Research Article

Application of Fuzzy C-Mean Clustering Based on Multi-Polar Fuzzy Entropy Improvement in Dynamic Truck Scale Cheating Recognition

Zhenyu Lu[1], Xianyun Huang[2],*

[1] *Artificial Intelligence Institute, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, 210044, China*
[2] *Scientific Research Post, Suzhou Institute of Metrology, Suzhou, Jiangsu, 215128, China*

ABSTRACT

In the big data background, the uncertainty of data is increasingly apparent. Multi-polar fuzzy feature of data has been more popularly used by the research community for the purpose of the classification of weighing cheating in dynamic truck scale characteristic and the clustering problem of multi-polar fuzzy feature information. Additionally, the traditional classification method leads to slow speed and inaccuracy because of its difficulties. Therefore, by considering a multi-polar fuzzy feature classification of defects, a fuzzy c-means (*FCM*) clustering algorithm based on multi-polar fuzzy entropy is proposed. Firstly, according to the evaluation of available characteristics, the characteristic value of clustering samples is established. Secondly, we calculated the multi-polar fuzzy entropy of clustering samples. Finally, an improved *FCM* clustering algorithm based on multi-polar fuzzy entropy is presented. The experimental results of the data set collected from 5 different types of weighing cheating cars demonstrate that the algorithm improves the classification accuracy of *FCM* with multi-polar fuzzy feature information clustering and reduces significantly both the number of iterations and the classification time.

1. INTRODUCTION

Artificial intelligence being rich in information, has been widely used in various research fields to solve the complex problems through computer simulation. Cayley bipolar fuzzy graphs theory is proposed by Alshehri *et al.* [1] to solve the real time system modeling problem, where the level of information inherent in the system varies with the different levels of precision, In 2014, Mesiarová *et al.* [2] expanded bipolar fuzzy sets to m-polar fuzzy sets and promoted the development of bipolar fuzzy theory. In 2016, Zhou *et al.* [3] discussed the problem of nonlinear optimization with bipolar fuzzy relation equation constraints and promoted the development of bipolar fuzzy theory. Hanying *et al.* however, successfully applied the Yin Yang bipolar fuzzy cognitive Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) method in the diagnosis of bipolar disorder.

FCM is a fuzzy clustering algorithm based on an objective function proposed by Dunn and promoted by Bezdek [4]. It has been widely used in the image segmentation, automatic speech recognition, fault diagnosis and customer classification. In the traditional fuzzy C-clustering algorithm, the eigenvalues of clustering objects are accurate by default, but the characteristic information has the essential attribute of fuzziness and the bipolarity or the multi-polarity [5]. In this paper, in order to solve the existing problem of FCM clustering, in the process of bipolar fuzzy characteristic information clustering analysis, an improved *FCM* algorithm based on multi-polar fuzzy information entropy is proposed.

2. MULTI-POLAR FUZZY SET THEORY

2.1. Fuzzy Set Theory

The fuzzy set was proposed by Zadeh in 1965 [6], which provides a form for dealing with less rigorous information.

Definition 1. Let Z be the set of elements, z denotes a class of elements of Z, namely $Z = z$. This set is called the domain. Fuzzy subset A in Z is represented by membership function $u_a(Z)$. The value is between [0, 1], and is shown as

$$A = \{Z, u_a(Z) | z \in Z\} \qquad (1)$$

where $0 \leqslant u_a(Z) \leqslant 1$ extending the concept in classical set theory to fuzzy sets. $u_A : Z \mapsto [0, 1]$ represent a fuzzy set, or a fuzzy subset called Z. Mapping (function) $u_A(\cdot)$ or shorthand for $A(\cdot)$ is the membership function of fuzzy sets A.

*Corresponding author. Email: huangxianyun06@163.com

2.2. Bipolar Fuzzy Set Theory

An extension of fuzzy set, called bipolar fuzzy set was first proposed by Zhang [7] in 1994. After more than 20 years of development, the bipolar fuzzy set was matured and then applied to the clinical diagnosis of mental health diseases of the World Health Organization [8].

Definition 2. Zhang [7] defined a bipolar fuzzy set as a pair (u^+, u^-), where, $u^+ : X \to [0, 1]$ and $u^- : X \to [-1, 0]$ are any mappings. Let $B = (u^+, u^-)$ be the bipolar fuzzy sets on X., in which, u^+ represents the membership that element z belongs to one character regarding bipolar fuzzy set B and u^- represents the membership that element z belongs to the opposite character and the set of all bipolar fuzzy sets on X is denoted by $BF(X)$.

Chinese medicine pays attention to Yin and Yang, and the astronomy is divided into day and night. There are positive and negative comments on the judgment of things. In fact, wide variety of human decision-making is based on the double-sided or bipolar judgmental thinking on a positive side and a negative side. In recent years, substantial developmental work has been reported on the bipolar fuzzy by many scholars [3,8–13].

2.3. Multi-Polar Fuzzy Set

The dual fuzzy set is a special case of multi-polar fuzzy sets. Most of the multi-polar problems are more comprehensive than the bipolar. Modeling of the real-world problems often involves multi-agents, multi-attributes, multi-objects and multi-indexes. If only one side is evaluated from the front and the back, a lot of useful information will be lost. If more information is given from the multi-polar feature information of the problem, it will make the description more reflective.

Definition 3. Let $A = \{x, u_A(x_1), u_A(x_2), \cdots, u_A(x_n) | x \in X\}$ and $B = \{x, u_B(x_1), u_B(x_2), \cdots, u_B(x_n) | x \in X\}$ are two multi-polar fuzzy set, then the Euclidean distance of the two multi-polar fuzzy sets is

$$D = d(A, B) \tag{2}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left((u_A(x_1) - u_B(x_1))^2 + (u_A(x_2) - u_B(x_2))^2 + \cdots \right.$$

$$\left. + (u_A(x_n) - u_B(x_n))^2 \right)$$

where $A : V \to [0, 1]^m$ is the m-polar fuzzy set on V, $B : E \to [0, 1]^m$ is the m-polar fuzzy set on E, $E \subseteq V * V$ satisfies the symmetry, scilicet $< x, y > \in E \iff < y, x > \in E$. If $\forall < x, y > \in E$ has $B(< x, y >) \leqslant \inf\{A(x), A(y)\}$, then the even pair (A, B) is an m-polar fuzzy graph whose base set is (V, E), a set of m-polar fuzzy vertices [3] where, A is V, and B is the m-pole of E Fuzzy edge set.

Let $G = (A, B)$ be an m-pole fuzzy graph whose base set is (V, E), if $B(< x, y >) = B(< y, x >)(\forall(x, y) \in E)$ and $B(< x, x >) = 0([0, 1]^m$ minimum element $(\forall x \in V)$, then G is a bipolar fuzzy graph [14] which is also a simple multi-pole fuzzy graph.

Let $G = (A, B)$ be a multi-polar fuzzy graph whose basis set is (V, E). If $\forall < x, y > \in E$, there is $B(< x, y >) = \inf A(x), A(y)$, then G is a strong multi-polar fuzzy graph. The complement of G is a basic set

of (V, E), multi-polar fuzzy graph $\overline{G} = (A, \overline{B})$, where $\overline{B} : E \to [0, 1]^m$ is defined as

$$\forall < x, y > \in E, i \in m, p_i^{\circ} \overline{B}(< x, y >)$$
$$= \begin{cases} 0, if p_i^{\circ} \overline{B}(< x, y >) < 0 \\ \inf p_i^{\circ} A(x), p_i^{\circ} A(y), if p_i^{\circ} \overline{B}(< x, y >) > 0 \end{cases} \tag{3}$$

2.4. Multi-Polar Fuzzy Entropy

Fuzzy entropy is not determined by a given value anymore. It comes from the fuzzy function of fuzzy value, which is closer to the real object, so it has a better fitting effect. Fuzzy entropy has fuzziness and is no longer determined by a given value, it comes from the fuzzy-valued function that is closer to the real thing therefore, it has a better fitting effect. The bipolar fuzzy entropy can be defined as follows:

Definition 4. Let the set $X = \{x_1, x_2, \ldots x_n\}$, the bipolar fuzzy set on X can be expressed as $A = \{< x_i, u_A^P(x_i), u_A^N(x_i) > | x_i \in X\}, i = 1, 2, \ldots, n$. In the formula, $u_A^P(x_i) \in [0, 1], u_A^N(x_i) \in [0, 1]$ denotes the anode and cathode membership of x_i to A, respectively.

Definition 5. Let the set $X = \{x_1, x_2, \ldots x_n\}$, the multi-polar fuzzy set on X can be expressed as $A = \{< x_i, u_A^1(x_i), u_A^2(x_i), \cdots, u_A^m(x_i) > | x_i \in X\}, i = 1, 2, \cdots, n$. In the formula, $u_A^j(x_i) \in [0, 1], j = 1 \cdots m$, denotes the membership degree of x_i belongs to A.

Definition 6. The set $X = \{x_1, x_2, \cdots x_n\}$ is a nonempty set, and E(A) is the set of bipolar fuzzy entropy.

$$E(A) = \frac{1}{2n} \sum_{i=1}^{n} (\pi_A(x) - \theta_A(x)) \tag{4}$$

In the formula, $\pi_A(x) = 1 - u_A^P(x_i) - u_A^N(x_i)$ is the uncertainty of x in A, $\theta_A(x) = 1 - |u_A^P(x_i) - u_A^N(x_i)|$ is the fuzziness of x in A. Let $X = \{x_1, x_2, \cdots x_n\}$ constitute the evaluation criterion, $C = \{c_1, c_2, \cdots, c_m\}$ is the condensed group.

Definition 7. The set $X = \{x_1, x_2, \cdots x_n\}$ is a nonempty set, and E(A) is the multi-polar fuzzy entropy of the set, then

$$A = \left(\frac{< c_{11}, c_{12}, \cdots, c_{1m} >}{x_1}, \frac{< c_{21}, c_{22}, \cdots, c_{2m} >}{x_2}, \cdots, \right. \tag{5}$$

$$\left. \frac{< c_{n1}, c_{n2}, \cdots, c_{nm} >}{x_n} \right)$$

$$E(c_j) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{< 0, 0, \cdots, c_{1j}, \cdots, 0, 0 >}{x_1} \right.$$

$$\left. + \frac{< 0, 0, \cdots, c_{2j}, \cdots, 0, 0 >}{x_2} + \cdots + \frac{< 0, 0, \cdots, c_{nj}, \cdots, 0, 0 >}{x_n} \right) \tag{6}$$

In the formula, j is $1, 2, \cdots, m$.

3. FCM CLUSTERING BASED ON MULTI-POLAR FUZZY ENTROPY

3.1. Limitations of FCM

Clustering analysis is the process of grouping a collection of elements into multiple classes, the classification of the elements is based on the large similarity of the elements in the same class and the large difference between the elements of the different classes. Traditional hard clustering is a hard partition, it has an either or quality. But there is an intermediation in the application of the actual genus and there is no definite boundary to draw. Ruspirti applied fuzzy set theory to the cluster analysis based on hard partition clustering and fuzzy membership knowledge in 1969 [15]. Based on this, Dunn [16] proposed *FCM* algorithm in 1974. FCM clustering improves the defect of hard clustering, it is basically a soft clustering algorithm, that improves the traditional hard clustering algorithm and uses the method of membership degree to determine each element belonging to a certain clustering degree. Fuzzy C-means clustering has the advantages of low complexity and low implementation difficulty, and has been widely used in many fields, such as data classification, image segmentation and cell analysis.

FCM clustering (*FCM*) algorithm, is going to use membership to determine each sample $X = \{x_1, x_2, \cdots x_n\}$, is a soft clustering algorithm express the degree to which it belongs to a class $C_{center} = \{c_1, c_2, \cdots c_n\}$ [17].

FCM algorithm is a flexible partitioning clustering algorithm. The basic principle of classification is to maximize the similarity of elements of the same class and minimize the similarity of elements of different classes.

An important concept in *FCM* is a membership function, which represents the degree to which a classification recognition object A belongs to a collection of sample categories x_i, generally do remember $u_A(x_i)$, the value range of the independent variable is all the objects that belong to the collection, the range of $u_A(x_i)$ is [0, 1]. Where $u_A(x_i) = 1$ represents x_i, completely subordinate to the set A; $u_A(x_i) = 0$ represents x_i, not A member of set A at all; $u_A(x_i) \in (0, 1)$ represents x_i, part of set A. A membership function defined in space $X = \{x_i | x_i \in X\}$ expresses a fuzzy set S.

FCM puts the sample data set, which is composed of sample data of n research objects divided into categories $X = \{x_1, x_2, \cdots x_n\}$ a number of c, and finds the clustering center c_i of each group, makes the sample similarity between different categories as small as possible, however, data in the same category are as similar as possible. The value of *FCM* adopts the membership degree of soft partition within the interval of [0, 1] to determine the degree to which each sample belongs to each category. $U = \{u_{ij} | i = 1, 2, \cdots, c \ j = 1, 2, \cdots, n\}$ represents the sample data x_j, and for the membership degree of category c_i, u_{ij} is the membership degree matrix and satisfies

$$\begin{cases} \sum_{i=1}^{c} u_{ij} = 1, 1 \leqslant j \leqslant n \\ \sum_{j=1}^{n} u_{ij} > 0, 1 \leqslant i \leqslant c \\ u_{ij} \leqslant 0, 1 \leqslant i \leqslant c, \ 1 \leqslant j \leqslant n \end{cases} \quad (7)$$

General form of the objective function of *FCM* is expressed as

$$Obj(U, c_1, c_2, \cdots, c_c) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \quad (8)$$

where $0 \leqslant u_{ij} \leqslant 1$, c_i is the center of the class. $d_{ij} = ||class_i - x_i||$ is the distance from the center c_i of the i-th class to the x_j-th, and m is the weighted exponent.

Construct the Lagrange function of the following objective function, the necessary condition for reaching the minimum value is obtained by using

$$\overline{Obj}(U, c_1, c_2, \cdots, c_c, \lambda_1, \lambda_2, \cdots, \lambda_n) \quad (9)$$

$$= \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 + \sum_{j=1}^{n} \left(\sum_{i=1}^{c} u_{ij} - 1 \right)$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ is the Lagrange multiplier. From derivation of Equation (9), we can get the objective function. The necessary condition for getting the minimum value is

$$Center_i = \frac{\sum_{j=1}^{n} u_{ij}^m X}{\sum_{j=1}^{n} u_{ij}^m} \quad (10)$$

$$= \sum_{j=1}^{n} \frac{u_{ij}^m}{\sum_{j=1}^{n} u_{ij}^m} X_j$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{M-1}}} \quad (11)$$

$$= \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{-m}}$$

The goal of the *FCM* clustering algorithm is to find the optimal cluster center and the membership matrix, and then use these cluster centers and membership matrix to classify the particles.

Traditionally FCM clustering has three defects. Firstly, the characteristic values of the clustering samples are all exact values, which are often difficult to be accurately given in practical problems; secondly, the algorithm is not stable, and finally the classification effect is often affected by the initial clustering center. For the existence of *FCM* high sensitivity to the cluster center initialization problem, yuan heuristic [18] is proposed based on chemical, and successfully obtain the optimal cluster center. However, the actual problems tend to be multi-polar, and needs multi-polar fuzzy theory to optimize the *FCM* algorithm.

3.2. Calculate the Eigenvalue Matrix

Fuzzy state describes the degree of fuzziness and uncertainty of fuzzy sets. The lesser the fuzzy set, the smaller is the weight and vice versa. The feature weight of sample wood must reflect the relative importance of each feature. It has a larger fuzzy state, in order to avoid subjective arbitrariness in the process of weighting. This

work uses multi-polar fuzzy entropy to determine the weight w_i of feature g_i, as follows:

$$E(g_j) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{<0,0,\cdots,g_{1j},\cdots,0,0>}{x_1}\right.$$

$$+\frac{<0,0,\cdots,g_{2j},\cdots,0,0>}{x_2}+\cdots,$$

$$\left.+\frac{<0,0,\cdots,g_{nj},\cdots,0,0>}{x_n}\right) \tag{12}$$

where j is $1, 2, \cdots, m$

According to the literature [19] formula (12) can be written as

$$E(g_j) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Mean<0,0,\cdots,g_{1j},\cdots,0,0>}{x_1}\right.$$

$$+\frac{Mean<0,0,\cdots,g_{2j},\cdots,0,0>}{x_2}+\cdots,$$

$$\left.+\frac{Mean<0,0,\cdots,g_nj,\cdots,0,0>}{x_n}\right) \tag{13}$$

$$w_j = \frac{1-E_{gi}}{\sum_{j=1}^{s}(1-E_{gi})} \tag{14}$$

where $E(g_j)$ reflects the fuzziness and uncertainty of the eigenvalues of the sample set Y at the feature g_j. The larger the value of g_j, the greater is the dependence of the clustering result on the feature g_j.

Weighting the eigenvalue matrix P to obtain an eigenvalue matrix $P' = (p'_{ij})_{(n \times s)}$, where

$$p'_{ij} = w_j * P_{ij} \tag{15}$$

3.3. Identify Cluster Centers

FCM clustering has been widely used from its inception in the fields of pattern recognition, fault diagnosis and image. Although, many samples are processed by filtering and other algorithms, they still contain a lot of noise. Since *FCM* is sensitive to the initial cluster center, the clustering results obtained by the different initial cluster centers are also quite different, especially, when FCM clustering is used to classify the samples [18]. In order to effectively avoid the noise area, the selected area of the initial cluster center can be subdivided into density, and the c points which are taken farthest from each other in the high-density area.

Let P_i be the position of the particle, define ρ_i as the area density and calculate n Euclidean distances using definition (3) centered on p'_i, $d(p'_i, p'_1)d(p'_i, p'_2)\cdots d(p'_i, p'_n)$, let them rearrange from big to small.

$$d\left(p'_i, p'_{(1)}\right) \leqslant d\left(p'_i, p'_{(2)}\right) \leqslant \cdots \leqslant d\left(p'_i, p'_{(n)}\right) \tag{16}$$

In the formula, subscript(i) refers the subscript rearrangement to satisfy the above condition. Due to $d(p'_i, p'_{(i)})$ m to 0, so after reordering, $p'_{(1)}$ is p'_i.

The p'_i area radius is the minimum Euclidean distance of C feature vectors including p'_i, recorded as $R(p'_i)$, by

$$R(p'_i) = d(p'_i, p'_{(c)}) \tag{17}$$

It is easy to know that the region where p'_i is located contains the regional density parameter of $p'_{(1)}p'_{(2)}, \cdots, p'_{(c)}$ total C feature vectors p'_i, ρ_i is

$$\rho_i = \frac{1}{R(p'_i)} = \frac{1}{d(p'_i, p'_{(c)})} \tag{18}$$

In the formula, $0 < c < n$ where c is an integer and takes specific value subject to the availability. The greater ρ_i is, explains that the regional density of p'_i is larger; on the contrary, the smaller ρ_i explains that the regional density of p'_i is smaller.

Calculate the regional density parameters of the feature vectors $p'_1, p'_2, \cdots p'_n$ using Equations (17) and (18). After comparison, the region with the highest density is selected as the first cluster center Z_1, and get a feature vector set of high-density regions $H = \{p'_{(1)}p'_{(2)}, \cdots, p'_{(c)}\}$; then, take the feature vector farthest from Z_1 in H as the first cluster center Z_2, recalculate all the feature vectors distance in H to Z_1, Z_2, take out in H to meet.

$$max\left(\min\left(d\left(p'_r, Z_1\right), \quad d\left(p'_r, Z_2\right) r = 1, 2\cdots, c\right.\right. \tag{19}$$

The feature vector is the third cluster center Z_3; at last, calculate the distance of all feature vectors s distance in H to $Z_1, Z_2, \cdots Z_{k-1}$, take out in H to meet

$$max\left(\min\left(d\left(p'_r, Z_1\right), d\left(p'_r, Z_2\right) \cdots d\left(p'_r, Z_{k-1}\right)\right)\right), r = 1, 2\cdots, c \tag{20}$$

The feature vector of d is the kth cluster center $Z_k(k = 1, 2, \cdots c)$. According to this, the clustering center $Z = (Z_1, Z_2, \cdots Z_k,)$ is obtained, where c cluster centers are taken from C feature vectors in the feature vector set H. When determining the initial cluster center, avoid excessive concentration, and the feature vector in H should be chosen to avoid noise points. In this work, we take $c \in ((n+c)/2, n)$ and are integers.

3.4. Update Cluster Center and Membership Matrix

In the *FCM* clustering algorithm based on multi-polar fuzzy entropy, it is necessary to calculate the corresponding membership matrix by using the cluster center.

(l) If $h, 1 < h < c$, such that $d(p'_i, Z_h) = 0$, then let

$$u_{ik} = \begin{cases} 1, & k = h \\ 0, & k \neq h \end{cases} \tag{21}$$

(2) If $h, 1 < h < c$, such that $d(p'_i, Z_h) > 0$, then let

$$u_{ik} = \frac{1}{\sum_{h=1}^{c}\left(\frac{d(p'_i, Z_k)}{d(p'_i, Z_h)}\right)^{\frac{2}{m-1}}} \tag{22}$$

In the formula, m is the fuzziness parameter.

Update the cluster center with the membership matrix, where the kth cluster center is recorded as Z_k,

$$Z_k = \{z_{k1}, z_{k2}, \cdots z_{ks}\} \tag{23}$$

At this time, the square of the generalized Euclidean distance of the sample set Y for the cluster center Z is

$$J(U, Z) = \sum_{i=1}^{n} \sum_{k=1}^{c} (u_{ik})^m \cdot (d(p_i', Z_h))^2 \tag{24}$$

3.5. MPFCM Algorithm Steps

The specific steps of the multi-polar fuzzy entropy optimization *FCM* clustering algorithm are as follows:

Step 1: Input the sample eigenvalue matrix P, the number of clusters c, the fuzziness parameter m and the threshold ϵ iteration number δ of the iteration is stopped.

Step 2: Calculate the feature weights by using Equations (13) and (14), and then calculate the weighted eigenvalue matrix P' using Equation (15).

Step 3: Let $t = 0$, determine the initial cluster center Z^t and calculate the membership matrix $U^{(t)}$ by using Equations (21) and (22).

Step 4: Determine if t is less than δ. If yes, then continue with step 5; If not, go to step 7.

Step 5: Let $t = t + 1$, update the clustering center $Z^{(t)}$ by using Equation (20), and then update the membership matrix $U^{(t)}$ by using equations.

Step 6: Uses Equation (17) to calculate $J(U^{(t-1)}, Z^{(t-1)})$ and if $J(U^{(t)}, Z^{(t)})$, $J(U^{(t-1)}, Z^{(t-1)}) - J(U^{(t)}, Z^{(t)}) < \epsilon$ continue with step 7; if not, go to step 4.

Step 7: Output the membership matrix U and the cluster center Z.

4. EXPERIMENTAL VERIFICATION

The algorithm FCM clustering was successfully applied to the diagnosis of diseases, making the diagnosis more accurate [20]. Then, we apply the algorithm which has been optimized to the dynamic truck scale for application analysis.

Expressway plays an important role in the transportation, but the phenomenon of the overload of freight vehicles occurs from time to time, which seriously affects the service life and transportation safety. However, traditional inspection is far from meeting the demand. Nowadays, the expressway has begun to operate on the Internet. We have realized the prejudgment of overload cheating of trucks based on the analysis of the characteristics of monitoring data of trucks and improved the correct rate of overload judgment.

High-speed jumping and towing weights are used when the vehicles pass through the scale. These have advantages of short reaction time, the effective data collected is less and inaccurate, the rear wheels brake the head and tail respectively to reduce weight and achieve the purpose of weight reduction cheating. These two cheating methods are dynamic and fast but most difficult to define and prevent from cheating at high speed.

There are five passing vehicles to be inspected, taking pictures of key points from different angles of vehicles weighed on scale, $X = \{x, y, z, u, v\}$ 5 system devices, from image data $G = \{g_1, g_2, g_3, g_4\}$. Evaluation of the three angles namely, loading capacity, alignment S-type, acceleration and deceleration, location of the four features as high-speed jumping weights, normal and towing weights, which can be combined with multi-polar fuzzy theory to give membership. The membership degree of the g_1 fault type is as follows: $A(x) =< 0.49, 0.46, 0.51 >$, $A(y) =< 0.45, 0.42, 0.59 >$, $A(z) =< 0.50, 0.40, 0.54 >$, $A(u) =< 0.40, 0.49, 0.60 >$ and $A(v) =< 0.51, 0.52, 0.50 >$.

Step 1: Similarly, the 3-pole fuzzy set $A : X \rightarrow [0, 1]^3$ of 5 vehicles is described as follows:

$$A(g_1) = \left\{ \frac{< 0.49, 0.46, 0.51 >}{x}, \frac{< 0.45, 0.42, 0.59 >}{y}, \frac{< 0.50, 0.40, 0.54 >}{z}, \frac{< 0.40, 0.49, 0.60 >}{u}, \frac{< 0.51, 0.52, 0.50 >}{v} \right\}$$

$$A(g_2) = \left\{ \frac{< 0.46, 0.51, 0.48 >}{x}, \frac{< 0.43, 0.49, 0.53 >}{y}, \frac{< 0.45, 0.51, 0.44 >}{z}, \frac{< 0.49, 0.51, 0.48 >}{u}, \frac{< 0.42, 0.45, 0.53 >}{v} \right\}$$

$$A(g_3) = \left\{ \frac{< 0.52, 0.47, 0.49 >}{x}, \frac{< 0.47, 0.45, 0.51 >}{y}, \frac{< 0.45, 0.54, 0.50 >}{z}, \frac{< 0.46, 0.53, 0.47 >}{u}, \frac{< 0.41, 0.48, 0.52 >}{v} \right\}$$

$$A(g_4) = \left\{ \frac{< 0.51, 0.46, 0.50 >}{x}, \frac{< 0.52, 0.47, 0.51 >}{y}, \frac{< 0.58, 0.47, 0.50 >}{z}, \frac{< 0.55, 0.51, 0.46 >}{u}, \frac{< 0.57, 0.52, 0.43 >}{v} \right\}$$

Step 2: Sample eigenvalue matrix

Table 1 | Sample eigenvalue matrix.

	g_1	g_2	g_3	g_4
x	$< 0.49, 0.46, 0.51 >$	$< 0.46, 0.51, 0.48 >$	$< 0.52, 0.47, 0.49 >$	$< 0.51, 0.46, 0.50 >$
y	$< 0.45, 0.42, 0.59 >$	$< 0.43, 0.49, 0.53 >$	$< 0.47, 0.45, 0.51 >$	$< 0.52, 0.47, 0.51 >$
z	$< 0.50, 0.40, 0.54 >$	$< 0.45, 0.51, 0.44 >$	$< 0.45, 0.54, 0.50 >$	$< 0.58, 0.47, 0.50 >$
u	$< 0.40, 0.49, 0.60 >$	$< 0.49, 0.51, 0.48 >$	$< 0.46, 0.53, 0.47 >$	$< 0.55, 0.51, 0.46 >$
v	$< 0.51, 0.52, 0.50 >$	$< 0.42, 0.45, 0.53 >$	$< 0.41, 0.48, 0.52 >$	$< 0.57, 0.52, 0.43 >$

Obtained weighted sample eigenvalue matrix P'

Table 2 | Sample eigenvalue matrix.

	g_1	g_2	g_3	g_4
x	$< 0.1230, 0.1155, 0.1281 >$	$< 0.1124, 0.1246, 0.1173 >$	$< 0.1286, 0.1163, 0.1212 >$	$< 0.1312, 0.1183, 0.1286 >$
y	$< 0.1130, 0.1055, 0.1482 >$	$< 0.1050, 0.1197, 0.1295 >$	$< 0.1163, 0.1113, 0.1262 >$	$< 0.1338, 0.1209, 0.1312 >$
z	$< 0.1256, 0.1004, 0.1356 >$	$< 0.1099, 0.1246, 0.1075 >$	$< 0.1113, 0.1336, 0.1237 >$	$< 0.1492, 0.1209, 0.1286 >$
u	$< 0.1004, 0.1230, 0.1507 >$	$< 0.1197, 0.1246, 0.1173 >$	$< 0.1138, 0.1311, 0.1163 >$	$< 0.1415, 0.1312, 0.1183 >$
v	$< 0.1281, 0.1306, 0.1256 >$	$< 0.1026, 0.1099, 0.1295 >$	$< 0.1014, 0.1187, 0.1286 >$	$< 0.1466, 0.1338, 0.1106 >$

Calculate the initial cluster center based on the weighted sample feature matrix

Table 3 | Initial cluster center.

$z_1^{(0)}$	$z_2^{(0)}$	$z_3^{(0)}$
$< 0.1230, 0.1155, 0.1281 >$	$< 0.1281, 0.1306, 0.1256 >$	$< 0.1256, 0.1004, 0.1356 >$
$< 0.1124, 0.1246, 0.1173 >$	$< 0.1026, 0.1099, 0.1295 >$	$< 0.1099, 0.1246, 0.1075 >$
$< 0.1286, 0.1163, 0.1212 >$	$< 0.1014, 0.1187, 0.1286 >$	$< 0.1113, 0.1336, 0.1237 >$
$< 0.1312, 0.1183, 0.1286 >$	$< 0.1466, 0.1338, 0.1106 >$	$< 0.1492, 0.1209, 0.1286 >$

After the initial clustering center, the membership matrix U is (Table 4):

Table 4 | The membership matrix $U^{(0)}$.

	x	y	z	u	v
$U_{i1}^{(0)}$	1	0.6295	0	0.3797	0
$U_{i2}^{(0)}$	0	0.1114	0	0.2094	1
$U_{i3}^{(0)}$	0	0.2591	1	0.4109	0

Sample feature clustering center after one iteration (Table 5)

Table 5 | Initial cluster center.

$z_1^{(1)}$	$z_2^{(1)}$	$z_3^{(1)}$
$< 0.1183, 0.1136, 0.1353 >$	$< 0.1267, 0.1300, 0.1269 >$	$< 0.1214, 0.1038, 0.1383 >$
$< 0.1112, 0.1233, 0.1204 >$	$< 0.1033, 0.1107, 0.1290 >$	$< 0.1110, 0.1243, 0.1100 >$
$< 0.1241, 0.1164, 0.1220 >$	$< 0.1021, 0.1192, 0.1281 >$	$< 0.1119, 0.1320, 0.1228 >$
$< 0.1328, 0.1202, 0.1283 >$	$< 0.1463, 0.1335, 0.1112 >$	$< 0.1473, 0.1223, 0.1273 >$

Membership matrix $U^{(1)}$ is (Table 6)

Table 6 | The membership matrix $U^{(1)}$.

	x	y	z	u	v
$U_{i1}^{(1)}$	0.8598	0.6179	0.0448	0.5039	0.0032
$U_{i2}^{(1)}$	0.0475	0.1353	0.0208	0.3301	0.9940
$U_{i3}^{(1)}$	0.0927	0.2469	0.9344	0.1660	0.0028

Sample feature clustering center after iteration two iteration (Table 7).

Table 7 | Clustering center $Z^{(2)}$.

$z_1^{(2)}$	$z_2^{(2)}$	$z_3^{(2)}$
< 0.1153, 0.1134, 0.1393 >	< 0.1221, 0.1261, 0.1318 >	< 0.1176, 0.1074, 0.1406 >
< 0.1114, 0.1230, 0.1210 >	< 0.1059, 0.1138, 0.1269 >	< 0.1117, 0.1238, 0.1136 >
< 0.1212, 0.1181, 0.1219 >	< 0.1058, 0.1201, 0.1262 >	< 0.1136, 0.1287, 0.1222 >
< 0.1345, 0.1218, 0.1274 >	< 0.1441, 0.1314, 0.1146 >	< 0.1473, 0.1103, 0.1041 >

Calculating the membership matrix $U^{(2)}$

Table 8 | The membership matrix $U^{(2)}$.

	x	y	z	u	v
$U_{i1}^{(2)}$	0.8598	0.6313	0.0311	0.5046	0.0116
$U_{i2}^{(2)}$	0.0426	0.1335	0.0220	0.3296	0.9866
$U_{i3}^{(2)}$	0.0776	0.2352	0.9469	0.1658	0.0018

According to parameter $c = 3, m = 2, \epsilon = 10^{-5}, \delta = 100$, we calculate each feature weight by Table 1, and then obtain the weighted sample eigenvalue matrix, as shown in Table 2. According to Table 2, the calculation method proposed in this work determines the initial cluster center, and the initial values of the three cluster centers are $Z_1^{(0)}, Z_2^{(0)}$ and $Z_3^{(0)}$, respectively. By using the degree calculation formula, we calculate the corresponding membership matrix of the initial cluster center. The cluster centers obtained after the iteration are $Z_1^{(2)}, Z_2^{(2)}$ and $Z_3^{(2)}$, as shown in the Table 3. After two iterations, we get $J(U^{(1)}, Z^{(1)}) - J(U^{(2)}, Z^{(2)})\epsilon$. The membership degree iteration results are shown in Table 8. As show in Table 8, the samples are classified into the first class is $z_1^{(0)}$, the second class is $z_2^{(0)}$ and the third category is $z_3^{(0)}$ according to the principle of maximum membership.

Table 9 shows that the iterative speed of the algorithm is superior to the other two kinds of algorithms in addition, the size of the numerical $J(U^{(0)}, Z^{(0)}) - J(U^{(n)}, Z^{(n)})$ reflects changes of the whole process of iteration of the objective function and reflection of the effectiveness based on multi-polar fuzzy entropy to initialize of the cluster center.

Table 9 | Algorithm comparison.

	Number of Iterations	$J(U^{(0)}, Z^{(0)}) - J(U^{(n)}, Z^{(n)})$
Literature [21] algorithm	8	0.8446
Literature [19] algorithm	4	0.7562
Algorithm in this paper	2	0.2635

5. CONCLUSION

This paper proposes a multi-polar fuzzy *FCM* based on the classification of multi-polar fuzzy features. The algorithm uses the sample multi-polar fuzzy membership degree to calculate the feature weights and obtain a new clustering center. The algorithm improves the problem of random initialization of the clustering center in the *FCM* clustering algorithm, which tends to cause slow convergence or even fall into the local minimum defects. Finally, the new clustering centers are obtained by calculating the feature weights with the multi-polar fuzzy membership degree of samples. The algorithm can effectively improve the ponderation on jump pounds and pound of the recognition accuracy and effectiveness of the increasing recognition reliability, reducing disputes.

CONFLICTS OF INTEREST

The authors declared that they have no conflicts of interest to this work.

ACKNOWLEDGMENTS

REFERENCES

[1] N.O. Alshehri, M. Akram, Cayley bipolar fuzzy graphs, Sci. World J. 2013 (2013), 156786.

[2] M.Z. Andrea, K. Ahmad, Extended multi-polarity and multi-polar-valued fuzzy sets, Fuzzy Sets Syst. 234 (2014), 61–78.

[3] J. Zhou, Y. Yu, Y. Liu, *et al.*, Solving nonlinear optimization problems with bipolar fuzzy relational equation constraints, J. Inequal. Appl. 2016 (2016), 1–10.

[4] J. Bezdek, R. Hathaway, M. Sobin, *et al.*, Convergence theory for fuzzy c-means: counterexamples and repairs, IEEE Trans. Syst. Man Cybern. 17 (1987), 873–877.

[5] N.R. Pal, J.C. Bezdek, On cluster validity for the fuzzy c-means model, IEEE Trans. Fuzzy Syst. 3 (2002), 370–379.

[6] L.A. Zadeh, Fuzzy sets, Inf. Control. 8 (1965), 338–353.

[7] Z. Wen-Ran, Bipolar fuzzy sets and relations: a computational framework for cognitive modeling and multiagent decision

analysis, in NAFIPS/IFIS/NASA 94 First International Joint Conference of the North American Fuzzy Information Processing Society Bi Conference the Industrial Fuzzy Control & Intellige, IEEE, San Antonio, TX, USA, 1994, pp. 305–309.

[8] Y. Han, Z. Lu, Z. Du, *et al.*, A YinYang bipolar fuzzy cognitive TOPSIS method to bipolar disorder diagnosis, Comput. Methods Prog. Biomed. 158 (2018), 1–10.

[9] Y. Han, P. Shi, S. Chen, Bipolar-valued rough fuzzy set and its applications to decision information system, IEEE Trans. Fuzzy Syst. 23 (2015), 2358–2370.

[10] H. Rashmanlou, S. Samanta, M. Pal, *et al.*, Bipolar fuzzy graphs with categorical properties, Fuzzy Inf. Eng. 7 (2015), 317–334.

[11] M.A. Alghamdi, A.N. Omair, A. Muhammad, Multi-criteria decision-making methods in bipolar fuzzy environment, Int. J. Fuzzy Syst. 20 (2018), 2057–2064.

[12] D. Hidekaru, K. Motoi, O. Kenji, *et al.*, A novel class of emitting amorphous molecular materials with bipolar character for electroluminescence, Chem. Mater. 15 (2016), 1080–1089.

[13] H. Gao, G. Wei, Y. Huang, Dual hesitant bipolar fuzzy hamacher prioritized aggregation operators in multiple attribute decision making, IEEE Access. 6 (2018), 11508–11522.

[14] H.L. Yang, S.G. Li, W.H. Yang, *et al.*, Notes on bipolar fuzzy graphs, Inf. Sci. 242 (2013), 113–121.

[15] E.H. Ruspini, A new approach to clustering, Inf. Control. 15 (1969), 22–32.

[16] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, J. Cybern. 3 (1974), 32–57.

[17] S.S. Khan, A. Ahmad, *et al.*, Cluster center initialization algorithm for k-means clustering, Pattern Recognit. Lett. 25 (2004), 1293–1302.

[18] J. Nayak, B. Naik, H.S. Behera, *et al.*, Hybrid chemical reaction based metaheuristic with fuzzy c-means algorithm for optimal cluster analysis, Expert Syst. Appl. 79 (2017), 282–295.

[19] P.Z. Wang, A factor spaces approach to knowledge representation, Fuzzy Sets Syst. 36 (1990), 113–124.

[20] M. Fraiwan, M. Fraiwan, H, Alkhalaf, *et al.*, Shannon entropy and fuzzy C-means weighting for AI-based diagnosis of vertebral column diseases, J. Ambient Intell. Humanized Comput. 11 (2020), 2557–2566.

[21] J.C. Bezdek, R. Ehrlich, W. Full, FCM: the fuzzy c-means clustering algorithm, Comput. Geosci. 10 (1984), 191–203.