Construction of Students’ Comprehensive Quality Evaluation Model based on Improved AHP

Cheng Yao, Qin Feng and Zheng Xiao
Construction of Students’ Comprehensive Quality Evaluation Model based on Improved AHP

QIN feng, CHENG Yao, ZHENG Xiao
School of Computer Science and Technology
Anhui University of Technology
Ma’anshan, Anhui 243032, China
chengyao@ahut.edu.cn

Abstract—Generally, the assessment of students in colleges is mostly based on academic performance ranking, and it is not conducive to the cultivation of all-round outstanding university students. In order to change the evaluation way of “the only achievement theory”, this paper constructed the comprehensive quality evaluation model (CQE) for college students based on the analytic hierarchy process (AHP). Focusing on the limitations of subjective arbitrariness in AHP, we improved the AHP based on the stakeholder perspective (IAHP_SP). Through the experimental verification of the students’ data in a university in Anhui, it shows that the comprehensive quality evaluation model of the students proposed by this study is feasible.

Keywords—College students evaluation, AHP, ranking, comprehensive quality evaluation

I. INTRODUCTION

The students’ evaluation is an indispensable part in higher education institutions, which has positive influence for promoting universities’ development. This is because the evaluation of students is one of the criteria for the audit evaluation and the engineering certification in China. By analyzing students’ comprehensive performance, a strategic personnel training programs can be well planned in their period of study in a college. Generally, most of higher education institutions are using the final grades to evaluate students’ performance [1][2]. However, “the only achievement theory” is one-sided that can’t evaluate students in the round. Thus, constructing the CQE model of college students is essential.

The undergraduate quality model is divided into four parts, Physical-Mental Quality, Moral Trait, Scientific Literacy and Professionalism[3]. Comprehensive quality model of college students is based on Ideological-Political Quality, Knowledge Quality, Physical-Mental Quality and Creative-Practice Ability [4]. Based on the principle of comprehensiveness, hierarchy, students’ dominant role and feasibility, the comprehensive quality evaluation (CQE) system of college students is grouped into four aspects in Fig. 1.

II. METHODOLOGY

A. Basic theory of AHP

AHP was first proposed by Thomas Saaty in the late 1970s [5], which is an operational research method combining quantitative and qualitative analysis. It can divide a complex problem into multiple components. Through the comparison of each other, the dominance of each factor is divided into orderly hierarchical structure, so as to determine the relative importance of various factors in each level [6]. The wide applicability of AHP is due to its simplicity, easy to use and great flexibility.

Applying AHP to make a structured decision making approach includes six steps: (i) define the problem and choose the criteria, (ii) construct hierarchies, (iii) use pair-wise comparison scales to have specification of numerical values in Table I [5], (iv) validate the inconsistencies in the decision process exist, including calculate maximum eigen value, consistency index (CI) CI=(λmax-n)/(n-1), random consistency index (RI) and consistency ratio (CR) CR=CI/RI, revise the

---

This work is supported by the project of Teaching Reform in Anhui University of Technology:2015siyey02

---

![Fig. 1. CQE System of College Students](image-url)
process till a consensus is reached and (v) adjust integration of weight values to reach an optimum decision.

### TABLE I. IMPORTANCE INTENSITY BETWEEN TWO PARAMETERS IN AHP[5]

<table>
<thead>
<tr>
<th>Importance intensity</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Criterion A has equal importance to criterion B</td>
</tr>
<tr>
<td>3</td>
<td>The importance of criterion A is moderately important to criterion B</td>
</tr>
<tr>
<td>5</td>
<td>The importance of criterion A is strongly important to criterion B</td>
</tr>
<tr>
<td>7</td>
<td>The importance of criterion A is very strongly important to criterion B</td>
</tr>
<tr>
<td>9</td>
<td>The importance of criterion A is extremely important to criterion B</td>
</tr>
<tr>
<td>2, 4, 6, 8</td>
<td>The importance of criterion A than criterion B is in the middle of the above description</td>
</tr>
</tbody>
</table>

Reciprocals Used for inverse comparison

Take evaluator $u_1$ as an example, the first-level evaluation matrix is shown in Table II.

### TABLE II. EVALUATOR $U_1$’S FIRST-LEVEL EVALUATION MATRIX

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1</td>
<td>1/5</td>
<td>1/3</td>
<td>1/2</td>
</tr>
<tr>
<td>X2</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>X3</td>
<td>3</td>
<td>1/3</td>
<td>1</td>
<td>1/2</td>
</tr>
<tr>
<td>X4</td>
<td>2</td>
<td>1/2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Calculate maximum eigen value $\lambda_{max}=4.1074$ and its corresponding eigenvector is $w_i=(0.0863,0.4909,0.2483,0.1745)^T$. Then validate the inconsistencies, according to RI index[6], $RI=0.90$ when $n=4$, $CR=0.0398<0.1$.

AHP provided objective criteria with cluster analysis for ranking suspect entities to evaluate suspicious medical claims[7]. The comprehensive evaluation model of reservoir resettlement was established by combining information entropy theory and fuzzy AHP[8].

But, evaluators are subjected to their personal experience and they are tend to take advantage of their interest in scoring process, resulting in unreasonable calculation of indicators. In this paper, we invited several evaluators to give their evaluation matrices. Teachers gave high emphasis on knowledge ability while academic staff preferred humanistic quality. If they help each other to advance their benefits, we called them stakeholders. In order to decrease the subjective effect from evaluators, we improved AHP based on stakeholder perspective (IAHP_SP).

#### B. The calculation of criteria’s comprehensive weight base on IAHP_SP

Let $A=[x_{ij}]_{m,n}$ be the criteria evaluation matrix, where $x_{ij} > 0 (i=1, 2, \ldots, m; j=1, 2, \ldots, n)$ is the score of $m$ factors’ given by $n$ evaluators.

We supposed $A$ has been min-max normalization processed by using:

$$x_{ij}^* = \frac{x_{ij} - m_j}{x_{ij} - M_j}$$

$M_j$ is the maximum value of the $j$-th criterion and $m_j$ is the minimum value of the $j$-th criterion.

**Definition 1:** Given $A$, the coefficient of the stakeholder $r_{ik}$ is given as follows:

$$r_{ik} = \frac{x_{ik} \cdot x_k}{\parallel x_i \parallel \parallel x_k \parallel}$$

where $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$, $x_k = (x_{k1}, x_{k2}, \ldots, x_{km})$, and $\parallel x_i \parallel$ denotes the norm of $x_i$.

**Definition 2:** For better describe the relationship between two evaluators, it is said to have a scalar $c_{ik}$ called net coefficient of the stakeholders, which is an expression of positive and negative effects to counteract each other.

$$c_{ik} = r_{ik} - (1-r_{ik}) = 2r_{ik} - 1$$

These evaluators not only maximize their own benefits, but also maximize the stakeholders’ benefits. These benefits are modeled as:

$$\max \sum_{k=1}^{n} c_{ik} \sum_{j=1}^{m} x_{ijk} w_{ji}$$

where $w_{ji}$ is the weight of criteria in perspective of the evaluator $u_i$.

**Lemma:** Criterion is assumed to be not dominant for the rest of the less important criteria, restrictive condition[11].

We defined $z_{ij} = \sum_{k=1}^{n} c_{ik} x_{ijk}$ and the model is described by maximizing:

$$\max \sum_{j=1}^{m} z_{ij} w_{ji}$$

subject to $0.5^{m-1} \leq w_{ji} \leq 0.5$ and $\sum_{j=1}^{m} w_{ji} = 1$.

**Output:** $W_i = (w_{i1}, w_{i2}, \ldots, w_{in})$
where $W_i$ is weight vector in perspective of the evaluator $u_i$.

Combining (6), (7) and (8), one can calculate the evaluation value matrix $Y$

$$y_{ki} = \sum_{j=1}^{m} x_{ji} W^*_{ji}$$  (6)

$$y_i = (y_{i1}, y_{i2}, \ldots, y_{in})^T$$  (7)

$$Y = (y_1, y_2, \ldots, y_n) = [y_{ki}]_{i,n} = 
\begin{bmatrix}
y_{11} & y_{12} & \cdots & y_{1n} 
y_{21} & y_{22} & \cdots & y_{2n} 
\vdots & \vdots & \ddots & \vdots 
y_{m1} & y_{m2} & \cdots & y_{mn}
\end{bmatrix}$$  (8)

Computing maximum eigenvalue $\lambda_{max}$ of $Y$, $y^* = (y_{11}^*, y_{12}^*, \ldots, y_{1n}^*)$ is the eigenvector corresponding to $\lambda_{max}$.

**Definition 3:** Let $\psi_i$ be the evaluators’ weight derived from $y^*$:

$$\psi_i = \frac{y_{i1}^*}{\sum_{i=1}^{n} y_{i1}^*}$$  (9)

subject to $\psi_i \geq 0$, $\sum_{i=1}^{n} \psi_i = 1$

Calculating global weight $W$ by formula $W = \psi \times U$, where $U$ is the weight of these criteria yielded by AHP method.

III. EXPERIMENT AND DISCUSSION

A. Dataset

We obtained our dataset from a university’s management systems, such as educational administration system data, scientific research management systems data and so on. The data have orderly structured and centrally classified. But it is inevitable that the data would be missing, multi-valued and invalid values. We have to preprocess the students’ data as follows:

- Deleting all data records for students who change their majors, drop out of school and cheat in the exams.
- Deleting duplicate data records in order to eliminate the interference caused by duplication of data records.
- Filling the corresponding supplementary examination results for students who are absent or applying for deferment of examinations. Delete all records of the students if there are no make-up results.

- For outlier values, choosing investigated data or mean value make up.
- For the student performance of multiple make-up and multiple reexaminations results, using the first valid results for the corresponding attribute value.

We chose three students in a class take an example in Table III. These scores come from the performance of the students participate in activities except Professional Courses’ Grade ($W_{21}$) and Elective Courses’ Grade ($W_{22}$) which are come from students’ academic grades.

<table>
<thead>
<tr>
<th>Number</th>
<th>Criteria</th>
<th>1402621</th>
<th>1402636</th>
<th>1402645</th>
</tr>
</thead>
<tbody>
<tr>
<td>W11</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>W12</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>W13</td>
<td>60</td>
<td>60</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>W14</td>
<td>60</td>
<td>60</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>W15</td>
<td>60</td>
<td>60</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>W21</td>
<td>90.1</td>
<td>89.63</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>W22</td>
<td>86</td>
<td>85</td>
<td>90.5</td>
<td></td>
</tr>
<tr>
<td>W23</td>
<td>60</td>
<td>70</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>W31</td>
<td>70</td>
<td>65</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>W32</td>
<td>60</td>
<td>70</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>W33</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>W34</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>W41</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>W42</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>W43</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

B. Experiment

![Model building flow chart](Fig. 2)
Firstly, we collected student dataset and evaluation matrices given by evaluators, then preprocessed these data. By determining the weight of each evaluator in the group, obtain the comprehensive weight of each criterion which weighted averages calculate with criteria weight by the evaluators scored.

We invited six evaluators to score the CQE system of college students. The group consists of teachers, students and academic staff.

The matrix A is calculated according to (2). And then we get the correlation coefficient \( r_{ik} \) between each evaluator:

\[
\begin{bmatrix}
1 & 0.5149 & 0.6985 & 0.8787 & 0.8355 & 0.8143 \\
0.5149 & 1 & 0.2221 & 0.6535 & 0.7937 & 0.7160 \\
0.6985 & 0.2221 & 1 & 0.7517 & 0.7562 & 0.8009 \\
0.8787 & 0.6535 & 0.7517 & 1 & 0.9388 & 0.9723 \\
0.8355 & 0.7937 & 0.7562 & 0.9388 & 1 & 0.9815 \\
0.8143 & 0.7160 & 0.8009 & 0.9723 & 0.9815 & 1 \\
\end{bmatrix}
\]

Calculate the net profit coefficient \( c_{ik} \) between each evaluator according to (3):

\[
\begin{bmatrix}
1 & 0.0299 & 0.3971 & 0.7574 & 0.6711 & 0.6286 \\
0.0299 & 1 & -0.5559 & 0.3070 & 0.5874 & 0.4320 \\
0.3971 & -0.5559 & 1 & 0.5034 & 0.5123 & 0.6019 \\
0.7574 & 0.3070 & 0.5034 & 1 & 0.8776 & 0.9445 \\
0.6711 & 0.5874 & 0.5123 & 0.8776 & 1 & 0.9629 \\
0.6286 & 0.4320 & 0.6019 & 0.9445 & 0.9629 & 1 \\
\end{bmatrix}
\]

According to (5) can obtain the value of the evaluator weight \( W_i \):

\[
w_{ij} = [0.375, 0.0625, 0.5, 0.0625]\]

Combining (6) (7) and (8) can obtain \( Y \) and \( y^* \):

\[
\begin{bmatrix}
0.4717 & 0.3720 & 0.3824 & 0.4717 & 0.4717 & 0.4717 \\
0.5 & 0.9375 & 0.5 & 0.5 & 0.5 & 0.5 \\
0.5260 & 0.0885 & 0.6562 & 0.5260 & 0.5260 & 0.5260 \\
0.5789 & 0.4368 & 0.4784 & 0.5789 & 0.5789 & 0.5789 \\
0.5086 & 0.4320 & 0.5310 & 0.5086 & 0.5086 & 0.5086 \\
0.7981 & 0.5884 & 0.7832 & 0.7981 & 0.7981 & 0.7981 \\
\end{bmatrix}
\]

\[
y^* = [0.3235, 0.4227, 0.3427, 0.3954, 0.3652, 0.5563]'
\]

According to (9) can obtain the value of the evaluator weight \( y_j \):

\[
y_j = (0.1345, 0.1757, 0.1425, 0.1644, 0.1518, 0.2312)
\]

The first-level criteria weight is the proportion of the first-level criterion in the student's comprehensive quality evaluation system, and the second-level criteria weight is the proportion of the second-level criterion in the corresponding first-level criterions. The comprehensive weight are the weighted average of the first-level criteria weight and the second-level criteria weight, which represents the proportion of each criterion in the entire evaluation system.

<table>
<thead>
<tr>
<th>Target</th>
<th>First-level Criteria</th>
<th>First-level Criteria Weight</th>
<th>Second-level Criteria</th>
<th>Second-level Criteria Weight</th>
<th>Comprehensive Criteria Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>W_1</td>
<td>0.1409</td>
<td>W_{11}</td>
<td>0.3278</td>
<td>0.0462</td>
<td></td>
</tr>
<tr>
<td>W_2</td>
<td>0.4554</td>
<td>W_{12}</td>
<td>0.3278</td>
<td>0.0462</td>
<td></td>
</tr>
<tr>
<td>W_3</td>
<td>0.2628</td>
<td>W_{13}</td>
<td>0.0610</td>
<td>0.0086</td>
<td></td>
</tr>
<tr>
<td>W_4</td>
<td>0.1409</td>
<td>W_{14}</td>
<td>0.1134</td>
<td>0.0160</td>
<td></td>
</tr>
<tr>
<td>W_5</td>
<td></td>
<td>W_{15}</td>
<td>0.1700</td>
<td>0.0239</td>
<td></td>
</tr>
</tbody>
</table>

Take a student numbered 1402621, his comprehensive grade is \( Y_{1402621} = [0.00462*80+0.00462*70+...+0.0141*60+0.0845*60=71.68.\) Take the class of the three students in Table III as an example, and sort data according to the “Weighted Average Score” column. The comparison results are shown in Table V. The three students are shown in bold.

There are 27 students in this class, and the proportion of the first, second, third prizes is 5%, 10% and 20%. There will be 1 first prize winner, 3 second prize winners, and 5 third prize winners. The weighted average score = “\( W_{11} * 80% + W_{12} * 20% \)”, “Award Prize-Score” according to the weighted evaluation score ranking, “Reward Prize-Comprehension” ranked by comprehensive quality score, the “Award Change” is a representation of the change in the prize.

It can be seen from Table V that two students numbered 1402636 and 1402647 have improved their scholarship levels, the ranking of the student numbered 1402633 has decreased. And the student numbered 1402625 has been reduced from third prize to no prize. The reason is because he/she got good academic scores while did not participate in any activities, resulting in lower comprehensive score. There is no change in the award level of the rest of the students.

After obtaining the student's quality scores in all aspects, the score was introduced into the radar chart of college student comprehensive quality. Taking the students in Table III as
examples, the quality scores of various aspects were introduced into the radar chart. The results are shown in Fig. 3. It can be seen that the student's Ideological-Mental Quality and Creative-Practice Quality of number 1402645 student are better than the other two students. The Knowledge Quality and Comprehensive Quality of student numbered 1402636 are better than the other two students. Three students scored the same in Humanistic Quality.

<table>
<thead>
<tr>
<th>Number</th>
<th>Weighted Average Score</th>
<th>Award Grade-Score</th>
<th>Compre hensive Quality Score</th>
<th>Award Grade-Comprehension</th>
<th>Award Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1402633</td>
<td>89.91</td>
<td>First Prize</td>
<td>70.71</td>
<td>Second Prize</td>
<td></td>
</tr>
<tr>
<td>1402621</td>
<td>89.28</td>
<td>Second Prize</td>
<td>71.68</td>
<td>Second Prize</td>
<td></td>
</tr>
<tr>
<td>1402636</td>
<td>88.71</td>
<td>Second Prize</td>
<td>72.61</td>
<td>First Prize</td>
<td></td>
</tr>
<tr>
<td>1402645</td>
<td>88.5</td>
<td>Second Prize</td>
<td>72.45</td>
<td>Second Prize</td>
<td></td>
</tr>
<tr>
<td>1402624</td>
<td>86.33</td>
<td>Third Prize</td>
<td>69.64</td>
<td>Third Prize</td>
<td></td>
</tr>
<tr>
<td>1402625</td>
<td>84.96</td>
<td>Third Prize</td>
<td>68.48</td>
<td>Third Prize</td>
<td></td>
</tr>
<tr>
<td>1402644</td>
<td>84.3</td>
<td>Third Prize</td>
<td>69</td>
<td>Third Prize</td>
<td></td>
</tr>
<tr>
<td>1402642</td>
<td>83.28</td>
<td>Third Prize</td>
<td>68.56</td>
<td>Third Prize</td>
<td></td>
</tr>
<tr>
<td>1402623</td>
<td>83.16</td>
<td>Third Prize</td>
<td>69.18</td>
<td>Third Prize</td>
<td></td>
</tr>
<tr>
<td>1402639</td>
<td>82.68</td>
<td></td>
<td>68.25</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402638</td>
<td>82.33</td>
<td></td>
<td>67.67</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402635</td>
<td>81.51</td>
<td></td>
<td>67.87</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402640</td>
<td>79.54</td>
<td></td>
<td>67.59</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402647</td>
<td>78.87</td>
<td></td>
<td>68.93</td>
<td>Third Prize</td>
<td>\</td>
</tr>
<tr>
<td>1402641</td>
<td>78.83</td>
<td></td>
<td>67.39</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402631</td>
<td>73.5</td>
<td></td>
<td>65.55</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402630</td>
<td>71.79</td>
<td></td>
<td>65.51</td>
<td>\</td>
<td>\</td>
</tr>
<tr>
<td>1402628</td>
<td>71.37</td>
<td></td>
<td>64.92</td>
<td>\</td>
<td>\</td>
</tr>
</tbody>
</table>

In the radar chart of college student comprehensive quality, students can find their own differences between other students in each aspect of quality, clarify their potential and advantages, and improve the deficiencies in order to become an all-roundly excellent college student. And the teacher can view the situation of each class of students and guide students to fill in the gaps. The higher educational departments can test the effectiveness of various quality education policies in order to formulate reasonable talent training plans and improve students' abilities in all aspects.

IV. RESULT

Evaluating students’ comprehensive quality performance is mostly useful to help the educators and students improving their teaching and learning process. This paper constructed a CQE model for university students. In evaluation process, we used IAHP_SP method to decrease the subjective of evaluators. Through experiment, it is proved that the weight distribution of the model is reasonable. The comprehensive quality score obtained from the model can deliver reference for education and teaching evaluation, such as class selection of outstanding student, audit evaluation, project certification and so on. It will help the educational apartments to evaluate the students’ comprehensive performance in a systemic way.

REFERENCES