



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

Zhenyu Lu and Xianyun Huang

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

March 11, 2020

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

Zhenyu Lu
Artificial Intelligence Institute
Nanjing University of Information Science and Technology
Nanjing, China
luzhenyu76@163.com

Xianyun Huang
Scientific Research Post
Suzhou Institute of Metrology
Suzhou, China
huangxianyun06@163.com

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-pole 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-pole 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.

Keywords—multi-polar fuzzy entropy, fuzzy c-means clustering, multi-polar fuzzy feature, dynamic truck scale

I. 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 multipolarity [5]. In this paper, in order to solve the existing problem of fuzzy c-means clustering, in the process of bipolar fuzzy characteristic information clustering analysis, an improved *FCM* algorithm based on multi-polar fuzzy information entropy is proposed.

II. MULTI-POLAR FUZZY SET THEORY

A. Fuzzy set theory

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

Defination 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:

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

where $0 \leq u_a(Z) \leq 1$ extending the concept in classical set theory to fuzzy sets. $u_A : Z \mapsto [0, 1]$ is represented 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 .

B. Bipolar fuzzy set theory

An extension of fuzzy set, called bipolar fuzzy set, 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].

Defination 2(see Zhang [7])

Zhang [7] defined a bipolar fuzzy set as a pair (u^+, u^-) , where $u^+ : X \rightarrow [0, 1]$ and $u^- : X \rightarrow [-1, 0]$ are any mappings. Let $B = (u^+, u^-)$ be the bipolar fuzzy sets on X . u^+ In which, 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. 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 [8-14].

C. multi-polar fuzzy set

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

$$\begin{aligned} D &= d(A, B) \\ &= \frac{1}{n} \sum_{i=1}^n ((u_A(x_1) - u_B(x_1))^2 \\ &\quad + (u_A(x_2) - u_B(x_2))^2 + \dots \\ &\quad + (u_A(x_n) - u_B(x_n))^2) \end{aligned} \quad (2)$$

Let $A : V \rightarrow [0, 1]^m$ is the m-pole fuzzy set on V , $B : E \rightarrow [0, 1]^m$ is the m-polar fuzzy set on E , $E \subseteq V * V$ satisfies the symmetry, scilicet $\langle x, y \rangle \in E \iff \langle y, x \rangle \in E$. if $\forall \langle x, y \rangle \in E$ has $B(\langle x, y \rangle) \leq \inf\{A(x), A(y)\}$, 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^[16] 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 $B(\langle x, y \rangle) = B(\langle y, x \rangle) (\forall (x, y) \in E)$ and $B(\langle x, x \rangle) = 0 ([0, 1]^m \text{ minimum element } (\forall x \in V))$, then G is a bipolar fuzzy graph[15] 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 $\forall \langle x, y \rangle \in E$, there is $B(\langle x, y \rangle) = \inf A(x), A(y)$, then G is a strong multi-polar fuzzy graph. The complement of G is a basic set of (V, E) , multi-polar

fuzzy graph $\bar{G} = (A, \bar{B})$, where $\bar{B} : E \rightarrow [0, 1]^m$ is defined as

$$\begin{aligned} \forall \langle x, y \rangle \in E, i \in m, p_i^\circ \bar{B}(\langle x, y \rangle) \\ = \begin{cases} 0, \text{ if } p_i^\circ \bar{B}(\langle x, y \rangle) < 0 \\ \text{inf } p_i^\circ A(x), p_i^\circ A(y), \text{ if } p_i^\circ \bar{B}(\langle x, y \rangle) > 0 \end{cases} \end{aligned} \quad (3)$$

D. 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 $X = \{x_1, x_2, \dots, x_n\}$, the bipolar fuzzy set on X can be expressed as

$$A = \{\langle x_i, u_A^P(x_i), u_A^N(x_i) \rangle | x_i \in X\}, i = 1, 2, \dots, n$$

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 $X = \{x_1, x_2, \dots, x_n\}$, the multipolar fuzzy set on X can be expressed as

$$A = \{\langle x_i, u_A^1(x_i), u_A^2(x_i), \dots, u_A^m(x_i) \rangle | x_i \in X\}, i = 1, 2, \dots, n$$

In the formula, $u_A^j(x_i) \in [0, 1], j = 1 \dots m$, denotes the membership degree of x_i belongs to A .

Definition 6

The set $X = \{x_1, x_2, \dots, x_n\}$ is a non-empty set, and $E(A)$ is the set of bipolar fuzzy entropy.

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

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 $X = \{x_1, x_2, \dots, x_n\}$ constitute the evaluation criterion, $C = \{c_1, c_2, \dots, c_m\}$ is the condensed group.

Definition 7

The set $X = \{x_1, x_2, \dots, x_n\}$ is a non-empty set, and $E(A)$ is the multipolar fuzzy entropy of the set, then

$$\begin{aligned} A = \left(\frac{\langle c_{11}, c_{12}, \dots, c_{1m} \rangle}{x_1}, \right. \\ \left. \frac{\langle c_{21}, c_{22}, \dots, c_{2m} \rangle}{x_2}, \dots, \right. \\ \left. \frac{\langle c_{n1}, c_{n2}, \dots, c_{nm} \rangle}{x_n} \right) \end{aligned} \quad (5)$$

$$\begin{aligned} E(c_j) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\langle 0, 0, \dots, c_{1j}, \dots, 0, 0 \rangle}{x_1} \right. \\ \left. + \frac{\langle 0, 0, \dots, c_{2j}, \dots, 0, 0 \rangle}{x_2} + \dots \right. \\ \left. + \frac{\langle 0, 0, \dots, c_{nj}, \dots, 0, 0 \rangle}{x_n} \right) \end{aligned} \quad (6)$$

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

III. FUZZY C MEAN CLUSTERING BASED ON MULTI-POLAR FUZZY ENTROPY

A. 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 [17]. Based on this, Dunn [18] proposed *FCM* algorithm in 1974. Fuzzy c-means 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, cell analysis.

Fuzzy c-means clustering (*FCM*) algorithm, is going to use membership to determine each sample $X = \{x_1, x_2, \dots, x_n\}$, is a soft clustering algorithm express the degree to which it belongs to a class $C_{center} = \{c_1, c_2, \dots, c_n\}$ ^[19].

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 membership function. this function that 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 objects that may 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 $X = \{x_i | x_i \in X\}$ defines a fuzzy set S .

FCM puts the sample data set, which is composed of sample data of n research objects divided into categories $X = \{x_1, x_2, \dots, x_n\}$ a number of c , and find the clustering center c_i of each group, make 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. $U = \{u_{ij} | i = 1, 2, \dots, c \quad j = 1, 2, \dots, n\}$ represents the sample data x_j , and for the membership degree of category c_i , u_{ij} is the membership degree matrix and satisfies:

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

General form of the objective function *Obj* of *FCM*:

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

Where, $0 \leq u_{ij} \leq 1$, 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:

$$\begin{aligned} \overline{Obj}(U, c_1, c_2, \dots, c_c, \lambda_1, \lambda_2, \dots, \lambda_n) \\ = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 + \sum_{j=1}^n \left(\sum_{i=1}^c u_{ij} - 1 \right) \end{aligned} \quad (9)$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ is the Lagrange multiplier. Derivation of equation (9), you can get the objective function. The necessary condition for getting the minimum value is:

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

$$\begin{aligned} u_{ij} &= \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}}} \\ &= \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{-m}} \end{aligned} \quad (11)$$

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

Traditional fuzzy c-means clustering has three defects. Firstly, the characteristic values of clustering samples are all exact values, which are often difficult to be accurately given in practical problems; secondly, the algorithm is not stable, and the classification effect is often affected by the initial clustering center. *FCM* the existence of high sensitivity to the cluster center initialization problem, literature [21] yuan heuristic 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.

B. 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:

$$E(g_j) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\langle 0, 0, \dots, g_{1j}, \dots, 0, 0 \rangle}{x_1} + \frac{\langle 0, 0, \dots, g_{2j}, \dots, 0, 0 \rangle}{x_2} + \dots + \frac{\langle 0, 0, \dots, g_{nj}, \dots, 0, 0 \rangle}{x_n} \right) \quad (12)$$

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

According to the literature [25] formula 5 can be written as:

$$E(g_j) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\text{Mean} \langle 0, 0, \dots, g_{1j}, \dots, 0, 0 \rangle}{x_1} + \frac{\text{Mean} \langle 0, 0, \dots, g_{2j}, \dots, 0, 0 \rangle}{x_2} + \dots + \frac{\text{Mean} \langle 0, 0, \dots, g_{nj}, \dots, 0, 0 \rangle}{x_n} \right) \quad (13)$$

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

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

$$p'_{ij} = w_j * P_{ij} \quad (15)$$

C. Identify cluster centers

Fuzzy C-means 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. When using fuzzy C-means clustering to classify the samples, since *FCM* is sensitive to the initial cluster center [21], the clustering results obtained by the different initial cluster centers are also quite different. 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 , $d(p'_i, p'_1)d(p'_i, p'_2) \dots d(p'_i, p'_n)$, let them rearrange from big to small.

$$d(p'_i, p'_{(1)}) \leq d(p'_i, p'_{(2)}) \leq \dots \leq d(p'_i, p'_{(n)}) \quad (16)$$

In the formula, subscript(i) refers the subscript rearrangement to satisfy the above condition. Due to $d(p'_i, p'_{(i)}) = 0$, 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

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

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

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

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, \dots, 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 $H = \{p'_{(1)}, p'_{(2)}, \dots, p'_{(c)}\}$; 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.

$$\max(\min(d(p'_r, Z_1), d(p'_r, Z_2))) \quad r = 1, 2 \dots, c \quad (19)$$

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, \dots, Z_{k-1} , take out in H to meet

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

The feature vector of d is the k^{th} cluster center $Z_k (k = 1, 2, \dots, c)$. According to this, the clustering center $Z = (Z_1, Z_2, \dots, Z_k)$ 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.

D. 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.

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

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

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

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

In the formula, m is the fuzziness parameter.

Update the cluster center with the membership matrix, where the k^{th} cluster center is recorded as Z_k ,

$$Z_k = \{z_{k1}, z_{k2}, \dots, z_{ks}\} \quad (23)$$

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

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

E. 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 (3.11) 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 .

IV. EXPERIMENTAL VERIFICATION

The bipolar fuzzy theory was used to improve the fuzzy C-means clustering [23]. It was successfully applied to the diagnosis of diseases, making the diagnosis of diseases more accurate. Then, we apply the algorithm 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 pre-judgment 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 multipolar fuzzy theory to give membership. The membership degree of the g_1 fault type is as follows: $A(x) = \langle 0.49, 0.46, 0.51 \rangle$, $A(y) = \langle 0.45, 0.42, 0.59 \rangle$, $A(z) = \langle 0.50, 0.40, 0.54 \rangle$, $A(u) = \langle 0.40, 0.49, 0.60 \rangle$ and $A(v) = \langle 0.51, 0.52, 0.50 \rangle$.

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.

V. 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

Step1 : Similarly, the 3-pole fuzzy set $A : X \rightarrow [0, 1]^3$ of 5 vehicles is described as follows:

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

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

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

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

Step2 : Sample eigenvalue matrix

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

TABLE I
SAMPLE EIGENVALUE MATRIX

Obtained weighted sample eigenvalue matrix P'

	g_1	g_2	g_3	g_4
x	$\langle 01230,01155,01281 \rangle$	$\langle 01124,01246,01173 \rangle$	$\langle 01286,01163,01212 \rangle$	$\langle 01312,01183,01286 \rangle$
y	$\langle 01130,01055,01482 \rangle$	$\langle 01050,01197,01295 \rangle$	$\langle 01163,01113,01262 \rangle$	$\langle 01338,01209,01312 \rangle$
z	$\langle 01256,01004,01356 \rangle$	$\langle 01099,01246,01075 \rangle$	$\langle 01113,01336,01237 \rangle$	$\langle 01492,01209,01286 \rangle$
u	$\langle 01004,01230,01507 \rangle$	$\langle 01197,01246,01173 \rangle$	$\langle 01138,01311,01163 \rangle$	$\langle 01415,01312,01183 \rangle$
v	$\langle 01281,01306,01256 \rangle$	$\langle 01026,01099,01295 \rangle$	$\langle 01014,01187,01286 \rangle$	$\langle 01466,01338,01106 \rangle$

TABLE II
SAMPLE EIGENVALUE MATRIX

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

$z_1^{(0)}$	$z_2^{(0)}$	$z_3^{(0)}$
$\langle 01230,01155,01281 \rangle$	$\langle 01281,01306,01256 \rangle$	$\langle 01256,01004,01356 \rangle$
$\langle 01124,01246,01173 \rangle$	$\langle 01026,01099,01295 \rangle$	$\langle 01099,01246,01075 \rangle$
$\langle 01286,01163,01212 \rangle$	$\langle 01014,01187,01286 \rangle$	$\langle 01113,01336,01237 \rangle$
$\langle 01312,01183,01286 \rangle$	$\langle 01466,01338,01106 \rangle$	$\langle 01492,01209,01286 \rangle$

TABLE III
INITIAL CLUSTER CENTER

After the initial clustering center, the membership matrix U is:

	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

TABLE IV
THE MEMBERSHIP MATRIX $U^{(0)}$

Sample feature clustering center after one iteration

$z_1^{(1)}$	$z_2^{(1)}$	$z_3^{(1)}$
< 01183,01136,01353 >	< 01267,01300,01269 >	< 01214,01038,01383 >
< 01112,01233,01204 >	< 01033,01107,01290 >	< 01110,01243,01100 >
< 01241,01164,01220 >	< 01021,01192,01281 >	< 01119,01320,01228 >
< 01328,01202,01283 >	< 01463,01335,01112 >	< 01473,01223,01273 >

TABLE V
INITIAL CLUSTER CENTER

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

	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

TABLE VI
THE MEMBERSHIP MATRIX $U^{(1)}$

Sample feature clustering center after iteration two iteration.

$z_1^{(2)}$	$z_2^{(2)}$	$z_3^{(2)}$
< 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 >

TABLE VII
CLUSTERING CENTER $Z^{(2)}$

Calculating 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

TABLE VIII
THE MEMBERSHIP MATRIX $U^{(2)}$

	Number of iterations	$J(U^{(0)}, Z^{(0)}) - J(U^{(n)}, Z^{(n)})$
Literature [20] algorithm	8	0.8446
Literature [24] algorithm	4	0.7562
Algorithm in this paper	2	0.2635

TABLE IX
ALGORITHM COMPARISON

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.

ACKNOWLEDGMENT

First and foremost, I would like to show my deepest gratitude to my supervisor, Prof Lu Zhenyu, a respectable, responsible and resourceful scholar, who has provided me with valuable guidance in every stage of the writing of this thesis. Without his enlightening instruction, impressive kindness and patience, I could not have completed my thesis. His keen and vigorous academic observation enlightens me not only in this thesis but also in my future study.

I shall extend my thanks to Weighing laboratory of Suzhou Institute of Metrology for providing data. I would also like to thank all members of research team who participated this study with great cooperation.

Last but not least, I'd like to thank all my family members, for their encouragement and support.

REFERENCES

- [1] Alshehri N O, Akram M. "Cayley bipolar fuzzy graphs", *The Scientific World Journal*, vol. 3, pp. 156786, 2013.
- [2] Mesiarová-Zemánková, Andrea, Ahmad K. "Extended multi-polarity and multi-polar-valued fuzzy sets", *Fuzzy Sets and Systems*, vol. 234, pp. 61-78, 2014.
- [3] Zhou J, Yu Y, Liu Y, et al. "Solving nonlinear optimization problems with bipolar fuzzy relational equation constraints". *Journal of Inequalities and Applications*, vol. 2016, pp. 1-10, 2016.
- [4] Bezdek J, Hathaway R, Sobin M, et al. "Convergence theory for fuzzy c-means: counterexamples and repairs". *IEEE Transactions on Systems Man and Cybernetics*, vol. 17, no. 5, pp. 873-877, 1987.
- [5] Pal N R, Bezdek J C. "On cluster validity for the fuzzy c-means model". *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370-379, 2002.
- [6] Zadeh L A. "Fuzzy sets". *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965.
- [7] Zhang, Wen-Ran. "Bipolar fuzzy sets and relations: a computational framework for cognitive modeling and multiagent decision analysis". *The Industrial Fuzzy Control and Intellige*, pp. 305-309, 1994.
- [8] Han Y, Lu Z, Du Z, et al. "A YinYang bipolar fuzzy cognitive TOPSIS method to bipolar disorder diagnosis". *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 1-10, 2018.
- [9] Han Y, Shi P, Chen S. "Bipolar-valued rough fuzzy set and its applications to decision information system". *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 6, pp. 2358-2370, 2015.
- [10] Rashmanlou H, Samanta S, Pal M, et al. "Bipolar fuzzy graphs with categorical properties". *Fuzzy Information and Engineering*, vol. 7, no. 3, pp. 317-334, 2015.
- [11] Alghamdi M A, Omair A N, Muhammad A. "Multi-criteria decision-making methods in bipolar fuzzy environment". *International Journal of Fuzzy Systems*, vol. 20, no. 6, pp. 2057-2064, 2018.
- [12] Zhou J, Yu Y, Liu Y, et al. "Solving nonlinear optimization problems with bipolar fuzzy relational equation constraints". *Journal of Inequalities and Applications*, vol. 2016, no.1, pp. 126, 2016.
- [13] Hidekaru D, Motoi K, Kenji O, et al. "A novel class of emitting amorphous molecular materials with bipolar character for electroluminescence". *Chemistry of Materials*, vol. 15, no. 5, pp. 1080-1089, 2016.
- [14] Gao H, Wei G, Huang Y. "Dual hesitant bipolar fuzzy hamacher prioritized aggregation operators in multiple attribute decision making". *IEEE Access*, vol. 6, pp. 11508-11522, 2018.
- [15] Yang H L, Li S G, Yang W H, et al. "Notes on bipolar fuzzy graphs. *Information sciences*", vol. 242, pp. 113-121, 2013.
- [16] Mandal P, Ranadive A S. "Hesitant bipolar-valued fuzzy sets and bipolar-valued hesitant fuzzy sets and their applications in multi-attribute group decision making". *Granular Computing*, vol. 4, no. 3, pp. 559-583, 2019.
- [17] Ruspini E H. "A new approach to clustering". *Information and Control*, vol. 15, no. 1, pp. 22-32, 1969.
- [18] Dunn J C. "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters". *Journal of Cybernetics*, vol. 3, no. 3, pp. 32-57, 1974.
- [19] Khan, Shehroz S, Ahmad, et al. "Cluster center initialization algorithm for k-means clustering". *Pattern Recognition Letters*, vol. 25, no. 11, pp. 1293-1302, 2004.
- [20] Bezdek J C, Ehrlich R, Full W. "FCM: The fuzzy c-means clustering algorithm". *Computers and Geosciences*, vol. 10, no. 2, pp. 191-203, 1984.
- [21] Nayak J, Naik B, Behera H S, et al. "Hybrid chemical reaction based metaheuristic with fuzzy c-means algorithm for optimal cluster analysis". *Expert Systems with Applications*, vol. 79, pp. 282-295, 2017.
- [22] Duan, Yanling, et al. "An improved initialization center k-means clustering algorithm based on distance and density". *AIP Conference Proceedings*, vol. 1955, no. 1, pp. 40046, 2018.
- [23] Salleh M N M, Ashikin Samat N. "FCMPSO: An imputation for missing data features in heart disease classification". *Materials Science and Engineering Conference Series*. vol. 226, no. 1, pp. 12102, 2017.
- [24] Wang P Z. "A factor spaces approach to knowledge representation". *Fuzzy Sets and Systems*, vol. 36, no. 1, pp. 113-124, 1990.