Developing an Online Game-Based Index System for Identifying Economically Disadvantaged College Students using Hesitant Fuzzy Attribute Reduction

Wenbin Yang[1*] and Lan Qiu[2]

[1]College of Artificial Intelligence, Jiangxi University of Technology, Nanchang, Jiangxi, China
yangwenbin2319@163.com
[2]College of Information Engineering, Jiangxi University of Technology, Nanchang, Jiangxi, China
hello.lanqiu@outlook.com

Corresponding author: Wenbin Yang, yangwenbin2319@163.com

Abstract. In the "post poverty alleviation era", the high-quality development of college student aid depends on the accurate judgement of the poverty status of students, but the academic community lacks the ability to scientifically identify relatively poor student groups. It is necessary to comprehensively consider family economic factors, special group factors, unexpected situations and students' daily consumption, and some indicator attribute values may be missing in the process of identifying poor college students. To objectively and quantitatively identify the poverty status of poor college students, this paper uses the hesitation fuzzy analysis method to construct an indicator system for identifying poor college students based on incomplete linguistic hesitant fuzziness and proposes an attribute reduction method based on the limited advantage relationship, which reduces the attributes of the identification indicator system for poor college students, selects useful indicators, and eliminates redundant indicators. Finally, the example proves that the method proposed in this paper can flexibly handle the diversity and incompleteness of identification indicators to obtain a simplified and scientific identification indicator system.

Keywords: Hesitant Fuzzy; Attribute Reduction; Identifying Poor College Students; Online Game-Based
DOI: https://doi.org/10.14733/cadaps.2024.S5.157-169

1 INTRODUCTION

Under the standards of 2020, China plans to lift all of its rural poor out of poverty. After solving absolute poverty, China will usher in the "post-poverty alleviation era". The crux of the poverty reduction strategy is to solve the relative poverty problem on the basis of consolidating the achievements of poverty alleviation [9]. Poverty alleviation through education is a key path to solving relative poverty and an important means to consolidate the poverty alleviation effect and comprehensively promote rural revitalization in the "post-poverty alleviation era" [17]. With the development of economy, society and higher education, China attaches great importance to the support of poor students. However, how to accurately identify poor college students and objectively calculate the degree of poverty has become a hot and difficult issue in the identification of poor college students for a long time, which has greatly affected the fairness and impartiality of aid to poor students [12].

A large number of education and teaching experts and scholars have conducted extensive research on how to scientifically formulate identification indicators and determine poverty levels and have provided various evaluation systems and methods. [16] developed an accurate measurement system for the poverty index of poor college students based on big data. This system has strong reliability, helps to achieve quantification in the identification process of poor students, and enhances the efficiency of the identification process and the accuracy of the identification results. [4] and others proposed that the index weight should be objectively determined according to the entropy method. They argued the system should be designed by combining entropy model analysis with manual review. This system can help colleges and universities effectively solve the problem that the identification of poor students is traditionally difficult to quantify and not objective. [15] and others revealed through empirical analysis of data simulation that the income method, expenditure method, income expenditure double index difference method and ratio method have certain limitations in the identification of poor college students, which shows that various methods have led to a certain rate of omission, leakage and miscalculation to varying degrees. Adjustment of the relative poverty line is expected to improve the recognition accuracy of the dual indicator ratio method. In the process of determining the degree of poverty of students with financial difficulties, considering the uncertainty of objective factors and the fuzziness of people's ideas and other factors, schools often cannot use accurate numbers to evaluate the indicators in the evaluation system. In these kinds of scenarios, fuzzy numbers can better describe the nature of things. Therefore, scholars have widely applied the fuzzy analysis method to the process of poverty identification and obtained a large number of evaluation methods based on fuzzy numbers. [8] constructed an identification index system of poor students from both personal and family conditions and applied the fuzzy comprehensive evaluation method to the identification of poor students to improve the accuracy and objectivity of the identification of poor students. [3] and others applied an improved a priori algorithm based on rough sets in the identification of poor students in colleges and universities, thus enhancing the accuracy of the identification of students with difficulties. [19] introduced the fuzzy analysis of linguistic hesitant into the identification of poor college students and established an incomplete fuzzy mixed information system of linguistic hesitant. the focus shifts to the practical implementation of the index system. It outlines the technical requirements, user interface design, and gaming elements incorporated into the system to engage and motivate students. The section also discusses the integration of data analytics and machine learning techniques to enhance the accuracy and efficiency of the identification process.

However, in the process of identifying poor college students, not all indicators can be quantified with general fuzzy numbers. If some indicators are quantified with general fuzzy numbers, they may distort the subjectivity of the indicators themselves, leading to untruthfulness of the results. In addition, in the identification index system of poor college students, there are many kinds of attribute values of each index, including information that students have not provided or is missing. This paper adopts the method of hesitation fuzzy analysis to build an indicator system for identifying poor college students based on incomplete linguistic hesitant fuzzy analysis to solve the problems of diversity of identification indicators and lack of attribute values. In addition, we propose an attribute reduction method based on the limited advantage relationship to reduce the attributes of the identification index system of poor college students to obtain a simplified and scientific identification index system.

2 HESITANT FUZZY ANALYSIS METHOD

The hesitation fuzzy analysis proposed by [10] is an improvement of general fuzzy analysis. For the indicator attributes in the evaluation process that are difficult to express with quantitative values, linguistic hesitant fuzzy analysis uses the qualitative natural language form of linguistic hesitant fuzzy sets as the attribute values, which are more flexible and practical than traditional fuzzy analysis in dealing with uncertain indicators.

Decision experts adopt appropriate linguistic evaluation scales when evaluating indicator attributes. For the convenience of discussion, Delgado M(1993)introduces the concept of the linguistic terminology set $\boldsymbol{S} = \{s_0, \cdots, s_{t-1}\}$, which satisfies the following conditions: If α>β then sα> sβ, when t is an uneven number.

By combining hesitant fuzzy sets with linguistic term sets, Meng et al.(2014) proposed the concept of linguistic hesitant fuzzy sets (LHFSs) and a comparison method between two linguistic hesitant fuzzy sets.

Definition 1 $\boldsymbol{S} = \{s_0, \cdots, s_{t-1}\}$ is the linguistic terminology set, and t is an uneven number. $LH = \left\{\left(s_{\theta(i)}, lh\left(s_{\theta(i)}\right)\right) \middle| s_{\theta(i)} \in \boldsymbol{S}\right\}$ is a linguistic hesitation fuzzy set. $lh\left(s_{\theta(i)}\right) = \{r_1, r_2 \cdots r_{m_i}\}$ indicates that element $s_{\theta(i)}$ belongs to all possible membership degrees of set LH, $r_1, r_2 \cdots r_{m_i} \in [0,1]$. LHF is used to represent all linguistic hesitant fuzzy sets.

Definition 2 $LH = \left\{\left(s_{\theta(i)}, lh\left(s_{\theta(i)}\right)\right) \middle| s_{\theta(i)} \in \boldsymbol{S}\right\}$ is a linguistic hesitant fuzzy set.

$$index(LH) = \left\{\theta(i) \middle| \left(s_{\theta(i)}, lh\left(s_{\theta(i)}\right)\right) \in \boldsymbol{LH}, lh\left(s_{\theta(i)}\right) \neq \{0\}\right\} \tag{1}$$

$E(LH) = s_{\mathrm{e}(LH)}$is the expected function of LH, as shown in Formula 2.$D(LH) = s_{\mathrm{v}(LH)}$ is the variance function of LH, as shown in Formula 3.

$$e(LH) = \frac{1}{|index(LH)|}\left(\sum_{\theta(i)\in index(LH)} \frac{\theta(i)}{\left|lh\left(s_{\theta(i)}\right)\right|}\left(\sum_{r\in lh\left(s_{\theta(i)}\right)} r\right)\right) \tag{2}$$

$$v(LH) = \frac{1}{|index(LH)|}\left(\sum_{\theta(i)\in index(LH)}\left(\frac{\theta(i)}{\left|lh\left(s_{\theta(i)}\right)\right|}\left(\sum_{r\in lh\left(s_{\theta(i)}\right)} r\right) - e(LH)\right)^2\right) \tag{3}$$

$\left|lh\left(s_{\theta(i)}\right)\right|$ and $|index(LH)|$ are cardinal numbers of $lh\left(s_{\theta(i)}\right)$ and $index(LH)$, respectively.

Definition 3 LH1 and LH2 are two linguistic hesitation fuzzy sets, so the size relationships between LH1 and LH2 are shown in Figure 1.

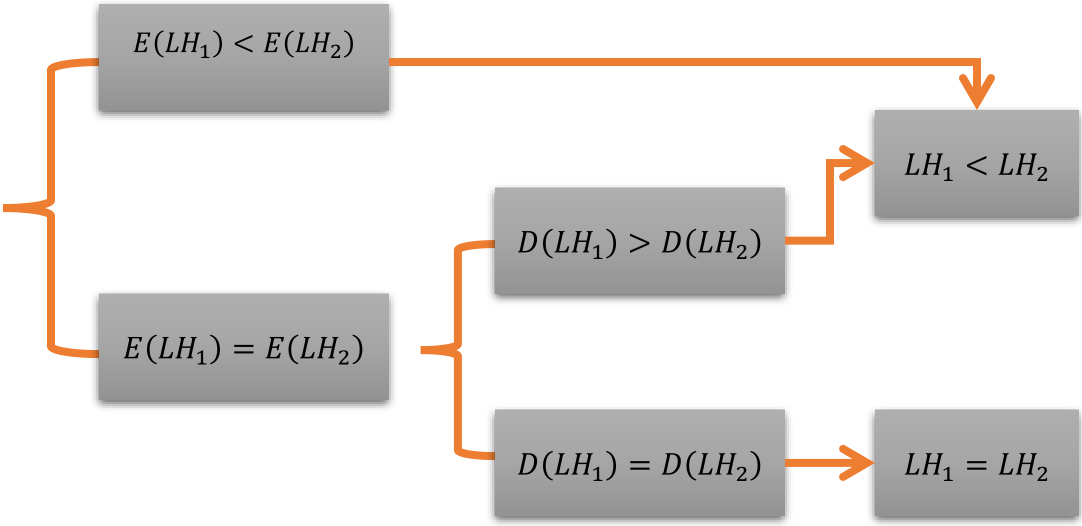

Figure 1: The size relationships between LH1 and LH2.

1.If $E(LH_1) < E(LH_2)$, then $LH_1 < LH_2$.

2.If $E(LH_1) = E(LH_2)$, then

a.If $D(LH_1) > D(LH_2)$ then $LH_1 < LH_2$.

b.If $D(LH_1) = D(LH_2)$ then $LH_1 = LH_2$.

For the convenience of discussion, this paper assumes that

$$LH_1 \leq LH_2 \Leftrightarrow LH_1 < LH_2 \vee LH_1 = LH_2$$

To facilitate calculation and avoid information loss, [13] defined an extended term set $\bar{\boldsymbol{S}} = \{s_\alpha | \alpha \in [0, t-1]\}$ based on the original term set $\boldsymbol{S} = \{s_0, \cdots, s_{t-1}\}$. In addition, we stipulate that for any $s_\alpha, s_\beta \in \bar{\boldsymbol{S}}$, if α>β then sα>sβ. In fact, other types of language term sets can also be set for different decision environments. For example, [14] proposed that decision-makers can set a set of language terms $\tilde{\boldsymbol{S}} = \{s_i | i = 1,2,\cdots,l\}$ for qualitative measurement, where l is odd, and if i>j then si>sj.

Therefore, in this paper, $LH = \left\{\left(s_{\theta(i)}, lh\left(s_{\theta(i)}\right)\right) \middle| s_{\theta(i)} \in \bar{\boldsymbol{S}}\right\}$ and $LH = \left\{\left(s_{\theta(i)}, lh\left(s_{\theta(i)}\right)\right) \middle| s_{\theta(i)} \in \tilde{\boldsymbol{S}}\right\}$ are both linguistic hesitant fuzzy sets. The comparison method of linguistic hesitant fuzzy sets in Definition 3 is still applicable.

3 ATTRIBUTE REDUCTION METHOD BASED ON THE RESTRICTIVE DOMINANCE RELATION

Rough set theory was first proposed by Polish scholar Pawlak(1982), which was used to deal with fuzziness in mathematics. It proposes the idea of attribute reduction, which can simplify problems and generate decisions without changing the classification ability. Attribute reduction is one of the key contents of rough set theory, which is based on binary relations. The classical binary relations are equivalence relations. Later, to meet the needs of practical problems, they were expanded to similarity relations, fault tolerance relations, dominance relations, etc. [1],[18],[5]Based on the attribute reduction method in rough set theory, this paper introduces the limited advantage relationship into the identification index system of poor college students with incomplete linguistic hesitant fuzziness and uses the discrimination matrix to obtain the attribute reduction of the poverty identification index system so that the subsequent poverty degree identification process can be optimized.

4 FUZZY DECISION SYSTEM WITH INCOMPLETE LINGUISTIC HESITANT

A linguistic hesitant fuzzy decision system can be defined as a quadruple (U, AT, V, f), as shown in Figure 2.

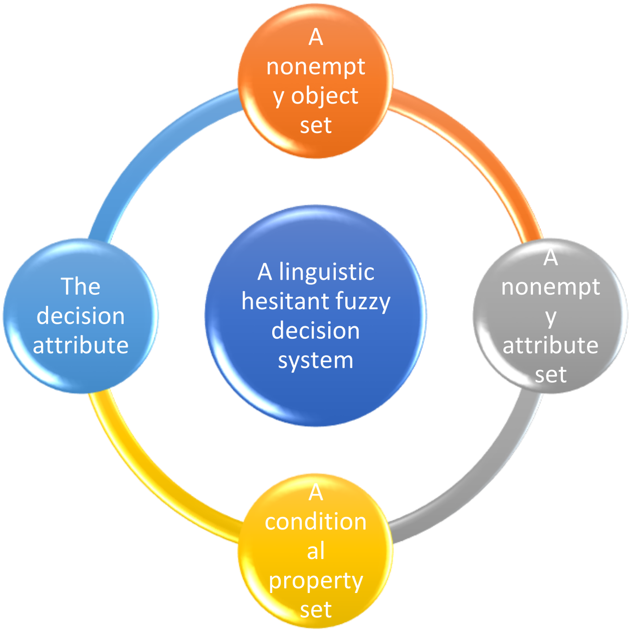

Figure 2: A linguistic hesitant fuzzy decision system.

where $\boldsymbol{U} = \{x_1, x_2 \cdots, x_n\}$ is a nonempty object set, $\boldsymbol{AT} = C \cup \{d\}$ is a nonempty attribute set, and $\boldsymbol{C} = \{c_1, c_2 \cdots, c_n\}$ is a conditional property set. d is the decision attribute, and V is the property value range. $f: U \times AT \to V$ is the information function.

$$V_c = \cup V_{c_j}, V_{c_j} = \{f(x_i, c_j) | i = 1,2,\cdots,n\} \tag{4}$$

$V_{c_j} \subseteq LHF \cup R$ or $f(x_i, c_j)$ are language terms. $f(x_i, d)$ is a language term. $V_d = \{f(x_i, d) | i = 1,2,\cdots,n\}$, $V = V_c \cup V_d$.When the conditional attribute values of some objects are unknown, the information system is called an incomplete linguistic hesitation fuzzy decision system, and the unknown attribute is represented by "*".

5 DOMINANCE RELATIONS OF FUZZY DECISION SYSTEMS WITH INCOMPLETE LINGUISTIC HESITATION

Based on the limited advantage relationship proposed by[20]and the specific background of identifying poor students, the advantage relationship of an incomplete linguistic hesitant fuzzy decision-making system is given.

Definition 4 (U, $C \cup \{d\}$, V, f) is an incomplete linguistic hesitant fuzzy decision system. For any $B \subseteq C$, the limiting advantage relationship determined by B is shown in Formula 5.

$$\tilde{\boldsymbol{R}}_{\boldsymbol{B}}^{\boldsymbol{L}} = \{x, y \in U^2 : \forall b \in B, (f(y,b) \leq f(x,b)) \vee (f(x,b) = maxV_b \wedge f(y,b) = *) \vee (f(x,b) = * \wedge f(y,b) = minV_b)\} \cup I_U \quad (5)$$

where $maxV_b = \{v \in V_b : \forall \bar{v} \in V_b, \bar{v} \leq v\}$, $minV_b = \{v \in V_b : \forall \bar{v} \in V_b, v \leq \bar{v}\}$, V_b is the set of all known attribute values of attribute b, $I_U = \{(x,x), x \in U\}$.

Under this limiting advantage relationship, the restrictive advantage class of x with respect to B is shown in Formula 6, where $\forall x \in U_x$.

$$[x]_B^L = \{y \in U : (y,x) \in \tilde{R}_B^L\} \quad (6)$$

The definition of the restricted dominance relation and restricted dominance class obviously has the following properties.

Property 1 (U, $C \cup \{d\}$, V, f) is an incomplete linguistic hesitant fuzzy decision system. If $B \subseteq C$, then

1.$\tilde{\boldsymbol{R}}_{\boldsymbol{C}}^{\boldsymbol{L}} \subseteq \tilde{\boldsymbol{R}}_{\boldsymbol{B}}^{\boldsymbol{L}}$

2. For any $x \in U$, there is $d_B(x) \leq d_C(x)$.

Definition 5 (U, $C \cup \{d\}$, V, f) is an incomplete linguistic hesitant fuzzy decision system. If $B \subseteq C$, then

1.For any $x \in U$, there is $d_B(x) = d_C(x)$.

2. For any $B' \subset B$, there is $x \in U$, so that $d_{B'}(x) \neq d_C(x)$. Then, B is called a reduction of (U, $C \cup \{d\}$, V, f).

Definition 6 (U, $C \cup \{d\}$, V, f) is an incomplete linguistic hesitant fuzzy decision system.

$$\tilde{R}_C(x,y) = \begin{cases} \{c \in C \mid y \notin [x]_{\{c\}}^L\}, d_C(x) \nleq f(y,d) \\ C, otherwise \end{cases} \quad (7)$$

The matrix $\tilde{R}_C = \left(\tilde{R}_C(x,y) \middle| x, y \in U\right)$ is the resolution matrix of the decision system.

Theorem 1 (U, $C \cup \{d\}$, V, f) is an incomplete linguistic hesitant fuzzy decision system. If $B \subseteq C$, then B is a reduction of the decision system. $\Leftrightarrow$ For any $x, y \in U$, $B \cap \tilde{R}_C(x,y) \neq \phi$.

The proof process of Theorem 1 is shown in Figure 3.

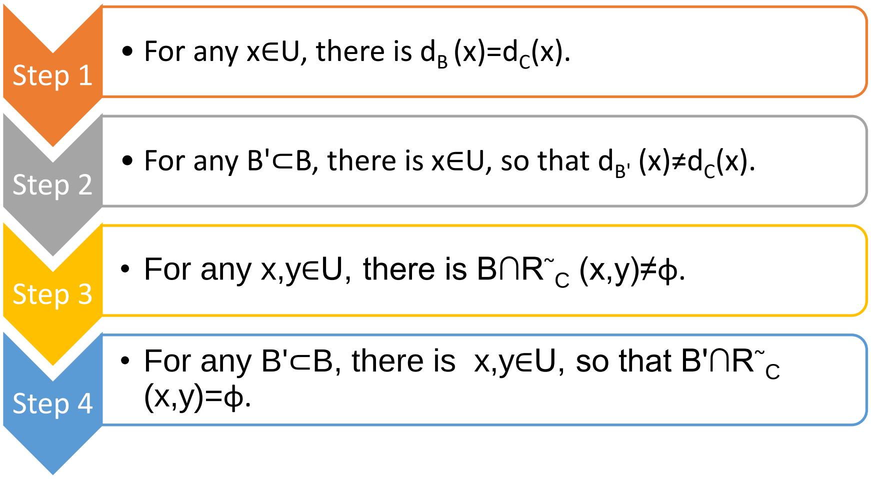

Figure 3: The proof process of Theorem 1.

Proof (1) " ⇐": For any $x \in U$, there is $d_B(x) = d_C(x)$.

By Property 1, it is only necessary to prove that for any $x \in U$, there is $d_C(x) \leq d_B(x)$. If it is not true, then there is $\bar{x} \in U$ such that $d_C(\bar{x}) \nleq d_B(\bar{x})$. So there is $y \in [x]_B^L$ such that $d_C(\bar{x}) \nleq f(y,d)$. At this time, because $B \cap \tilde{R}_C(\bar{x},y) \neq \phi$, so there is $b \in B \cap \tilde{R}_C(\bar{x},y)$. According to the definition of $\tilde{R}_C(\bar{x},y)$, $y \notin [x]_{\{b\}}^L$ and $y \in [x]_B^L$ are contradictory. So the assumption is wrong and for any $x \in U$, there is $d_C(x) \leq d_B(x)$.

Proof: For any $B' \subset B$, there is $x \in U$, so that $d_{B'}(x) \neq d_C(x)$.

For any $B' \subset B$,from the conditions: there is $x, y \in U$, so that $B' \cap \tilde{R}_C(x,y) = \phi$.So $d_C(x) \nleq f(y,d)$. For any $b \in B'$, there is $y \in [x]_{\{b\}}^L$,so $y \in [x]_{B'}^L$.Therefore, $d_{B'}(x) \leq f(y,d)$ and $d_{B'}(x) \neq \ d_C(x)$.

Proof (2) " ⇒" : B is a reduction of (U, $C \cup \{d\}$, V, f). For any $x, y \in U$, there is $B \cap \tilde{R}_C(x,y) \neq \phi$. There are 2 cases:

a. If $d_C(x) \leq f(y,d)$,then $\tilde{R}_C(x,y) = C$.So $B \cap \tilde{R}_C(x,y) \neq \phi$.

b. If $d_C(x) \nleq f(y,d)$, using the method of contradiction. Assume $B \cap \tilde{R}_C(x,y) = \phi$, for any $b \in B$, there is $y \in [x]_{\{b\}}^L$ that is $y \in [x]_B^L$. Then, $d_B(x) \leq f(y,d)$ and $d_B(x)$=$d_C(x) \nleq f(y,d)$ are contradictory. Therefore, the assumption is wrong. At this time, there is still $\cap \tilde{R}_C(x,y) \neq \phi$.

Proof: For any $B' \subset B$, there is $x, y \in U$, so that $B' \cap \tilde{R}_C(x,y) = \phi$.

For any $B' \subset B$,from the conditions: there is $x \in U$, so that $d_{B'}(x) \neq d_C(x)$. From Property 1, we can get $d_{B'}(x) < d_C(x)$. So there is $y \in [x]_B^L$, so that $(y,d) < d_C(x)$. According to the definition of $\tilde{R}_C(\bar{x},y)$, $B' \cap \tilde{R}_C(x,y) = \phi$.

6 THE IDENTIFICATION INDEX SYSTEM OF POOR COLLEGE STUDENTS BASED ON INCOMPLETE LINGUISTIC HESITATION AND FUZZINESS

To further improve the accuracy of student funding, in October 2018, the Ministry of Education and six other departments issued the Guidance on Doing a Good Job in Identifying Students from Economically Difficult Families. This paper refers to the guidance, through consulting relevant materials and expert opinions, following the principle of comprehensive, feasible, qualitative and quantitative combination, taking into account family economic factors, special group factors, emergencies and students' daily consumption outlook, etc., and constructs an identification index system of poor college students based on incomplete linguistic hesitant fuzziness, as shown in Figure 4.

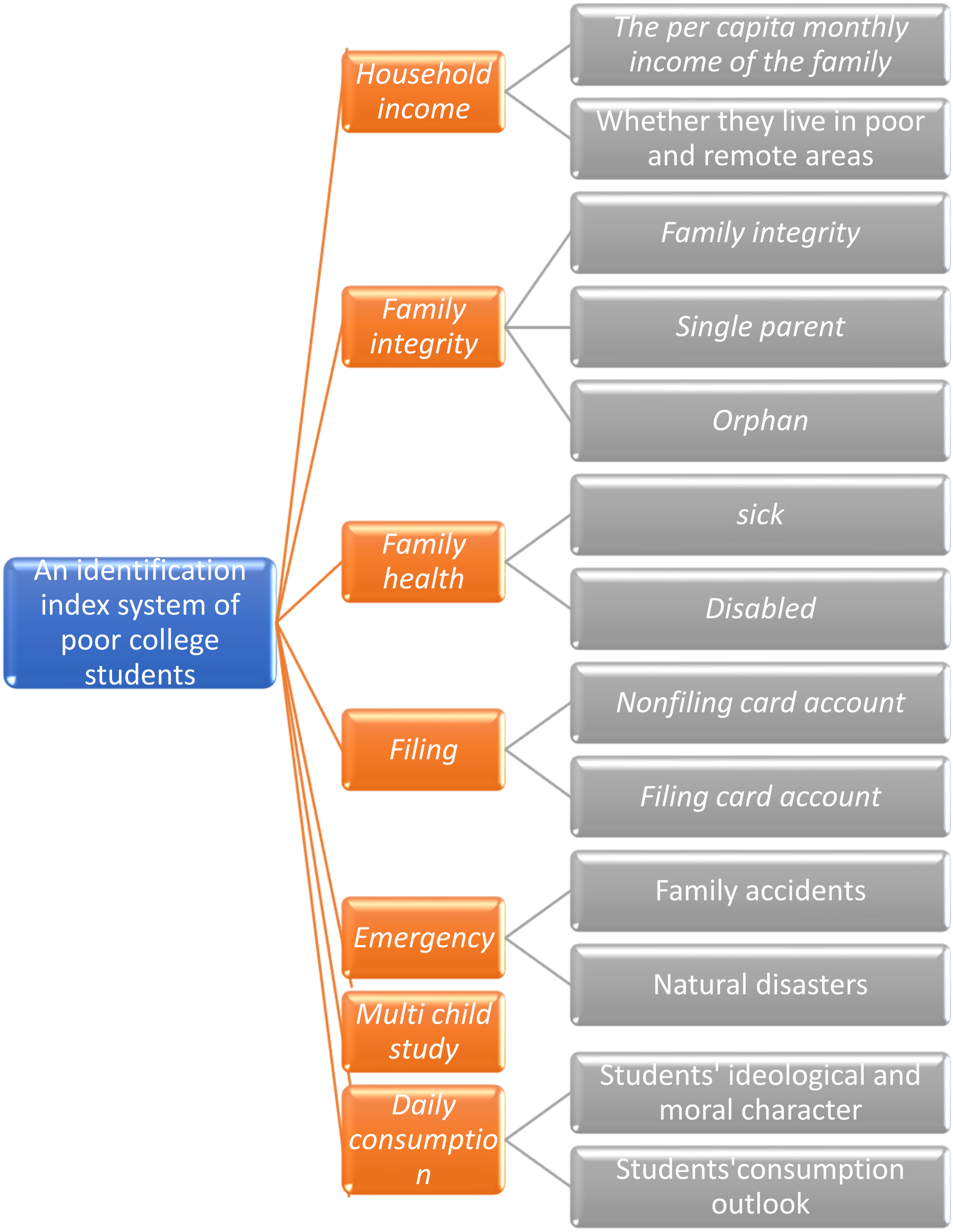

Figure 4: an identification index system of poor college students.

The glossary of basic language evaluation terms corresponding to the identification index system of poor college students is shown in Table 1.

Index	*Glossary of basic language evaluation terms*
Household income c_1	*s_1: The per capita monthly income of the family is more than 600 yuan, s_2: Average monthly household income is between 350 and 600 yuan, s_3: Average monthly household income ＜350 yuan*
Family integrity c_2	*s_0: Family integrity, s_1: single parent, s_2: orphan*
Family health c_3	*s_0: 0 people are sick or disabled, s_1: 1 person is sick or disabled, s_2: 2 or more people are sick or disabled*
Filing c_4	*0 refers to nonfiling card account, 1 refers to filing card account*
Emergency c_5	*s_0: No emergency, s_1: Emergency loss less than 10000 yuan, s_2: Loss of more than 10000 yuan in an emergency*
Multi child study c_6	*s_0: 0 brothers and sisters receive noncompulsory education, s_1: 1 sibling received noncompulsory education, s_2: Two or more brothers and sisters receive noncompulsory education*
Daily consumption c_7	*s_1: Common moral character and consumption outlook, s_2: Good moral character and consumption outlook, s_3: Excellent moral character and consumption outlook*

Table 1: The identification index system of poor college students based on incomplete linguistic hesitation and fuzziness.

The indicators of family income include the per capita monthly income and whether they live in poor and remote areas. The per capita monthly income directly reflects the economic status of the family. Although the family location cannot directly describe the economic status of the family, for families living in poor and remote areas with relatively backwards economic levels, there are few opportunities to improve their economic income in the short term, and the poverty status is difficult to improve. The indicator of family integrity examines whether family members are missing, since the financial resources of students from orphan families are limited, and the economic statuses of such students is generally more difficult than those of students from non-missing families. Family health indicators examine two aspects. First, large treatment fees for sick or disabled family members will cause sharp increases in family expenditure. Second, illness or disability will lead to a decline in the ability of family members to work, resulting in a sharp decline in family income. The index of filing and card establishment is used to check whether students have a filing and card establishment account. Poor households with registered cards are targeted for poverty alleviation, which shows that their economic conditions have been dire for an extended period of time. The indicators of emergencies include family accidents and natural disasters. When family members encounter incidents such as car accidents and civil disputes or suddenly encounter natural disasters such as earthquakes and floods, family income will decrease, and family burden will increase. The indicator of multiple children's schooling mainly focuses on the education expenditure of brothers and sisters. Families with multiple children receiving noncompulsory education have a large expenditure, and such families are more vulnerable than other normal families. The daily consumption indicators reflect the students' ideological and moral character and consumption outlook, so the evaluation of students' usual performance by the members of the evaluation group is also one of the indicators that needs to be considered when identifying students with financial difficulties.

7 ATTRIBUTE REDUCTION OF THE IDENTIFICATION INDEX SYSTEM FOR POOR COLLEGE STUDENTS

7.1 Fuzzy Decision System with Incomplete Linguistic Hesitation

Among the 2019 undergraduate students of the Jiangxi University of Technology, 9 poor students were randomly selected, relevant data were obtained, and an linguistic hesitant fuzzy decision-making system (U, AT, V, f) was established, as shown in Table 2.

U	c_1	c_2	c_3	c_4	c_5	c_6	c_7	d
x_1	$\{(s_2, 0.3)\}$	*	s_0	0	s_2	s_1	$\{(s_2, 0.13), (s_3, 0.87)\}$	s_1
x_2	$\{(s_2, 0.3)\}$	s_0	s_2	0	s_0	s_1	$\{(s_2, 0.38), (s_3, 0.62)\}$	s_1
x_3	$\{(s_2, 0.3)\}$	s_0	s_1	1	s_0	s_2	$\{(s_1, 0.06), (s_3, 0.94)\}$	s_2
x_4	$\{(s_3, 0.3)\}$	s_1	s_1	1	s_1	s_0	$\{(s_2, 0.15), (s_3, 0.85)\}$	s_2
x_5	$\{(s_1, 0.6)\}$	s_0	s_2	1	s_0	*	$\{(s_2, 0.18), (s_3, 0.82)\}$	s_1
x_6	$\{(s_1, 0.3)\}$	s_0	*	0	s_1	s_1	$\{(s_2, 0.33), (s_3, 0.67)\}$	s_0
x_7	$\{(s_1, 0.6)\}$	s_0	*	0	s_0	s_1	$\{(s_1, 0.19), (s_3, 0.81)\}$	s_0
x_8	$\{(s_1, 0.3)\}$	s_0	s_0	0	s_1	s_0	$\{(s_2, 0.27), (s_3, 0.73)\}$	s_0
x_9	$\{(s_2, 0.3)\}$	s_0	s_0	0	s_0	s_1	$\{(s_1, 0.18), (s_3, 0.82)\}$	s_1

Table 2: Fuzzy decision system with incomplete linguistic hesitation.

$U = \{x_1, x_2 \cdots, x_9\}$ is the object set, and $x_1, x_2 \cdots, x_9$ represent the 9 students.

$AT = C \cup \{d\}$ is a nonempty attribute set, and the conditional attribute set $C = \{c_1, c_2 \cdots, c_7\}$ is the set of identification indices of seven poor college students in Table 1. Decision attribute d is the poverty level, which is generally divided into three levels: special difficulty, general difficulty and difficulty.

$V = V_c \cup V_d$is the attribute value range. $f: U \times AT \rightarrow V$is an information function. The value range of decision attribute d is $V_d = \{f(x_i, d) \in S_0 | i = 1,2, \cdots ,9\}$.

S_0=$\{s_0\text{: difficulty},\ s_1\text{: general difficulty},\ s_2\text{: special difficulty}\}$ is the set of basic speech evaluation terms. $V_c = \cup_{j=1}^{7} \{f(x_i, c_j) | i = 1,2, \cdots ,9\}$. The condition attribute values $f(x_i, c_j)$ include real numbers, language terms, and the form of language hesitant fuzzy sets. The data of an indicator of an individual student is marked with an asterisk if it is not provided or missing. The value $f(x_i, c_1)$ of index c1 is determined by the linguistic hesitant fuzzy set $\{(s_{\theta(i)}, lh(s_{\theta(i)})) | s_{\theta(i)} \in \tilde{S}_1\}$, where $\tilde{S}_1$is shown in Table 1. $lh(s_{\theta(i)}) \in H$ indicates the degree of subordination of student $x_i$to comment $s_{\theta(i)}$. $H = \{0.3, 0.6\}$. If $lh(s_{\theta(i)})$ is 0.3, student x_i lives in nonpoor remote areas; if $lh(s_{\theta(i)})$ is 0.6, student $x_i$lives in poor and remote areas. When students live in poverty-stricken and remote areas, the possibility of their family's economic status being in $lh(s_{\theta(i)})$ for a short time is set as 0.6. The probability that the family's economic situation will be in $lh(s_{\theta(i)})$ for a short time when it is greater than that when the family lives in nonpoor remote areas is 0.3. The value of indicator c2 is $f(x_i, c_2) \in S_2$, and S_2 is the basic language evaluation term set of indicator c2. The value of indicator c3 is $f(x_i, c_3) \in S_3$, and S_3 is the basic language evaluation term set of indicator c3. The value of indicator c4 is $f(x_i, c_4) \in \{0,\ 1\}$. The value of indicator c5 is $f(x_i, c_5) \in S_5$, and S_5 is the basic language evaluation term set of indicator c5. The value of indicator c6 is $f(x_i, c_6) \in S_6$, and S_6 is the basic language evaluation term set of indicator c6. The value $f(x_i, c_7)$ of index c7 is determined by the linguistic hesitation fuzzy set$\{(s_{\theta(i)}, lh(s_{\theta(i)})) | s_{\theta(i)} \in \tilde{S}_7\}$, where $\tilde{S}_7$is shown in Table 1. $lh(s_{\theta(i)}) \in [0,\ 1]$ indicates the degree of subordination of student $x_i$to comment $s_{\theta(i)}$. Details for all indicators are given in Table 1.

With the first student $x_1$as an example, $f(x_1, c_1)= \{(s_2, 0.3)\}$ represents student x_1's family living in nonpoor remote areas, and the per person monthly revenue of the family is between 350 and 600 yuan. $f(x_1, c_2) =*$ indicates that the attribute value of student x_1 on indicator c_2 is not provided or missing. $f(x_1, c_3) = s_0$ means student x_1's family has no illness or disability. $f(x_1, c_4) = 0$ means that student x_1 does not have a registered account. $(x_1, c_5) = s_2$ means that there is an emergency at student x_1's home, and the loss is more than 10000 yuan. $(x_1, c_6) = s_1$ indicates that student x_1 has one brother or sister receiving noncompulsory education. $f(x_1, c_7) = \{(s_2, 0.13), (s_3, 0.87)\}$ means that 13% of the members of the evaluation group think that student x_1's ideological and moral character and consumption outlook are good, while 87% of the members think that student x_1's ideological and moral character and consumption outlook are excellent. $f(x_1, d)=s_1$ means that the poverty level determined by the school for student x_1 is generally difficult.

7.2 Attribute Reduction

Based on the data of the incomplete linguistic hesitant fuzzy decision-making system in Table 2, this paper performs attribute reduction on the identification index system of poor college students constructed in Table 1. Formula 8 can be obtained according to the definition of the limiting dominance relationship.

$$\widetilde{\boldsymbol{R}}_{\boldsymbol{C}}^{\boldsymbol{L}}=\{(x_1,x_1),(x_2,x_2),(x_3,x_3),(x_4,x_4),(x_5,x_5),(x_6,x_6),(x_7,x_7),(x_1,x_8),(x_8,x_8),$$
$$(x_3,x_9),(x_8,x_9),(x_9,x_9)\} \tag{8}$$

Among them, $(x_1, x_8) \in \widetilde{\boldsymbol{R}}_{\boldsymbol{C}}^{\boldsymbol{L}}$ indicates that all the conditional attribute values of student x_1 have limited advantages over student x_8. $[x_1]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_1\}$, $[x_2]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_2\}$, $[x_3]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_3\}$, $[x_4]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_4\}$, $[x_5]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_5\}$, $[x_6]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_6\}$, $[x_7]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_7\}$, $[x_8]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_1, x_8\}$, $[x_9]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_3, x_8, x_9\}$. $[x_8]_{\boldsymbol{C}}^{\boldsymbol{L}}=\{x_1, x_8\}$ represents the restrictive advantage class of x_8 with respect to attribute C.

The resolution matrix of the decision system is obtained from definition 6, as shown in Formula 9.

$$\widetilde{\boldsymbol{R}}_{\boldsymbol{C}}=\begin{bmatrix} \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & c_1c_2c_5c_7 & c_2c_5c_7 & c_1c_2c_5c_6c_7 & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & c_1c_3 & c_3 & c_1c_3c_6 & \boldsymbol{C} \\ c_3c_4c_6 & c_4c_6c_7 & \boldsymbol{C} & \boldsymbol{C} & c_6 & c_1c_3c_4c_6c_7 & c_3c_4c_6c_7 & c_1c_3c_4c_6c_7 & c_3c_4c_6c_7 \\ c_1c_2c_3c_4 & c_1c_2c_4c_5c_7 & \boldsymbol{C} & \boldsymbol{C} & c_1c_2c_5c_7 & c_1c_2c_3c_4c_7 & c_1c_2c_3c_4c_5c_7 & c_1c_2c_3c_4c_7 & c_1c_2c_3c_4c_5c_7 \\ \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & c_1c_3c_4c_6c_7 & c_3c_4c_6c_7 & c_1c_3c_4c_6c_7 & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & \boldsymbol{C} & c_1 & c_7 & c_1c_6 & \boldsymbol{C} \end{bmatrix} \tag{9}$$

where $\widetilde{\boldsymbol{R}}_{\boldsymbol{C}}(x_1, x_6)$ in the discrimination matrix $\widetilde{\boldsymbol{R}}_{\boldsymbol{C}}$ is $c_1c_2c_5c_7$, which means that the decision attribute value $f(x_6, d)=s_0$ of student x_6 is not greater than the decision attribute value $f(x_1, d)=s_1$ of student x_1. This is mainly because under the condition indicator attribute c_1, c_2, c_5, c_7, the value of student x_6 is no more restrictive than that of student x_1. According to Theorem 1, the reduction of the information system is $B=\{c_1, c_3, c_6, c_7\}$. This reduction shows that in the process of identifying poor college students, family income, family health, additional children's schooling and daily consumption are the four key indicators.

8 CONCLUSION

To achieve the precise support of colleges and universities for poor students, improve the effectiveness of funding and education, and solve the problems of diversity, fuzziness and lack of attribute values of identification indicators for poor college students, this paper first constructs an indicator system for the identification of poor college students based on incomplete linguistic hesitation and fuzziness. The indicator system includes seven indicators: family income, family integrity, family health, filing and card registration, emergencies, additional children's schooling and daily consumption. Then, based on the survey data, an incomplete linguistic hesitant fuzzy decision system is established. Finally, based on the attribute reduction method of the limited dominance relationship, according to the attribute discrimination matrix, the attribute reduction of the identification index system of poor college students is obtained. The method proposed in this paper can fully restore the essential characteristics of each index using multiple types of index attribute values. With the help of linguistic hesitant fuzzy sets, it well reflects the dynamic of family economic conditions and the students' general performance in ordinary times and truly "lets the data speak". The redundant indicator attributes in the identification system of poor students are eliminated through attribute reduction, which improves the operability of the identification process of poor students and optimizes the subsequent identification of poverty levels.

Wenbin Yang, https://orcid.org/0009-0005-2622-0700
Lan Qiu, https://orcid.org/0009-0005-6120-460X

ACKNOWLEDGEMENTS

Wenbin Yang and Lan Qiu would like to acknowledge the financial support of grants from Jiangxi Humanities and Social Sciences Project (No. GL18222) and Humanities and Social Sciences Project of Jiangxi University of Technology (No. RW2015).

FUNDING DETAILS:

This work was supported by Jiangxi Humanities and Social Sciences Project under Grant GL18222 and Humanities and Social Sciences Project of Jiangxi University of Technology under Grant RW2015.

REFERENCES

[1] An, L. P.; Tong, L. Y.: Binary Relations as a Basis for Rule Induction in Presence of Quantitative Attributes, Journal of Computers, 5(3), 2010, 440 - 447. https://doi.org/10.4304/jcp.5.3.440-447
[2] Delgado, M.; Verdegay, J. L.; Vila, M. A.: Linguistic Decision-Making Models, International Journal of Intelligent Systems, 1993(7), 1993, 479-492. https://doi.org/10.1002/int.4550070507
[3] Fan, W.; Liu, H.; Kuang, Y.: Application of Improved Apriori Algorithm Based on Rough Set in the Identification of Poor College Students, Computer Knowledge and Technology, 15(21), 2019, 3-6.
[4] Ji, C.; Wang, H.; Song, T.: Research on the Identification System of Poor College Students Based on Entropy Model, Journal of Southwest Normal University Natural Science Edition,46(4), 2021, 122-129.
[5] Jiang, H.: Research on the Application of Fault Tolerant Rough Sets in Natural Language Processing, Chongqing University of Posts and Telecommunications, 2020

[6] Meng, F. Y.; Chen, X. H.; Zhang, Q.: Multi-Attribute Decision Analysis Under a Linguistic Hesitant Fuzzy Environment, In-formation Sciences, 2014(267), 2014, 287-305. https://doi.org/10.1016/j.ins.2014.02.012
[7] Pawlak, Z.: Rough sets, International Journal of Information and Computer Sciences,1982(11), 1982, 341 - 356. https://doi.org/10.1007/BF01001956
[8] Song, M.: Research on the Identification of Poor College Students Based on Fuzzy Comprehensive Evaluation Method, Heilongjiang Higher Education Research, 2016(7), 2016, 16-20.
[9] Tao, P.; Qi, J.: Research on the Status Quo and Countermeasures of the Identification of Poor College Students from the Perspective of Self cognition, Journal of Hubei University of Economics (Humanities and Social Sciences Edition), 18(9), 2021, 20-23.
[10] Torra, V.: Hesitant fuzzy sets, International Journal of Intelligent Systems, 2010(25), 2010, 529- 539. https://doi.org/10.1109/FUZZY.2009.5276884
[11] Torra, V.; Narukawa, Y.: On Hesitant Fuzzy Sets and Decision, In The 18th IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, 2009, 1378-1382. https://doi.org/10.1002/int.20418
[12] Xiao, F.; Feng, X.: The Purpose, Principle and Strategy of Precise Funding for Poor College Students, Modern Education Management, 2020(3), 2020, 117-122.
[13] Xu, Z. S.: A Method for Multiple Attribute Decision Making with Incomplete Weight Information in Linguistic Setting, Knowledge-Based System, 2007(20), 2007, 719-725. https://doi.org/10.1016/j.knosys.2006.10.002
[14] Xu, Z.: Research on Pure Language Multi-Attribute Group Decision-Making Method, Control and Decision, 2004 (7), 2004, 778-781+786.
[15] Yang, P.; Jin, H.; Liu, H.: Research on the Identification Strategy of Poor Students in the Construction of Double First-Class Universities from the Perspective of Relative Poverty, China Higher Education Research, 2022(3), 2022, 45-51.
[16] Yuan, Y.: Design Principle and Reliability Analysis of the Precision Measurement System for Poverty Index of Poor College Students, Education and Teaching Forum, 2020 (52), 2020,104-106.
[17] Zhou, C.; Zhang, B.; Zhang, Q.: Exploring the High-Quality Development Path of College Student Funding in the Post Poverty Alleviation Era, Educational Economic Review, 5(6), 2020, 108-115.
[18] Zhou, T.; Lu, H.; Ren, H.; Huo, B.: Overview of Attribute Reduction Algorithms Based on Rough Sets, Journal of Electronics,49(7), 2021, 1439-1449
[19] Zhu, L.: Application of Hesitant Fuzzy Information Reduction Method in Identification of Poor College Students, Journal of Kunming Metallurgical College, 37, 2021, 594-99+114.
[20] Zhu, L.; Zhu, C.; Zhang, X.: Attribute Reduction of Incomplete Hesitant Fuzzy Hybrid Information Systems, Fuzzy Systems and Mathematics, 29(2), 2015, 150-156.