

Several new results based on the study of distance measures of intuitionistic fuzzy sets

C. Chen¹ and X. Deng²

^{1,2}*School of Mathematics, South China University of Technology, Guangzhou, 510640, China*

2080200747@qq.com, dxue@scut.edu.cn

Abstract

It is doubtless that intuitionistic fuzzy set (IFS) theory plays an increasingly important role in solving the problems under uncertain situation. As one of the most critical members in the theory, distance measure is widely used in many aspects. Nevertheless, it is a pity that part of the existing distance measures has some drawbacks in practical significance and accuracy. To make up for their drawbacks and pursue more accuracy and effectiveness, in this paper, we propose a new inclusion relation of IFSs and a new definition called strict distance measure. Based on this new relation, an analysis is given to point out that the common shortcoming of Hamming distance measure and Euclidean distance measure is the mishandling of hesitancy degree. Therefore, the role of hesitancy degree in distance measure is studied deeply and then three strict distance measures are put forward to overcome the above shortcoming. In addition, a novel definition called the characteristic function of distance measure is defined to describe the character of strict distance measure. On this basis, a theorem is presented to illustrate the inevitability of the occurrence of unrecognized result in pattern recognition problems in some special cases. This theorem also shows that the problem cannot be entirely attributed to distance measures. In view of this condition, we provide an appropriate solution. Compared with other existing distance measures in some examples, the superiorities of our improved distance measures are demonstrated to be more effective and more significant.

Keywords: Intuitionistic fuzzy set, distance measure, inclusion relation, hesitancy degree, characteristic function.

1 Introduction

Fuzzy set theory, proposed by Zadeh [33] in 1965, received great concern owe to its function to describe the uncertainty in mathematical language. Although there were a few limits with it, it is doubtless that fuzzy set theory provided us a logical way to deal with uncertainty problems and led the research on uncertainty to a new domain with much achievement. One of the most noticeable and remarkable achievements is Atanassov's theory [1] that called intuitionistic fuzzy set (IFS). The theory introduced non-membership degree, the counterpart but no the complement of membership degree, into fuzzy set theory. Since it can describe uncertainties not only from the positive side but also from the opposite side, it is proved that IFS can deal with uncertainty problems more practicability and effectively. For this reason, IFS is widely applied in many practical fields in recent years. For example, Nguyen [22] introduced a new measure of amount of knowledge related to information provided in terms of IFSs and applied it into multiple attribute group decision making; Based on the proposed intuitionistic fuzzy support function, risk function and credibility function, Shen et al. [26] put forward a novel outranking sorting method for multi-criteria group decision making; Botía et al. [4] proposed a new method to analyze optical frequency comb behavior and the spectral shape based on fuzzy cellular automata and IFS; Yin and Li [32] proposed an efficient and practical matching decision making method for matching management of green building technologies supply and demand based on IFS and considering interactions among aspiration criteria; Schitea et al. [24] proposed an multi-criteria decision making method, which is based on IFS and includes WASPAS, COPRAS and EDAS, to select the best location for hydrogen mobility roll-up site selection in Romania. In addition,

there are also other extensions of fuzzy set, such as interval-valued fuzzy set, interval-valued intuitionistic fuzzy set [3], L-fuzzy set [11] and L-intuitionistic fuzzy set [2]. These extensions also promote the rapid development of fuzzy set theory and make it be applied to more fields.

Apart from the extensions of fuzzy set, many researchers prefer to pay their attention to the properties of fuzzy set or its extensions in order to make the theory more practicable. Based on their work, various properties and definitions with respect to fuzzy set or its extensions had been presented in recent years, such as a variety of fuzzy entropy [13, 19, 20, 30], weighted operators with different characteristics [7, 9, 27] and similarity measures [13, 16, 17]. Distance measure, which can describe the deviation degree of IFSs, is also one of the most critical roles in intuitionistic fuzzy set theory. As it is not only frequently used in the comparison of two IFSs for pattern recognition [6] or the ranking of alternatives [10, 25], but also active in other fields, it is necessary to pursue the accuracy of distance measure. Szmidt and Kacprzyk [28] were the first to define the distance measures for two IFSs, which were generalized from the definition of distance measure for fuzzy sets. Then, Grzegorzewski [12], Yang and Chiclana [31], Wang and Xin [29], Jin et al. [15], Shen et al. [25] and other researchers had proposed several kinds of distance measures for IFSs. Nevertheless, it is a pity that part of the existing distance measures has some drawbacks. Chen [5] gave some counterexamples to state the errors that exist in Grzegorzewski's distance measure [12] and the invalid of Euclidean distance [28], Chen and Chang [6] pointed out the distance measure proposed by Zhang et al. [34] has the drawback of "the division by zero", Shen et al. [25] used an instance to show the ineffectiveness of the distance measure proposed by Chen et al. [8] in some special cases, Ngan et al. [21] listed a lot of existing distance measures and pointed out the drawbacks of them by comparison in some examples. Although Shen et al. [25] and Luo et al. [18] had proposed their new distance measures, these distance measures are complex in their form and time-consuming in calculating. Even H-max distance measure, proposed by Ngan et al. [21], lacks of a rational mathematical exposition. In consideration of the importance of distance measure for intuitionistic fuzzy set theory, it is imperative to search for a reasonable and realistic distance measure.

The innovation of this paper lies in the following aspects. First, a new relation between two IFSs, which called complete inclusion, is proposed and the distance measure axioms are refined according to the new relation. Secondly, based on the analysis about the shortcomings of Hamming distance measure and Euclidean distance measure, three kinds of improved distance measures are given. Moreover, the definition of the characteristic function of distance measure is presented to further study some properties of distance measure. Finally, we put forward a theorem to describe the inevitability as to the invalidity of some existing distance measure.

The rest of this paper is organized as follows. In Section 2, the basic conceptions of IFS and some existing distance measures are introduced. In Section 3, a new relation between two IFSs that called complete inclusion is proposed, and then the axioms for distance measures are refined. The shortcomings of Hamming distance measure and Euclidean distance measure are analyzed in Section 4, while three improved distance measures are given and compared with other existing distance measures. In Section 5, we present a definition called the characteristic function of distance measure and make some relevant proofs. Finally, the conclusions are drawn in Section 6.

2 Preliminaries

In this section, some basic concepts of IFS and several relevant distance measures are provided as follows.

Definition 2.1. [33] Let X be a fixed set, a fuzzy set A in X has a form as:

$$A = \{\langle x, \mu_A(x) \rangle | x \in X\}, \quad (1)$$

where the function $\mu_A(x)$ represents the membership degree of the element x to the set A with a range in $[0, 1]$.

Definition 2.2. [1] Let X be a fixed set, an intuitionistic fuzzy set (IFS) A in X has a form as:

$$A = \{\langle x, \mu_A(x), \nu_A(x) \rangle | x \in X\}, \quad (2)$$

where the functions $\mu_A(x)$ and $\nu_A(x)$ should satisfy $0 \leq \mu_A(x) + \nu_A(x) \leq 1$ and represent the membership degree and non-membership degree of the element x to the set A in the range of $[0, 1]$, respectively, $\forall x \in X$.

It is obvious that if we let $\pi_A(x) = 1 - (\mu_A(x) + \nu_A(x))$, we can get the result of $0 \leq \pi_A(x) \leq 1$, $\forall x \in X$. We consider that $\pi_A(x)$ is the hesitancy degree of x to the IFS A , which is always determined by $\mu_A(x)$ and $\nu_A(x)$. Thus, A can also be described as the following form:

$$A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle | x \in X\}.$$

For the sake of convenience, if there is just one element x in X , i.e. $A = \{\langle x, \mu_A(x), \nu_A(x) \rangle\}$, IFS A is called simple intuitionistic fuzzy set (SIFS). In addition, $IFS(X)$ will be used to denote the collection of all IFSs in X in this paper.

Definition 2.3. [1] Let X be a fixed set, A and B be two IFSs in X , where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle | x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle | x \in X\}$. Atanassov defined the inclusion relation \subseteq between A and B as the following form:

$$A \subseteq B \Leftrightarrow \mu_A(x) \leq \mu_B(x), \nu_A(x) \geq \nu_B(x), \forall x \in X. \tag{4}$$

In 2009, Park et al. [23] defined a modified inclusion relation \subseteq^2 between A and B as the following form:

$$A \subseteq^2 B \Leftrightarrow \mu_A(x) \leq \mu_B(x), \nu_A(x) \geq \nu_B(x), \pi_A(x) \geq \pi_B(x), \forall x \in X, \tag{5}$$

where $\pi_A(x) = 1 - (\mu_A(x) + \nu_A(x))$ and $\pi_B(x) = 1 - (\mu_B(x) + \nu_B(x))$, respectively.

Definition 2.4. [29] Let X be a fixed set, d be a mapping from $IFS(X) \times IFS(X)$ to R . We call d a distance measure of IFSs if it can satisfy the axioms listed as follows:

- (i) $d(A, B) \geq 0, \forall A, B \in IFS(X)$;
- (ii) $d(A, B) = d(B, A), \forall A, B \in IFS(X)$;
- (iii) $d(A, B) = 0 \Leftrightarrow A = B, \forall A, B \in IFS(X)$;
- (iv) If $A \subseteq B \subseteq C$, there will be $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C), \forall A, B, C \in IFS(X)$.

To be convenient, the above axioms are called DMA, i.e. the distance measure axioms for intuitionistic fuzzy sets. Let $X = \{x_1, x_2, \dots, x_n\}$ be a finite fixed set, A and B be two IFSs in X , where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle | x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle | x \in X\}$. Partial existing distance measures of IFSs are introduced as follows:
Hamming distance measure of two IFSs [28]:

$$d_H(A, B) = \frac{1}{2n} \sum_{i=1}^n (|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|), \tag{6}$$

Euclidean distance measure of two IFSs [28]:

$$d_E(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + (\pi_A(x_i) - \pi_B(x_i))^2)}, \tag{7}$$

The distance measure of two IFSs proposed by Park et al. [23]:

$$d_{Pa}(A, B) = \frac{1}{4n} \sum_{i=1}^n \left(\begin{array}{l} |\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)| + \\ 2 \max \{ |\mu_A(x_i) - \mu_B(x_i)|, |\nu_A(x_i) - \nu_B(x_i)|, |\pi_A(x_i) - \pi_B(x_i)| \} \end{array} \right), \tag{8}$$

The distance measure of two IFSs proposed by Shen et al. [25]:

$$d_S(A, B) = \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{1}{2} \left\{ \begin{array}{l} [\mu_A(x_i) (1 + \frac{2}{3}\pi_A(x_i) (1 + \pi_A(x_i))) - \mu_B(x_i) (1 + \frac{2}{3}\pi_B(x_i) (1 + \pi_B(x_i)))]^2 + \\ [\nu_A(x_i) (1 + \frac{2}{3}\pi_A(x_i) (1 + \pi_A(x_i))) - \nu_B(x_i) (1 + \frac{2}{3}\pi_B(x_i) (1 + \pi_B(x_i)))]^2 \end{array} \right\}}, \tag{9}$$

H-max distance measure of two IFSs [21]:

$$d_{Hm}(A, B) = \frac{1}{3n} \sum_{i=1}^n (|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\max\{\mu_A(x_i), \nu_B(x_i)\} - \max\{\mu_B(x_i), \nu_A(x_i)\}|), \tag{10}$$

The distance measure of two IFSs proposed by Jiang et al. [14]:

$$d_J(A, B) = \frac{1}{2n} \sum_{i=1}^n \left[\begin{array}{l} \left| \frac{2(\mu_A(x_i)\pi_B(x_i) - \mu_B(x_i)\pi_A(x_i)) - 4(\mu_A(x_i) - \mu_B(x_i))}{4 - \pi_A(x_i)\pi_B(x_i)} \right| + \\ \left| \frac{4(\nu_A(x_i) - \nu_B(x_i)) + 2(\nu_A(x_i)\pi_B(x_i) - \nu_B(x_i)\pi_A(x_i)) + 2(\pi_A(x_i) - \pi_B(x_i))}{4 - \pi_A(x_i)\pi_B(x_i)} \right| \end{array} \right]. \tag{11}$$

3 The new inclusion relation of IFSs and the modified distance measure axioms

Before analyzing the shortcomings of Hamming distance measure and Euclidean distance measure, one thing should be noticed is that the modified inclusion relation \subseteq^2 in formula (5) is not effective enough in some cases [31], so that it is less meaningful than the original inclusion relation \subseteq showed in formula (4). Then, according to the original inclusion relation, DMA (iv) can be rethought as the following form:

(iv') $\forall A, B, C \in IFS(X)$, $A = \{\langle x, \mu_A(x), \nu_A(x) \rangle | x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x) \rangle | x \in X\}$, $C = \{\langle x, \mu_C(x), \nu_C(x) \rangle | x \in X\}$. If $\mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$, $\nu_A(x) \geq \nu_B(x) \geq \nu_C(x)$, $\forall x \in X$, there should be $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$.

However, the original inclusion relation is also not perfect enough. Let's consider the following examples:

Example 3.1. Let A , B and C be three different IFSs in the fixed set $X = \{x_1, x_2, x_3\}$, where

$$\begin{aligned} A &= \{\langle x_1, 0.1, 0.7, 0.2 \rangle, \langle x_2, 0.2, 0.7, 0.1 \rangle, \langle x_3, 0.1, 0.8, 0.1 \rangle\}, \\ B &= \{\langle x_1, 0.2, 0.6, 0.2 \rangle, \langle x_2, 0.3, 0.7, 0 \rangle, \langle x_3, 0.3, 0.5, 0.2 \rangle\}, \\ C &= \{\langle x_1, 0.3, 0.6, 0.1 \rangle, \langle x_2, 0.4, 0.4, 0.2 \rangle, \langle x_3, 0.4, 0.5, 0.1 \rangle\}. \end{aligned}$$

It is apparent that $A \subseteq B \subseteq C$, especially there is $\mu_A(x) < \mu_B(x) < \mu_C(x)$, $\forall x \in X$. To explain the disparity between them better, assume that there are three ballots which are denoted as x_1, x_2, x_3 , while A, B and C represent three sets of the voting results that come from ten elder, ten youths and ten children, respectively. Then, A can be interpreted as: for x_1 , one elder voted for support, seven elders voted against and two elders abstained; for x_2 , two elders voted for support, seven elders voted against and one elder abstained; for x_3 , one elder voted for support, eight elders voted against and one elder abstained. Similarly, B and C can be explained in this way. If we use distance measure $d(A, B)$ and $d(A, C)$ to imply the disparities that between elders and youths, elders and children with respect to the support to all the ballots, respectively, it is evident that $d(A, B) < d(A, C)$ but not $d(A, B) \leq d(A, C)$. In addition, we have $d(B, C) < d(A, C)$ but not $d(B, C) \leq d(A, C)$.

Example 3.2. Let D, E and F be three different IFSs in the fixed set $X = \{x_1, x_2, x_3\}$, where

$$\begin{aligned} D &= \{\langle x_1, 0.1, 0.6, 0.3 \rangle, \langle x_2, 0.3, 0.6, 0.1 \rangle, \langle x_3, 0.2, 0.7, 0.1 \rangle\} \\ E &= \{\langle x_1, 0.1, 0.5, 0.4 \rangle, \langle x_2, 0.4, 0.5, 0.1 \rangle, \langle x_3, 0.3, 0.6, 0.1 \rangle\} \\ F &= \{\langle x_1, 0.3, 0.2, 0.5 \rangle, \langle x_2, 0.4, 0.4, 0.2 \rangle, \langle x_3, 0.3, 0.3, 0.4 \rangle\} \end{aligned}$$

It is apparent that $D \subseteq E \subseteq F$, especially there is $\nu_D(x) > \nu_E(x) > \nu_F(x)$, $\forall x \in X$. We can still use the voting model to explain these IFSs just as in example 3.1. Assume that there are three ballots which are denoted as x_1, x_2, x_3 , while D, E and F represent three sets of the voting results that come from ten elder, ten youths and ten children, respectively. If we use distance measure $d(D, E)$ and $d(D, F)$ to imply the disparities that between elders and youths, elders and children with respect to the opposition of all ballots, respectively, it is evident that $d(D, E) < d(D, F)$ but not $d(D, E) \leq d(D, F)$. In addition, we also have $d(E, F) < d(D, F)$ but not $d(E, F) \leq d(D, F)$.

From example 3.1 and 3.2, we can find that DMA (iv) is not strictly enough and may make the result inconsistent with the fact due to the non-strict inclusion relation of IFSs. In order to make the inclusion relation of IFSs more strict and practical while make the distance measure for IFSs more accurate, the new relation called complete inclusion of IFSs is proposed as following:

Definition 3.3. Let X be a fixed set, A and B be two IFSs in X , where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle | x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle | x \in X\}$. The complete inclusion relation \subset^C between A and B is defined as follows:

$$A \subset^C B \Leftrightarrow \begin{cases} \mu_A(x) < \mu_B(x), \nu_A(x) \geq \nu_B(x), \forall x \in X, \text{ or} \\ \mu_A(x) \leq \mu_B(x), \nu_A(x) > \nu_B(x), \forall x \in X \end{cases} \quad (12)$$

For convenience, if there are several IFSs A_1, A_2, \dots, A_m in X , the complete inclusion relation \subset^C among them is stipulated as follows:

$$\begin{aligned} &A_1 \subset^C A_2 \subset^C \dots \subset^C A_m \\ \Leftrightarrow &\begin{cases} \mu_1(x) < \mu_2(x) < \dots < \mu_m(x), \nu_1(x) \geq \nu_2(x) \geq \dots \geq \nu_m(x), \forall x \in X, \text{ or} \\ \mu_1(x) \leq \mu_2(x) \leq \dots \leq \mu_m(x), \nu_1(x) > \nu_2(x) > \dots > \nu_m(x), \forall x \in X \end{cases} \end{aligned} \quad (13)$$

According to the new definition, DMA (iv) should be refined as the following form:

(iv'') If $A \subseteq B \subseteq C$, there should be $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$, $\forall A, B, C \in IFS(X)$. Especially, If $A \subset^C B \subset^C C$, there should be $d(A, B) < d(A, C)$ and $d(B, C) < d(A, C)$, $\forall A, B, C \in IFS(X)$.

Obviously, DMA (iv'') is practicable and more reasonable than DMA (iv). Therefore, we can get more strict distance measure for IFSs.

Definition 3.4. Let X be a fixed set, d be a mapping from $IFS(X) \times IFS(X)$ to R . We call d a strict distance measure of IFSs if it can satisfy the axioms listed as follows:

- (i) $d(A, B) \geq 0, \forall A, B \in IFS(X)$;
- (ii) $d(A, B) = d(B, A), \forall A, B \in IFS(X)$;
- (iii) $d(A, B) = 0 \Leftrightarrow A = B, \forall A, B \in IFS(X)$;
- (iv) If $A \subseteq B \subseteq C$, there will be $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$, $\forall A, B, C \in IFS(X)$. Especially, If $A \subset^C B \subset^C C$, there should be $d(A, B) < d(A, C)$ and $d(B, C) < d(A, C)$, $\forall A, B, C \in IFS(X)$.

To be convenient, the above modified axioms are called MDMA. Just as DMA (iv') is equivalent to DMA (iv), MDMA (iv) is equivalent to the following form:

(iv*) $\forall A, B, C \in IFS(X)$, $A = \{ \langle x, \mu_A(x), \nu_A(x) \rangle | x \in X \}$, $B = \{ \langle x, \mu_B(x), \nu_B(x) \rangle | x \in X \}$, $C = \{ \langle x, \mu_C(x), \nu_C(x) \rangle | x \in X \}$. If $\mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$, $\nu_A(x) \geq \nu_B(x) \geq \nu_C(x)$, $\forall x \in X$, there should be $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$. Especially, if $\mu_A(x) < \mu_B(x) < \mu_C(x)$, $\nu_A(x) \geq \nu_B(x) \geq \nu_C(x)$, $\forall x \in X$ or $\mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$, $\nu_A(x) > \nu_B(x) > \nu_C(x)$, $\forall x \in X$, there should be $d(A, B) < d(A, C)$ and $d(B, C) < d(A, C)$.

4 Three improved distance measures for intuitionistic fuzzy sets

In this section, we will use some examples to analyze the shortcomings of Hamming distance measure and Euclidean distance is measure. And then, after the other analysis with respect to the role of hesitancy degree in the distance measure of IFSs, three improved distances will be given.

4.1 The shortcomings of Hamming distance measure and Euclidean distance measure

Let us look at the following instances:

Example 4.1. If A, B and C are three SIFSs in the fixed set $X = \{x\}$, where $A = \{ \langle x, 0.3, 0.3 \rangle \}$, $B = \{ \langle x, 0.5, 0.25 \rangle \}$ and $C = \{ \langle x, 0.5, 0.1 \rangle \}$, then, by formula (6), the Hamming distance between A and B is $d_H(A, B) = 0.2$ while it between A and C is also $d_H(A, C) = 0.2$, i.e. $d_H(A, B) = d_H(A, C)$. But it is easily to find out that $A \subset^C B \subset^C C$, so Hamming distance measure cannot satisfy MDMA (iv), which indicates that it is not effective enough in some situations.

Example 4.2. If D, E and F are three SIFSs in the fixed set $X = \{x\}$, where $D = \{ \langle x, 0, 1, 0 \rangle \}$, $E = \{ \langle x, 0, 0, 1 \rangle \}$ and $F = \{ \langle x, 1, 0, 0 \rangle \}$, then, by formulas (6) and (7), the Hamming distance and the Euclidean distance between D and E are $d_H(D, E) = 1$ and $d_E(D, E) = 1$, respectively, while the distances between D and F are $d_H(D, F) = 1$ and $d_E(D, F) = 1$, respectively. In addition, we also have $d_H(E, F) = 1$ and $d_E(E, F) = 1$, i.e.

$$d_H(D, E) = d_H(D, F) = d_H(E, F), \quad d_E(D, E) = d_E(D, F) = d_E(E, F).$$

However, it is easily to find out that $D \subset^C E \subset^C F$, which points out that both Hamming distance measure and Euclidean distance measure cannot satisfy MDMA (iv), and it indicates that they are not effective enough in some situations.

In fact, if we make a modification on Euclidean distance and transform it to the following form:

$$d_{E'}(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^n \left((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 \right)},$$

It could be easily proved that the modified Euclidean distance measure can satisfy DMA (iv), while the original one cannot for its mishandling of hesitancy degree. Even the original Euclidean distance is not a distance measure. Although Hamming distance measure can satisfy DMA (iv), it does not take more practical significance of hesitancy degree into account. The following example given by Ngan et al. [21] also indicates this drawback of Hamming distance measure, so that it should be paid an intense attention.

Example 4.3. [21] If there are four SIFSs G , H , I and J in the fixed set $X = \{x\}$, where $G = \{\langle x, 0.5, 0.5, 0 \rangle\}$, $H = \{\langle x, 0.4, 0.6, 0 \rangle\}$, $I = \{\langle x, 0.6, 0.4, 0 \rangle\}$ and $J = \{\langle x, 0.4, 0.4, 0.2 \rangle\}$. By using the ten-people-voting model to interpret, G represents the number of people that vote for x is equal to the number of people that vote against and there is not any person who hesitates; H represents four people vote for x , six people vote against and there is not any person who hesitates; I represents six people vote for x , four people vote against and there is not one who hesitates; whereas J represents the number of people vote for x is equal to the number of people vote against and there are only two people who hesitate. Ngan et al. pointed out that the difference between G and J should be less than those between G and H , G and I . However, we have

$$d_H(G, J) = 0.2 > d_H(G, H) = d_H(G, I) = 0.1,$$

which is contrary to our expectation.

From the above example, we can find that the unreasonable result came from Hamming distance measure is also caused by the mishandling of hesitancy degree. Thereby, it is imperative to research for a new way to introduce hesitancy degree appropriately.

4.2 The role of hesitancy degree for distance measure of IFSs

In order to introduce hesitancy degree appropriately, it is rational to analyze the role of hesitancy degree for distance measure of IFSs in the first place.

To make the analysis simpler and clearer, we consider two SIFSs A and B in the fixed set $X = \{x\}$, where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle\}$, respectively. Then, there are two different conditions that should be treated separately: $\Delta\pi = 0$ and $\Delta\pi > 0$, where $\Delta\pi = |\pi_A(x) - \pi_B(x)|$.

(i) When $\Delta\pi = 0$.

Suppose that $\mu_A(x) + \nu_A(x) = \mu_B(x) + \nu_B(x) = 1 - \theta$, where $0 \leq \theta \leq 1$. Then, there are $\pi_A(x) = \pi_B(x) = \theta$, i.e. $\Delta\pi = 0$. For convenience, $IFS(X; \theta)$ is used to denote the subset of $IFS(X)$, which satisfies that for any SIFS $U = \{\langle x, \mu_U(x), \nu_U(x) \rangle\}$ in X , $U \in IFS(X; \theta) \Leftrightarrow \mu_U(x) + \nu_U(x) = 1 - \theta$. Thereby, we have $A \in IFS(X; \theta)$ and $B \in IFS(X; \theta)$.

Consider any fuzzy set $A' = \{\langle x, \mu_{A'}(x) \rangle\}$ in X , in accordance with Atanassov's theory [1], A' can be expressed as the form $A' = \{\langle x, \mu_{A'}(x), 1 - \mu_{A'}(x) \rangle\}$. Then, compared with A' , it is doubtless that both $A = \{\langle x, \mu_A(x), 1 - \theta - \mu_A(x) \rangle\}$ and $B = \{\langle x, \mu_B(x), 1 - \theta - \mu_B(x) \rangle\}$ can be regarded as two special fuzzy sets in $IFS(X; \theta)$. On this basis, the Hamming distance $d_H(A, B)$ and the Euclidean distance $d_E(A, B)$ between the two special fuzzy sets A and B are:

$$d_H(A, B) = |\mu_A(x) - \mu_B(x)|, \quad (14)$$

$$d_E(A, B) = \sqrt{(\mu_A(x) - \mu_B(x))^2}. \quad (15)$$

Moreover, it should be noticed that $|\mu_A(x) - \mu_B(x)| = |\nu_A(x) - \nu_B(x)|$ and $(\mu_A(x) - \mu_B(x))^2 = (\nu_A(x) - \nu_B(x))^2$, thus formula (14) is identical with:

$$d_H(A, B) = \frac{1}{2} (|\mu_A(x) - \mu_B(x)| + |\nu_A(x) - \nu_B(x)|). \quad (16)$$

Similarity, formula (15) is identical with:

$$d_E(A, B) = \sqrt{\frac{1}{2} \left((\mu_A(x) - \mu_B(x))^2 + (\nu_A(x) - \nu_B(x))^2 \right)}. \quad (17)$$

To introduce hesitancy degree while keep both the form of formula and calculation as simple as possible, two different methods are provided as follows:

Case 1. Assume that $f(\Delta\pi)$ is a function of $\Delta\pi$ and it satisfies $f(0) = 0$. The Hamming distance of two SIFSs may have the following form:

$$d_H(A, B) = \frac{1}{2} (|\mu_A(x) - \mu_B(x)| + |\nu_A(x) - \nu_B(x)| + f(\Delta\pi)), \quad (18)$$

or the Euclidean distance of two SIFSs may have the following form:

$$d_E(A, B) = \sqrt{\frac{1}{2} \left((\mu_A(x) - \mu_B(x))^2 + (\nu_A(x) - \nu_B(x))^2 + f(\Delta\pi) \right)}.$$

Case 2. Assume that $g(\Delta\pi)$ is a function of $\Delta\pi$ and it satisfies $g(0) = 1$. The Hamming distance of two SIFSSs may have the following form:

$$d(A, B) = \frac{1}{2} (|\mu_A(x) - \mu_B(x)| + |\nu_A(x) - \nu_B(x)|) \times g(\Delta\pi), \tag{20}$$

or the Euclidean distance of two SIFSSs may have the following form:

$$d_E(A, B) = \sqrt{\frac{1}{2} \left((\mu_A(x) - \mu_B(x))^2 + (\nu_A(x) - \nu_B(x))^2 \right)} \times g(\Delta\pi). \tag{21}$$

Both of Hamming distance measure and Euclidean distance measures are adopted the first method, but they met failure. Hence, we attempt to adopt the second one. Before finalizing the function $g(\Delta\pi)$, let us analyze the other condition as to $\Delta\pi$:

(ii) When $\Delta\pi > 0$.

Under this condition, the analysis in reference to Hamming distance of IFSs is held based on example 4.3. From example 4.3, we know that:

$$\begin{aligned} |\mu_G(x) - \mu_H(x)| &= |\mu_G(x) - \mu_I(x)| = |\mu_G(x) - \mu_J(x)| = 0.1, \\ |\nu_G(x) - \nu_H(x)| &= |\nu_G(x) - \nu_I(x)| = |\nu_G(x) - \nu_J(x)| = 0.1, \\ \Delta\pi_{GH}(x) &= |\pi_G(x) - \pi_H(x)| = 0 = |\pi_G(x) - \pi_I(x)| = \Delta\pi_{GI}(x), \\ \Delta\pi_{GJ}(x) &= |\pi_G(x) - \pi_J(x)| = 0.1 > 0 = \Delta\pi_{GH}(x) = \Delta\pi_{GI}(x), \end{aligned}$$

while it is pointed out that there should be $d(G, J) < d(G, H)$ and $d(G, J) < d(G, I)$. Consequently, we can suppose that $\Delta\pi$ has a weakening effect on the Hamming distance between two SIFSSs, and this effect is positively correlated with the value of $\Delta\pi$. Based on this hypothesis, we consider $g(\Delta\pi)$ is a decreasing function of $\Delta\pi$ under the premise that formula (20) is adopted to measure the Hamming distance of SIFSSs.

Next, we will demonstrate the validity of the hypothesis with another instance.

Example 4.4. If there are three different SIFSSs K , M and N in the fixed set $X = \{x\}$, where $K = \{x, 0.5, 0.5, 0\}$, $M = \{x, 0.4, 0.3, 0.3\}$ and $N = \{x, 0.3, 0.6, 0.1\}$, respectively. By using the ten-people-voting model again, we can get an explanation which is similar to example 4.3 and then regard that K is comparatively neutral. For the sake of making the contrast more distinct, we consider the Hamming distance between K and M represents the minimal difficulty in transforming the voting result of ten people in M into a comparatively neutral state, and the Hamming distance between K and N represents the minimal difficulty in transforming the voting result of ten people in N into a comparatively neutral state. Without question, the difficulty that to persuade one who holds opposing view to vote for support is much higher than persuade a neutral one to vote for agreement or opposition, which clearly shows that the Hamming distance between K and M is more compact. To put it another way, we have the following result:

$$\begin{aligned} |\mu_K(x) - \mu_M(x)| + |\nu_K(x) - \nu_M(x)| &= 0.3 = |\mu_K(x) - \mu_N(x)| + |\nu_K(x) - \nu_N(x)|, \\ \Delta\pi_{KM}(x) &= |\pi_K(x) - \pi_M(x)| = 0.3 > 0.1 = |\pi_K(x) - \pi_N(x)| = \Delta\pi_{KN}(x), \\ d(K, M) &< d(K, N) \Rightarrow g(\Delta\pi_{KM}(x)) < g(\Delta\pi_{KN}(x)), \end{aligned}$$

where d is a distance measure of IFSs and $g(\Delta\pi)$ is a function of $\Delta\pi$. Both of them are presented in formula (20). This result indicates the aforementioned hypothesis that $g(\Delta\pi)$ is a decreasing function of $\Delta\pi$ is feasible.

In fact, under the condition that

$$|\mu_K(x) - \mu_M(x)| + |\nu_K(x) - \nu_M(x)| = |\mu_K(x) - \mu_N(x)| + |\nu_K(x) - \nu_N(x)|,$$

if we increase the value of $\Delta\pi_{KM}(x)$, the number of neutral people will increase. It denotes that the difficulty to transform the voting result into a comparatively neutral state will be decreased, which is equivalent that the Hamming distance between K and M will be reduced. Thus, it shows that our hypothesis is practicable and rational.

Since any IFS can be viewed as the composition of a certain amount of SIFSSs, it is trivial to prove that the hypothesis is also feasible for the Hamming distance between any two IFSs that in $IFS(X)$.

On the other hand, to ensure formula (20) can satisfy MDMA(i), $g(\Delta\pi)$ should meet the demand that $g(\Delta\pi) \geq 0$, $\forall \Delta\pi \geq 0$; To make the formula satisfy MDMA (iii), $g(\Delta\pi)$ should meet the demand that $g(\Delta\pi) \neq 0$, $\forall \Delta\pi \geq 0$. In summary, $g(\Delta\pi)$ should meet the demand that $g(\Delta\pi) > 0$, $\forall \Delta\pi \geq 0$.

The above conclusion is also applicable to formula (21).

Table 1: The values of d_{Pa} , d_S , d_{Hm} , d_J , d_{Hh} , d_{Hc} , d_{Eh} .

distance		d_{Pa}	d_S	d_{Hm}	d_J	d_{Hh}	d_{Hc}	d_{Eh}
Example 4.1	(A, B)	0.2000	0.1565	0.1500	0.1282	0.1156	0.1246	0.1348
	(B, C)	0.1500	0.1303	0.0500	0.0769	0.0694	0.0748	0.0981
	(A, C)	0.2000	0.2747	0.2000	0.2083	0.2000	0.2000	0.2000
Example 4.2	(D, E)	0.1000	0.7071	0.6667	0.5000	0.2500	0.4330	0.3536
	(E, F)	0.1000	0.7071	0.6667	0.5000	0.2500	0.4330	0.3536
	(D, F)	0.1000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Example 4.3	(G, H)	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
	(G, I)	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000
	(G, J)	0.2000	0.0360	0.0667	0.0500	0.0900	0.0995	0.0900
Example 4.4	(K, M)	0.3000	0.0863	0.1000	0.0750	0.1275	0.1482	0.1344
	(K, N)	0.2000	0.1619	0.1333	0.1500	0.1425	0.1498	0.1502

¹The numbers in blue color denote that the corresponding distance measure cannot make a clear distinction between some IFSs.

²The numbers in red color denote that the corresponding distance measure made an incorrect distinction between some IFSs.

4.3 Three improved distance measures for IFSs

Now, two improved distance measures for IFSs based on Hamming distance measure and one improved distance measure for IFSs based on Euclidean distance measure are given as follows:

Definition 4.5. Let $X = \{x_1, x_2, \dots, x_n\}$ be a finite fixed set, A and B be two IFSs in X , where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle \mid x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle \mid x \in X\}$. Their Hamming-hesitation distance is defined as:

$$d_{Hh}(A, B) = \frac{1}{2n} \sum_{i=1}^n \left[(|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)|) \times \left(1 - \frac{1}{2} |\pi_A(x_i) - \pi_B(x_i)| \right) \right]. \tag{22}$$

Definition 4.6. Let $X = \{x_1, x_2, \dots, x_n\}$ be a finite fixed set, A and B be two IFSs in X , where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle \mid x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle \mid x \in X\}$. Their Hamming-cosine distance is defined as:

$$d_{Hc}(A, B) = \frac{1}{2n} \sum_{i=1}^n \left[(|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)|) \times \cos \left(\frac{\pi}{6} |\pi_A(x_i) - \pi_B(x_i)| \right) \right]. \tag{23}$$

Definition 4.7. Let $X = \{x_1, x_2, \dots, x_n\}$ be a finite fixed set, A and B be two IFSs in X , where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle \mid x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle \mid x \in X\}$. Their Euclidean-hesitation distance is defined as:

$$d_{Eh}(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^n \left[\left((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 \right) \times \left(1 - \frac{1}{2} |\pi_A(x_i) - \pi_B(x_i)| \right)^2 \right]}. \tag{24}$$

The relevant proofs are left in the next section. To illustrate the effectiveness and advantage of the proposed distance measures, let us review the examples 4.1 to 4.4 and do some comparison between the proposed distance measures with partial existing distance measures that mentioned in formulas (8) to (11). The results are shown in Table 1.

From Table 1, we can find that the results from three proposed distance measures are not only in line with our analysis, but also consistent with the results which are from the distance measure d_S [25], H-max distance measure d_{Hm} [21] and the distance measure d_J [14]. In addition, the distance measure d_{Pa} proposed by Park et al. [23] cannot satisfy MDMA (iv) and gets some contrary results. In briefly, the proposed distance measures are as effective as d_S , d_{Hm} , and d_J while they are superior than d_{Pa} .

Example 4.8 is given to show the validity of three improved distance measures in pattern recognition.

Example 4.8. (Pattern recognition). Let $X = \{x_1, x_2\}$ be a fixed set. Suppose P , Q , and R are three IFSs in X , which represent three known patterns and can be expressed as follows:

$$P = \{\langle x_1, 0.33, 0.57 \rangle, \langle x_2, 0.12, 0.45 \rangle\},$$

$$Q = \{\langle x_1, 0.42, 0.44 \rangle, \langle x_2, 0.23, 0.52 \rangle\},$$

Table 2: The values of $d_H, d_E, d_{Pa}, d_S, d_{Hm}, d_J, d_{Hh}, d_{Hc}, d_{Eh}$.

distance	(P, U)	(Q, U)	(R, U)	ranking of similarity
d_H	0.1500	0.0400	0.1500	$Q \succ P = R$
d_E	0.1371	0.0361	0.1400	$Q \succ P \succ R$
d_{Pa}	0.1500	0.0400	0.1500	$Q \succ P = R$
d_S	0.0701	0.0333	0.0988	$Q \succ P \succ R$
d_{Hm}	0.0900	0.0300	0.1167	$Q \succ P \succ R$
d_J	0.0692	0.0303	0.1004	$Q \succ P \succ R$
d_{Hh}	0.0789	0.0297	0.0948	$Q \succ P \succ R$
d_{Hc}	0.0847	0.0300	0.0998	$Q \succ P \succ R$
d_{Eh}	0.0823	0.0328	0.1142	$Q \succ P \succ R$

Table 3: The values of $d_S, d_{Hm}, d_J, d_{Hh}, d_{Hc}, d_{Eh}$ in Example 4.9.

distance	(S, A_1)	(S, A_2)	(S, A_3)	ranking of the candidates
d_S	0.0347	0.0198	0.0117	$A_3 \succ A_2 \succ A_1$
d_{Hm}	0.0367	0.0500	0.0367	$A_1 = A_3 \succ A_2$
d_J	0.0351	0.0151	0.0201	$A_2 \succ A_3 \succ A_1$
d_{Hh}	0.0348	0.0430	0.0291	$A_3 \succ A_1 \succ A_2$
d_{Hc}	0.0350	0.0450	0.0300	$A_3 \succ A_1 \succ A_2$
d_{Eh}	0.0352	0.0453	0.0350	$A_3 \succ A_1 \succ A_2$

¹The numbers in blue color denote that the corresponding distance measure cannot make a clear distinction between the SIFs.

²The numbers in red color denote that the corresponding distance measure made an incorrect distinction between the SIFs.

$$R = \{ \langle x_1, 0.28, 0.67 \rangle, \langle x_2, 0.21, 0.65 \rangle \},$$

Assume that there is an unknown pattern which needs to be classified and can be expressed as $U = \{ \langle x_1, 0.37, 0.47 \rangle, \langle x_2, 0.22, 0.55 \rangle \}$, by using the distance measures, we can get the results shown in Table 2.

From Table 2, we can find that d_H and d_{Pa} have exactly the same results and cannot rank P and R by the similarity to U . Apart from them, all the other distance measures show that Q is the best classification for U , and the identical ranking results also imply that the proposed distance measures are accurate and effective enough in practical application.

In addition, it should be noticed that the superiorities of the proposed distance measures are lie not only in their simple form, but also in the consideration of the practical significance as to hesitancy degree for the distances of IFSs. Conversely, these are the disadvantages of d_S, d_{Hm} , and d_J .

In Section 4.4, a numerical application is provided to further illustrate the superiority of our proposed distance measures.

4.4 Numerical application

Example 4.9. Suppose there are two vacant positions in a company and some candidates are ready to compete for these positions. After the written exam and the first round of interview, there are still three candidates remain. Considering the characteristics of the position, the interviewer set a standard S for the position, where $S = \{ \langle x, 0.82, 0.09 \rangle \}$ is a SIFS and describes the personality that the position requires. The membership degree and the non-membership degree of S represent extroversion degree and introversion degree, respectively. And then, each candidate's personality was assessed according to their performance. To keep consistency with the standard, all the assessments are given in terms of SIFS. For the sake of convenience, the candidates are denoted as A_1, A_2, A_3 , and their assessments are shown as follows:

$$A_1 = \{ \langle x, 0.78, 0.12 \rangle \}, \quad A_2 = \{ \langle x, 0.76, 0.06 \rangle \} \quad \text{and} \quad A_3 = \{ \langle x, 0.77, 0.08 \rangle \}.$$

To choose two better candidates for the vacant positions, the interviewer adopts distance measure to reflect the differences between each assessment and the standard. It is rational that the smaller the distance is, the more suitable the candidate is. Considering that the distance measures d_H, d_E , and d_{Pa} have been proved to have some drawbacks, we just use the existing distance measures d_S, d_{Hm}, d_J and our proposed distance measures d_{Hh}, d_{Hc}, d_{Eh} to calculate the differences between each assessment and the standard, and the results are shown in Table 3.

Table 4: The comparisons between our proposed distance measures and some existing distance measures.

Distance measure	Authors	Example	Result
d_H	Szmidt and Kacprzyk[28] (2000)	Examples 4.3 (Page 6)	irrational
d_E	Szmidt and Kacprzyk[28] (2000)	Examples 4.2 (Page 5)	ineffective
d_{Pa}	Park et al. [23] (2009)	Examples 4.1 (Page 5)	ineffective
d_S	Shen et al. [25] (2018)	Example 4.9 (Page 9)	irrational
d_{Hm}	Ngan et al. [21] (2018)	Example 4.9 (Page 9)	ineffective
d_J	Jiang et al. [14] (2019)	Example 4.9 (Page 9)	irrational
Our proposed methods	—	The examples above	rational

¹The numbers in red color denote that the corresponding distance measure made an incorrect distinction between some IFSs.

From Table 3, we can find that:

- (i) The distance between S and A_1 is equal to that between S and A_3 if we use the distance d_{Hm} measure (Both of them are $\frac{11}{300} \approx 0.0367$). Thus, if there is just only vacant position, it would be difficult for the interviewer to decide who is more suitable for the job. However, if we adopt distance measures d_S , d_{Hh} , d_{Hc} , d_{Eh} , we can get the same result that A_3 is better than A_1 .
- (ii) In addition, we find that the results of d_S and d_J are different from the results of our proposed methods, especially in the ranking of A_1 and A_2 . In fact, in aspect of introversion degree, the difference between S and A_1 is equal to that between S and A_2 , but in aspect of extroversion degree, the difference between S and A_1 is less than that between S and A_2 . Therefore, it is rational to think the distance between S and A_1 is less than that between S and A_2 . That is, A_1 is better than A_2 , and the results of our proposed methods are more rational than that of d_S and d_J .
- (iii) By (i) and (ii), we can draw the conclusion that the appropriate ranking of candidates A_1, A_2, A_3 is $A_3 \succ A_1 \succ A_2$, which is in line with the results of our proposed methods.

In summary, A_1 and A_3 are more suitable for the vacant positions, which implies that our proposed distance measures are more accurate than the distance measures d_S , d_{Hm} , and d_J .

Table 4 shows the comparisons between our proposed distance measures and some existing distance measures. It is doubtless that our distance measures are superior to the existing methods.

5 The novel definitions of the characteristic function to distance measure

In this section, we will proposed a novel definition called the characteristic function of distance measure, which will be used to make some proofs.

5.1 The characteristic function of distance measure

In order to describe the character of the strict distance measure better, the following new definition is proposed:

Definition 5.1. Suppose d is a distance measure for IFSs and $d(A, U)$ represents the distance between A and U , where $A = \{\langle x, \mu_{Ai}, \nu_{Ai} \rangle | i = 1, 2, \dots, n\}$ is a specified IFS and $U = \{\langle x_i, \mu_i, \nu_i \rangle | i = 1, 2, \dots, n\}$ is an arbitrary IFS in the finite fixed set $X = \{x_1, x_2, \dots, x_n\}$. It is easy to prove that $d(A, U)$ is a function of both μ and ν , where $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ and $\nu = (\nu_1, \nu_2, \dots, \nu_n)$, respectively. Thus, we can denote $d(A, U)$ by a binary function $f_A(\mu, \nu)$. $f_A(\mu, \nu)$ is called the characteristic function of distance measure d .

The characteristic function of any strict distance measure should satisfy the following theorem:

Theorem 5.2. Suppose d is a strict distance measure of IFSs and $A = \{\langle x, \mu_{Ai}, \nu_{Ai} \rangle | i = 1, 2, \dots, n\}$ is a specified IFS in the finite fixed set $X = \{x_1, x_2, \dots, x_n\}$. If $f_A(\mu, \nu)$ is the characteristic function of d , where $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ and $\nu = (\nu_1, \nu_2, \dots, \nu_n)$, respectively. There will be

- (i) When $\mu \geq \mu_A$ and $\nu \leq \nu_A$, $f_A(\mu, \nu)$ is a strict increasing function of μ and a strict decreasing function of ν .

(ii) When $\mu \leq \mu_A$ and $\nu \geq \nu_A$, $f_A(\mu, \nu)$ is a strict decreasing function of μ and a strict increasing function of ν .

Where $\mu_A = (\mu_{A1}, \mu_{A2}, \dots, \mu_{An})$ and $\nu_A = (\nu_{A1}, \nu_{A2}, \dots, \nu_{An})$.

Proof. Under condition (i), suppose there are two arbitrary IFSs B and C in X , where $B = \{ \langle x, \mu_{Bi}, \nu_{Bi} \rangle \mid i = 1, 2, \dots, n \}$, $C = \{ \langle x, \mu_{Ci}, \nu_{Ci} \rangle \mid i = 1, 2, \dots, n \}$ and there are $\mu_B \geq \mu_A$, $\nu_B \leq \nu_A$, $\mu_C \geq \mu_A$, $\nu_C \leq \nu_A$, where $\mu_B = (\mu_{B1}, \mu_{B2}, \dots, \mu_{Bn})$, $\nu_B = (\nu_{B1}, \nu_{B2}, \dots, \nu_{Bn})$ and $\mu_C = (\mu_{C1}, \mu_{C2}, \dots, \mu_{Cn})$, $\nu_C = (\nu_{C1}, \nu_{C2}, \dots, \nu_{Cn})$.

Without loss of generality, we assume that $\mu_C \geq \mu_B$, $\nu_C = \nu_B$, then, we have $A \subseteq B \subseteq C$. According to DMA (iv), there should be $d(A, B) \leq d(A, C)$, i.e. $f_A(\mu_B, \nu_B) \leq f_A(\mu_C, \nu_C)$. Especially, if $\mu_C > \mu_B \geq \mu_A$, let $D = \{ \langle x, \mu_{Di}, \nu_{Di} \rangle \mid i = 1, 2, \dots, n \}$ and $\mu_D = (\mu_{D1}, \mu_{D2}, \dots, \mu_{Dn})$, $\nu_D = (\nu_{D1}, \nu_{D2}, \dots, \nu_{Dn})$, where $\mu_D = \frac{\mu_C + \mu_B}{2}$, $\nu_D = \nu_B$. Then, we have $A \subseteq B \subseteq D$ and $A \subset^C D \subset^C C$. According to MDMA (iv), there should be $d(A, B) \leq d(A, D) < d(A, C)$, i.e. $f_A(\mu_B, \nu_B) < f_A(\mu_C, \nu_C)$, so $f_A(\mu, \nu)$ is a strict increasing function of μ when $\mu \geq \mu_A$ and $\nu \leq \nu_A$.

If $\mu_C = \mu_B$, $\nu_C \leq \nu_B$, we can also prove that $f_A(\mu, \nu)$ is a strict decreasing function of ν under condition (i). In summary, $f_A(\mu, \nu)$ is a strict increasing function of μ and a strict decreasing function of ν when $\mu \geq \mu_A$ and $\nu \leq \nu_A$.

Condition (ii) can be proved in the same way. □

Conversely, if the characteristic function of distance measure d of IFSs can satisfy Theorem 5.2, it is easy to prove that d is a strict distance measure which can satisfy MDMA.

5.2 The proofs for the proposed distance measures

Now let us prove that the improved distance measures proposed in Section 4 are strict distance measures. It is apparent that they can satisfy MDMA (i), MDMA (ii) and MDMA (iii), thus we just need to prove that the distances can satisfy MDMA (iv). In other words, we will prove that their characteristic function can satisfy MDMA (iv).

Theorem 5.3. Both Hamming-hesitation distance measure and Hamming-cosine distance measure satisfy MDMA (iv).

Proof. In consideration of the limited space, we just prove the former while the latter can be proved in the same way.

Let $X = \{x\}$ be a fixed set, $U = \{ \langle x, \mu, \nu \rangle \}$ be an arbitrary SIFS and $A = \{ \langle x, \mu_A, \nu_A \rangle \}$ be a specified SIFS in X . We can get the characteristic function $f_A(\mu, \nu)$ of Hamming-hesitation distance measure, where

$$f_A(\mu, \nu) = \frac{1}{2} (|\mu - \mu_A| + |\nu - \nu_A|) \times \left(1 - \frac{1}{2} |(1 - \mu - \nu) - (1 - \mu_A - \nu_A)| \right).$$

Then, there are two conditions that should be treated separately:

- (i) $\mu + \nu \geq \mu_A + \nu_A$;
- (ii) $\mu + \nu < \mu_A + \nu_A$.

When $\mu \geq \mu_A$ and $\nu \leq \nu_A$, under the first condition, we have

$$f_A(\mu, \nu) = \frac{1}{2} (\mu - \mu_A - \nu + \nu_A) \times \left(1 - \frac{1}{2} (\mu + \nu - \mu_A - \nu_A) \right),$$

$$f'_\mu = \frac{\partial f_A(\mu, \nu)}{\partial \mu} = \frac{1}{2} \left[\left(1 - \frac{1}{2} (\mu + \nu - \mu_A - \nu_A) \right) + (\mu - \mu_A - \nu + \nu_A) \times \left(-\frac{1}{2} \right) \right] = \frac{1}{2} [(1 - \mu) + \mu_A],$$

$$f'_\nu = \frac{\partial f_A(\mu, \nu)}{\partial \nu} = \frac{1}{2} \left[(-1) \times \left(1 - \frac{1}{2} (\mu + \nu - \mu_A - \nu_A) \right) + (\mu - \mu_A - \nu + \nu_A) \times \left(-\frac{1}{2} \right) \right] = \frac{1}{2} [(\nu - 1) - \nu_A],$$

and $\mu + \nu \leq 1$, $\mu \geq 0$, $\nu \geq 0$, $\mu_A \geq 0$, $\nu_A \geq 0$, then we can get $f'_\mu \geq 0$, $f'_\nu \leq 0$.

Suppose there are four arbitrary real numbers μ_1, μ_2, μ_3 and ν' that satisfy $\mu_A \leq \mu_1 < \mu_2 \leq 1$, $\mu_3 = \frac{\mu_1 + \mu_2}{2}$ and $0 \leq \nu' \leq \nu_A$, we can easily get $f'_\mu(\mu_1) > 0$ and $\mu_A \leq \mu_1 < \mu_3 < \mu_2 \leq 1$, then we have $f_A(\mu_1, \nu') < f_A(\mu_3, \nu') \leq f_A(\mu_2, \nu')$, i.e. $f_A(\mu_1, \nu') < f_A(\mu_2, \nu')$. Thereby, $f_A(\mu, \nu)$ is a strict increasing function of μ when $\mu \geq \mu_A$ and $\nu \leq \nu_A$.

In the same manner, we can prove that $f_A(\mu, \nu)$ is a strict decreasing function of ν under condition (i), i.e. $f_A(\mu, \nu)$ is a strict increasing function of μ and a strict decreasing function of ν when $\mu \geq \mu_A$ and $\nu \leq \nu_A$.

When $\mu \geq \mu_A$ and $\nu \leq \nu_A$, under the second condition, we have

$$f(\mu, \nu) = \frac{1}{2} (\mu - \mu_A - \nu + \nu_A) \times \left(1 - \frac{1}{2} (\mu_A + \nu_A - \mu - \nu) \right),$$

$$f'_\mu = \frac{1}{2} \left[\left(1 - \frac{1}{2} (\mu_A + \nu_A - \mu - \nu) \right) + (\mu - \mu_A - \nu + \nu_A) \times \frac{1}{2} \right] = \frac{1}{2} [(1 - \mu_A) + \mu],$$

$f'_\nu = \frac{1}{2} \left[(-1) \times \left(1 - \frac{1}{2} (\mu_A + \nu_A - \mu - \nu) \right) + (\mu - \mu_A - \nu + \nu_A) \times \frac{1}{2} \right] = \frac{1}{2} [(\nu_A - 1) - \nu]$,
and $\mu + \nu \leq 1$, $\mu \geq 0$, $\nu \geq 0$, $\mu_A \geq 0$, $\nu_A \geq 0$, then we can also get $f'_\mu \geq 0$, $f'_\nu \leq 0$.

Like the course of the proof under condition (i), we can prove that $f_A(\mu, \nu)$ is a strict increasing function of μ and a strict decreasing function of ν when $\mu \geq \mu_A$ and $\nu \leq \nu_A$.

Moreover, it is not difficult to find that $f_A(\mu, \nu)$ is a continuous function of both μ and ν , i.e. when $\mu \geq \mu_A$ and $\nu \leq \nu_A$, $f_A(\mu, \nu)$ is an increasing function of μ and a decreasing function of ν .

When $\mu \leq \mu_A$ and $\nu \geq \nu_A$, by the similar approach, we can draw the conclusion that $f_A(\mu, \nu)$ is a strict decreasing function of μ and a strict increasing function of ν .

It is not difficult to extend the above conclusions to any X .

In summary, the characteristic function of Hamming-hesitation distance measure can satisfy Theorem 5.2, which indicates that Hamming-hesitation distance measure is a strict distance measure of IFSs. \square

Remark 5.4. We use the form as formula (20) to express two improved Hamming distance measures, then we have the following results: for Hamming-hesitation distance measure, $g_{Hh}(\Delta\pi) = 1 - \frac{1}{2}\Delta\pi$; for Hamming-cosine distance measure, $g_{Hc}(\Delta\pi) = \cos\left(\frac{\pi}{6} \times \Delta\pi\right)$. By making a contrast between them, we can find that $g_{Hh}(\Delta\pi)$ is a linear function of $\Delta\pi$, thus the influence of $\Delta\pi$ in the distance of two IFSs will keep a steady change if $\Delta\pi$ changes; However, $g_{Hc}(\Delta\pi)$ is a decreasing concave function of $\Delta\pi$ in its domain, which tells us that for two IFSs, the greater the difference of their hesitancy degree, the greater the change of influence to $\Delta\pi$ in their distance. On this basis, we can choose Hamming-hesitation distance measure or Hamming-cosine distance measure in accordance with the influence of $\Delta\pi$ in the distance that we need.

Theorem 5.5. Euclidean-hesitation distance measure can satisfy MDMA (iv).

Proof. Let $X = \{x\}$ be a fixed set, $U = \{x, \mu, \nu\}$ be an arbitrary SIFS and $A = \{x, \mu_A, \nu_A\}$ be a specified SIFS in X . We can get the characteristic function $f_A(\mu, \nu)$ of Euclidean-hesitation distance measure, where

$$f_A(\mu, \nu) = \sqrt{\frac{1}{2} \left((\mu - \mu_A)^2 + (\nu - \nu_A)^2 \right) \times \left(1 - \frac{1}{2} |(1 - \mu - \nu) - (1 - \mu_A - \nu_A)| \right)^2}.$$

To be convenience, let $g(\mu, \nu) = \left((\mu - \mu_A)^2 + (\nu - \nu_A)^2 \right) \times \left(1 - \frac{1}{2} |\mu + \nu - \mu_A - \nu_A| \right)^2$, it is not difficult to find that $g(\mu, \nu) \geq 0$ and $f_A(\mu, \nu)$ is a strict increasing function of $g(\mu, \nu)$. Then, there are two conditions that should be treated separately:

- (i) $\mu \geq \mu_A$ and $\nu \leq \nu_A$;
- (ii) $\mu \leq \mu_A$ and $\nu \geq \nu_A$.

Since $g(\mu, \nu) = \left((\mu - \mu_A)^2 + (\nu - \nu_A)^2 \right) \times \left(1 - \frac{\alpha}{2} (\mu + \nu - \mu_A - \nu_A) \right)^2$, where $\alpha = \begin{cases} 1, & \text{if } \mu + \nu \geq \mu_A + \nu_A \\ -1, & \text{if } \mu + \nu < \mu_A + \nu_A \end{cases}$, there is

$$\begin{aligned} g'_\mu &= \frac{\partial g(\mu, \nu)}{\partial \mu} \\ &= 2(\mu - \mu_A) \times \left(1 - \frac{\alpha}{2} (\mu - \mu_A + \nu - \nu_A) \right)^2 - \alpha \left((\mu - \mu_A)^2 + (\nu - \nu_A)^2 \right) \times \left(1 - \frac{\alpha}{2} (\mu - \mu_A + \nu - \nu_A) \right) \\ &= \left[(\mu - \mu_A) \times (2 - \alpha(\mu - \mu_A + \nu - \nu_A)) - \alpha \left((\mu - \mu_A)^2 + (\nu - \nu_A)^2 \right) \right] \times \left(1 - \frac{1}{2} |\mu + \nu - \mu_A - \nu_A| \right) \\ &= [2(\mu - \mu_A)(1 - \alpha(\mu - \mu_A)) - \alpha(\nu - \nu_A)((\mu + \nu) - (\mu_A + \nu_A))] \times \left(1 - \frac{1}{2} |(\mu + \nu) - (\mu_A + \nu_A)| \right) \\ &= [2(\mu - \mu_A)(1 - \alpha(\mu - \mu_A)) + (\nu_A - \nu)|\mu + \nu - \mu_A - \nu_A|] \times \left(1 - \frac{1}{2} |(\mu + \nu) - (\mu_A + \nu_A)| \right). \end{aligned}$$

By $\mu + \nu \leq 1$, $\mu_A + \nu_A \leq 1$, $\mu \geq 0$, $\nu \geq 0$, $\mu_A \geq 0$, $\nu_A \geq 0$, we have $|(\mu + \nu) - (\mu_A + \nu_A)| \leq 1$, $1 - \frac{1}{2} |(\mu + \nu) - (\mu_A + \nu_A)| \geq \frac{1}{2}$, $1 - \alpha(\mu - \mu_A) \geq 1 - |\alpha| \geq 0$.

(i) When $\mu \geq \mu_A$ and $\nu \leq \nu_A$.

Since $2(\mu - \mu_A)(1 - a(\mu - \mu_A)) + (\nu_A - \nu)|\mu + \nu - \mu_A - \nu_A| \geq 0$, then we have $g'_\mu \geq 0$.

Suppose there are four arbitrary real numbers μ_1, μ_2, μ_3 and ν' that satisfy $\mu_A \leq \mu_1 < \mu_2 \leq 1, \mu_3 = \frac{\mu_1 + \mu_2}{2}$ and $0 \leq \nu' \leq \nu_A$. Then, we can easily get that $\mu_A \leq \mu_1 < \mu_3 < \mu_2 \leq 1, 0 < \mu_3 - \mu_A < 1$ and $g'_\mu(\mu_3, \nu') > 0$. On this basis, we have $g(\mu_1, \nu') < g(\mu_3, \nu') < g(\mu_2, \nu')$, i.e. $g(\mu_1, \nu') < g(\mu_2, \nu')$. Hence, $g(\mu, \nu)$ is a strict increasing function of μ when $\mu \geq \mu_A$ and $\nu \leq \nu_A$. Since $f_A(\mu, \nu)$ is a strict increasing function of $g(\mu, \nu)$, then, when $\mu \geq \mu_A$ and $\nu \leq \nu_A$, $f_A(\mu, \nu)$ is a strict increasing function of μ .

(ii) When $\mu \leq \mu_A$ and $\nu \geq \nu_A$.

Since $2(\mu - \mu_A)(1 - a(\mu - \mu_A)) + (\nu_A - \nu)|\mu + \nu - \mu_A - \nu_A| \leq 0$, then we have $g'_\mu \leq 0$. In the same way as the proof under condition (i), it can be proved that $f_A(\mu, \nu)$ is a strict decreasing function of μ , when $\mu \leq \mu_A$ and $\nu \geq \nu_A$.

Moreover, we can also prove the corresponding property of $f_A(\mu, \nu)$ as to ν and have the following conclusion:

When $\mu \geq \mu_A$ and $\nu \leq \nu_A$, $f_A(\mu, \nu)$ is a strict increasing function of μ and a strict decreasing function of ν ; When $\mu \leq \mu_A$ and $\nu \geq \nu_A$, $f_A(\mu, \nu)$ is a strict decreasing function of μ and a strict increasing function of ν .

This conclusion can be easily extended to any X .

In summary, the characteristic function of Euclidean-hesitation distance measure can satisfy Theorem 5.2, so it is a strict distance measure of IFSs. □

Theorem 5.6. Suppose $X = \{x_1, x_2, \dots, x_n\}$ is a finite fixed set. $\forall A, B \in IFS(X)$, where $A = \{\langle x, \mu_A(x), \nu_A(x), \pi_A(x) \rangle | x \in X\}$, $B = \{\langle x, \mu_B(x), \nu_B(x), \pi_B(x) \rangle | x \in X\}$, there is $d_{Hh}(A, B) \leq 1$.

Proof. $\forall i \in [1, n]$, we have $0 \leq |\mu_A(x_i) - \mu_B(x_i)| \leq 1, 0 \leq |\nu_A(x_i) - \nu_B(x_i)| \leq 1, \frac{1}{2} \leq 1 - \frac{1}{2} |\pi_A(x_i) - \pi_B(x_i)| \leq 1$. Then, we obtain

$$\frac{1}{2} \left[(|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)|) \times \left(1 - \frac{1}{2} |\pi_A(x_i) - \pi_B(x_i)| \right) \right] \leq \frac{1}{2} [(1 + 1) \times 1] = 1,$$

i.e. $d_{Hh}(A, B) \leq 1$. □

Remark 5.7. It can also prove that $d_{Hc}(A, B) \leq 1$ and $d_{Eh}(A, B) \leq 1$ in the same way, $\forall A, B \in IFS(X)$.

5.3 An analysis of the special case for distance measures and the solution

To illustrate some special cases for distance measures better, the following example, which is similar to the example given by Shen et al. [25], is presented and should be noticed.

Example 5.8. (Pattern recognition). Let $X = \{x\}$ be a fixed set. Suppose A and B are two IFSs in X , which represent known patterns and can be expressed as follows:

$$A = \{\langle x_1, 0.3, 0.5, 0.2 \rangle\}, \quad B = \{\langle x_1, 0.4, 0.4, 0.2 \rangle\}.$$

If there is an unknown pattern which needs to be classified and can be expressed as $U = \{\langle x_1, 0.1, 0.1, 0.8 \rangle\}$, by using the proposed Hamming distance measures, we have

$$d_{Hh}(U, A) = d_{Hh}(U, B) = 0.21 \quad \text{and} \quad d_{Hc}(U, A) = d_{Hc}(U, B) = 0.2853.$$

Thus, we cannot identify which of A and B is the better classification for U .

It seems that the proposed improved Hamming distance measures lose their validity in this instance. But, in reality, there is always something cannot get a determined classification. Let us analyze the example just like the analysis in example 4.4. First of all, notice that

$$\pi_U(x) - \pi_A(x) = \pi_U(x) - \pi_B(x),$$

$$|\mu_U(x) - \mu_A(x)| + |\nu_U(x) - \nu_A(x)| = |\mu_U(x) - \mu_B(x)| + |\nu_U(x) - \nu_B(x)|$$

are the direct reasons for the terrible results. The special case means that under the model of ten-people-voting, whether we want to transform the voting result of U to A or B , we should persuade six persons who keep neutral to vote for agreement or opposition. And then, under the condition that lacks of extra information, it is rational to think the difficulty of persuading a neutral one to vote for agreement is equivalent to the difficulty that persuading a neutral one

to vote against, i.e. the Hamming distance between U and A should be identical with the Hamming distance between U and B , which is accordance to the results.

Now, we will step in the further analysis. As a matter of fact, in order to obtain some conveniences which came from the relevant properties of continuity, the characteristic functions of most of existing distance measures are continuous. For instance, the characteristic functions of formulas (6) to (11) and the proposed distance measures are all continuous. For this kind of characteristic function, we have the following theorem:

Theorem 5.9. *Let $X = \{x\}$ be a fixed set. Suppose d is a distance measure over X and $f(\mu, \nu)$ is the characteristic function of it. If $f(\mu, \nu)$ is continuous, we can get the conclusion as follows:*

$$\forall A \in IFS(X), \exists B, C \in IFS(X), B = \{\langle x, \mu_B, \nu_B \rangle\}, C = \{\langle x, \mu_C, \nu_C \rangle\}, B \neq C, d(B, A) = d(C, A) > 0.$$

It is not difficult to extend the above conclusion to any X .

Proof. There are three conditions that should be treated separately:

- (i) $0 \leq \mu_A + \nu_A < 1$.
- (ii) $\mu_A = 1, \nu_A = 0$ or $\mu_A = 0, \nu_A = 1$.
- (iii) $\mu_A, \nu_A \in (0, 1)$ while $\mu_A = 1 - \nu_A$.

The proofs as to these three conditions are given as follows:

(i) When $0 \leq \mu_A + \nu_A < 1$, let $B' = \{\langle x, 1 - \nu_A, \nu_A \rangle\}$ and $C' = \{\langle x, \mu_A, 1 - \mu_A \rangle\}$ be two SIFSs in X and note $d(B', A) = p$ and $d(C', A) = q$. If $B = C$, we can get $1 - \nu_A = \mu_A$ and $\nu_A = 1 - \mu_A$, i.e. $\mu_A + \nu_A = 1$, which is contrary to the condition. Therefore, It is evident that $1 - \nu_A > \mu_A$ and $1 - \mu_A > \nu_A$, i.e. $A \neq B \neq C$ and $p, q > 0$.

If $p = q$, the theorem is proved under this condition. Otherwise, without loss of generality, considering $p > q$, i.e. $f_A(1 - \nu_A, \nu_A) > f_A(\mu_A, 1 - \mu_A) > f_A(\mu_A, \nu_A) = 0$. Since $f(\mu, \nu)$ is a continuous function of μ , there should be $\mu_B \in (\mu_A, 1 - \nu_A)$ that can satisfy $f_A(\mu_B, \nu_A) = f_A(\mu_A, 1 - \mu_A)$. Let $B = \{\langle x, \mu_B, \nu_A \rangle\}$ and $C = C'$, we have:

$$B \neq C, d(B, A) = d(C, A) > 0.$$

(ii) When $\mu_A = 1, \nu_A = 0$, let $B' = \{\langle x, 0.5, 0.5 \rangle\}$ and $C' = \{\langle x, 0, 0 \rangle\}$ be two SIFSs in X . If $f_A(0.5, 0.5) = f_A(0, 0)$, i.e. $B' \neq C', d(B', A) = d(C', A) > 0$, the theorem is proved under this condition.

Otherwise, according to part of Theorem 5.2, we know that $f_A(\mu, \nu)$ is a decreasing function of μ under this condition. Then, if $f_A(0.5, 0.5) > f_A(0, 0)$, we have $f_A(0, 0) < f_A(0.5, 0.5) \leq f_A(0, 0.5)$. Since $f_A(0, \nu)$ is a continuous strict increasing function of ν , there should be $\nu_C \in (0, 0.5]$ that can satisfy $f_A(0, \nu_C) = f_A(0.5, 0.5)$. Let $B = \{\langle x, 0.5, 0.5 \rangle\}$ and $C = \{\langle x, 0, \nu_C \rangle\}$, we have:

$$B \neq C, d(B, A) = d(C, A) > 0.$$

If $f_A(0.5, 0.5) < f_A(0, 0)$, we also have $0 = f_A(1, 0) < f_A(0.5, 0.5) < f_A(0, 0)$. Since $f_A(\mu, 0)$ is a continuous function of μ , there should be $\mu_C \in (0, 1)$ that can satisfy $f_A(\mu_C, 0) = f_A(0.5, 0.5)$. Let $B = \{\langle x, 0.5, 0.5 \rangle\}$ and $C = \{\langle x, \mu_C, 0 \rangle\}$, we have:

$$B \neq C, d(B, A) = d(C, A) > 0.$$

When $\mu_A = 0, \nu_A = 1$, the theorem can be proved in the same way.

(iii) When $\mu_A, \nu_A \in (0, 1)$ and $\mu_A = 1 - \nu_A$, let $B' = \{\langle x, 1, 0 \rangle\}$ and $C' = \{\langle x, 0, 1 \rangle\}$ be two SIFSs in X . If $f_A(1, 0) = f_A(0, 1)$, i.e. $d(B', A) = d(C', A) > 0$, the theorem is proved under this condition.

Otherwise, without loss of generality, considering $f_A(1, 0) > f_A(0, 1)$. According to part of Theorem 5.2, we know that Since $f_A(\mu, \nu)$ is a decreasing function of μ under this condition, we have:

$$0 = f_A(\mu_A, \nu_A) < f_A(\mu_A, 0) \leq f_A(1, 0),$$

$$0 = f_A(\mu_A, \nu_A) < f_A(0, \nu_A) \leq f_A(0, 1).$$

Then:

If $f_A(\mu_A, 0) = f_A(0, 1)$, Let $B = \{\langle x, \mu_A, 0 \rangle\}$ and $C = C'$, the theorem is proved.

If $f_A(\mu_A, 0) > f_A(0, 1)$, i.e. $0 = f_A(\mu_A, \nu_A) < f_A(0, 1) < f_A(\mu_A, 0)$. Since $f_A(\mu_A, \nu)$ is a continuous function of ν , there should be $\nu_B \in (0, \nu_A)$ that can satisfy $f_A(\mu_A, \nu_B) = f_A(0, 1)$. Let $B = \{\langle x, \mu_A, \nu_B \rangle\}$ and $C = C'$, we have:

$$B \neq C, d(B, A) = d(C, A) > 0.$$

If $f_A(\mu_A, 0) < f_A(0, 1)$, i.e. $f_A(\mu_A, 0) < f_A(0, 1) < f_A(1, 0)$, since $f_A(\mu, 0)$ is a continuous function of μ , there should be $\mu_B \in (\mu_A, 1)$ that can satisfy $f_A(\mu_B, 0) = f_A(0, 1)$. Let $B = \{x, \mu_B, 0\}$ and $C = C'$, we have:

$$B \neq C, \quad d(B, A) = d(C, A) > 0.$$

The theorem tells us that a distance measure cannot apply to all the pattern recognition cases if its characteristic function is continuous. □

Remark 5.10. *Still considering the voting model, if the vote can bring more benefits than risk, it is regarded positive. It is sure that in a positive environment, persuading a neutral one to vote for agreement is easier than persuading him/her to vote against. On this basis, we can decrease the weight for the difference of membership degree while increase the weight for the difference of non-membership degree to keep the result in line with the circumstance. Then, we have the following formulas:*

Weighted Hamming-hesitation distance measure:

$$d_{Hh}(A, B) = \frac{1}{2n} \sum_{i=1}^n \left[(\alpha |\mu_A(x_i) - \mu_B(x_i)| + \beta |\nu_A(x_i) - \nu_B(x_i)|) \times \left(1 - \frac{1}{2} |\pi_A(x_i) - \pi_B(x_i)| \right) \right] \quad (25)$$

Weighted Hamming-cosine distance measure:

$$d_{Hc}(A, B) = \frac{1}{2n} \sum_{i=1}^n \left[(\alpha |\mu_A(x_i) - \mu_B(x_i)| + \beta |\nu_A(x_i) - \nu_B(x_i)|) \times \cos \left(\frac{\pi}{6} |\pi_A(x_i) - \pi_B(x_i)| \right) \right] \quad (26)$$

Weighted Euclidean-hesitation distance measure:

$$d_{Eh}(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^n \left[\left(\alpha (\mu_A(x_i) - \mu_B(x_i))^2 + \beta (\nu_A(x_i) - \nu_B(x_i))^2 \right) \times \left(1 - \frac{1}{2} |\pi_A(x_i) - \pi_B(x_i)| \right)^2 \right]} \quad (27)$$

Where there are $\alpha < \beta$ in positive circumstance or $\alpha > \beta$ in pessimistic circumstance, $\alpha + \beta = 2$, $\alpha, \beta \geq 0.5$. Obviously, if $\alpha = \beta = 1$, we can get the general Hamming-hesitation distance measure, the general Hamming-cosine distance measure and the general Euclidean-hesitation distance measure in neutral circumstance.

For instance, in example 5.8, if the circumstance is highly positive and $\alpha = 0.5$, $\beta = 1.5$, we have

$$d_{Hh}(U, A) = 0.245 > 0.21 = d_{Hh}(U, B) \quad \text{and} \quad d_{Hc}(U, A) = 0.3329 > 0.2853 = d_{Hc}(U, B);$$

If the circumstance is highly pessimistic and $\alpha = 1.5$, $\beta = 0.5$, we have

$$d_{Hh}(U, A) = 0.175 < 0.21 = d_{Hh}(U, B) \quad \text{and} \quad d_{Hc}(U, A) = 0.2378 < 0.2853 = d_{Hc}(U, B).$$

These results are consistent with both our intuition and the cognition of the voting model.

6 Conclusions

Distance measure is an important tool to measure the deviation of IFSs so that it plays a significant role in many fields, such as pattern recognition or the ranking of alternatives. Nevertheless, there are some drawbacks with partial existing distance measures of IFSs, which will cause a terrible damage if we apply them to practice. To avoid the problem, a new inclusion relationship called complete inclusion is proposed in this paper and the distance measure axioms for IFSs are refined. Then, two improved distance measures for IFSs based on Hamming distance measure and one improved distance measures for IFSs based on Euclidean distance measure are presented according to a rational analysis in reference to the practical significance of hesitancy degree. These improved distance measures are tested in some instances and to demonstrate their effectiveness and superiority.

To make a further study, the definition called the characteristic function of distance measure is put forward. Then, an instance is given to illustrate the invalid condition for the improved Hamming distance measures and a theorem is presented to point out the inevitability of the condition that the occurrence of unrecognized result in pattern recognition problems in some cases. In the future work, the distance measure for IFSs will be made some further research from the basic operation of IFSs, which can make it more practical and meaningful.

Acknowledgement

This research was supported by “the Humanities and social sciences research and planning fund of the Ministry of Education of China, No. x2lxY9180090”, “the Natural Science Foundation of Guangdong Province, No. 2016A030313545”, “the Fundamental Research Funds for the Central Universities of China, No. x2lxC2180170” and “the Graduate Education Innovation Projects of Guangdong Province, No. 2016QTLXXM-19”. The authors are highly grateful to the referees and editor in-chief for their very helpful comments and suggestions.

References

- [1] K. T. Atanassov, *Intuitionistic fuzzy sets*, Fuzzy Sets and Systems, **20**(1) (1986), 87-96.
- [2] K. T. Atanassov, *Intuitionistic fuzzy sets: Theory and applications*, Physica-Verlag HD, **35**(1) (1999), 1-137.
- [3] K. T. Atanassov, G. Gargov, *Interval-valued intuitionistic fuzzy sets*, Fuzzy Sets and Systems, **31**(3) (1989), 343-349.
- [4] J. F. Botía, A. M. Cárdenas, C. M. Sierra, *Fuzzy cellular automata and intuitionistic fuzzy sets applied to an optical frequency comb spectral shape*, Engineering Applications of Artificial Intelligence, **62** (2017), 181-194.
- [5] T. Y. Chen, *A note on distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric*, Fuzzy Sets and Systems, **158**(22) (2007), 2523-2525.
- [6] S. M. Chen, C. H. Chang, *A novel similarity measure between atanassovs intuitionistic fuzzy sets based on transformation techniques with applications to pattern recognition*, Information Sciences, **291** (2015), 96-114.
- [7] S. M. Chen, C. H. Chang, *Fuzzy multiattribute decision making based on transformation techniques of intuitionistic fuzzy values and intuitionistic fuzzy geometric averaging operators*, Information Sciences, **352-353** (2016), 133-149.
- [8] S. M. Chen, S. H. Cheng, T. C. Lan, *A novel similarity measure between intuitionistic fuzzy sets based on the centroid points of transformed fuzzy numbers with applications to pattern recognition*, Information Sciences, **343-344** (2016), 115-40.
- [9] M. Citil, F. Turul, *Some intuitionistic fuzzy operators*, Ifscom Mersin, Turkey, **22**(4) (2016), 1-5.
- [10] G. Dwivedi, R. K. Srivastava, S. K. Srivastava, *A generalised fuzzy TOPSIS with improved closeness coefficient*, Expert Systems with Applications, **96** (2017), 185-195.
- [11] J. A. Goguen, *L-fuzzy sets*, Journal of Mathematical Analysis and Applications, **18**(1) (1967), 145-174.
- [12] P. Grzegorzewski, *Distances between intuitionistic fuzzy sets and interval-valued fuzzy sets based on the hausdorff metric*, Fuzzy Sets and Systems, **148**(2) (2004), 319-328.
- [13] E. D. Han, P. Guo, J. Zhao, *Multi-attribute group decision making method based on similarity and intuitionistic fuzzy entropy*, Modernization of Management, **36**(1) (2016), 105-107.
- [14] Q. Jiang, X. Jin, S. J. Lee, S. W. Yao, *A new similarity/distance measure between intuitionistic fuzzy sets based on the transformed isosceles triangles and its applications to pattern recognition*, Expert Systems with Applications, **116** (2019), 439-453.
- [15] H. P. Jin, K. M. Lim, Y. C. Kwun, *Distance measure between intuitionistic fuzzy sets and its application to pattern recognition*, South Korean Intelligence System Society, **19**(4) (2009), 556-561.
- [16] H. Lee-Kwang, Y. S. Song, K. M. Lee, *Similarity measure between fuzzy sets and between elements*, Fuzzy Sets and Systems, **62**(3) (1994), 291-293.
- [17] X. Liu, *Entropy, distance measure and similarity measure of fuzzy sets and their relations*, Fuzzy Sets and Systems, **52**(3) (1992), 305-318.
- [18] M. Luo, R. Zhao, *A distance measure between intuitionistic fuzzy sets and its application in medical diagnosis*, Artificial Intelligence in Medicine, **89** (2018), 34-39.

- [19] Y. C. Lv, S. Z. Guo, *Entropy of intuitionistic fuzzy sets and its general forms*, Computer Engineering and Applications, **47**(28) (2011), 52-55.
- [20] J. Mao, D. Yao, C. Wang, *A novel cross-entropy and entropy measures of IFSs and their applications*, Knowledge-Based Systems, **48**(2) (2013), 37-45.
- [21] R. T. Ngan, H. S. Le, B. C. Cuong, M. Ali, *H-max distance measure of intuitionistic fuzzy sets in decision making*, Applied Soft Computing, **69** (2018), 393-425.
- [22] H. Nguyen, *A new knowledge-based measure for intuitionistic fuzzy sets and its application in multiple attribute group decision making*, Expert Systems with Applications, **42**(22) (2015), 8766-8774.
- [23] J. H. Park, K. M. Lim, Y. C. Kwun, *Distance measure between intuitionistic fuzzy sets and its application to pattern recognition*, Journal of The Korean Institute of Intelligent Systems, **19**(4) (2009), 556-561.
- [24] D. Schitea, M. Deveci, M. Iordache, K. Bilgili, I. ZekiAkyurt, L. Iordache, *Hydrogen mobility roll-up site selection using intuitionistic fuzzy sets based WASPAS, COPRAS and EDAS*, International Journal of Hydrogen Energy, **44**(16) (2019), 8585-8600.
- [25] F. Shen, X. Ma, Z. Li, Z. Xu, D. Cai, *An extended intuitionistic fuzzy TOPSIS method based on a new distance measure with an application to credit risk evaluation*, Information Sciences, **428** (2017), 105-119.
- [26] F. Shen, J. P. Xu, Z. S. Xu, *An outranking sorting method for multi-criteria group decision making using intuitionistic fuzzy sets*, Information Sciences, **334** (2016), 338-353.
- [27] G. Sirbiladze, A. Sikharulidze, *Extensions of probability intuitionistic fuzzy aggregation operators in fuzzy MCDM/MADM*, International Journal of Information Technology and Decision Making, **17**(2) (2018), 621-655.
- [28] E. Szmidt, J. Kacprzyk, *Distances between intuitionistic fuzzy sets*, Fuzzy Sets and Systems, **114**(3) (2000), 505-518.
- [29] W. Wang, X. Xin, *Distance measure between intuitionistic fuzzy sets*, Pattern Recognition Letters, **26**(13) (2005), 2063-2069.
- [30] C. P. Wei, Z. H. Gao, T. T. Guo, *An intuitionistic fuzzy entropy measure based on trigonometric function*, Control and Decision, **27**(4) (2012), 571-574.
- [31] Y. Yang, F. Chiclana, *Consistency of 2D and 3D distances of intuitionistic fuzzy sets*, Expert Systems with Applications, **39**(10) (2012), 8665-8670.
- [32] S. Yin, B. Z. Li, *Matching management of supply and demand of green building technologies based on a novel matching method with intuitionistic fuzzy sets*, Journal of Cleaner Production, **201** (2018), 748-763.
- [33] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8** (1965), 338-353.
- [34] H. Zhang, L. Yu, *New distance measures between intuitionistic fuzzy sets and interval-valued fuzzy sets*, Information Sciences, **245** (2013), 181-196.

Several new results based on the study of distance measures of intuitionistic fuzzy sets

C. Chen and X. Deng

نتایج جدید متعدد بر اساس مطالعه‌ی اندازه‌های فاصله مجموعه‌های فازی شهودی

چکیده. بدون شک نظریه مجموعه‌ی فازی شهودی (IFS) نقش مهم فزاینده‌ای در حل مسایل در شرایط نامعلوم ایفا می‌کند. به عنوان یکی از حیاتی‌ترین عناصر در نظریه، اندازه فاصله در بسیاری جهات بطور گسترده‌ای به کار برده می‌شود. با این وجود، متأسفانه بخشی از اندازه‌های فاصله دارای نقاط ضعفی در مفهوم و دقت می‌باشد. برای جبران نقاط ضعف و دقت بیشتر و کارایی، در این مقاله، یک رابطه شمول جدید از IFS ها و یک تابع جدید که اندازه فاصله اکید نامیده می‌شود، پیشنهاد می‌کنیم. بر اساس این رابطه جدید، تحلیلی ارائه گردیده که خاطر نشان می‌کند، نقطه ضعف مشترک اندازه فاصله Hamming و اندازه فاصله اقلیدسی بد بکار بردن درجه تردید است. بنابراین، نقش درجه تردید در اندازه فاصله بطور عمیق بررسی شده و سپس سه اندازه فاصله اکید جهت غلبه بر نقطه ضعف بالا ارائه گردیده‌است. بعلاوه، یک تعریف جدید، که تابع مشخصه اندازه فاصله نامیده می‌شود جهت توصیف خاصیت اندازه فاصله اکید تعریف شده‌است. بر این اساس، قضیه‌ای ارائه گردیده، تا رخداد غیرقابل تشخیص در مسایل تشخیص الگو، در برخی از حالات خاص را نشان دهد. این مسئله همچنین نشان می‌دهد که مسئله نمی‌تواند بطور کامل به اندازه‌های فاصله نسبت داده شود. در نتیجه، یک راه حل مناسب ارائه می‌کنیم. در مقایسه با اندازه‌های فاصله موجود دیگر، در برخی مثال‌ها، اولویت‌های اندازه‌های فاصله بهبود یافته جهت نشان دادن تأثیرپذیری بیشتر و اهمیت عملی بودن بیشتر، نشان داده شده‌است.