

Research of Fair Grading Mechanism in Youth Singer Competition*

—Based on the Commentators Affective Filter

Weiwei Li, Yajun Guo, Pingtao Yi

School of Business Administration, Northeastern University, Shenyang
Email: liweiwei198558@163.com

Received: Jan. 1st, 2013; revised: Jan. 17th, 2013; accepted: Feb. 28th, 2013

Abstract: There is always a dispute about scoring mechanism of the Youth Singer Competition which is held by CCTV in China, because the scoring objectivity is related to competition fairness directly. To this problem, we present one scoring model associated with the commentators' affective filter. The basic principle of this model is as follows. First, each competitor is given a fictitious scoring vector. Second, the fictitious scoring vector is optimized repeatedly through the maximal approaching degree between the fictitious and the group scoring vector. At last, a simulation example is given to illustrate the validity of the affective filter scoring model.

Keywords: Comprehensive Evaluation; Group Evaluation; Weight Modulus; Affective Filter; Circulation Optimization

青歌赛评分机制的公平性研究*

—基于评委情感因素过滤的视角

李伟伟, 郭亚军, 易平涛

东北大学工商管理学院, 沈阳
Email: liweiwei198558@163.com

收稿日期: 2013 年 1 月 1 日; 修回日期: 2013 年 1 月 17 日; 录用日期: 2013 年 2 月 28 日

摘要: CCTV 青年歌手电视大奖赛自举办以来, 评分机制一直是争议颇多的一大问题, 评分的客观公正与否直接关系到比赛的公平性。针对这一问题, 从评委情感因素过滤的视角提出了一种评价模型。该模型的基本原理是首先给定参赛选手一个待定的虚拟得分值向量, 然后以虚拟得分值与评委群体评分向量的贴近期最大为基准, 对虚拟得分值进行循环优化, 使其不断接近于参赛选手的客观表现。最后, 通过一个模拟算例验证了带有情感过滤特征评价模型的有效性。

关键词: 综合评价; 群组评价; 权重系数; 情感过滤; 循环优化

1. 引言

青年歌手电视大奖赛(简称青歌赛)由 CCTV 自 1984 年首次举办, 迄今为止, 已经为中国歌坛输送了许多优秀的音乐人才。经过 20 多年的发展, 青歌赛

的发展水平有了显著的提高, 其影响力也日益显著。在比赛中, 对于优秀音乐人才的选拔, 首先是事先确定若干考核指标, 例如第十七届的评分细则包括: 基本要求部分(主要指音准、节奏、咬字和吐字等方面)、气息控制部分、声音控制部分、音乐表现部分和综合印象部分; 然后多个评委对参赛选手的现场表演进行

*资助信息: 国家自然科学基金资助项目(71071030, 71071031)。

打分, 通过层层筛选来确定最终的获胜者。

青歌赛的评分机制是评委依据事先规定的评分规则对参赛选手进行主观打分, 然后按照一定的方式综合各评委的评分给出参赛选手的最终得分。由于参赛选手的得分是由评委主观给出的, 因而青歌赛评分机制的公平性一直以来也是争议颇多的一个话题。例如, 第十三届青歌赛中, 某参赛选手得分的最高分和最低分相差了 13 分, 遭到广大观众的质疑。评分机制是公平的基础, 评分的客观公正与否直接关系到整个比赛的公平性。因而, 对于青歌赛评分机制的公平性研究十分必要, 且具有现实意义。

2. 传统评价模型

目前, 关于青歌赛评分机制的研究理论界鲜有涉及, 但在比赛过程中, 常用的有两种模型: 一是将各评委的打分进行简单平均, 取各评委打分值的平均值作为参赛选手的最终得分(简称取均值的评价模型, 见模型(M1)); 二是对于各评委的打分值, 去掉一个最高分, 去掉一个最低分, 取剩下打分值的平均值作为各参赛选手的最终得分(简称去掉极端打分后取均值的评价模型, 见模型(M2))。设青歌赛的参赛选手的集合为 $O = \{o_1, o_2, \dots, o_n\}$, 评分指标项为

$X = \{x_1, x_2, \dots, x_m\}$, 评委的集合为 $S = \{s_1, s_2, \dots, s_l\}$ 。为方便起见, 分别记 $N = \{1, 2, \dots, n\}$, $M = \{1, 2, \dots, m\}$, $L = \{1, 2, \dots, l\}$ 。设 $x_{ij}^{(k)}$ 为评委 s_k ($k \in L$) 给出的参赛选手 o_i ($i \in N$) 在第 j ($j \in M$) 个评分指标项上的主观打分值, 则上述两种评分模型可分别描述为

$$x_{ij} = \frac{1}{l} \sum_{k=1}^l x_{ij}^{(k)} \quad (i \in N, j \in M) \quad (M1)$$

$$x_{ij} = \frac{1}{l-2} \left(\sum_{k=1}^l x_{ij}^{(k)} - \max \{x_{ij}^{(k)}\} - \min \{x_{ij}^{(k)}\} \right) \quad (i \in N, j \in M) \quad (M2)$$

其中, x_{ij} 表示参赛选手 o_i ($i \in N$) 在评分指标项 x_j ($j \in M$) 上的最终得分, 模型(M2)中 $\max \{x_{ij}^{(k)}\}$ 和 $\min \{x_{ij}^{(k)}\}$ 分别表示 l 个评委给出的关于参赛选手 o_i ($i \in N$) 的评分指标项 x_j ($j \in M$) 打分值的最大值和最小值。

设比赛评分指标项的权重向量为 $\omega = (\omega_1, \omega_2, \dots, \omega_m)^T$, 一般依据比赛规则事先已知,

则参赛选手 o_i ($i \in N$) 的最终得分值(设为 y_i)可采用如下表达式求得:

$$y_i = \sum_{j=1}^m \omega_j x_{ij}, \quad i \in N \quad (1)$$

依据式(1)求得的各参赛选手的最终得分值 y_i ($i \in N$), 并按 y_i 值由大到小对各参赛选手进行排序并择优, 从而判断出各参赛选手之间的优劣关系。

由以上分析可知, 青歌赛的评分机制是典型的群体评价(群体决策)^[1]的过程, 目前对于群体评价的理论研究比较丰富^[2-16], 但是对于群体评价中专家的主观情感因素的研究鲜有涉及。在青歌赛中, 评委按照自己的主观判断对参赛选手打分, 难免会有主观情感因素的参与, 因而观众对评委评分的客观公正性一直存在质疑。针对这一问题, 本文从评委主观情感因素过滤的视角对青歌赛的评分机制进行分析, 提出了一种新的评分模型。该模型的基本原理是依据评分指标项给每个参赛选手一个待定虚拟得分向量, 然后以虚拟得分向量与评委群体评分向量的贴进度最大为准则, 对参赛选手的虚拟得分值进行循环优化, 使其不断接近于参赛选手的客观表现, 从而保证了评分过程的客观公正性。

3. 带有情感过滤特征的评价模型

3.1. 评委主观情感因素的界定

评委在打分时, 受其心情因素、主观偏好因素甚至利益因素或首因效应、晕轮效应等各种因素的影响, 会使其情感因素有意或无意的参与到评分过程中, 从而使得参赛选手的得分偏离其客观表现, 评分过程的客观公正性难以保证。针对这一问题, 本文对评委在评分过程中的主观情感因素进行抽象, 给出如下假设。

假设 1 每个参赛选手都存在一个客观真实的表现, 并对应着一个客观真实的得分。

假设 2 评委在评分过程中, 受其正向情感因素或利益因素等的影响, 会对参赛选手给出较其客观表现偏高的打分。

假设 3 评委在评分过程中, 受其负向情感因素或利益因素等的影响, 会对参赛选手给出较其客观表现偏低的打分。

假设 4 评委基于其评分更易被其他评委和广大观众接受的目的, 会对与其情感或利益不发生作用关系的参赛选手给出与其客观表现相一致的打分。

3.2. 参赛选手虚拟得分值的确定

设 $x_{ij}^{(k)}$ 为评委 $s_k (k \in L)$ 给出的参赛选手 $o_i (i \in N)$ 在第 $j (j \in M)$ 个评分指标项上的主观打分。为使参赛选手的最终得分值具有可比性, 需要对各评委给出的打分值进行规范化处理, 即将所有打分值规范化到相同的分值区间 $[M^0, M^*]$, 从而保证参赛选手的得分值具有可比性。具体方法如下^[17]:

$$x_{ij}^{(k)} = M^0 + (M^* - M^0) \frac{x_{ij}^{(k)} - \min\{x_{ij}^{(k)}\}}{\max\{x_{ij}^{(k)}\} - \min\{x_{ij}^{(k)}\}} \quad (2)$$

其中, M^0 和 M^* 的取值不同对得分结果并无影响, 只要保证评委的打分值都规范到该区间就行。为方便起见, 将规范化后的打分值仍记为 $x_{ij}^{(k)}$ ($i \in N, j \in M, k \in L$)。

对于参赛选手 $o_i (i \in N)$, 设其待定的虚拟得分值向量为 $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{im}^*)$, 其中 $x_{ij}^* \in [M^0, M^*]$, $j \in M$ 。需要说明的是, 虚拟得分值是待求变量, 作用是存储参赛选手在各指标项上的得分值并使其不断接近于参赛选手的客观表现。

定义 1 称

$$d_{ij}^{(k)} = x_{ij}^{(k)} - x_{ij}^*, \quad i \in N, j \in M, k \in L \quad (3)$$

为评委 $s_k (k \in L)$ 给出的关于参赛选手 $o_i (i \in N)$ 的第 $j (j \in M)$ 个评分指标项的打分值与其虚拟得分值的偏差, 其中 $d_{ij}^{(j)} \in (-\infty, +\infty)$ 。

定义 2 称

$$w_{ij}^{(k)} = \ln(a_i |d_{ij}^{(k)}| + b_i) / \sum_{k=1}^l \ln(a_i |d_{ij}^{(k)}| + b_i) \quad (4)$$

为评委 $s_k (k \in L)$ 在参赛选手 $o_i (i \in N)$ 第 $j (j \in M)$ 个评分指标项上的权重, 满足 $w_{ij}^{(k)} \in (0, 1)$ 且 $\sum w_{ij}^{(k)} = 1$ 。其中 a_i 和 b_i 是与参赛选手 $o_i (i \in N)$ 有关的待定变量, 变量 a_i 的设置是为了实现评委的权重系数与其主观情感因素的投入成反比变化关系, 即通过权重分配达到主观情感过滤的目的, 变量 b_i 的设置是为了满足权重非负的条件, 因而有 $a_i < 0$, $b_i > 1 - a_i \max\{d_{ij}^{(k)}\}$ ($i \in N, j \in M$)。

对于参赛选手的待定虚拟得分值向量, 本文通过循环优化使其不断接近于参赛选手的客观表现, 因此参赛选手 $o_i (i \in N)$ 第 $j (j \in M)$ 个评分指标项, 评委 $s_k (k \in L)$ 给出的打分值与其虚拟得分值的偏差绝对值 $|d_{ij}^{(k)}|$ 越大, 说明评委 s_k 的情感因素参与越多, 其权重应该越小, 即 $w_{ij}^{(k)}$ 是 $|d_{ij}^{(k)}|$ 的减函数, 因而定义 2 中规定 $a_i < 0$ 。为了满足 $w_{ij}^{(k)} \in (0, 1)$, 定义 2 中规定与参赛选手 $o_i (i \in N)$ 有关的所有 $k \in L, j \in M$, 都要有 $a_i |d_{ij}^{(k)}| + b_i > 1$, 即 $b_i > 1 - a_i \max\{d_{ij}^{(k)}\}$ (a_i, b_i 的值由规划模型(M3)求出)。

定义 3 对于参赛选手 $o_i (i \in N)$ 的第 $j (j \in M)$ 个评分指标项, 称

$$x_{ij}^\# = \sum_{k=1}^l w_{ij}^{(k)} x_{ij}^{(k)}, \quad i \in N, j \in M, k \in L \quad (5)$$

为参赛选手 $o_i (i \in N)$ 的第 $j (j \in M)$ 个评分指标项上的群体得分值。

设参赛选手 $o_i (i \in N)$ 在所有评分指标项上的群体得分值向量为 $x_i^\#$, 则 $x_i^\# = (x_{i1}^\#, x_{i2}^\#, \dots, x_{im}^\#)$ 。

定义 4 称

$$\gamma_i = \sum_{j=1}^m x_{ij}^* x_{ij}^\# / \sqrt{\sum_{j=1}^m (x_{ij}^*)^2 \sum_{j=1}^m (x_{ij}^\#)^2}, \quad i \in N, j \in M \quad (6)$$

为参赛选手 $o_i (i \in N)$ 的群体得分值向量与其相应的虚拟得分值向量的贴适度, $\gamma_i \in [0, 1]$ 。

通过以上分析, 基于参赛选手的群体得分值向量与其虚拟得分值向量贴适度最大的思想, 下面给出确定参赛选手 $o_i (i \in N)$ 的虚拟得分值向量 $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{im}^*)$ 的规划模型。

$$\begin{aligned} \max \quad & \gamma_i = \sum_{j=1}^m x_{ij}^* x_{ij}^\# / \sqrt{\sum_{j=1}^m (x_{ij}^*)^2 \sum_{j=1}^m (x_{ij}^\#)^2} \\ \text{s.t.} \quad & \begin{cases} x_{ij}^\# = \sum_{k=1}^l w_{ij}^{(k)} x_{ij}^{(k)}, \quad i \in N, j \in M, k \in L \\ w_{ij}^{(k)} = \ln(a_i |d_{ij}^{(k)}| + b_i) / \sum_{k=1}^l \ln(a_i |d_{ij}^{(k)}| + b_i) \\ a_i < 0, \quad b_i > 1 - a_i \max\{d_{ij}^{(k)}\} \\ x_{ij}^* \in [M^0, M^*], \quad i \in N, j \in M \end{cases} \quad (\text{M3}) \end{aligned}$$

3.3. 带有情感过滤特征的评价过程

对青歌赛评委打分过程中的主观情感因素进行

过滤, 本文采用的思路是: 依据评委给出的参赛选手评分值与其虚拟得分值的偏差确定评委在该参赛选手各评分指标项上的情感参与系数; 在此基础上, 基于评委的情感参与系数对其给出的打分值进行调整, 以降低其给出的打分值的主观情感因素。

定义 5 称

$$\mu_{ij}^{(k)} = \left| d_{ij}^{(k)} \right| / \left| \sum_{k=1}^l d_{ij}^{(k)} \right|, \quad i \in N, j \in M, k \in L \quad (7)$$

为评委 $s_k (k \in L)$ 在参赛选手 $o_i (i \in N)$ 的第 $j (j \in M)$ 个评分指标项上的情感参与系数, 满足 $\mu_{ij}^{(k)} \in [0, 1]$ 且 $\sum_{k=1}^l \mu_{ij}^{(k)} = 1$ 。

基于情感参与系数对评委给出的关于参赛选手 $o_i (i \in N)$ 的各评分指标项的评分值进行调整的原则为:

- 1) 当 $d_{ij}^{(k)} \geq 0$ 时, $x_{ij}^{(k)} = x_{ij}^{(k)} (1 - \mu_{ij}^{(k)})$
($i \in N, j \in M, k \in L$);
- 2) 当 $d_{ij}^{(k)} < 0$ 时, $x_{ij}^{(k)} = x_{ij}^{(k)} (1 + \mu_{ij}^{(k)})$
($i \in N, j \in M, k \in L$)。

通过以上分析, 下面给出基于情感过滤特征对参赛选手的虚拟得分值向量进行循环优化的具体过程。

步骤 1: 对于参赛选手 $o_i (i \in N)$ 的第 $j (j \in M)$ 个评分指标项, 依据式(2)对所有评委的打分值进行规范化处理;

步骤 2: 对于参赛选手 $o_i (i \in N)$, 给定其待定的虚拟得分值向量 $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{im}^*)$, 依据式(3)和(4)确定各评委在参赛选手 o_i 的各评分指标项上的权重, 在此基础上, 依据式(5)确定参赛选手 o_i 的群体得分值向量 $x_i^\# = (x_{i1}^\#, x_{i2}^\#, \dots, x_{im}^\#)$;

步骤 3: 利用规划模型(M3)确定参赛选手的虚拟得分值向量 $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{im}^*)$, 据此利用式(7)求得评委在参赛选手 $o_i (i \in N)$ 各评分指标项上的情感参与系数;

步骤 4: 基于步骤 3 求得的情感参与系数, 依据评分指标项得分的调整原则对评委给出的参赛选手 $o_i (i \in N)$ 各指标项的得分值进行调整;

步骤 5: 针对调整后的得分值, 重新给定参赛选手 $o_i (i \in N)$ 虚拟指标值向量 $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{im}^*)$, 重复步骤 2 和步骤 3, 求得参赛选手 o_i 调整后的虚拟指标值向量 $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{im}^*)$;

步骤 6: 给定一任意小的正数 ε (如 $\varepsilon = 0.1$), 若 $|x_i^* - x_i^*| = \sqrt{\sum_{j=1}^m (x_{ij}^* - x_{ij}^*)^2} < \varepsilon$, 则停止计算, 参赛选手 $o_i (i \in N)$ 的最终得分值向量为 x_i^* ; 否则, 令 $x_i^* = x_i^*$, 转入步骤 4。

依据比赛评分指标项的权重向量为

$\omega = (\omega_1, \omega_2, \dots, \omega_m)^T$, 按照式(1)求得各参赛选手 $o_i (i \in N)$ 的最终得分值 y_i , 并按 y_i 由大到小的顺序对各参赛选手进行排序并择优, 从而判断出各参赛选手之间的优劣关系。

4. 模拟算例

本文通过一个算例对青歌赛的评分过程进行模拟。设有 4 名参赛选手, 4 名评委就以下 5 个评分指标项对参赛选手的表现进行打分: 基本要求(主要指音准、节奏、咬字和吐字等方面)、气息控制、声音控制、音乐表现力和综合印象。打分标准: 表现很差得 1 分, 表现差得 3 分, 表现一般得 5 分, 表现好得 7 分, 表现很好得 9 分; 2, 4, 6, 8 分的含义介于 1, 3, 5, 7, 9 分之间。所有参赛选手的原始得分如表 1 所示。

需要说明的是, 由于参赛选手的现场表现客观存在, 但能够真实体现其客观表现的分值却不是显而易见的, 因而才会有多个评委共同打分的处理方式, 也有了本文研究的必要。这里, 为了便于比较验证本文方法的有效性, 对于能够真实体现 4 名参赛选手在 5 个评分指标项上客观表现的分值向量假设已知, 分别

Table 1. Initial sores on every index of objects
表 1. 参赛对象在各评分指标项上的原始得分

	o_1					o_2					o_3					o_4				
s_1	9	8	7	8	9	2	3	2	2	1	4	5	4	3	5	7	8	6	5	6
s_2	2	3	1	3	2	8	7	9	8	9	4	4	4	3	5	8	9	8	9	9
s_3	5	4	3	7	6	6	5	3	4	4	9	8	9	7	8	2	1	3	2	2
s_4	6	3	4	6	7	8	8	7	9	8	2	3	1	1	2	9	8	8	7	9

为(5,3,4,7,6); (6,5,3,4,4); (4,5,4,3,5); (7,8,6,5,6)。评委打分时若没有情感参与,其给出的打分值应与参赛选手的客观分值相同或很接近(选手的客观分值随机给出,表1中的原始打分值则依据客观分值上下波动随机产生)。为简便起见,取各评分指标项的权重向量为 $\omega=(0.2,0.2,0.2,0.2,0.2)^T$ 。

依据式(2)对参赛选手的原始得分数据进行规范化处理(规定 $M^0=1, M^*=9$),处理后的参赛选手在各评分指标项上的标准得分如表2所示。

4.1. 传统取均值评价模型的评价结果

依据模型(M1)求得参赛选手在各评分指标项上的得分如表3所示。

依据式(1)求得各参赛选手的最终得分值向量为 $y=(4.89,5.445,4.44,6.055)^T$,则参赛选手的排序为 $o_4 > o_2 > o_1 > o_3$,即参赛选手4的综合表现最好,参赛选手3的综合表现最差。

4.2. 传统去掉极端打分后取均值评价模型的评价结果

依据模型(M2)求得参赛选手在各评分指标项上的得分如表4所示。

依据式(1)求得各参赛选手的最终得分值向量为 $y=(4.78,5.89,3.88,7.11)^T$,则参赛选手的排序为 $o_4 > o_2 > o_1 > o_3$,即参赛选手4的综合表现最好,参赛选手3的综合表现最差。

4.3. 带有情感过滤特征的评价模型的评价结果

依据带有情感过滤特征的参赛对象虚拟得分值向量的循环优化步骤,对各评委给出的打分值进行调整,得到基于评委情感过滤之后的参赛对象的最终得分值如表5所示。

依据参赛选手虚拟得分值向量的循环优化步骤,求得各评委在参赛对象各评分指标项上的最终权重以及参赛选手的最终得分值如表6所示($\varepsilon=0.1$)。

依据式(1)求得各参赛选手的最终得分值向量为 $y=(5.058,4.486,4.298,6.458)^T$,则参赛选手的排序为 $o_4 > o_1 > o_2 > o_3$,即参赛选手4的综合表现最好,参赛选手3的综合表现最差。

4.4. 结果分析

依据式(1)和假设已知的各参赛选手的客观得分值向量,求得各参赛选手的最终得分值向量为 $y=(5,4.4,4.2,6.4)^T$,则参赛选手的排序为

$o_4 > o_1 > o_2 > o_3$,即参赛选手4的综合表现最好,参赛选手3的综合表现最差。

设由评价模型(M1),(M2)和(M3)求得参赛选手 $o_i(i=1,2,3,4)$ 的最终得分值分别记为 $y_i^{(M1)}, y_i^{(M2)}$ 和 $y_i^{(M3)}$,并将参赛选手的得分值向量分别记为 $y^{(Mp)}=(y_1^{(Mp)}, y_2^{(Mp)}, y_3^{(Mp)}, y_4^{(Mp)})^T(p=1,2,3)$;将通过参赛选手的客观得分值向量求得的参赛选手 $o_i(i=1,2,3,4)$ 的最终客观得分值记为 y_i^* ,客观得分值

Table 2. Standard sores on every index of objects
表 2. 参赛选手在各评分指标项上的标准得分

	o_1					o_2					o_3					o_4				
s_1	9.0	9.0	9.0	9.0	9.0	1.0	1.0	1.0	1.0	1.0	3.3	4.2	4.0	3.7	5.0	6.7	8.0	5.8	4.4	5.6
s_2	1.0	1.0	1.0	1.0	1.0	9.0	7.4	9.0	7.9	9.0	3.3	2.6	4.0	3.7	5.0	7.9	9.0	9.0	9.0	9.0
s_3	4.4	2.6	3.7	7.4	5.6	6.3	4.2	2.1	3.3	4.0	9.0	9.0	9.0	9.0	9.0	1.0	1.0	1.0	1.0	1.0
s_4	5.6	1.0	5.0	5.8	6.7	9.0	9.0	6.7	9.0	8.0	1.0	1.0	1.0	1.0	1.0	9.0	8.0	9.0	6.7	9.0

Table 3. The scores on every index of objects by averaging model
表 3. 评委打分取均值模型中参赛选手在各评分指标项上的得分

	x_1	x_2	x_3	x_4	x_5
o_1	5	3.4	4.675	5.8	5.575
o_2	6.325	5.4	4.7	5.3	5.5
o_3	4.15	4.2	4.5	4.35	5
o_4	6.15	6.5	6.2	5.275	6.15

Table 4. The scores on every index of objects by averaging model by moving extreme scores
表 4. 去掉极端打分的均值模型中参赛选手在各评分指标项上的得分

	x_1	x_2	x_3	x_4	x_5
o_1	5	1.8	4.35	6.6	6.15
o_2	7.65	5.8	4.4	5.6	6
o_3	3.3	3.4	4	3.7	5
o_4	7.3	8	7.4	5.55	7.3

Table 5. The final values of objects after affective filter
表 5. 基于评委情感过滤之后的参赛对象的最终得分值

	o_1					o_2					o_3					o_4				
s_1	5.43	2.72	4.94	6.61	6.30	2.33	2.10	1.64	1.85	2.01	3.50	4.11	4.18	3.87	5.06	6.56	7.31	6.60	5.21	6.47
s_2	5.10	2.73	4.75	5.86	5.67	5.62	4.39	3.91	4.42	4.57	3.50	2.98	4.18	3.87	5.06	6.52	7.31	6.60	5.21	6.47
s_3	5.42	2.72	4.94	6.61	6.29	5.67	4.41	2.82	3.93	4.43	4.02	4.31	4.29	4.16	4.70	3.08	3.29	2.69	2.44	2.62
s_4	5.43	2.73	4.95	6.60	6.30	5.62	4.41	3.91	4.43	4.56	1.31	1.31	1.39	1.35	1.50	6.55	7.31	6.57	5.21	6.47

Table 6. Evaluators' weights on evaluation index and the final score of every object ($\varepsilon = 0.1$)
表 6. 评委在参赛对象各评分指标项上的最终权重以及参赛选手的最终得分值 ($\varepsilon = 0.1$)

	o_1					o_2					o_3					o_4				
	$w_{11}^{(i)}$	$w_{12}^{(i)}$	$w_{13}^{(i)}$	$w_{14}^{(i)}$	$w_{15}^{(i)}$	$w_{21}^{(i)}$	$w_{22}^{(i)}$	$w_{23}^{(i)}$	$w_{24}^{(i)}$	$w_{25}^{(i)}$	$w_{31}^{(i)}$	$w_{32}^{(i)}$	$w_{33}^{(i)}$	$w_{34}^{(i)}$	$w_{35}^{(i)}$	$w_{41}^{(i)}$	$w_{42}^{(i)}$	$w_{43}^{(i)}$	$w_{44}^{(i)}$	$w_{45}^{(i)}$
s_1	0.251	0.248	0.251	0.251	0.251	0.220	0.230	0.238	0.231	0.229	0.255	0.259	0.257	0.256	0.260	0.254	0.254	0.254	0.253	0.254
s_2	0.248	0.251	0.248	0.248	0.248	0.260	0.256	0.256	0.257	0.257	0.255	0.251	0.257	0.256	0.260	0.254	0.254	0.264	0.253	0.254
s_3	0.250	0.248	0.251	0.251	0.251	0.260	0.257	0.251	0.255	0.257	0.257	0.257	0.258	0.258	0.256	0.239	0.237	0.238	0.241	0.238
s_4	0.251	0.251	0.251	0.250	0.251	0.260	0.257	0.256	0.257	0.257	0.233	0.233	0.230	0.230	0.223	0.254	0.254	0.254	0.253	0.254
a_j	-0.211					-0.435					-0.382					-0.263				
b_j	4.999					5.990					5.991					7.998				
x_j^*	(5.30, 2.70, 4.85, 6.36, 6.08)					(5.64, 4.45, 3.56, 4.25, 4.53)					(3.85, 3.97, 4.39, 4.14, 5.13)					(6.63, 7.38, 6.56, 5.27, 6.45)				

向量记为 $y^* = (y_1^*, y_2^*, y_3^*, y_4^*)$ 。依据下式求得各评价模型与参赛选手的客观表现的贴近程度分别为

$$r_{M1,*} = 0.99452, \quad r_{M2,*} = 0.99124, \quad r_{M3,*} = 0.99999。$$

$$r_{Mp,*} = \sum_{i=1}^4 y_i^{(Mp)} y_i^* / \sqrt{\sum_{i=1}^4 (y_i^{(Mp)})^2 \sum_{i=1}^4 (y_i^*)^2}, \quad p = 1, 2, 3$$

通过以上分析可以得出如下结论：1) 带有情感过滤特征的评价模型给出的最终评分值最能体现参赛选手的客观表现，评委打分取均值的评分模型其次，去掉极端打分的均值模型最差；2) 比较各参赛选手的最终排序情况，可以发现虽然由各评价模型得到的最好和最差的参赛选手都相同，但是依据带有情感过滤特征的评价模型得到的排序与依据参赛选手的客观表现得到的排序完全相同。由此可见，虽然传统的评价模型具有简单直观的特点，但其评分精度不够，而

带有情感过滤特征的评价模型能够更好地体现参赛选手的客观表现。原因在于带有情感过滤特征的评价模型，通过对评委在评分过程中的主观情感因素进行过滤，使评委给出的同一参赛选手的打分值逐渐向参赛选手的客观表现靠拢，对于情感过滤后的打分值与参赛选手的客观表现(体现在群体偏好上)仍有所偏离的评委，赋予其较小的权重而减弱其影响力，从而进一步保证了最终评分结果的客观公正性。

5. 结束语

本文针对青歌赛中评委情感因素参与评分的问题，从情感过滤的视角提出了一种低成本、高效率的评价模型。该模型具有如下特点：1) 给定参赛选手一待虚拟得分值向量，并对其循环优化使其不断接近

于参赛选手的客观表现,从而保证了比赛过程与比赛结果的客观性,可信度较高;2) 利用情感参与系数对评委给出的打分值进行调整,能够在一定程度降低评委主观情感因素对评分结果的影响,并将其与参赛选手的虚拟得分值向量的循环优化相结合,能够对评分结果起到双重优化的效果,进一步保证了评分结果的客观公正性;3) 在虚拟得分值向量的循环优化过程中,对于与群体偏好有所偏离的评委,说明其仍有情感因素参与评分,赋予其较小的权重,降低其对评分结果的影响力,从而加快了循环优化的收敛过程;4) 对参赛选手的得分值作为一个向量来考虑,对整个向量进行循环优化,不仅简化了循环优化过程,而且保证了评分结果的客观性;5) 本文提出的带有情感过滤特征的评价模型,原理简单,便于通过计算机模拟实现。

参考文献 (References)

- [1] D. Black. The theory of committed and elections. Cambridge: Cambridge University Press, 1958.
- [2] C. L. Hwang. Group decision making under multiple criteria, methods and applications. Berlin, New York: Springer-Verlag, 1987.
- [3] F. Herrera, E. Herrera-Viedma and J. L. Verdegay. A model of consensus in group decision making under linguistic assessments. *Fuzzy Sets and Systems*, 1996, 78(1): 73-87.
- [4] K. H. Chen and H. H. Lin. Interactive group decision-making. *Modeling and Application Socioeconomic Planning Sciences*, 1998, 32(2): 113-121.
- [5] H. K. Soung, H. C. Sang and K. K. Jae. An interactive procedure for multiple attribute group decision making with incomplete information: Range-based approach. *European Journal of Operational Research*, 1999, 118(1): 139-152.
- [6] F. Herrera, E. Herrera-Viedma and F. Chiclana. Multiperson decision-making based on multiplicative preference relations. *European Journal of Operational Research*, 2001, 129(2): 372-385.
- [7] E. Bendoly, D. G. Bachrach. A process-based model for priority convergence in multi-period group decision-making. *European Journal of Operational Research*, 2003, 148(3): 534-545.
- [8] I. Takehiro. Cluster ability of groups and information exchange in group decision making with approval voting system. *Applied Mathematics and Computation*, 2003, 136: 1-15.
- [9] Z. S. Xu, P. R. Yager. Power-geometric operators and their use in group decision making. *IEEE Transactions on Fuzzy Systems*, 2010, 18(1): 94-105.
- [10] 苏波, 王焯尘. 群决策研究的评述[J]. *决策与决策支持系统*, 1995, 5(3): 115-124.
- [11] 魏存平, 邱苑华. 群体决策基本理论评述[J]. *北京航空航天大学学报(社会科学版)*, 2000, 13(2): 24-28.
- [12] 易平涛. 多属性综合评价的理论与应用研究[D]. 东北大学博士学位论文, 2006: 1.
- [13] 徐玖平. 群决策理论与方法及实现[M]. 清华大学出版社, 2009: 54.
- [14] 彭怡. 动态群体决策理论及其应用研究[D]. 西南交通大学博士研究生学位论文, 2006.
- [15] 徐选华, 陈晓红. 基于矢量空间的群体聚类方法研究[J]. *系统工程与电子技术*, 2005, 27(6): 1034-1037.
- [16] 王丹力, 戴汝为. 专家群体思维收敛的研究[J]. *管理科学学报*, 2002, 5(2): 1-5.
- [17] 岳超源. 决策理论与方法[M]. 北京: 科学出版社, 2003.