估值。以正确性和准确性为例,设正确性包含的格式正确、值域正确、值合理、拼写正确、非空性等五项分指标的评估值分别为 0.683、0.743、0.952、0.983、0.831;设准确性包含的位置准确、数值准确、描述准确等三项分指标的评估值分别为 0.712、0.523、0.346。

加权以后的正确性指标的评估值应为: 0.82225; 加权以后的准确性指标的评估值应为: 0.56846。

同样方法,对一致性、完整性等另外两项指标,进行评估值计算,分别得到:

$$I_{CQA} = 0.7982 \quad I_{CQF} = 0.6893$$

假设正确性维度权重为 0.4; 准确性维度权重为 0.3; 一致性维度权重为 0.2; 完整性维度权重为 0.1。按照传统层次分析法,综合评估的结果应该是:

$$I = I_{CQC} \times 0.4 + I_{CQR} \times 0.3 + I_{CQA} \times 0.2 + I_{CQF} \times 0.1 = 0.7279$$

如果以综合评估结果小于 0.6 为差、0.6 至 0.7 之间为

一般,0.7至0.8之间为良好,0.8以上为优秀为标准,依据传统层次分析法得出的数据维护阶段的数据质量综合评价结果为良好。

显然 0.7279 这一综合评价结果并未体现最终决策者对正确性、准确性、一致性和完整性这四项维度的实测值的喜好。例如,准确性是核心影响因子,所以对它的实测值,应该更多地给予关注,尤其是如果该影响因子的实测值比较“差”的时候,应该给予惩罚。那么通过传统层次分析法得到的综合评价结果显然无法体现这一奖励和惩罚原则。

为此,使用基于指数变权的层次分析改进评价方法,对影响因子的实测向量进行变权。假设对其中的正确性、准确性、一致性维度的实测向量进行局部变权处理,对完整性维度的实测向量进行非变权处理。而在做变权处理时,对正确性维度实测向量进行的是奖励惩罚并举的变权处理;对准确性维度实测向量进行的是只惩罚的变权处理;对一致性维度实测向量进行的是只奖励的变权处理。

由  $a^{\alpha_i} = \gamma$ ,  $b^{\beta_i} \times m = \gamma$ ,  $b \times m = 1$ , 可以解出  $a$ ,  $b$  和  $m$ , 从

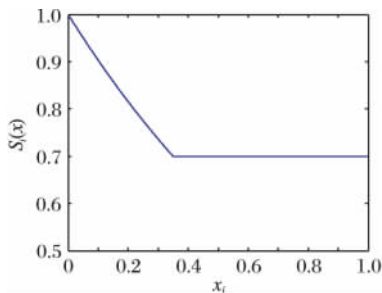


图2 局部变权函数(惩罚型)

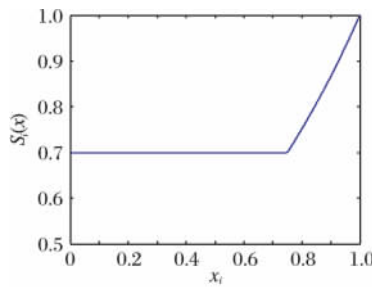


图3 局部变权函数(奖励型)

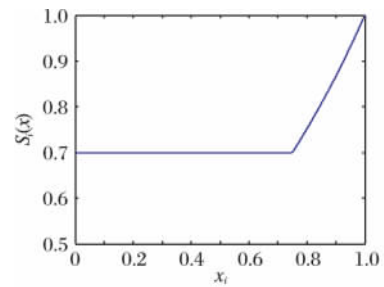


图4 局部变权函数(惩罚奖励并举型)

$\alpha_i$  为第  $i$  项指标所有数据的 35% 分位数;  $\beta_i$  为第  $i$  项指标所有数据的 (1% ~ 25%) 分位数。

按照之前得出的正确性维度的评估值  $I_{CQC} = 0.82225$ , 准确性维度的评估值  $I_{CQR} = 0.56846$ , 一致性维度的评估值  $I_{CQA} = 0.7982$ , 完整性维度的评估值  $I_{CQF} = 0.6893$ 。即:  $x_1 = 0.82225$ ,  $x_2 = 0.56846$ ,  $x_3 = 0.7982$ ,  $x_4 = 0.6893$ 。

分别对应根据式 (1) ~ (3), 得到:

$$S_1(x) = 0.7760, S_2(x) = 0.7, S_3(x) = 0.7498.$$

$$S = [0.7760 \quad 0.7 \quad 0.7498 \quad 0.6893]^T.$$

将影响因子的常权向量与  $S$  进行合成, 得到变权向量:

$$S = [0.3104 \quad 0.21 \quad 0.14996 \quad 0.06893]^T$$

对  $S$  进行归一化处理得  $Q$ :

$$Q = [0.4200 \quad 0.2840 \quad 0.2028 \quad 0.0932]^T$$

由归一化得出的  $Q$ , 即为二次变权后正确性、准确性、一致性和完整性等四项指标的最终权重。从权重的数值来看, 正确性的权重由传统层次分析法下的 0.4 提升至 0.42, 体现了对其实测值的奖励, 一致性的权重由传统层次分析法下的 0.2 提升至 0.2028, 也体现了对其实测值的奖励, 相应的准确性、完整性的权重则有所降低。

由正确性、准确性、一致性和完整性等四项维度分别对应的评估值  $I$ :

$$I = [0.82225 \quad 0.56846 \quad 0.7982 \quad 0.6893]^T$$

则最后的综合评估结果为:  $I = 0.7328$ 。

与之前采用传统层次分析法得到的综合评估结果相比,

而有  $a^{0.35} = 0.7 b^{0.75} \times m = 0.7 b \times m = 1$ ; 计算:  $a = e^{(\ln(0.7)/0.35)} = 0.3609 b = e^{(\ln(0.7)/(-0.25))} = 4.1649 m = 0.2401$ 。

所以, 只体现惩罚原则的惩罚型局部变权函数如下所示:

$$S_i(x) = \begin{cases} 0.3609^{x_i} & 0 \leq x_i < \alpha_i \\ 0.7, & \alpha_i \leq x_i \leq 1 \end{cases} \quad (1)$$

函数  $S_i(x)$  如图 2 所示。

只体现奖励原则的奖励型的局部变权函数如下所示:

$$S_i(x) = \begin{cases} 0.7, & 0 \leq x_i < \beta_i \\ 4.1649^{x_i} \times 0.2401, & \beta_i \leq x_i \leq 1 \end{cases} \quad (2)$$

函数  $S_i(x)$  如图 3 所示。

奖励与惩罚并举的局部变权函数如下:

$$S_i(x) = \begin{cases} 0.3609^{x_i}, & 0 \leq x_i < \alpha_i \\ 0.7, & \alpha_i \leq x_i < \beta_i \\ 4.1649^{x_i} \times 0.2401, & \beta_i \leq x_i < 1 \end{cases} \quad (3)$$

函数  $S_i(x)$  如图 4 所示。

采用层次分析法改进方法得到的综合评估结果在数值上有了提高, 其提高的部分主要体现了对部分影响因子实测值的奖励原则。

## 4 结语

针对数据开展质量综合评估方法研究, 是提高数据质量的基础和必要前提。然而, 在评估过程中, 除了对评估指标赋予权重外, 不同的决策者对评估指标实测值也会提出喜好要求, 而传统的层次分析法难以满足这方面需要。为此, 本文引入变权思想对传统层次分析法进行了改进, 通过变权体现对指标值不同状态的激励或惩罚。同时, 考虑到指数函数在放大或缩小的幅度上比线性函数具有更好的优势, 更能体现这种激励或惩罚的作用, 所以提出了一种基于指数函数变权的层次分析法的改进评价方法, 并在数据质量综合评价中进行了应用。案例分析结果显示本改进方法具有一定的可行性和实用性。

参考文献:

- [1] 杨青云, 赵培英, 杨冬青, 等. 数据质量评估方法研究[J]. 计算机工程与应用, 2004, 40(9): 3-4.
- [2] 邱均平, 文庭孝. 评价学: 理论·方法·实践[M]. 北京: 科学出版社, 2010.
- [3] 卢代军, 夏学知, 张子鹤, 等. 目标信息的时效性分析[J]. 火力与指挥控制, 2007, 32(1): 38-41.
- [4] Evidence Based Research, Inc. Network centric operations conceptual framework version 1.0[R]. Vienna: Office of Force Transformation, 2003.

(下转第 297 页)

网络语音质量合格率、TD 语音业务感知接通率、TD 语音业务端到端接续时长、TD 语音业务感知掉话率、TD 网络语音质量、2G 到 2G 入局切换成功率、寻呼成功率、正常 Inter-VLR 位置更新成功率、GSM 语音业务网内接通率、GSM 语音业务端到端接续时长、IMSI 附着成功率、系统接通率、TD 网络寻呼成功率、TD 语音业务网内接通率、MSS 处理器平均负荷、VLR 利用率,超过阈值即认为异常;8:00-23:00 内指标异常次数大于等于 3 次定义为“不健康”,异常次数大于 1 次并小于 3 次定义为“亚健康”,异常次数小于等于 1 次定义为“健康”。

7) RNC 健康度分析。每日定时检查网元前一日的:切换成功率(TD)、码资源利用率(TD)、3G 小区退服率、TD 语音业务无线接通率、TD 语音业务无线掉话率,超过阈值即认为异常;8:00-23:00 内指标异常次数大于等于 3 次定义为“不健康”,异常次数大于 1 次并小于 3 次定义为“亚健康”,异常次数小于等于 1 次定义为“健康”。

8) MGW 健康度分析。每日定时检查网元前一日的:SCTP 偶联数据块重传率,超过阈值即认为异常;8:00-23:00 内指标异常次数大于等于 3 次定义为“不健康”,异常次数大于 1 次并小于 3 次定义为“亚健康”,异常次数小于等于 1 次定义为“健康”。

基于上述分析结果:

1) 生成省级、地市级的网络监控报告,可按照报表格式进行呈现,按照切片地图进行呈现。通过直观的图形界面对数据进行展示,可以提升用户的感知度与对数据的理解程度,效果远远好于传统的报表展示<sup>[8]</sup>。

2) 生成网元级的网络健康报告,可按照报表格式进行呈现,支持钻取。

3) 实施网元级的网络健康度督办,将不健康的网元名称与指标详情通过接口派发到 EOMS 系统进行自动督办,督办工单在 EOMS 系统中流转完毕后,将处理结果反馈回本系统(网络质量智能分析督办系统,下同),本系统根据工单中的网元名称去核查该网元当前体检值,如果正常则将该工单落地,完成此次督办工作;反之则继续将此工单派发到 EOMS 系统进行督办。

另外,对 CELL 的分析督办有:

1) 每日定时分析全省小区感知类告警排名,排名 TOP20 以内的小区被认定为告警异常网元,派发工单进行督办,并生成督办告警发送到综合监控系统。

2) 每日定时分析各分公司小区感知类告警排名,排名 TOP3 以内的小区的被认定为告警异常网元,派发工单进行督办,并生成督办告警发送到综合监控系统。

## 4 结语

本文对网络质量智能分析督办系统从设计思想、设计方案、应用实例等方面作了详细介绍。所研制的网络质量智能分析督办系统可承担电信网络 80% 的网络质量管控工作,压缩 90% 的网络质量管控人员数量,维护频度提高 3 倍。相信网络质量智能分析督办系统将成为 OSS 体系的基础平台之一,这将是所有运营商的必然选择。

尽管本系统的使用状况良好,但是还有一些需要扩展和改进的地方,例如:系统只能下发督办任务工单给维护人员,让维护人员去执行具体的优化工作;系统尚不能自主自动执行优化工作。因此,系统后续发展的趋势是将分析结果通过自动接口与设备进行对接,根据指令库去执行维护指令,完成系统的闭环操作<sup>[9]</sup>。另外,还要进一步扩大系统的应用范围,逐步推进自动化系统的应用率,从而降低企业的工作负荷<sup>[10]</sup>。还有,通过系统进行自动维护与优化工作是今后的发展趋势,也是技术变革的趋势<sup>[11]</sup>,将增加系统的智能性,通过系统的自我学习来完善知识库,以最大限度替代人工的维护方式<sup>[12]</sup>。

参考文献:

- [1] 陈明. 软件工程[M]. 北京: 中国铁道出版社, 2012.
- [2] WEINBERG G M, WEINBERG D. 系统设计的一般原理[M]. 张锐, 王佳, 译. 北京: 清华大学出版社, 2004.
- [3] van BON J. IT 服务管理基于 ITIL 的全球最佳实践[M]. 章斌, 译. 北京: 清华大学出版社, 2012.
- [4] CSDN FORUM. 系统引擎技术汇总[EB/OL]. [2012-02-10]. <http://bbs.csdn.net>.
- [5] JAVAEEY FORUM. 系统设计论坛. 分析引擎中规则库详解[EB/OL]. [2013-04-11]. <http://www.iteye.com/forums>.
- [6] CHINAITLAB FORUM. UML 设计论坛. 工作流与系统设计[EB/OL]. [2011-01-15]. <http://bbs.chinaitlab.com>.
- [7] UMLORG FORUM. 系统设计. 任务与调度设计[EB/OL]. [2011-04-14]. <http://www.umlchina.com>.
- [8] MSDN FORUM. 分析系统设计. 分析系统界面展示说明[EB/OL]. [2010-01-13]. <http://social.msdn.microsoft.com>.
- [9] 中国网信论坛. 网管系统设计. 网管系统后期演进[EB/OL]. [2012-01-17]. <http://bbs.bitscn.com>.
- [10] IBM FORUM. 系统设计. 智能系统[EB/OL]. [2011-11-14]. <http://www-900.ibm.com>.
- [11] Oracle FORUM. 系统设计讨论. 智能系统发展趋势汇总[EB/OL]. [2009-12-12]. <http://www.oracle.com/cn/community>.
- [12] Google FORUM. 开发者论坛. 人工智能的研究[EB/OL]. [2010-11-07]. <https://groups.google.com>.

(上接第 290 页)

- [5] PARKER M B, MOLESHE V, De la HARPE R, *et al.* An evaluation of information quality frameworks for the World Wide Web [C]// Proceedings of 8th Annual Conference on WWW Applications. Bloemfontein: [s. n.], 2006: 1-11.
- [6] NAUMANN F, ROLKER C. Assessment methods for information quality criteria[C]// Proceedings of 5th International Conference on Information Quality. 2000: 148-162.
- [7] LEE Y W, STRONG D M, KAHN B K, *et al.* AIMQ: a methodology for information quality assessment[J]. Information & Management,

2002, 40(2): 133-146.

- [8] 郭亚军. 综合评价理论、方法及应用[M]. 北京: 科学出版社, 2007.
- [9] 谢剑锋, 宋华文, 李元左. 装备质量管理风险评估指标体系及评估模型研究[J]. 价值工程, 2011, 30(28): 6-8.
- [10] 陈根忠, 薛业飞. 基于层次分析法的引导式电子对抗系统效能评估[J]. 电子对抗, 2006(1): 26-28.
- [11] 郑雪冰, 陈慧敏. 模糊层次分析法在航空电子系统作战效能评估中的应用[J]. 光学技术, 2006, 32(Z1): 247-249.