Fuzzy Cluster Analysis of Undergraduates’ Employment Quality and Its Application

Yong-ping LIU
Faculty of Information Technology, Beijing University of Technology, Beijing, P.R. China

Keywords: Employment quality, Index system, analytic hierarchy process, fuzzy assessment, classified threshold.

Abstract. Assessment index reflects assessed contents and the choice of assessment method. A scientific assessment index system is in favor of exactly assessing the quality of talent cultivation, as well as guiding the enhancement of the quality. This paper intends to establish an index system of fuzzy assessment of undergraduates’ employment quality through the theoretical approach of management science and engineering and apply it into the guidance of undergraduates’ employment from Beijing University of Technology.

Introduction
Since China’s college entrance exam has been resumed in 1978, the scale of China’s colleges has been expanding geometrically. Especially since 1990’s, there formed a huge gap between talents’ supply from China’s colleges and social demand. In order to bridge the gap to balance the supply-demand relationship in highly educated talents’ market, the authority has embarked on reform of education system, with enlarging the scale of higher education and increasing the number of college admission becoming the reform’s priorities. In 1999 in particular, the implementation of college admission expansion policy has transformed higher education from elite mode to public mode. Confronted with the rapid outspreading of colleges’ scales, policy makers have to answer a question: whether the economic demand for highly educated talents and the talent supply from college has reached a balance? An intuitional index for answering this question is the employment situation of college undergraduates. Before college admission expansion in 1999, fresh graduates always had multiple employment choices; however, after the expansion in 1999, especially after 2001, Ministry of Education started to pay attention to undergraduates’ employment rate—their first employment rate basically remained at 75% or so, which meant that 25% fresh graduates were unemployable. It is worth noting that in 2008 and 2009, the first employment rate, affected by worldwide financial crisis, fell down below 70% for the first time—it remained at around 68%[1]. Thus, the grim picture for fresh graduates’ employment could be seen and key universities, as a result, had to consider the subsequent influence on their own enrollment exerted by the employment rate and employment quality.

Literature Review
In recent years, there are three aspects of assessing the quality of university’s talent cultivation.

Survey and Statistics
First, to assess via mathematical statistics according to survey and statistics. Wang Qin and Zhang Shulian combine qualitative research and quantitative analysis through factor analysis to assess the quality from the perspective of employee’s demand[2]; Hong Meixiang and He Zhenrui assess the quality via AHP from the perspectives of cultivation condition, process, and results[3]; Xue Wei assesses the quality with fuzzy and comprehensive assessment method[4]; Dai Ying, in line with employee’s actual demand for talents, builds up indexes from three dimensions of talents’ basic quality, competency, and potential to assess the quality by combining fuzzy and comprehensive assessment, data envelopment analysis, and hierarchy analysis[5].
Contents and Standards
Second, to assess by expert method or management model in accordance with the contents and standards of university’s talent cultivation quality. Li Zhonghua assesses the quality through expert method from the perspectives of talent cultivation’s ideas and conditions, events, and results [6]; Ma Wanmin analyzes the quality via talent cultivation assessment model from the influenced factors such as knowledge, capability, and quality[7]; Zhao Peihua and Hu Po assess the quality from the perspective of pluralist interaction by setting up a pluralistic expert team and quality assessment group[8]; Wu Haiying assesses the quality of talents from Engineering Cost Department via “PDCA” (put forward by Dai Minghuan) [9].

Other Methods
Third, some scholars apply other methods due to different perspectives. For instance, Wang Dongling and Shi Junxia apply input-output method[10]; Tang Liguo and Guo Qing assess the quality by taking output derived from learning as standard[11]; and Li Xingguo assesses the quality by evidence-reasoning method[12].

Further Study
To assess the quality via mathematical statistics can avoid subjectivity and objectively and impartially reflect the actual condition of talent cultivation, but the result could be affected due to accessibility of reliable statistics and all-round statistics. And expert method (an expert’s assessing experience and knowledge) and management model can scientifically and eventually assess the quality, which will be inevitably affected by expert’s subjectivity though. Hence, a scientific and reasonable combination of the two methods is a more effective one with the assessment cost could be influenced. From the perspective of adjusting to or meeting social demands, it is reasonable that the quality should be assessed by the third party. However, according to the basic standard of conducting quality-oriented education, promoting students’ all-round development and adjusting to the social demands, only when social assessment and the third party’s assessment are scientifically integrated can the quality be assessed exactly so as to constantly improve the quality.

Data Source of Employment Quality’s Fuzzy Cluster
Quality is the eternal theme and the lifeline of higher education’s development. Improving talent cultivation quality is the core issue of higher education’s development and the fundamental requirement of enhancing the international competitiveness of China’s higher education. So the standards for college talent cultivation quality and the idea of it become a focus of many experts and scholars [11]. This paper intends to establish an index system of fuzzy assessment of undergraduates’ employment quality through the theoretical approach of management science and engineering and apply it into the guidance of undergraduates’ employment from Beijing University of Technology (BJUT for short).

Index Set
In the light of aspects of undergraduate’ employment quality, this paper categorizes various elements that influence assessed objects into primary and secondary elements. The element set mentioned above is the index set to be assessed, which can be represented in the same manner as \( U = \{u_1, u_2, \ldots, u_m\} \). The set, consisting of multiple primary and secondary elements, indicates from which aspects to assess objects. The ultimate goal of this paper’s fuzzy assessment system is to feedback through undergraduate’ employment and to make comprehensive assessment on the employment quality of undergraduates from different majors (or from different Colleges). For that reason, contributions to general objective made by all basic elements needs to be considered.

In order to assess undergraduates’ employment quality scientifically, this paper, on the basis of referring a large amount of literature, puts forward a new set of assessment index system for the quality of undergraduates’ employment. This system consists of five first-class indexes: students’
satisfaction degree of employees, students’ satisfaction degree of employment guidance, preparation of each student’s job hunting, teaching contents’ practicability, and students’ satisfaction degree of teaching. Such a design aims to investigate the teaching and employment guiding of colleges and their Colleges before their undergraduates finish their study, consider these students’ subjective initiative during their job hunting, and assess the overall condition of employment quality from employees’ perspective. As a result, the assessment index system of employment quality designed by this paper gives consideration to the logicality, the timeliness, and the scientifi city of employment’s related work and process.

To exactly and easily collect and assess statistics, this paper classifies each first-class index into several second-class index, as shown in Tab.1. When selecting second-class indexes, this paper should consider if these second-class indexes belong with the first-class indexes’ concepts, and concepts of second-class indexes, when the indexes belong to the same first-class index, cannot repeat, which means each second-class index is a clear explanation of their first-class index in certain aspect.

Table 1. Remark set of assessment index of BJUT undergraduates’ employment quality.

<table>
<thead>
<tr>
<th>First-class index</th>
<th>First-class weights</th>
<th>Second-class index</th>
<th>Second-class weights</th>
<th>Ideal</th>
<th>Good</th>
<th>Average</th>
<th>Sub-standard</th>
<th>Disappointing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students’ satisfaction degree of employees</td>
<td>0.198</td>
<td>Salary</td>
<td>0.033</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Welfare</td>
<td>0.033</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Work environment</td>
<td>0.033</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relationships among colleagues</td>
<td>0.033</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Management level</td>
<td>0.033</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Promotion system</td>
<td>0.033</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td>Students’ satisfaction degree of employment guidance</td>
<td>0.247</td>
<td>University’s guidance</td>
<td>0.185</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td></td>
<td></td>
<td>College’s guidance</td>
<td>0.062</td>
<td>Quite satisfied</td>
<td>Satisfied</td>
<td>Average</td>
<td>Dissatisfied</td>
<td>Quite dissatisfied</td>
</tr>
<tr>
<td>Preparation of each student’s job hunting</td>
<td>0.198</td>
<td>Has the student made career plan?</td>
<td>0.101</td>
<td>The student has quite clear plans.</td>
<td>The student has clear plans.</td>
<td>The student has simple plans.</td>
<td>The student plays it by ear.</td>
<td>The student never makes plans.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How much does the student understand current employment’s policies and process?</td>
<td>0.026</td>
<td>The student understands them quite good.</td>
<td>The student understands them.</td>
<td>Average</td>
<td>The student Understands them a little.</td>
<td>The student doesn’t understand them at all.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How much does the student understand and pays</td>
<td>0.052</td>
<td>The student quite understands</td>
<td>The student understands it a little and</td>
<td>Average</td>
<td>The student doesn’t understand it</td>
<td>100</td>
</tr>
<tr>
<td>Understand and pay attention to employment service website?</td>
<td>The student knows them.</td>
<td>The student knows a little about them.</td>
<td>The student doesn’t know about them at all.</td>
<td>sometimes pays attention to it.</td>
<td>at all and pays no attention to it.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much does the student know about student unions of employment?</td>
<td>The student knows about them quite good.</td>
<td>The student sticks to creditability as much as possible.</td>
<td>It depends on different occasion.</td>
<td>It doesn’t matter to them.</td>
<td>The student stick to creditability as much as possible.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is the student creditable during job hunting?</td>
<td>The student sticks to creditability.</td>
<td>The student sticks to creditability.</td>
<td>It is unnecessary to be creditable.</td>
<td>It depends on different occasion.</td>
<td>It doesn’t matter to them.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching contents’ practicability</td>
<td>0.010</td>
<td>0.008</td>
<td>0.052</td>
<td>0.101</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge structure</td>
<td>0.007</td>
<td>Quite satisfactory</td>
<td>Satisfactory</td>
<td>Average</td>
<td>Unsatisfactory</td>
<td>Quite unsatisfactory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional knowledge</td>
<td>0.013</td>
<td>Quite satisfactory</td>
<td>Satisfactory</td>
<td>Average</td>
<td>Unsatisfactory</td>
<td>Quite unsatisfactory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic skills</td>
<td>0.026</td>
<td>Quite satisfactory</td>
<td>Satisfactory</td>
<td>Average</td>
<td>Unsatisfactory</td>
<td>Quite unsatisfactory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability to work independently</td>
<td>0.052</td>
<td>Quite satisfactory</td>
<td>Satisfactory</td>
<td>Average</td>
<td>Unsatisfactory</td>
<td>Quite unsatisfactory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability to deal with interpersonal relationship</td>
<td>0.101</td>
<td>Quite satisfactory</td>
<td>Satisfactory</td>
<td>Average</td>
<td>Unsatisfactory</td>
<td>Quite unsatisfactory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students’ satisfaction degree of teaching</td>
<td>0.198</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultivation of analyzing and solving problems</td>
<td>0.022</td>
<td>Quite good</td>
<td>Good</td>
<td>Average</td>
<td>Bad</td>
<td>Quite bad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultivation of creativity</td>
<td>0.005</td>
<td>Quite good</td>
<td>Good</td>
<td>Average</td>
<td>Bad</td>
<td>Quite bad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge broadening</td>
<td>0.005</td>
<td>Quite good</td>
<td>Good</td>
<td>Average</td>
<td>Bad</td>
<td>Quite bad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth and scope of professional knowledge</td>
<td>0.011</td>
<td>Quite good</td>
<td>Good</td>
<td>Average</td>
<td>Bad</td>
<td>Quite bad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultivation of basic skills and manipulative ability</td>
<td>0.041</td>
<td>Quite good</td>
<td>Good</td>
<td>Average</td>
<td>Bad</td>
<td>Quite bad</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formation of world and life views</td>
<td>0.075</td>
<td>Quite good</td>
<td>Good</td>
<td>Average</td>
<td>Bad</td>
<td>Quite bad</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Remark Set

Remark set is represented as Set $V$, i.e. $V = (v_1, v_2, \ldots, v_k)$, in which $k$ stands for the number of remarks and is actually a division of variable intervals of assessed objects. This paper divides the quality of undergraduates’ employment into five levels: disappointing, sub-standard, average, good, and ideal. Employment quality has been divided into intervals by remark set, but in actual counting, especially in questionnaire investigation, remarks of every index must be approachable and accord with context. Hence this paper resets remark set of every second-class index, as shown in Tab. 1.

Fuzzy Assessment Matrix

Each factor $u_i$ is counted as an assessment $f(u_i)$, so it can be regarded as fuzzy mapping $\tilde{f}$ from $U$ to $V$.

$$f : U \rightarrow \wp(V)$$  \hspace{1cm} (1)

$$u_i \mapsto f(u_i) \in \wp(V)$$  \hspace{1cm} (2)

From $\tilde{f}$, an induced fuzzy linear transformation $T_f$ from $U$ to $V$ can be regarded as the mathematic model of the comprehensive assessment $B$ resulted from weight $A$.

Consequently, fuzzy mapping $\tilde{f}$ can induce fuzzy relation $R_j \in \wp(U \times V)$, i.e. $R_j(u, v_h) = f(u)(v_h) = r_{nh}$, among which $R_j$ can be represented as fuzzy matrix $R \in \mu_{mok}$:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mk} \end{bmatrix}$$  \hspace{1cm} (3)

$R$ is called as single-factor assessment matrix, among which $r_{nh}$ means starting from factor $x_i$ and this assessed object can be counted as membership degree of $v_h$. To put it specifically, $r_{nh}$ stands for the frequency distribution of the $i$th factor $x_i$ in the $h$th assessment $v_h$, and the frequency distribution is generally normalized to meet the demand of $\sum_{i=1}^{k} r_{nh} = 1$. As such, matrix $R$ itself doesn’t have dimension so particular processing is unnecessary.

Original Data Matrix

To set domain $U = \{x_1, x_2, \ldots, x_n\}$ as the investigation objects of the graduates’ employment quality survey, and each object’s various aspects of employment is presented $m$ indexes, i.e. $x_j = (x_{j1}, x_{j2}, \ldots, x_{jm}) \ (j = 1, 2, \ldots, n)$. This paper has built a fuzzy matrix $R \in \mu_{mok}$ according to fuzzy relation $R_j \in \wp(U \times V)$, and the data of the matrix come from the feedback of every object. To make a cluster analysis on every object, this paper applies the following formula to convert $n$ fuzzy matrix $R$:

$$A(x_j) = R(x_j) \cdot F$$  \hspace{1cm} (4)

$A(x_j)$ stands for fuzzy assessment matrix of the $j$th object; $F$ converts and integrates fuzzy assessment’s score set, i.e. $F = (f_1, f_2, \ldots, f_k)$ which is a column vector with the value of $k$ equaling to that of assessment set and $f_h$ representing the assessment’s score of the $h$th level. With 100 as the full marks, through arithmetic scoring this paper can get $F = (100, 80, 60, 40, 20); A(x_j) \ (j = 1, 2, \ldots, n)$
stands for the index column vector of the \( j \)th object, of which the inter element \( a_i(x_j) = \sum_{h=1}^{k} f_{ih} \) \((i=1,2,\ldots,m)\). As described above, fuzzy matrix data and the data converting formula here can lead to original data matrix of fuzzy assessment’s cluster analysis of the employment quality.

This paper, in line with Formula (4), converts the membership degree of the employment quality into \( A(x_i) \), i.e. each college’s score of every index. And the scores are integrated into the original data matrix \( A \) of fuzzy cluster analysis.

To make the following cluster analysis easier, matrix \( A \) is transposed and \( A' = A^T \) is the original data matrix for following algorithm.

**Data Processing of Employment Quality’s Fuzzy Cluster**

In science and technology and economic management, data should be classified according to certain standards and principles. Scientific classification can make research more exact and operable. Classification of researched objects in the light of certain standards is called cluster analysis in mathematics. Cluster analysis is a method of clustering things according to their kinds in multivariate statistics analysis. In research of science and technology and economic management, indistinct-defined classification is not in favor of the research. Therefore, it is practical to apply fuzzy cluster method.

**Dimension Processing of Original Data**

In solving actual problems, to have different quantities of dimensions for standard comparison, data should be dealt with via dimensions in order to eliminate affects from distinct calibers. The dimensions here are what we call data standardization [15], the converting way of which is:

**Translation and Standard Deviation Conversion**

\[
x_i' = \frac{x_i - \bar{x}_j}{s_i}, (j = 1,2,\ldots,n; i = 1,2,\ldots,m) \tag{5}
\]

\[
\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}, s_i = \sqrt{\frac{1}{n} \sum_{j=1}^{n} (x_{ij} - \bar{x}_j)^2}
\]

Where,

After conversion, every variable’s value is 0 with 1 as standard deviation and dimensions’ affects have been eliminated. But \( x_{ik}' \) obtained in this way does not surely belong to interval \([0, 1]\).

**Translation and range conversion**

\[
x_i'^* = \frac{x_i' - \min_{1 \leq j \leq n} \{x_i'\}}{\max_{1 \leq j \leq n} \{x_i'\} - \min_{1 \leq j \leq n} \{x_i'\}}, (i = 1,2,\ldots,m) \tag{6}
\]

Evidently, \( 0 \leq x_i'^* \leq 1 \), and dimensions’ affects have been eliminated.

**Logarithm conversion**

\[
x_i' = \lg x_{ij}, (j = 1,2,\ldots,n; i = 1,2,\ldots,m) \tag{7}
\]

Selecting range to decrease orders of magnitude among variables.

In consideration of the data quality of matrix \( A^T \), this paper applies the method of translation and range conversion to standardize the data by standardizing matrix \( A^T \) of each department’s employment quality fuzzy assessment data. Thus, matrix \( A'^* \) is obtained.
Fuzzy Similar Matrix

In actual application, it is hard to establish a fuzzy equivalence relation or fuzzy equivalence matrix, the reason of which lies in the difficulty in reaching matrix’s transitivity. Yet it is easy to set up a self-reflective and symmetrical fuzzy matrix, i.e. fuzzy similar matrix. After being improved as a fuzzy equivalence relation or a fuzzy equivalence matrix that can reach transitivity and also keep self-reflexivity and symmetry, this matrix can be assessed through fuzzy cluster analysis [15]. This paper sets domain $U = \{x_1, x_2, \ldots, x_n\}$, $x_j = \{x_{j1}, x_{j2}, \ldots, x_{jm}\}$, and determines the similarity of $x_j$ and $x_k$ by coefficient method $r_{jk} = R(x_j, x_k)$.

**Related coefficient method**

\[
r_{jk} = \frac{\sum_{i=1}^{m} |x_{ji} - \bar{x}_j||x_{ki} - \bar{x}_k|}{\sqrt{\sum_{i=1}^{n} (x_{pi} - \bar{x}_j)^2} \cdot \sqrt{\sum_{i=1}^{n} (x_{pi} - \bar{x}_k)^2}}
\]

(8)

Where, \(\bar{x}_j = \frac{1}{m} \sum_{i=1}^{m} x_{ji}\), \(\bar{x}_k = \frac{1}{m} \sum_{i=1}^{m} x_{ki}\),

**Exponential similarity coefficient method**

\[
r_{jk} = \frac{1}{m} \sum_{i=1}^{m} \exp \left( -\frac{3}{4} \frac{(x_{pi} - x_{pi})^2}{s_i^2} \right)
\]

(9)

Where, \(s_i = \frac{1}{n} \sum_{j=1}^{n} (x_{pj} - \bar{x}_j)^2\), \(\bar{x}_j = \frac{1}{n} \sum_{j=1}^{n} x_{pj}\) (\(i = 1, 2, \ldots, m\)).

In related coefficient method, $m$ coordinates of $x_j = (x_{j1}, x_{j2}, \ldots, x_{jm})$ are from $m$ samples of the same $X_j$ with $r_{jk}$ standing for the similarity between $X_j$ and $X_k$. When reflected in original data matrix, different rows from different parent matrixes. In exponential similarity coefficient method, $x_1, x_2, \ldots, x_n$ are selected from $n$ samples of $m$ dimensions of the same $m$-dimension dimension parent matrix $X = (x_1, x_2, \ldots, x_n)$. So $r_{jk}$ reflects the similarity between two samples. When reflected in original data matrix, different columns from different parent matrixes while different columns are from the same parent matrix. Therefore, exponential similarity coefficient method is the only choice of the two above algorithm to describe the similarity between two samples. [16]
Fuzzy Equivalence Matrix

Fuzzy similarity matrix $R$ based on standards does not surely possess transitivity, which means $R$ is not definitely fuzzy equivalence matrix. For the sake of classification, $R$ should be converted into fuzzy equivalence matrix $R^\prime$. When $R \in \mu_{[0,1]}$ is set as fuzzy equivalence matrix, there exists a minimal natural number $i (i \leq n)$; therefore transitive closure $t(R) = R^\prime$ always has the equation $R^\prime = R^i = t(R)$ with the condition that any natural number $l$ that is greater than $i$. As a result, $t(R)$ is a fuzzy equivalence matrix. Based on the above theorem and algorithm, the transitive closure $t(R)$ of $R$ can be obtained via successive square to convert fuzzy similarity matrix into fuzzy equivalence matrix, which indicates successive calculation of $R \rightarrow R^2 \rightarrow R^4 \rightarrow \cdots R^{2^n}$ until $R^{2^k} = R^{2^{k+1}}$. Consequently, $R^{2^k} = R^{2^{k+1}}$ is fuzzy equivalence matrix $R^\prime$, i.e. $t(R) = R^\prime$.

Result Analysis of Employment Quality Fuzzy Cluster

The Best Threshold

In fuzzy cluster analysis, different $\lambda \in [0,1]$ leads to different classifications that form dynamic clusters. Yet in dealing with problems, it is necessary to choose proper threshold $\lambda$ in line with current situations, so that is a question of how to determine threshold $\lambda$. In order to guarantee the scientificity and the accuracy of classification analysis, this paper takes advantage of statistical magnitude $F$ to determine classified threshold $\lambda$.

When domain $U = \{x_1, x_2, \ldots, x_n\}$ is set as sample space with the total number of samples being $n$, and each sample $x_j$ has $m$ features, which results from experiment or observation: $x_j = (x_{j1}, x_{j2}, \ldots, x_{jm})$ ($j = 1, 2, \ldots, n$), the original data matrix can be got, which is called central vector of general samples.

Number of category corresponding to $\lambda$ is set as $r$, and the number of category of the $k$ th category is $n_k$, so the sample of the $k$ th category is: $x_{k1}, x_{k2}, \ldots, x_{kn_k}$. The cluster center of the $k$ th category is vector $\bar{x}^{(k)} = (\bar{x}_1^{(k)}, \bar{x}_2^{(k)}, \ldots, \bar{x}_m^{(k)})$, in which $\bar{x}_i^{(k)}$ is the average value of the $i$ th feature.
\[ x_{k}^{(i)} = \frac{1}{n} \sum_{j=1}^{n} x_{j}^{(i)}, \quad (k = 1, 2, \ldots, m) \] (10)

Statistical magnitude \( F \) is introduced with the formula as:

\[
F = \frac{\sum_{k=1}^{r} \sum_{i=1}^{n} n_k \|x_{k}^{(i)} - \overline{x}\|^2 / (r-1)}{\sum_{k=1}^{r} \sum_{i=1}^{n} n_k \|x_{k}^{(i)} - \overline{x}^{(i)}\|^2 / (n-r)}
\] (11)

Where, \( \|x_{k}^{(i)} - \overline{x}\|^2 \) represents the distance between \( x_{k}^{(i)} \) and \( \overline{x} \), and \( \|x_{k}^{(i)} - \overline{x}^{(i)}\|^2 \) means the distance between the \( j \)th sample \( x_{j}^{(k)} \) and its center \( \overline{x}^{(k)} \). Formula (11) is statistical magnitude \( F \), which distribution \( F \) complying with free degree \( r-1, n-r \). Molecule measures the distance between categories while denominator measures the distance between samples within categories. Therefore, the bigger the value of \( F \) is, the larger the distance between categories is, and the more different between categories, the better categories are.

If \( F > F_{\alpha}(r-1, n-r) \) (\( \alpha = 0.05 \)), the differences between categories will be obvious based on mathematical statistics variance analysis theory, indicating that categories are reasonable. If there are more than one value of \( F \) that meet the demand of inequation \( F > F_{\alpha}(r-1, n-r) \), the values of difference proportion can be compared to select one \( F \) that meets the requirements among \( F \)s of the larger value.

This paper, in accordance with original data matrix \( A' \), calculates the values of statistical magnitude \( F \) corresponding to part of categories, relevant critical value \( F_{\alpha}(\alpha = 0.05) \), and relative difference \( (F - F_{\alpha}) / F_{\alpha} \).

**Cluster Results**

This paper, after giving an overall consideration to \( F, F_{\alpha}, (F - F_{\alpha}) / F_{\alpha} \) and the numbers of categories, select 7 as the number of categories for further cluster analysis, and the threshold \( \lambda \) corresponding to exponential similarity coefficient method is 0.275. The cluster results are shown in Tab.2.

<table>
<thead>
<tr>
<th>Fuzzy clusters</th>
<th>Names of colleges</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first category</td>
<td>College of Material, College of Environment and Energy, College of Experiment, and College of Foreign Languages and Literature</td>
</tr>
<tr>
<td>The second category</td>
<td>College of Electronic Control and College of Computing</td>
</tr>
<tr>
<td>The third category</td>
<td>College of Mechanical and Electrical Technology, College of Architectural Planning, College of Civil Engineering, and College of Humanities</td>
</tr>
<tr>
<td>The fourth category</td>
<td>College of Economics and Management, College of Arts, and College of Life Science</td>
</tr>
<tr>
<td>The fifth category</td>
<td>College of Software</td>
</tr>
<tr>
<td>The sixth category</td>
<td>College of Mathematics and Physics</td>
</tr>
<tr>
<td>The seventh category</td>
<td>College of Transportation</td>
</tr>
</tbody>
</table>

In terms of the cluster results of BJUT, if several colleges are in the same category, it indicates that the employment qualities of these colleges have similarities, and then it is convenient to analyze...
common problems in students’ satisfaction degree of employees, students’ satisfaction degree of employment guidance, preparation of each student’s job hunting, teaching contents’ practicability, and students’ satisfaction degree of teaching, or to promote these colleges’ positive experience. As a result, next research needs to view colleges of the same category as an entire entity, which will reduce 16 colleges into 7 college groups; thus, statistical analysis of each index results in the discovery of critical points that demand attention concerning BJUT’s employment.

Common Problems

The investigation has assist the paper to approximately summarize three common problems which basically reflect employment issues confronting every college.

Current employment policy increases the proportion of flexible employment

Most of graduates expect to find a job in Beijing. However, as the policy of Beijing household register has been tightened up and other policies are faced with large fluctuations, graduates who do not come from Beijing have more difficulty in finding a job in Beijing. Consequently, the progress of contract settlement and has delayed and postponed, the proportion of flexible employment has been rising, and the problems of contract breaking and the detaining of household register and personal files has been increasing.

Graduates’ reaction to employment market cannot catch up with changes of the market

In 2015, both the number and the quality of posts for employers in Beijing have been decreasing. The authority, local governments, and universities’ administrators are all anxious about it, so they have taken all kinds of measures to create employment conditions for graduates to find jobs. However, graduates still don’t take the thing seriously and they just fall in with the crowd which leads them to lose lots of opportunities. Therefore, there should be more propaganda and lectures about graduates’ employment.

Graduates’ awareness of employment and entrepreneurship should be strengthened since they are freshmen.

Most students have no such a kind of awareness, and still, teachers from some colleges deem that employment and entrepreneurship education are only related to graduates or juniors. So the whole university has not reached agreement in employment and entrepreneurship education. To address this issue, it needs propaganda and guidance, training, and system construction.

References


