

A new similarity measure between intuitionistic fuzzy sets and the positive definiteness of the similarity matrix

Yafei Song¹ · Xiaodan Wang¹

Received: 9 January 2015 / Accepted: 7 June 2015 / Published online: 21 June 2015
© Springer-Verlag London 2015

Abstract As a generation of fuzzy set theory, intuitionistic fuzzy (IF) set theory has received considerable attention for its capability on dealing with uncertainty. Similarity measures of IF sets are used to indicate the degree of commonality between IF sets. Although several similarity measures for IF sets have been proposed in previous studies, some of those cannot satisfy the axiomatic definitions of similarity, or provide counter-intuitive cases. In this paper, a new similarity measure between IF sets is proposed. The definition of similarity matrix is also presented to depict the relations among more than two IF sets. It is proved that the proposed similarity measures satisfy the properties of the axiomatic definition for similarity measures. Comparison between the previous similarity measures and the proposed similarity measure indicates that the proposed similarity measure does not provide any counter-intuitive cases. Moreover, it is demonstrated that the proposed similarity measure can be applied to define a positive definite similarity matrix.

Keywords Intuitionistic fuzzy set · Distance measure · Similarity measure · Similarity matrix

1 Introduction

Since it was proposed by Zadeh [1], the theory of fuzzy set (FS) has achieved a great success due to its capability of handling uncertainty [2, 3]. Therefore, over the last decades, several higher order fuzzy sets have been introduced in the literature. Intuitionistic fuzzy (IF) set, as one of the higher order FSs, was proposed by Atanassov [4] to deal with vagueness. The main advantage of IF set is its capability on coping with the uncertainty that may exist due to information impression. Because it assigns to each element a membership degree, a non-membership degree and a hesitation degree, and thus, IF set constitutes an extension of Zadeh's fuzzy set which only assigns to each element a membership degree, taking 1 minus it as the degree of non-membership [5]. So IF set is regarded as a more effective way to deal with vagueness than FS [6].

Similarity measure is of great significance in almost every scientific field. Similarity measures between two IF sets are related to their commonality on the information conveyed by them. Since it is difficult to measure the amount of information hidden in IF sets, we cannot compare two IF sets directly. Therefore, similarity measure and its counterpart, distance measure, play an important role in discriminating IF sets. With the development of IF set theory, the definition of similarity measure for IF sets has also received considerable attention in recent years [7, 8]. It has developed and will continue to develop into an important tool for decision making, fault detection, pattern recognition, machine learning, image processing, etc.

For its fundamental importance in application, many similarity measures have been proposed. The first study was carried out by Szmidt and Kacprzyk [9]. They applied Hamming distance and Euclidian distance to IF environment and comparing them with the approaches used for

✉ Xiaodan Wang
afeu_wang@163.com

Yafei Song
yafei_song@163.com

¹ Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, People's Republic of China

ordinary FSs. Following this work, many researchers presented similarity measures for IF sets by extending the well-known distance measures, such as Hamming distance, Euclidian distance, and Hausdorff distance [10–16]. Meanwhile, some studies defined new similarity measures for IF sets by defining some intermediate variables based on membership and non-membership degrees [7, 17–21]. For example, Li and Cheng [17] suggested a new similarity measure for IF sets based on the definition of φ_A . Recently, many novel similarity measures are emerging in an endless stream. This trend can be illustrated by similarity measures defined based on cosine similarity [8], Sugeno integral [22], interval comparison [23], intuitionistic entropy [24], and so on. In addition, Boran and Akay [25] proposed a new general type of similarity measure for IF sets with two parameters, expressing norm and the level of uncertainty, respectively. This similarity measure can also make sense in terms of counter-intuitive cases. As a comprehensive study on similarity measures of IF sets, Papakostas et al. [26] investigated the main theoretical and computational properties of the measures, as well as the relationships between them. Moreover, a comparison of the distance and similarity measures was carried out from a pattern recognition point of view.

Among the proposed similarity measures between IF sets, some of those, however, cannot satisfy the axioms of similarity, or provide counter-intuitive cases, or are produced by complex formats. Therefore, we propose a new similarity measures between IF sets in this paper, based on the cosine similarity and Euclidean distance between IF sets. Axiomatic definitions of similarity and the distance measures are first presented. Then the relation between similarity and distance measure follows. Similarity matrix is also defined to describe relationships among more than two IF sets. The properties of similarity matrix are examined to explore the performance of our new similarity measure, which is defined after critically analyzing the cosine similarity and similarity generated by Euclidean distance. Properties and performances of the proposed similarity measure are indicated by both mathematical proofs and illustrative examples.

The remainder of this paper is organized as follows. Section 2 recalls the definitions related to the IF sets. In Sect. 3, distance measure, similarity measure, similarity matrix, and their properties with proofs are proposed. The new similarity measure is defined in Sect. 4. Its properties are also proved in this section. Comparison between similarity measures and illustration of the positive definiteness of similarity matrix is carried out in Sect. 5. We come to the conclusion of this paper in Sect. 6.

2 Preliminaries

In this section, we briefly recall the basic concepts related to IF set, and then list the properties of the axiomatic definition for similarity measures.

Definition 1 [1] Let $X = \{x_1, x_2, \dots, x_n\}$ be a universe of discourse, then a FS A in X is defined as follows:

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

where $\mu_A(x) : X \rightarrow [0, 1]$ is the membership degree.

Definition 2 [4] An IF set A in X defined by Atanassov can be written as

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

where $\mu_A(x) : X \rightarrow [0, 1]$ and $\nu_A(x) : X \rightarrow [0, 1]$ are membership degree and non-membership degree, respectively, with the condition:

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1 \quad (3)$$

$\pi_A(x)$ determined by the following expression:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \quad (4)$$

is called the hesitancy degree of the element $x \in X$ to the set A , and $\pi_A(x) \in [0, 1]$, $\forall x \in X$.

$\pi_A(x)$ is also called the intuitionistic index of x to A . Greater $\pi_A(x)$ indicates more vagueness on x . Obviously, when $\pi_A(x) = 0$, $\forall x \in X$, the IF set degenerates into an ordinary FS.

In the sequel, the couple $\langle \mu_A(x), \nu_A(x) \rangle$ is called an IF set or IF value for clarity. Let $IFSs(X)$ denote the set of all IF sets in X .

Definition 3 For $A \in IFSs(X)$ and $B \in IFSs(X)$, some relations between them are defined as:

- (R1) $A \subset B$ iff $\forall x \in X \mu_A(x) \leq \mu_B(x), \nu_A(x) \geq \nu_B(x)$
- (R2) $A = B$ iff $\forall x \in X \mu_A(x) = \mu_B(x), \nu_A(x) = \nu_B(x)$
- (R3) $A^C = \{ \langle x, \nu_A(x), \mu_A(x) \rangle | x \in X \}$, where A^C is the complement of A .

It is worth noting that besides Definition 2 there are other possible representations of IF sets proposed in the literature. Hong and Choi [27] proposed to use an interval representation $[\mu_A(x), 1 - \nu_A(x)]$ of IF set A in X instead of pair $\langle \mu_A(x), \nu_A(x) \rangle$. This approach is equivalent to the interval valued FSs interpretation of IF set, where the interval $[\mu_A(x), 1 - \nu_A(x)]$ represents the membership degree of $x \in X$ to the set A . Obviously, $[\mu_A(x), 1 - \nu_A(x)]$ is a valid interval, since $\mu_A(x) \leq 1 - \nu_A(x)$ always holds for $\mu_A(x) + \nu_A(x) \leq 1$.

3 Distance and similarity measures for IF sets

3.1 Related definitions and properties

Generally, a distance is a measure of the difference between two elements of a set. For the case of IF sets, the distance between them must satisfy the axiomatic definitions of a metric distance. Moreover, the distance should not counter-intuitive analysis, i.e., the distance measure must be capable of reflecting the similarity among the IF sets. So we have next definition.

Definition 4 Let D denote a mapping $D : IFS \times IFS \rightarrow [0, 1]$, if $D(A, B)$ satisfies the following properties, is called a distance between $A \in IFSs(X)$ and $B \in IFSs(X)$.

- (D1) $0 \leq D(A, B) \leq 1$,
- (D2) $D(A, B) = 0 \Leftrightarrow A = B$,
- (D3) $D(A, B) = D(B, A)$,
- (D4) If $A \subseteq B \subseteq C$, then $D(A, B) \leq D(A, C)$, and $D(B, C) \leq D(A, C)$,
- (D5) $D(A, B) + D(B, C) \geq D(A, C)$.

As the complementary concept of distance measure, the similarity measure between two IF sets can be described in next definition.

Definition 5 A mapping $S : IFS \times IFS \rightarrow [0, 1]$ is called a degree of similarity between $A \in IFSs(X)$ and $B \in IFSs(X)$ a, if $S(A, B)$ satisfies the following properties:

- (S1) $0 \leq S(A, B) \leq 1$,
- (S2) $S(A, B) = 1 \Leftrightarrow A = B$,
- (S3) $S(A, B) = S(B, A)$,
- (S4) If $A \subseteq B \subseteq C$, then $S(A, B) \geq S(A, C)$ and $S(B, C) \geq S(A, C)$.

Theorem 1 Let D denote the distance measure between IF sets. Then, $S_D = 1 - D$ is the similarity measure between IF sets.

Proof As the distance measure between IF sets, D satisfies the conditions in Definition 4 as:

$$0 \leq D(A, B) \leq 1, \\ D(A, B) = 0 \Leftrightarrow A = B, D(A, B) = D(B, A)$$

Considering $S_D = 1 - D$, we can get the following expressions straightforwardly:

$$0 \leq S_D(A, B) \leq 1, \\ S_D(A, B) = 1 \Leftrightarrow D(A, B) = 0 \Leftrightarrow A = B, \\ S_D(A, B) = S_D(B, A).$$

Given IF sets A, B and C satisfying $A \subseteq B \subseteq C$, we have $D(A, B) \leq D(A, C)$ and $D(B, C) \leq D(A, C)$.

So we get:

$$1 - D(A, B) \geq 1 - D(A, C), \quad 1 - D(B, C) \geq 1 - D(A, C).$$

Thus, $S_D(A, B) \geq S_D(A, C), S_D(B, C) \geq S_D(A, C)$.

Hence, $S_D(A, B) = 1 - D$ is the similarity measure between IF sets. \square

Theorem 1 says that the distance measure can be applied to define its complementary concept, similarity measure. Given a similarity measure S , we can learn from the proof of Theorem 1 that $D = 1 - S$ satisfies all the properties in Definition 4, excluding D5, the triangle inequality. So all the distance measures can be transformed to a similarity measure, but not vice versa. We can say that the axiomatic definition for distance measure is stricter than that of the similarity measure.

Since the inception of IF sets, many similarity measures between IF sets have been proposed in the technical literature. Table 1 summarizes several well-known similarity measures that will be analyzed in this paper. In this table, we let $X = \{x_1, x_2, \dots, x_n\}$ be a universe of discourse, $A \in IFSs(X)$ and $B \in IFSs(X)$ be two IF sets in X , denoted by $A = \{(x, \mu_A(x), \nu_A(x)) | x \in X\}$ and $B = \{(x, \mu_B(x), \nu_B(x)) | x \in X\}$, respectively. For clarity, we only give the expressions of similarity measures, with an absence of the interpretations of other intermediate variables, which can be found in related references.

3.2 Similarity matrix for IF sets

The distance and similarity measures can demonstrate the relationships between two IF sets. However, under most circumstances, more than two IF sets will we are confronted with, where the distance and similarity measures are not suitable to cope with the relationships among them. So we define the similarity matrix for IF sets.

Definition 6 Let A_1, A_2, \dots, A_N denote $N(N \geq 2)$ IF sets in the universe of discourse $X = \{x_1, x_2, \dots, x_n\}$. S denotes the similarity measure between IF sets. The similarity matrix between them is defined as:

$$S = \begin{bmatrix} S(A_1, A_1) & S(A_1, A_2) & \dots & S(A_1, A_N) \\ S(A_2, A_1) & S(A_2, A_2) & \dots & S(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ S(A_N, A_1) & S(A_N, A_2) & \dots & S(A_N, A_N) \end{bmatrix} \quad (5)$$

Since $S(A_i, A_j) = S(A_j, A_i)$ for $i = 1, 2, \dots, N$ and $S(A_i, A_i) = 1$, the similarity matrix S is a square and symmetric matrix, with 1 as its diagonal elements.

Theorem 2 The similarity matrix S between IF sets is a non-singular matrix.

Table 1 Existing similarity measures

Author(s)	Similarity measure
Li and Cheng [17]	$S_{DC}(A, B) = 1 - \sqrt[p]{\frac{\sum_{i=1}^n \varphi_A(x_i) - \varphi_B(x_i) ^p}{n}}$
Mitchell [18]	$S_{HB}(A, B) = \frac{1}{2} (\rho_\mu(A, B) + \rho_\nu(A, B))$
Liang and Shi [19]	$S_e^p(A, B) = 1 - \sqrt[p]{\frac{\sum_{i=1}^n (\phi_\mu(x_i) + \phi_\nu(x_i))^p}{n}}$
Liang and Shi [19]	$S_s^p(A, B) = 1 - \sqrt[p]{\frac{\sum_{i=1}^n \psi_{s1}(x_i) + \psi_{s2}(x_i) ^p}{n}}$
Liang and Shi [19]	$S_h^p(A, B) = 1 - \sqrt[p]{\frac{\sum_{i=1}^n (\eta_1(x_i) + \eta_2(x_i) + \eta_3(x_i))^p}{3n}}$
Chen [21]	$S_C(A, B) = 1 - \frac{\sum_{i=1}^n (\mu_A(x_i) - \nu_A(x_i)) - (\mu_B(x_i) - \nu_B(x_i)) }{2n}$
Hong and Kim [14]	$S_H(A, B) = 1 - \frac{\sum_{i=1}^n (\mu_A(x_i) - \mu_B(x_i)) - (\nu_A(x_i) - \nu_B(x_i)) }{2n}$
Ye [8]	$C_{IFS}(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{\mu_A(x_i)\mu_B(x_i) + \nu_A(x_i)\nu_B(x_i)}{\sqrt{(\mu_A(x_i))^2 + (\nu_A(x_i))^2} \sqrt{(\mu_B(x_i))^2 + (\nu_B(x_i))^2}}$
Li et al. [16]	$S_O(A, B) = 1 - \sqrt{\frac{\sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2)}{2n}}$
Hung and Yang [13]	$S_{HY}^1(A, B) = 1 - d_H(A, B), S_{HY}^2(A, B) = \frac{e^{-d_H(A,B)} - e^{-1}}{1 - e^{-1}}, S_{HY}^3(A, B) = \frac{1 - d_H(A,B)}{1 + d_H(A,B)}$
Li and Xu [15]	$S_L(A, B) = 1 - \frac{\sum_{i=1}^n (\mu_A(x_i) - \nu_A(x_i)) - (\mu_B(x_i) - \nu_B(x_i)) }{4n} - \frac{\sum_{i=1}^n (\mu_A(x_i) - \mu_B(x_i) + \nu_A(x_i) - \nu_B(x_i))}{4n}$
Zhang and Yu [23]	$S_t^p(A, B) = 1 - \sqrt[p]{\sum_{i=1}^n \frac{1}{2n(1+p)} \{ t(\mu_A(x_i) - \mu_B(x_i)) - (\nu_A(x_i) - \nu_B(x_i)) ^p + t(\nu_A(x_i) - \nu_B(x_i)) - (\mu_A(x_i) - \mu_B(x_i)) ^p \}}$

Proof Let’s suppose that S is a singular matrix. Then at least two of its column vectors are linearly dependent. Without any loss of generality, we can assume x_j and x_k are linearly dependent. So we have:

$$x_j = t \cdot x_k.$$

Hence,

$$x_{jp} = t \cdot x_{kp} \quad \text{for all } p = 1, 2, \dots, N.$$

Given $p = j$ and $p = k$, we have:

$$x_{jj} = t \cdot x_{kj} = 1, \quad x_{jk} = t \cdot x_{kk} = t$$

And then, two contradictory equations can be achieved as:

$$t = 1/x_{kj} > 1 \quad \text{and} \quad t = x_{jk} < 1.$$

Subsequently, the assumption that S is a singular matrix cannot stand up.

So the similarity matrix S between IF sets is a non-singular matrix. \square

Theorem 3 Let D be a metric distance measure between IF sets. The similarity matrix S_D defined by $S_D = 1 - D$ according Eq. (5) is a positive definite matrix.

Proof Theorem 1 indicates that $S_D = 1 - D$ is a similarity measure between IF sets. The similarity defined by matrix can be expressed as:

$$S_D = \begin{bmatrix} 1 - D(A_1, A_1) & 1 - D(A_1, A_2) & \dots & 1 - D(A_1, A_N) \\ 1 - D(A_2, A_1) & 1 - D(A_2, A_2) & \dots & 1 - D(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 - D(A_N, A_1) & 1 - D(A_N, A_2) & \dots & 1 - D(A_N, A_N) \end{bmatrix}$$

Given $D(A_i, A_i) = 0$ and $D(A_i, A_j) = D(A_j, A_i)$ for $i = 1, 2, \dots, N$, we also get:

$$S_D = S_D^T = \begin{bmatrix} 1 & 1 - D(A_1, A_2) & \dots & 1 - D(A_1, A_N) \\ 1 - D(A_2, A_1) & 1 & \dots & 1 - D(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 - D(A_N, A_1) & 1 - D(A_N, A_2) & \dots & 1 \end{bmatrix}$$

Since the eigenvalues of symmetric matrix are all real numbers, the similarity matrix S_D has N real eigenvalues (including repeated eigenvalues), denoted by $\lambda_1, \lambda_2, \dots, \lambda_N$.

Let λ be an arbitrary eigenvalue of S , i.e., $\lambda \in \{\lambda_1, \lambda_2, \dots, \lambda_N\}$. Following the Gerschgorin Theorem, we can get:

$$|\lambda - 1| \leq \sum_{\substack{j=1 \\ j \neq i}}^N (1 - D(A_i, A_j)), \quad \exists i \in \{1, 2, \dots, N\}$$

Thus,

$$\begin{aligned}
 |\lambda - 1| &\leq \sum_{\substack{j=1 \\ j \neq i}}^N (1 - D(A_i, A_j)) \\
 &= (N - 1) - \sum_{\substack{j=1 \\ j \neq i}}^N D(A_i, A_j), \quad \exists i \in \{1, 2, \dots, N\}
 \end{aligned}$$

Considering the following relations:

$$\sum_{k=1}^N \lambda_k = \sum_{i=1}^N S_{ii} = N \quad \text{and} \quad N \cdot \lambda_{\max} \geq \sum_{k=1}^N \lambda_k \geq N \cdot \lambda_{\min},$$

where λ_{\min} and λ_{\max} are the minimum and maximum eigenvalues, respectively, we have $\lambda_{\min} \leq 1$ and $\lambda_{\max} \geq 1$.

Then, can get:

$$\begin{aligned}
 |\lambda_{\min} - 1| &= 1 - \lambda_{\min} \\
 &\leq (N - 1) - \sum_{\substack{j=1 \\ j \neq i}}^N D(A_i, A_j), \quad \exists i \in \{1, 2, \dots, N\}, \\
 \lambda_{\min} &\geq 2 + \sum_{\substack{j=1 \\ j \neq i}}^N D(A_i, A_j) - N, \quad \exists i \in \{1, 2, \dots, N\}.
 \end{aligned}$$

No generality will be lost by considering $i = 1$, i.e., λ_{\min} is in the first Gerschgorin circle of S_D . So we have:

$$\lambda_{\min} \geq 2 + \sum_{j=2}^N D(A_1, A_j) - N.$$

(i) For $N = 2$, we have:

$$\lambda_{\min} \geq 2 + D(A_1, A_2) - 2 = D(A_1, A_2) > 0.$$

(ii) For $N = 3$, we have:

$$\begin{aligned}
 \lambda_{\min} &\geq 2 + D(A_1, A_2) + D(A_1, A_3) - 3 \\
 &= D(A_1, A_2) + D(A_1, A_3) - 1 \\
 &\geq D(A_2, A_3) - 1.
 \end{aligned}$$

Considering the randomness of $D(A_2, A_3)$ and $D(A_2, A_3) - 1 \leq 0$, we can conclude that only $\lambda_{\min} \geq 0$ can make $\lambda_{\min} \geq D(A_2, A_3) - 1$ holds for any $D(A_2, A_3)$. So we have $\lambda_{\min} \geq 0$.

(iii) Given $N \geq 4$ and $D(A_1, A_2) + D(A_1, A_3) \geq D(A_2, A_3)$, we have:

$$\begin{aligned}
 \lambda_{\min} &\geq 2 + \sum_{j=2}^N D(A_1, A_j) - N \\
 &= 2 - N + (D(A_1, A_2) + D(A_1, A_3)) + \sum_{j=4}^N D(A_1, A_j) \\
 &\geq 2 - N + D(A_2, A_3) + \sum_{j=4}^N D(A_1, A_j).
 \end{aligned}$$

We can also get

$$\begin{aligned}
 2 - N + D(A_2, A_3) + \sum_{j=4}^N D(A_1, A_j) \\
 \leq 2 - N + 1 + (N - 3) = 0.
 \end{aligned}$$

Since $\lambda_{\min} \geq 2 - N + D(A_2, A_3) + \sum_{j=4}^N D(A_1, A_j)$ is constant for any arbitrary $D(A_2, A_3)$ and $D(A_1, A_j)$ ($j = 4, 5, \dots, N$), λ_{\min} should be not less than the maximum of the right hand side. Then we have $\lambda_{\min} \geq 0$.

Considering (i)–(iii), we can conclude that all the eigenvalues of S_D are nonnegative. So the similarity matrix S_D is positive semidefinite.

Taking $\prod_{i=1}^N \lambda_i = \det S_D$ and $\det S_D \neq 0$ (S_D is non-singular) into consideration, we can know there is no zero eigenvalue in the eigenvalues of S_D , i.e., all the eigenvalues of S_D are strictly positive.

So the similarity matrix S_D defined by $S_D = 1 - D$ is positive definite. □

4 A new similarity measure between IF sets

In the similarity measures shown in Table 1, some of them are based on the well-known distances measures, such as the Hamming distance and the Euclidian distance, while others are characterized by the linear or non-linear combinations of the membership and non-membership functions of the IF sets, respectively. Since the similarity measures based on the Euclidian distance are of definitude physical meaning, we will review the generation of Euclidian distance in IF environment [11]. Moreover, some of its properties will be presented along with their proofs.

Definition 7 [11] The distance between two IF sets A and B in $X = \{x_1, x_2, \dots, x_n\}$ can be defined as:

$$D_o(A, B) = \sqrt{\frac{\sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2)}{2n}} \tag{6}$$

Theorem 4 $d_o(A, B)$ is a metric distance measure between IF sets.

Proof (1) Since $\mu(x), v(x) \in [0, 1]$, we have:

$$-1 \leq \mu_A(x_i) - \mu_B(x_i) \leq 1 \quad \text{and} \quad -1 \leq v_A(x_i) - v_B(x_i) \leq 1,$$

Hence,

$$0 \leq (\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2 \leq 1 + 1 = 2$$

So we get:

$$0 \leq \frac{\sum_{i=1}^n \left((\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2 \right)}{2n} \leq \frac{\sum_{i=1}^n 2}{2n} = 1$$

Thus, $0 \leq D_o(A, B) \leq 1$.

(2) $D_o(A, A) = 0$ is straightforward. Conversely, if $D_o(A, B) = 0$, according to the definition of D_o , it must follow that $\mu_A(x) = \mu_B(x)$ and $v_A(x) = v_B(x)$ for all $x \in X$, i.e., $A = B$. So we have $D_o(A, B) = 0 \Leftrightarrow A = B$.

(3) It is evident that $D_o(A, B) = D_o(B, A)$.

(4) Given $A \subseteq B \subseteq C$, we have $\mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$ and $v_A(x) \geq v_B(x) \geq v_C(x)$ for all $x \in X$. Therefore,

$$\begin{aligned} 0 &\leq \mu_B(x_i) - \mu_A(x_i) \leq \mu_C(x_i) - \mu_A(x_i), \\ 0 &\leq \mu_C(x_i) - \mu_B(x_i) \leq \mu_C(x_i) - \mu_A(x_i), \\ 0 &\leq v_A(x_i) - v_B(x_i) \leq v_A(x_i) - v_C(x_i), \\ 0 &\leq v_B(x_i) - v_C(x_i) \leq v_A(x_i) - v_C(x_i), \quad \text{for } \forall x_i \in X. \end{aligned}$$

Consequently,

$$\begin{aligned} &\sqrt{\frac{\sum_{i=1}^n \left((\mu_B(x_i) - \mu_A(x_i))^2 + (v_A(x_i) - v_B(x_i))^2 \right)}{2n}} \\ &\leq \sqrt{\frac{\sum_{i=1}^n \left((\mu_C(x_i) - \mu_A(x_i))^2 + (v_A(x_i) - v_C(x_i))^2 \right)}{2n}}, \\ &\sqrt{\frac{\sum_{i=1}^n \left((\mu_C(x_i) - \mu_B(x_i))^2 + (v_B(x_i) - v_C(x_i))^2 \right)}{2n}} \\ &\leq \sqrt{\frac{\sum_{i=1}^n \left((\mu_C(x_i) - \mu_A(x_i))^2 + (v_A(x_i) - v_C(x_i))^2 \right)}{2n}}. \end{aligned}$$

According to Eq. (6), we can get: $D_o(A, B) \leq D_o(A, C)$ and $D_o(B, C) \leq D_o(A, C)$.

(5) By the Cauchy inequality, $(\sum_{i=1}^n a_i b_i)^2 \leq (\sum_{i=1}^n a_i^2) (\sum_{i=1}^n b_i^2)$, we can get:

$$\begin{aligned} \sum_{i=1}^n (a_i + b_i)^2 &= \sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2 + 2 \sum_{i=1}^n a_i b_i \\ &\leq \sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2 + 2 \left(\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2 \right)^{1/2} \\ &= \left(\sqrt{\sum_{i=1}^n a_i^2} + \sqrt{\sum_{i=1}^n b_i^2} \right)^2. \end{aligned}$$

Making the following assignments:

$$\begin{aligned} a &= \sum_{i=1}^n (\mu_A(x_i) - \mu_B(x_i))^2, & b &= \sum_{i=1}^n (\mu_B(x_i) - \mu_C(x_i))^2 \\ g &= \sum_{i=1}^n (v_A(x_i) - v_B(x_i))^2, & h &= \sum_{i=1}^n (v_B(x_i) - v_C(x_i))^2 \end{aligned}$$

we have:

$$\begin{aligned} &2n(D_o(A, C))^2 \\ &= \sum_{i=1}^n \left((\mu_A(x_i) - \mu_C(x_i))^2 + (v_A(x_i) - v_C(x_i))^2 \right) \\ &= \sum_{i=1}^n \left((\mu_A(x_i) - \mu_B(x_i) + \mu_B(x_i) - \mu_C(x_i))^2 + (v_A(x_i) - v_B(x_i) + v_B(x_i) - v_C(x_i))^2 \right) \\ &\leq \left(\sqrt{\sum_{i=1}^n (\mu_A(x_i) - \mu_B(x_i))^2} + \sqrt{\sum_{i=1}^n (\mu_B(x_i) - \mu_C(x_i))^2} \right)^2 \\ &\quad + \left(\sqrt{\sum_{i=1}^n (v_A(x_i) - v_B(x_i))^2} + \sqrt{\sum_{i=1}^n (v_B(x_i) - v_C(x_i))^2} \right)^2 \\ &\leq (\sqrt{a} + \sqrt{b})^2 + (\sqrt{g} + \sqrt{h})^2 \\ &= a + b + g + h + 2\sqrt{ab} + 2\sqrt{gh}. \end{aligned}$$

$$\begin{aligned} &2n(D_o(A, B) + D_o(B, C))^2 \\ &= \left(\sqrt{\sum_{i=1}^n \left((\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2 \right)} + \sqrt{\sum_{i=1}^n \left((\mu_B(x_i) - \mu_C(x_i))^2 + (v_B(x_i) - v_C(x_i))^2 \right)} \right)^2 \\ &= (\sqrt{a+g} + \sqrt{b+h})^2 \\ &= a + b + g + h + 2\sqrt{(a+g)(b+h)}. \end{aligned}$$

Considering such inequality:

$$\begin{aligned} &(a+g)(b+h) - (\sqrt{ab} + \sqrt{gh})^2 \\ &= ah + bg - 2\sqrt{abgh} \\ &= (\sqrt{ah} - \sqrt{bg})^2 \geq 0, \end{aligned}$$

we get:

$$\begin{aligned} &\sqrt{(a+g)(b+h)} \geq \sqrt{ab} + \sqrt{gh}, \\ &a + b + g + h + 2\sqrt{(a+g)(b+h)} \\ &\geq a + b + g + h + 2\sqrt{ab} + 2\sqrt{gh}. \end{aligned}$$

And then

$$2n(D_o(A, C))^2 \leq 2n(D_o(A, B) + D_o(B, C))^2.$$

Finally we obtain $D_o(A, C) \leq D_o(A, B) + D_o(B, C)$. So

$D_o(A, B)$ is a metric distance measure between IF sets. \square

According to Theorem 1, $S_o = 1 - D_o(A, B)$ is a similarity measure between IF sets. This similarity has been proposed by Li et al. [16]. However, this similarity measure has a drawback that it is not sensitive to the change of IF sets. Considering an example where $A = \{ \langle x, 0.3, 0.2 \rangle \}$, $B = \{ \langle x, 0.4, 0.3 \rangle \}$, $C = \{ \langle x, 0.4, 0.1 \rangle \}$, $D = \{ \langle x, 0.2, 0.1 \rangle \}$, $E = \{ \langle x, 0.2, 0.3 \rangle \}$, we can find $S_o(A, B) = S_o(A, C) = S_o(A, D) = S_o(A, E)$. So S_o is not capable of discriminating the difference between IF sets.

Besides, there is another interesting similarity measure between IF sets, cosine similarity measure, defined by Ye [8]. He has proved that $C_{IFS}(A, B)$ satisfied properties (S1) and (S3) in Definition 5. He only illustrated that $C_{IFS}(A, B) = 1$ if $A = B$. However, for two different IF sets $A = \{ \langle x, 0.3, 0.3 \rangle \}$ and $B = \{ \langle x, 0.4, 0.4 \rangle \}$, we can get $C_{IFS}(A, B) = 1$. That is, the condition (S2) in Definition 5 is not satisfied. So we can say $C_{IFS}(A, B)$ is not a genuine similarity measure.

Taking a close examine on $C_{IFS}(A, B)$ and $D_o(A, B)$, we can get that the cosine similarity measure indicates the angle which quantifies how orthogonal two IF sets are, while the distance between IF sets can quantify how close two IF sets are from each other. Therefore, we can combine the cosine similarity measure and the distance measure together to define a new similarity measure for IF sets.

Definition 8 Let A and B be two IF sets in $X = \{x_1, x_2, \dots, x_n\}$. A new similarity measure between them can be defined as:

$$S_F(A, B) = \frac{1}{2}(C_{IFS}(A, B) + 1 - D_o(A, B)) \tag{7}$$

Theorem 5 The measure $S_F(A, B)$ is a similarity measure between IF sets A and B .

Proof (1) $C_{IFS}(A, B)$ can be taken as the cosine value between vectors, so $0 \leq C_{IFS}(A, B) \leq 1$. According to Theorems 1 and 3, it is evident that $0 \leq 1 - D_o(A, B) \leq 1$. Then we can get: $0 \leq S_F(A, B) \leq 1$.

(2) $S_F(A, A) = 1$ is straightforward. Since $0 \leq C_{IFS}(A, B) \leq 1$ and $0 \leq 1 - D_o(A, B) \leq 1$, $C_{IFS}(A, B) = 1 - D_o(A, B) = 1$ must holds when $S_F(A, B) = 1$.

$C_{IFS}(A, B) = 1$ indicates that $\mu_A(x) = k \cdot v_A(x)$, $\mu_B(x) = k \cdot v_B(x)$ with $k \in [0, +\infty)$ for all $x \in X$. Considering

$D_o(A, B) = 0 \Rightarrow A = B$ shown in Theorem 3, we can get the converse proposition $S_F(A, B) = 1 \Rightarrow A = B$.

Finally, we get $S_F(A, B) = 1 \Leftrightarrow A = B$.

(3) $S_F(A, B) = S_F(B, A)$ can be obtain straightforwardly.

(4) For three IF sets A, B and C satisfying $A \subseteq B \subseteq C$, we have $\mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$ and $v_A(x) \geq v_B(x) \geq v_C(x)$ for all $x \in X$.

A function $f(y, z)$ can be defined as:

$$f(y, z) = \frac{ay + bz}{\sqrt{a^2 + b^2}\sqrt{y^2 + z^2}}$$

We can calculate its derivatives as:

$$\begin{aligned} \frac{\partial f}{\partial y} &= \frac{a\sqrt{y^2 + z^2} - (ay + bz)\frac{y}{\sqrt{y^2 + z^2}}}{\sqrt{a^2 + b^2}(y^2 + z^2)} \\ &= \frac{az^2 - byz}{\sqrt{a^2 + b^2}\sqrt{y^2 + z^2}(y^2 + z^2)} \\ &= \frac{z(az - by)}{\sqrt{a^2 + b^2}\sqrt{y^2 + z^2}(y^2 + z^2)} \\ \frac{\partial f}{\partial z} &= \frac{b\sqrt{y^2 + z^2} - (ay + bz)\frac{z}{\sqrt{y^2 + z^2}}}{\sqrt{a^2 + b^2}(y^2 + z^2)} \\ &= \frac{by^2 - ayz}{\sqrt{a^2 + b^2}\sqrt{y^2 + z^2}(y^2 + z^2)} \\ &= \frac{-y(az - by)}{\sqrt{a^2 + b^2}\sqrt{y^2 + z^2}(y^2 + z^2)} \end{aligned}$$

If $y \geq a, z \leq b$, we have $\frac{\partial f}{\partial y} \leq 0, \frac{\partial f}{\partial z} \geq 0$. So for $a = \mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$ and $b = v_A(x) \geq v_B(x) \geq v_C(x)$, we have $f(\mu_C(x), v_C(x)) \leq f(\mu_B(x), v_B(x))$, which can be written as:

$$\begin{aligned} &\frac{\mu_A(x)\mu_C(x) + v_A(x)v_C(x)}{\sqrt{(\mu_A(x))^2 + (v_A(x))^2}\sqrt{(\mu_C(x))^2 + (v_C(x))^2}} \\ &\leq \frac{\mu_A(x)\mu_B(x) + v_A(x)v_B(x)}{\sqrt{(\mu_A(x))^2 + (v_A(x))^2}\sqrt{(\mu_B(x))^2 + (v_B(x))^2}} \tag{8} \end{aligned}$$

If $y \leq a, z \geq b$, we have $\frac{\partial f}{\partial y} \geq 0, \frac{\partial f}{\partial z} \leq 0$. So for $a = \mu_C(x) \geq \mu_B(x) \geq \mu_A(x)$ and $b = v_C(x) \leq v_B(x) \leq v_A(x)$, we have $f(\mu_A(x), v_A(x)) \leq f(\mu_B(x), v_B(x))$, which can be written as:

$$\frac{\mu_A(x)\mu_C(x) + v_A(x)v_C(x)}{\sqrt{(\mu_A(x))^2 + (v_A(x))^2} \sqrt{(\mu_C(x))^2 + (v_C(x))^2}} \leq \frac{\mu_C(x)\mu_B(x) + v_C(x)v_B(x)}{\sqrt{(\mu_C(x))^2 + (v_C(x))^2} \sqrt{(\mu_B(x))^2 + (v_B(x))^2}} \tag{9}$$

Since $\mu_A(x) \leq \mu_B(x) \leq \mu_C(x)$ and $v_A(x) \geq v_B(x) \geq v_C(x)$ for all $x \in X$, (8) and (9) holds for all $x \in X$. So we get: $C_{IFS}(A, C) \leq C_{IFS}(A, B)$, $C_{IFS}(A, C) \leq C_{IFS}(B, C)$.

Moreover, $D_o(A, B) \leq D_o(A, C)$ and $D_o(B, C) \leq D_o(A, C)$ have been proved in the proof of Theorem 3. So we have $1 - D_o(A, C) \leq 1 - D_o(A, B)$ and $1 - D_o(A, C) \leq 1 - D_o(B, C)$. We can finally get: $S_F(A, C) \leq S_F(A, B)$ and $S_F(A, C) \leq S_F(B, C)$.

Thus, $S_F(A, B)$ satisfies all the properties in Definition 5, and it is a similarity measure between IF sets A and B . \square

Theorem 6 Let A_1, A_2, \dots, A_N denote N ($N \geq 2$) IF sets in the universe of discourse X . The similarity matrix S_F defined by S_F is a positive definite matrix.

Proof The similarity matrix S_F can be decomposed as:

$$S_F = \frac{1}{2}S_o + \frac{1}{2}S_C$$

where

$$S_o = \begin{bmatrix} S_o(A_1, A_1) & S_o(A_1, A_2) & \dots & S_o(A_1, A_N) \\ S_o(A_2, A_1) & S_o(A_2, A_2) & \dots & S_o(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ S_o(A_N, A_1) & S_o(A_N, A_2) & \dots & S_o(A_N, A_N) \end{bmatrix}$$

$$= \begin{bmatrix} 1 - D_o(A_1, A_1) & 1 - D_o(A_1, A_2) & \dots & 1 - D_o(A_1, A_N) \\ 1 - D_o(A_2, A_1) & 1 - D_o(A_2, A_2) & \dots & 1 - D_o(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 - D_o(A_N, A_1) & 1 - D_o(A_N, A_2) & \dots & 1 - D_o(A_N, A_N) \end{bmatrix}$$

$$S_C = \begin{bmatrix} C_{IFS}(A_1, A_1) & C_{IFS}(A_1, A_2) & \dots & C_{IFS}(A_1, A_N) \\ C_{IFS}(A_2, A_1) & C_{IFS}(A_2, A_2) & \dots & C_{IFS}(A_2, A_N) \\ \vdots & \vdots & \ddots & \vdots \\ C_{IFS}(A_N, A_1) & C_{IFS}(A_N, A_2) & \dots & C_{IFS}(A_N, A_N) \end{bmatrix}$$

The cosine similarity can be also written as:

$$C_{IFS}(A_j, A_k) = \frac{1}{n} \sum_{i=1}^n \frac{\mu_{A_j}(x_i)\mu_{A_k}(x_i) + v_{A_j}(x_i)v_{A_k}(x_i)}{\sqrt{(\mu_{A_j}(x_i))^2 + (v_{A_j}(x_i))^2} \sqrt{(\mu_{A_k}(x_i))^2 + (v_{A_k}(x_i))^2}}$$

$$= \frac{1}{n} \sum_{i=1}^n \left(\frac{\frac{\mu_{A_j}(x_i)}{\sqrt{(\mu_{A_j}(x_i))^2 + (v_{A_j}(x_i))^2}} \cdot \frac{\mu_{A_k}(x_i)}{\sqrt{(\mu_{A_k}(x_i))^2 + (v_{A_k}(x_i))^2}}}{1} + \frac{\frac{v_{A_j}(x_i)}{\sqrt{(\mu_{A_j}(x_i))^2 + (v_{A_j}(x_i))^2}} \cdot \frac{v_{A_k}(x_i)}{\sqrt{(\mu_{A_k}(x_i))^2 + (v_{A_k}(x_i))^2}}}{1} \right)$$

So S_C can be further decomposed as:

$$S_C = \frac{1}{n} \sum_{i=1}^n S_\mu(x_i) + \frac{1}{n} \sum_{i=1}^n S_\nu(x_i)$$

where for $i = 1, 2, \dots, n$,

$$S_\mu(x_i) = \begin{bmatrix} F_1(x_i) \cdot F_1(x_i) & F_1(x_i) \cdot F_2(x_i) & \dots & F_1(x_i) \cdot F_N(x_i) \\ F_2(x_i) \cdot F_1(x_i) & F_2(x_i) \cdot F_2(x_i) & \dots & F_2(x_i) \cdot F_N(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ F_N(x_i) \cdot F_1(x_i) & F_N(x_i) \cdot F_2(x_i) & \dots & F_N(x_i) \cdot F_N(x_i) \end{bmatrix},$$

$$S_\nu(x_i) = \begin{bmatrix} G_1(x_i) \cdot G_1(x_i) & G_1(x_i) \cdot G_2(x_i) & \dots & G_1(x_i) \cdot G_N(x_i) \\ G_2(x_i) \cdot G_1(x_i) & G_2(x_i) \cdot G_2(x_i) & \dots & G_2(x_i) \cdot G_N(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ G_N(x_i) \cdot G_1(x_i) & G_N(x_i) \cdot G_2(x_i) & \dots & G_N(x_i) \cdot G_N(x_i) \end{bmatrix},$$

with $F_j(x_i) = F(\mu_{A_j}(x_i), v_{A_j}(x_i)) = \frac{\mu_{A_j}(x_i)}{\sqrt{(\mu_{A_j}(x_i))^2 + (v_{A_j}(x_i))^2}}$,

and $G_j(x_i) = G(\mu_{A_j}(x_i), v_{A_j}(x_i)) = \frac{v_{A_j}(x_i)}{\sqrt{(\mu_{A_j}(x_i))^2 + (v_{A_j}(x_i))^2}}$,

for $j = 1, 2, \dots, N$.

Let $\mu_i = [F_1(x_i), F_2(x_i), \dots, F_N(x_i)]$ and $\nu_i = [G_1(x_i), G_2(x_i), \dots, G_N(x_i)]$, then we can get:

$$S_\mu(x_i) = \mu_i^T \mu_i, \quad S_\nu(x_i) = \nu_i^T \nu_i.$$

So $S_\mu(x_i)$ and $S_\nu(x_i)$ are positive semidefinite.

And then $S_C = \frac{1}{n} \sum_{i=1}^n S_\mu(x_i) + \frac{1}{n} \sum_{i=1}^n S_\nu(x_i)$ is also positive semidefinite.

Considering S_o is positive definite and S_C is positive semidefinite, we can conclude that $S_F = \frac{1}{2}S_o + \frac{1}{2}S_C$ is positive definite. \square

5 Illustrative examples

5.1 Performance of similarity measure S_F

To illustrate the superiority of the proposed similarity measure, a comparison between the proposed similarity measure and all the existing similarity measures is conducted. The comparison is implemented based on the widely used counter-intuitive examples. Table 2 presents the result with $p = 1$ for $S_{HB}, S_e^p, S_s^p, S_h^p$ and $p = 1, t = 2$ for S_t^p .

From Table 2, We can see that $S_C(A, B) = S_{DC}(A, B) = C_{IFS}(A, B) = 1$ for two different IF sets $A = \langle 0.3, 0.3 \rangle$ and $B = \langle 0.4, 0.4 \rangle$. This indicates that the second axiom of similarity measure (S2) is not satisfied by $S_C(A, B)$, $S_{DC}(A, B)$ and $C_{IFS}(A, B)$. This also can be illustrated by $S_C(A, B) = S_{DC}(A, B) = 1$ when $A = \langle 0.5, 0.5 \rangle$, $B = \langle 0, 0 \rangle$ and $A = \langle 0.4, 0.2 \rangle$, $B = \langle 0.5, 0.3 \rangle$. As for $S_H, S_O, S_{HB}, S_e^p, S_s^p$ and S_h^p , different pairs of A, B may provide the identical results,

Table 2 The comparison of similarity measures (counter-intuitive cases are in bold type)

	1	2	3	4	5	6
$A = \langle \mu(x), \nu(x) \rangle$	$\langle 0.3, 0.3 \rangle$	$\langle 0.3, 0.4 \rangle$	$\langle 1, 0 \rangle$	$\langle 0.5, 0.5 \rangle$	$\langle 0.4, 0.2 \rangle$	$\langle 0.4, 0.2 \rangle$
$B = \langle \mu(x), \nu(x) \rangle$	$\langle 0.4, 0.4 \rangle$	$\langle 0.4, 0.3 \rangle$	$\langle 0, 0 \rangle$	$\langle 0, 0 \rangle$	$\langle 0.5, 0.3 \rangle$	$\langle 0.5, 0.2 \rangle$
S_C	1	0.9	0.5	1	1	0.95
S_H	0.9	0.9	0.5	0.5	0.9	0.95
S_L	0.95	0.9	0.5	0.75	0.95	0.95
S_O	0.9	0.9	0.3	0.5	0.9	0.93
S_{DC}	1	0.9	0.5	1	1	0.95
S_{HB}	0.9	0.9	0.5	0.5	0.9	0.95
S_e^p	0.9	0.9	0.5	0.5	0.9	0.95
S_s^p	0.95	0.9	0.5	0.75	0.95	0.95
S_h^p	0.933	0.933	0.5	0.67	0.933	0.95
S_{HY}^1	0.9	0.9	0	0.5	0.9	0.9
S_{HY}^2	0.85	0.85	0	0.38	0.85	0.85
S_{HY}^3	0.82	0.82	0	0.33	0.82	0.82
C_{IFS}	1	0.96	0	0	0.9971	0.9965
S_t^p	0.967	0.9	0.5	0.833	0.967	0.95
S_F	0.95	0.93	0.15	0.25	0.948	0.963

which cannot satisfy the application of pattern recognition. It can be read from Table 2 that $S_{HB} = 0.9$ for both $A = \langle 0.3, 0.3 \rangle$, $B = \langle 0.4, 0.4 \rangle$ and $A = \langle 0.3, 0.4 \rangle$, $B = \langle 0.4, 0.3 \rangle$. Such situation seems to be worse for S_{HY}^1, S_{HY}^2 and S_{HY}^3 , where all the cases take the same similarity degree except case 3 and case 4. S_t^p seems to be reasonable without any counter-intuitive results, but it bring new problem with the choice of parameters p and t , which is still an open problem. Moreover, we can notice an interesting situation when comparing case 3 and case 4. For three IF sets $A = \langle 1, 0 \rangle$, $B = \langle 0.5, 0.5 \rangle$ and $C = \langle 0, 0 \rangle$, intuitively, it is more reasonable to take the similarity degree between them as: $S_F(A, C) = 0.15$, $S_F(B, C) = 0.25$ than taking $S_t^p(A, C) = 0.5$ and $S_t^p(B, C) = 0.833$. In such a sense, the proposed similarity measure is the most reasonable one with a relative simple expression, and has none of the counter-intuitive cases. Three IF sets $A = \langle 0.4, 0.2 \rangle$, $B = \langle 0.5, 0.3 \rangle$ and $C = \langle 0.5, 0.2 \rangle$ can be written in forms of interval values as: $A = [0.4, 0.8]$, $B = [0.5, 0.7]$ and $C = [0.5, 0.8]$, respectively. In such a sense, we can say that the similarity degree between A and C is greater than the similarity degree between A and B .

Therefore, our proposed similarity measure is in agreement with this analysis. The proposed similarity measure is the most reasonable similarity measure without any counter-intuitive cases. That is because that our proposed similarity measure combines cosine similarity C_{IFS} and distance-inducing similarity $(1 - D_o)$. If we take IF sets $A = \langle \mu_A(x), \nu_A(x) \rangle$ and $B = \langle \mu_B(x), \nu_B(x) \rangle$ as two vectors, $C_{IFS}(A, B)$ and $D_o(A, B)$ represent the angle which quantifies how orthogonal two IF sets are and the distance between IF sets can quantify how close two IF sets are from each other, respectively. Generally, the relationship between two

vectors can be determined by angle and distance. So the combination of $C_{IFS}(A, B)$ and $(1 - D_o(A, B))$ can be applied to define similarity measure for IF sets without any counter-intuitive cases. Moreover, there is no need to determine other parameters. Such analysis is also potentially applicable to any IF sets on arbitrary universe $X = \{x_1, x_2, \dots, x_n\}$.

Besides the satisfaction of the definitional axioms and the avoidance of counter-intuitive cases, the discrimination capability of a measure is another important property for similarity measures, which is very useful in pattern recognition applications. To study the effectiveness of the proposed similarity measure for IF sets in the application of pattern recognition, the widely used pattern recognition problem discussed in [8, 17] will be considered.

Suppose there are m patterns, which can be represented by IF sets $A_j = \left\{ \langle x_i, \mu_{A_j}(x_i), \nu_{A_j}(x_i) \rangle \mid x_i \in X \right\}$, $A_j \in IFSs(X)$, $j = 1, 2, \dots, m$. Let the sample to be recognized be denoted as $B = \{ \langle x_i, \mu_B(x_i), \nu_B(x_i) \rangle \mid x_i \in X \}$. According to the recognition principle of maximum degree of similarity between IF sets, the process of assigning B to A_k is described by:

$$k = \arg \max_{j=1,2,\dots,m} \{S(A_j, B)\} \tag{10}$$

To illustrate the discrimination capability of our proposed similarity measure, comparisons with the measures proposed earlier by other authors will be carried out based on the example analyzed in [22].

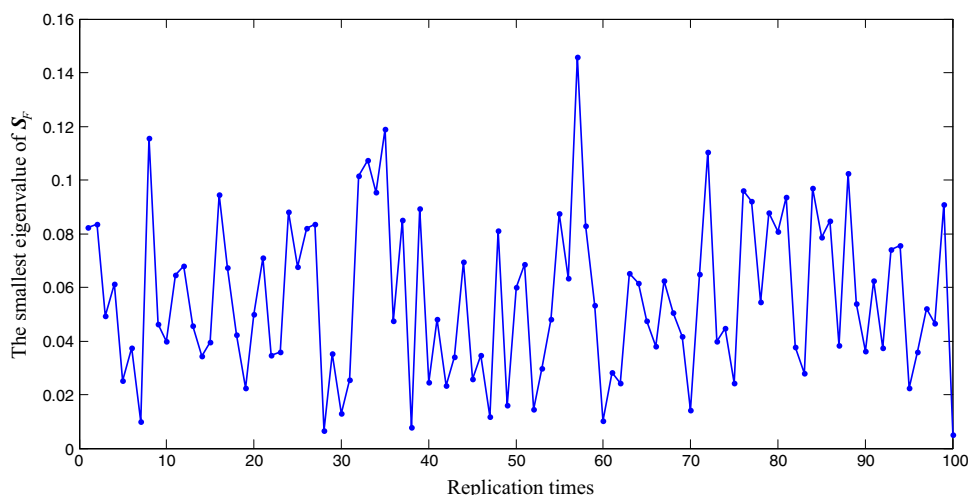
Example 1 Assume that there are three IF sets in $X = \{x_1, x_2, x_3\}$ representing three patterns. The three patterns are written as follows:

Table 3 The similarity measures between the known patterns and the unknown pattern in **Example** (Patterns not discriminated are in bold type)

	$S(A_1, B)$	$S(A_2, B)$	$S(A_3, B)$		$S(A_1, B)$	$S(A_2, B)$	$S(A_3, B)$
S_C	1	1	1	S_s^p	1	0.967	0.900
S_H	1	1	1	S_h^p	1	0.956	0.867
S_L	1	0.967	0.9	S_{HY}^1	1	0.967	0.8
S_O	1	0.918	0.784	S_{HY}^2	1	0.898	0.713
S_{DC}	1	1	1	S_{HY}^3	1	0.875	0.667
S_{HB}	1	0.933	0.8	C_{IFS}	1	1	1
S_e^p	1	0.933	0.8	S_t^p	1	0.978	0.933

$p = 1$ for $S_{HB}, S_e^p, S_s^p, S_h^p$ and $p = 1, t = 2$ for S_t^p

Fig. 1 Evolution of the smallest eigenvalue of S_F with 100 replications



$$A_1 = \{\langle x_1, 0.3, 0.3 \rangle, \langle x_2, 0.2, 0.2 \rangle, \langle x_3, 0.1, 0.1 \rangle\},$$

$$A_2 = \{\langle x_1, 0.2, 0.2 \rangle, \langle x_2, 0.2, 0.2 \rangle, \langle x_3, 0.2, 0.2 \rangle\},$$

$$A_3 = \{\langle x_1, 0.4, 0.4 \rangle, \langle x_2, 0.4, 0.4 \rangle, \langle x_3, 0.4, 0.4 \rangle\}.$$

Assume that a sample $B = \{\langle x_1, 0.3, 0.3 \rangle, \langle x_2, 0.2, 0.2 \rangle, \langle x_3, 0.1, 0.1 \rangle\}$ is to be classified.

The similarity degrees of $S(A_1, B)$, $S(A_2, B)$ and $S(A_3, B)$ calculated for all similarity measures listed in Table 1 are shown in Table 3.

The proposed similarity measure S_F can be calculated by Eq. 7 as:

$$S_F(A_1, B) = 1, S_F(A_2, B) = 0.959, S_F(A_3, B) = 0.892.$$

It is obvious that B is equal to A_1 , which indicates that sample B should be classified to A_1 . However, the similarity degrees of $S(A_1, B)$, $S(A_2, B)$ and $S(A_3, B)$ are equal to each other when S_C, S_H, S_{DC} and C_{IFS} are employed. These four similarity measures cannot be capable of discriminating difference between the three patterns.

Fortunately, the results of $S_F(A_i, B)$ ($i = 1, 2, 3$) can be used to make correct classification conclusion. This means that the proposed similarity measure shows an identical discrimination capability with majority of the existing measures.

5.2 Properties of similarity matrix S_F

As an illustration of Theorem 6, the positive definiteness of the similarity matrix S_F will be verified by number of IF sets. Without any loss of generality, we investigate the similarity matrix S_F among a group of five IF sets. For the sake of persuasiveness, all IF sets are generated randomly, with 100 replications. The evolution of the smallest eigenvalue of S_F is displayed in Fig. 1. We observe that all the smallest eigenvalues are strictly positive, which indicates that similarity matrices between random IF sets are all positive definite. Such result is identical to the proof of Theorem 6. So we have enough confidence to claim that the similarity matrix defined based on our proposed similarity measure is positive definite.

6 Conclusion

The definition of similarity measure between two IF sets has been researched for decades. Even though researchers have defined mounts of similarity measures to depict the similarity degree between IF sets, most of them are stuck with counter-intuitive results. After analyzing these similarity measures critically, a new similarity measure is proposed in this paper. To explore the properties of our proposed similarity measure, the similarity matrix is also defined. A comparison between our proposed similarity measure and other existing measures is carried out based on the widely used counter-intuitive cases. It is illustrated that the proposed similarity measure is more reasonable, without any counter-intuitive results. It has also been proved that the similarity matrix defined by our proposed similarity measure is positive definite, which is significant for the application of similarity matrix.

One case worth mentioning is that our proposed similarity measure is not the only similarity measure that can be used to define positive definite similarity matrix. Besides the combination of cosine similarity and Euclidean distance, cosine similarity can be also combined with Hamming distance to define a new similarity in a similar way. So much work remains to be done for a better exploration and exploitation of IF set theory.

Appendix 1: Definitions and properties about positive definite matrix

When proving theorems concerning positive definite matrices, we use some properties of positive definite matrix. For ease of reference, some background knowledge related to positive definite matrix is shown in this part. Since all the results given below are well known, we mainly present definitions and theorems with an absence of their proofs.

Definition 9 An $n \times n$ real symmetric matrix A is positive semidefinite (PSD) if it holds that $\mathbf{x}^T A \mathbf{x} \geq 0$ for every $n \times 1$ column vector $\mathbf{x} \neq \mathbf{0}$. It is (strictly) positive definite (PD) if additionally: $\mathbf{x}^T A \mathbf{x} = 0 \Rightarrow \mathbf{x} = \mathbf{0}$.

Theorem 7 Eigenvalues of real symmetric matrix A are all real numbers.

Theorem 8 A is PSD iff its eigenvalues are nonnegative and A is PD iff its eigenvalues are strictly positive.

Theorem 9 If two matrices A and B are both PSD, $A + B$ is also PSD.

Consequently, the sum of PSD matrices is PSD. The result is PD if there is at least one PD matrix among the PSD matrices.

Theorem 10 A set of necessary and sufficient conditions for an $n \times n$ symmetric matrix A to be PSD is that all the principal leading minors $\Delta_{pp}(p = 1, 2, \dots, n)$ of A must be nonnegative. Additionally, A is PD iff $\Delta_{pp}(p = 1, 2, \dots, n)$ is strictly positive.

Theorem 11 Let \mathbf{x} be an $n \times 1$ column vector, $\mathbf{x} \neq \mathbf{0}$. Then $A = \mathbf{x}\mathbf{x}^T$ is PSD.

Theorem 12 If A is square, symmetric and positive definite, then A can be uniquely factorized as $A = U^T U$ where U is upper triangular with positive diagonal entries (called Cholesky decomposition).

Definition 10 Let $A = (a_{ij})_{n \times n}$ denote an $n \times n$ real matrix, the area defined by $|z - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}|$ is called the i th Gerschgorin circle of A .

Theorem 13 Let $A = (a_{ij})_{n \times n}$ denote an $n \times n$ real matrix. Then all of its eigenvalues are in the union of its n Gerschgorin circles (known as Gerschgorin theorem).

Let λ be an arbitrary eigenvalue of A , we have:

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1, \\ j \neq i}}^n |a_{ij}|, \quad \exists i \in \{1, 2, \dots, n\}.$$

References

- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8:338–353
- Saeedi J, Faez K (2013) A classification and fuzzy-based approach for digital multi-focus image fusion. *Pattern Anal Appl* 16(3):365–379
- Moustakidis SP, Theocharis JB (2012) A fast SVM-based wrapper feature selection method driven by a fuzzy complementary criterion. *Pattern Anal Appl* 15(4):365–379
- Atanassov KT (1986) Intuitionistic fuzzy-sets. *Fuzzy Set Syst* 20:87–96
- Xu Z, Yager RR (2008) Dynamic intuitionistic fuzzy multi-attribute decision making. *Int J Approx Reason* 48(1):246–262
- Burduk R (2012) Imprecise information in Bayes classifier. *Pattern Anal Appl* 15(2):147–153
- Xia M, Xu Z (2010) Some new similarity measures for intuitionistic fuzzy values and their application in group decision making. *J Syst Sci Syst Eng* 19(4):430–452
- Ye J (2011) Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Math Comput Model* 53(1–2):91–97
- Szmidt E, Kacprzyk J (2000) Distances between intuitionistic fuzzy sets. *Fuzzy Set Syst* 114(3):505–518
- Wang W, Xin X (2005) Distance measure between intuitionistic fuzzy sets. *Pattern Recogn Lett* 26(13):2063–2069
- Grzegorzewski P (2004) Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric. *Fuzzy Set Syst* 148(2):319–328
- Chen T-Y (2007) A note on distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric. *Fuzzy Set Syst* 158(22):2523–2525
- Hung WL, Yang MS (2004) Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. *Pattern Recogn Lett* 25(14):1603–1611

14. Hong DH, Kim C (1999) A note on similarity measures between vague sets and between elements. *Inf Sci* 115(1–4):83–96
15. Li F, Xu Z (2001) Similarity measures between vague sets. *J Softw* 12(6):922–927
16. Li Y, Chi Z, Yan D (2002) Similarity measures between vague sets and vague entropy. *J Comput Sci* 29(12):129–132
17. Li D, Cheng C (2002) New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions. *Pattern Recogn Lett* 23(1–3):221–225
18. Mitchell HB (2003) On the Dengfeng-Chuntian similarity measure and its application to pattern recognition. *Pattern Recogn Lett* 24(16):3101–3104
19. Liang Z, Shi P (2003) Similarity measures on intuitionistic fuzzy sets. *Pattern Recogn Lett* 24(15):2687–2693
20. Li Y, Olson DL, Qin Z (2007) Similarity measures between intuitionistic fuzzy (vague) sets: a comparative analysis. *Pattern Recogn Lett* 28(2):278–285
21. Chen SM (1995) Measures of similarity between vague sets. *Fuzzy Set Syst* 74(2):217–223
22. Hwang CM, Yang MS, Hung WL et al (2012) A similarity measure of intuitionistic fuzzy sets based on the Sugeno integral with its application to pattern recognition. *Inf Sci* 189:93–109
23. Zhang H, Yu L (2013) New distance measures between intuitionistic fuzzy sets and interval-valued fuzzy sets. *Inf Sci* 245:181–196
24. Li J, Deng G (2012) The relationship between similarity measure and entropy of intuitionistic fuzzy sets. *Inf Sci* 188:314–321
25. Boran FE, Akay D (2014) A biparametric similarity measure on intuitionistic fuzzy sets with applications to pattern recognition. *Inf Sci* 255:45–57
26. Papakostas GA, Hatzimichailidis AG, Kaburlasos VG (2013) Distance and similarity measures between intuitionistic fuzzy sets: a comparative analysis from a pattern recognition point of view. *Pattern Recogn Lett* 34(14):1609–1622
27. Hong DH, Choi C-H (2000) Multicriteria fuzzy decision-making problems based on vague set theory. *Fuzzy Set Syst* 114(1):103–113