A similarity measure of intuitionistic fuzzy sets based on the Sugeno integral with its application to pattern recognition

Chao-Ming Hwang, Miin-Shen Yang, Wen-Liang Hung, Ming-Gay Lee

A Department of Applied Mathematics, Chinese Culture University, Yangminshan, Taipei, Taiwan
b Department of Applied Mathematics, Chung Yuan Christian University, Chung-Li, Taiwan
c Department of Applied Mathematics, National Hsinchu University of Education, Hsinchu, Taiwan

Article info
Article history:
Received 2 December 2008
Received in revised form 12 April 2010
Accepted 15 November 2011
Available online 23 November 2011

Keywords:
Fuzzy set
Intuitionistic fuzzy set
Similarity measure
Sugeno integral
Robust clustering method

Abstract

Similarity measures of intuitionistic fuzzy sets are used to indicate the similarity degree between intuitionistic fuzzy sets. Although several similarity measures for intuitionistic fuzzy sets have been proposed in previous studies, no one has considered the use of the Sugeno integral to define them. Since the Sugeno integral provides an expected-value-like operation, it can be a useful tool in defining the expected total similarity degree between two intuitionistic fuzzy sets. In this paper, we propose a new similarity measure formula for intuitionistic fuzzy sets induced by the Sugeno integral. Some examples are illustrated to compare the proposed method with several existing methods. Numerical results show that the proposed similarity measure is more reasonable than those existing methods. On the other hand, measuring the similarity between intuitionistic fuzzy sets is also important in pattern recognition. Finally, the proposed similarity measure uses a robust clustering method to recognize the patterns of intuitionistic fuzzy sets.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Fuzzy sets, first introduced by Zadeh [45], provide an approach for treating fuzziness, which is a type of uncertainty different from randomness. In an ordinary set, an element of a universe either belongs to or does not belong to the set; that is, the membership value of an element in an ordinary set is 0 or 1. A fuzzy set is a generalization of an ordinary set, which allows the membership value of an element in a fuzzy set to range from 0 and 1. The idea of fuzzy sets effectively conveys partial degrees of belongingness described by membership functions.

A membership function in fuzzy sets assigns a number from the unit interval to each element in a universe of discourse to indicate the degree of belongingness to the set under consideration. The degree of nonbelongingness in fuzzy sets is simply the complement to 1 of the membership degree. However, humans who express the degree of membership of a given element in a fuzzy set very often do not express a corresponding degree of nonmembership as the complement to 1. Thus, Atanassov [1] introduced the concept of an intuitionistic fuzzy set (IFS), which is a generalization of a fuzzy set. Since an IFS can present the degrees of both membership and nonmembership with a degree of hesitancy, knowledge and semantic representation becomes more meaningful and applicable [1–3]. These intuitionistic fuzzy sets (IFSs) have been widely studied and applied in various areas such as logic programming and reasoning [4,25], decision-making problems [33,37–39], medical diagnosis [10], algebra [9,44], pattern recognition [22,24] and clustering [35,41,46].
Similarity measures are an important tool for determining the degree of similarity between two objects. Kaufman and Rousseuw [26] presented some examples to illustrate traditional similarity measure applications in hierarchical cluster analysis. Different similarity measures between fuzzy sets have been proposed in numerous studies [8,15,34,47]. Moreover, similarity measures of IFSs have also been widely studied. For example, Dengfeng and Chuntian [11] proposed similarity measures used in pattern recognition problems. In [28], similarity measures for IFSs were proposed and the relationships between these measures with applications to pattern recognition were discussed. Numerical comparisons were then used to show that these similarity measures were more reasonable than previous ones in [11]. Mitchell [29] interpreted IFSs as ensembles of ordered fuzzy sets from the statistical point of view to modify Dengfeng and Chuntian’s methods [11]. Hung and Yang [22] proposed several similarity measures of IFSs based on Hausdorff distance which can effectively be used with linguistic variables. More recently, Xu and Chen [40] gave a comprehensive overview of distance and similarity measures of IFSs where they also extended those existing measures based on a geometric distance model. In addition, Li et al. [27] gave a comparative analysis of several similarity measures for IFSs.

Although many similarity measures of IFSs have been proposed, no studies have considered using the Sugeno integral to define similarity measures of IFSs. Since the Sugeno integral provides an operation similar to “expected value”, it should be effective for defining expected total difference between two IFSs. It can then be transferred to the expected total similarity degree between the two IFSs. In this paper, a new similarity measure formula of IFSs induced by the Sugeno integral is proposed. Some examples are illustrated to make comparisons between the proposed similarity measure and several existing measures used in pattern recognition problems. In [28], similarity measures for IFSs were proposed and the relationships between these measures with applications to pattern recognition were discussed. Numerical comparisons were then used to show that these similarity measures were more reasonable than previous ones in [11]. Mitchell [29] interpreted IFSs as ensembles of ordered fuzzy sets from the statistical point of view to modify Dengfeng and Chuntian’s methods [11]. Hung and Yang [22] proposed several similarity measures of IFSs based on Hausdorff distance which can effectively be used with linguistic variables. More recently, Xu and Chen [40] gave a comprehensive overview of distance and similarity measures of IFSs where they also extended those existing measures based on a geometric distance model. In addition, Li et al. [27] gave a comparative analysis of several similarity measures for IFSs.

2. Intuitionistic fuzzy set and similarity measures

We first describe the aspects of intuitionistic fuzzy sets (IFSs) discussed by Atanassov [1].

**Definition 2.1** (Atanassov [1]). An intuitionistic fuzzy set (IFS) $\tilde{A}$ in $X$ is defined as $\tilde{A} = \{(x, \mu_{\tilde{A}}(x), \nu_{\tilde{A}}(x)) | x \in X\}$ where $\mu_{\tilde{A}}(x) : X \rightarrow [0, 1]$ and $\nu_{\tilde{A}}(x) : X \rightarrow [0, 1]$ with the condition $0 \leq \mu_{\tilde{A}}(x) + \nu_{\tilde{A}}(x) \leq 1$, $\forall x \in X$. The numbers $\mu_{\tilde{A}}(x)$ and $\nu_{\tilde{A}}(x)$ denote the degree of membership and non-membership of $x$ to $\tilde{A}$, respectively.

For each IFS $\tilde{A}$ in $X$, the number $\pi_{\tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x) - \nu_{\tilde{A}}(x)$ denotes a hesitation degree of $x$ in $\tilde{A}$, which is called the intuitionistic index $x$ to $\tilde{A}$. Obviously, $0 \leq \pi_{\tilde{A}}(x) \leq 1$, for each $x \in X$. Two other extended fuzzy sets have been proposed, including vague sets [16] and interval-valued fuzzy sets (IVFSs) [36]. However, Bustince and Burillo [7] showed that vague sets are intuitionistic fuzzy sets and Deschrijver and Kerre [12] showed that there is an isomorphism between IVFSs and IFSs. Moreover, Deschrijver and Kerre [13] showed the mathematical relationships between IFSs and other models of imprecision. We also note that Dubois et al. [14] pointed out a terminological clash between the “intuitionistic” in IFSs and what is currently understood as intuitionistic logic. However, Atanassov [5] answered Dubois et al. [14] about the relation between IFS and intuitionistic logic, and interested readers may refer to previous studies [5,14,18].

In this paper, we use $IFS(X)$ to denote the class of all IFSs of $X$. Next, we present some basic operations on IFSs which are needed in the following discussion.

**Definition 2.2.** If $\tilde{A}$ and $\tilde{B}$ are in $IFS(X)$, then

(i) $\tilde{A} \subseteq \tilde{B}$ if and only if $\forall x \in X$, $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$ and $\nu_{\tilde{A}}(x) \geq \nu_{\tilde{B}}(x)$;

(ii) $\tilde{A} = \tilde{B}$ if and only if $\forall x \in X$, $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$ and $\nu_{\tilde{A}}(x) = \nu_{\tilde{B}}(x)$.

Before we consider similarity measures between IFSs, let us review some correlations of IFSs proposed in previous studies. We know that the correlation coefficient is an important concept in statistics. Correlation analysis can be used to study relations between (random) variables. Murthy et al. [31] initially proposed a measure of correlation between two membership functions for fuzzy sets. Gerstekorn and Manko [17] then considered correlations of IFSs in a finite space. Hong and Hwang [20] defined the correlations of IFSs in a probability space. Bustince and Burillo [6] extended the concept of correlations of IFSs to interval-valued IFSs. Hong [19] then generalized correlations of IFSs to interval-valued IFSs in a general probability space. Hung and Wu [21] provided a method to calculate correlations of IFSs and also interval-valued IFSs by means of a centroid. Mitchell [30] defined the correlations of IFSs by interpreting IFSs as the ensembles of ordinary membership functions. Recently, in [41,46], the authors proposed clustering methods for IFSs based on the correlation matrix of IFSs. As a whole, these correlation calculations of IFSs are defined by considering IFSs as statistical data from a statistical viewpoint. On the other hand, to present a similarity degree between them, similarity measures between fuzzy sets or IFSs are always based on their data types. Similarity measures are conceptually quite different from correlations. They can be seen as the
result distance measures. Measuring the similarity between IFSs has become important in pattern recognition research. Some methods have previously been advanced to calculate the degree of similarity between IFSs [11,28]. In this regard, Dengfeng and Chuntian [11] introduced the following definition.

**Definition 2.3.** A mapping \( S : IFS(X) \times IFS(X) \rightarrow [0, 1] \) is said to be the degree of similarity between \( \tilde{A} \) and \( \tilde{B} \) in \( IFS(X) \) if \( S(\tilde{A}, \tilde{B}) \) satisfies the following properties (P1-P4):

(P1) \( 0 \leq S(\tilde{A}, \tilde{B}) \leq 1 \);
(P2) \( S(\tilde{A}, \tilde{B}) = 1 \) if and only if \( \tilde{A} = \tilde{B} \);
(P3) \( S(\tilde{A}, \tilde{B}) = S(\tilde{B}, \tilde{A}) \);
(P4) \( S(\tilde{A}, \tilde{C}) \leq S(\tilde{A}, \tilde{B}) \) and \( S(\tilde{A}, \tilde{C}) \leq S(\tilde{B}, \tilde{C}) \) if \( \tilde{A} \subseteq \tilde{B} \subseteq \tilde{C}, \tilde{C} \in IFS(X) \).

Let \( X = \{x_1, \ldots, x_n\} \) be a discrete set of universe. Consider two IFSs \( \tilde{A} \) and \( \tilde{B} \) in \( IFS(X) \). Dengfeng and Chuntian [11] proposed a similarity measure between \( \tilde{A} \) and \( \tilde{B} \) as follows:

\[
S_\delta(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} |m^\mu_A(i) - m^\mu_B(i)|^p \right)^{1/p},
\]

where \( m^\mu_A(i) = \frac{\mu_A(x_i) + 1 - \nu_A(x_i)}{2} \), \( m^\mu_B(i) = \frac{\mu_B(x_i) + 1 - \nu_B(x_i)}{2} \) and \( 1 \leq p \leq \infty \). Liang and Shi [28] proposed a similarity measure between \( \tilde{A} \) and \( \tilde{B} \) as follows:

\[
S_\phi(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} \left( \phi_{\text{min}}(i) + \phi_{\text{max}}(i) \right)^p \right)^{1/p},
\]

where \( \phi_{\text{min}}(i) = \frac{m^\mu_A(i) - m^\mu_B(i)}{2} \) and \( \phi_{\text{max}}(i) = \frac{m^\mu_A(i) + m^\mu_B(i)}{2} \) with \( m^\mu_A(i) = \frac{\mu_A(x_i) + \nu_A(x_i)}{2}, m^\mu_B(i) = \frac{\mu_B(x_i) + \nu_B(x_i)}{2} \) and \( 1 \leq p \leq \infty \). To obtain more information on IFSs, Liang and Shi [28] gave another similarity measure as follows:

\[
S_{\delta}(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} \left( \frac{3}{2} \omega \phi_{\text{min}}(i) \right)^p \right)^{1/p},
\]

where \( \phi_{\text{min}}(i) = \phi_{\text{max}}(i) = \frac{m^\mu_A(i) - m^\mu_B(i)}{2} \) and \( \phi_{\text{max}}(i) = \frac{m^\mu_A(i) + m^\mu_B(i)}{2} \) with \( \phi_{\text{min}}(i) = \frac{\mu_A(x_i) + 1 - \nu_A(x_i)}{2}, m^\mu_B(i) = \frac{\mu_B(x_i) + 1 - \nu_B(x_i)}{2} \). Liang and Shi [28] also gave another similarity measure as follows:

\[
S_{\phi}(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} \frac{3}{2} \omega \phi_{\text{min}}(i) \right)^p,
\]

where \( \phi_{\text{min}}(i) = \phi_{\text{max}}(i) = \frac{m^\mu_A(i) - m^\mu_B(i)}{2} \) and \( \phi_{\text{max}}(i) = \frac{m^\mu_A(i) + m^\mu_B(i)}{2} \) with \( \phi_{\text{min}}(i) = \frac{\mu_A(x_i) + 1 - \nu_A(x_i)}{2}, m^\mu_B(i) = \frac{\mu_B(x_i) + 1 - \nu_B(x_i)}{2} \). Mitchell [29] interpreted IFSs as ensembles of ordered fuzzy sets from a statistical viewpoint to modify Dengfeng and Chuntian’s methods as follows:

\[
S_{\phi}(\tilde{A}, \tilde{B}) = \frac{1}{2} (\rho_\mu(\tilde{A}, \tilde{B}) + \rho_\nu(\tilde{A}, \tilde{B})),
\]

where

\[
\rho_\mu(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|^p \right)^{1/p}
\]

and

\[
\rho_\nu(\tilde{A}, \tilde{B}) = 1 - \frac{1}{\sqrt{n}} \left( \sum_{i=1}^{n} |\nu_A(x_i) - \nu_B(x_i)|^p \right)^{1/p}.
\]

Hung and Yang [22] proposed several similarity measures of IFSs based on Hausdorff distance as follows. They first defined two intervals \( l^\mu_A(x_i) = [\mu_A(x_i), 1 - \nu_A(x_i)] \) and \( l^\nu_A(x_i) = [\nu_A(x_i), 1 - \mu_A(x_i)] \), \( i = 1, \ldots, n \). The Hausdorff distance \( H(l^\mu_A(x_i), l^\mu_B(x_i)) \) between \( l^\mu_A(x_i) \) and \( l^\mu_B(x_i) \) was then defined as follows:
H(I_\bar{A}(x_i), I_\bar{B}(x_i)) = \max\{|\mu_\bar{A}(x_i) - \mu_\bar{B}(x_i)|, |\nu_\bar{A}(x_i) - \nu_\bar{B}(x_i)|\}.

Thus, they define the distance \(d_H(\bar{A}, \bar{B})\) between \(\bar{A}\) and \(\bar{B}\) with

\[d_H(\bar{A}, \bar{B}) = \frac{1}{n} \sum_{i=1}^{n} H(I_\bar{A}(x_i), I_\bar{B}(x_i)).\]

Based on this distance, they proposed the following three similarity measures:

\[
S_1(\bar{A}, \bar{B}) = 1 - d_H(\bar{A}, \bar{B}),
\]

\[
S_2(\bar{A}, \bar{B}) = \frac{\exp(-d_H(\bar{A}, \bar{B})) - \exp(-1)}{1 - \exp(-1)},
\]

\[
S_3(\bar{A}, \bar{B}) = \frac{1 - d_H(\bar{A}, \bar{B})}{1 + d_H(\bar{A}, \bar{B})}.
\]

Recently, Hung and Yang \cite{23} proposed two exponential-type similarity measures between IFSs, the first one being

\[
S_{n1}(\bar{A}, \bar{B}) = 1 - \frac{1 - \exp\left(\sum_{i=1}^{n} (|\mu_\bar{A}(x_i) - \mu_\bar{B}(x_i)| + |\nu_\bar{A}(x_i) - \nu_\bar{B}(x_i)|)/2\right)}{1 - \exp(-n)}
\]

and second one

\[
S_{n2}(\bar{A}, \bar{B}) = 1 - \frac{1 - \exp\left(\sum_{i=1}^{n} (|\sqrt{\mu_\bar{A}(x_i)} - \sqrt{\mu_\bar{B}(x_i)}| + |\sqrt{\nu_\bar{A}(x_i)} - \sqrt{\nu_\bar{B}(x_i)}|)/2\right)}{1 - \exp(-n)}.
\]

3. Similarity measure of IFSs based on the Sugeno integral

In this section, we first review fuzzy measures and the Sugeno integral. Let \(X\) be a nonempty set and \(\mathcal{F}\) be a \(\sigma\)-field of subsets of \(X\). Let \(m : \mathcal{F} \to [0, 1]\) be a nonnegative and real-valued set function defined by \(\mathcal{F}\).

**Definition 3.1.** The measure \(m\) is called a fuzzy measure on \((X, \mathcal{F})\) if it satisfies the following conditions:

(FM1) \(m(\emptyset) = 0\) (vanishing at \(\emptyset\)).

(FM2) \(E, F \in \mathcal{F}\) and \(E \subseteq F\) imply \(m(E) \leq m(F)\) (monotonicity).

(FM3) \(\{E_i\} \in \mathcal{F}, E_1 \subseteq E_2 \subseteq \cdots\), and \(\cup_{i=1}^{\infty} E_i \in \mathcal{F}\) imply

\(\lim m(E_i) = m(\bigcup_{i=1}^{\infty} E_i)\) (continuity from below).

(FM4) \(\{E_i\} \in \mathcal{F}, E_1 \supset E_2 \supset \cdots\), and \(\cap_{i=1}^{\infty} E_i \in \mathcal{F}\) imply

\(\lim m(E_i) = m(\bigcap_{i=1}^{\infty} E_i)\) (continuity from above).

On the other hand, \(m\) is called a lower or upper semi-continuous fuzzy measure if it satisfies the above conditions of either (FM1), (FM2), and (FM3) or (FM1), (FM2), and (FM4), respectively. Both of these are called semi-continuous fuzzy measures. Furthermore, we say that a fuzzy measure or semi-continuous fuzzy measure \(m\) is regular if and only if \(X \in \mathcal{F}\) and \(m(X) = 1\). Usually, we consider a monotone class, semi-ring, ring, algebra, \(\sigma\)-ring, \(\sigma\)-algebra, plump class, or power set as the class \(\mathcal{F}\) on which \(m\) is defined. We call \((X, \mathcal{F}, m)\) a fuzzy measure space (or semi-continuous fuzzy measure space) if \(m\) is a fuzzy measure (or semi-continuous fuzzy measure) on a measurable space \((X, \mathcal{F})\). On a semi-ring, in comparison with the classical measure, the fuzzy measure (or semi-continuous fuzzy measure) abandons additivity, but reserves monotonicity, continuity (or partial continuity), and vanishing on the empty set.

Let \(\mathcal{F}\) be the class of all finite nonnegative measurable functions defined in \((X, \mathcal{F})\). For any given \(f \in \mathcal{F}\), we write \(f_\geq \{x|f(x) \geq x\}\), \(f_\leq \{x|f(x) \leq x\}\), where \(x \in [0, \infty]\). The sets \(f_\geq\) and \(f_\leq\) are called \(\geq\)-cut and \(\leq\)-cut sets, respectively. For the sake of simplicity, the range of the function is \([0, 1]\).

**Definition 3.2.** Let \(A \in \mathcal{F}\) and \(f \in \mathcal{F}\). A fuzzy integral of \(f\) on \(A\) with respect to \(m\), denoted by \(\int_A f dm\), is defined by

\[
\int_A f dm = \sup_{x \in [0, 1]} [x \wedge m(A \cap f_\geq)].
\]

When \(A = X\), the fuzzy integral of \(f\) with respect to \(m\), denoted by \(\int f dm\), is defined by

\[
\int f dm = \sup_{x \in [0, 1]} [x \wedge m(f_\geq)].
\]
Sometimes, the fuzzy integral is also called the Sugeno integral, which provides an “expected value” like operation. In addition, with respect to fuzzy measures, the Choquet integral is another popular type of integral. The Choquet integral of $f$ with respect to fuzzy measure $m$ is defined by

\[
(C) \int f \, dm = \int_{0}^{1} m(f_x) \, dx.
\]

For a finite set $X$, we denote the value of function $f$ at point $x_i$ in $X$ by $f_i$. If the values of the function $f$ satisfy $f_1 < f_2 < \cdots < f_n$, then the Choquet integral of $f$ with respect to $m$ will be

\[
(C) \int f \, dm = \sum_{i=1}^{n} f_i (m(A_i) - m(A_{i-1}))
\]

and the Sugeno integral of $f$ with respect to $m$ will be

\[
\int f \, dm = \vee_{i=1}^{n} (f_i \land m(A_i))
\]

where $A_i = \{x_1, x_2, \ldots, x_i\}$, $i = 1, \ldots, n$ and $A_{n+1} = \phi$. We can see the parallelism between the Choquet and Sugeno integrals in that both integrals fit the same formal model, and they only differ in the operators used. The Choquet integral uses the sum and product, while the Sugeno integral uses the maximum and minimum. Narukawa and Torra [32] recently considered the Choquet integral for IFSs, and defined a non-monotonic fuzzy measure induced by an IFS. Based on the Choquet integral with respect to non-monotonic fuzzy measures, Narukawa and Torra [32] proposed a weighted distance between two IFSs. They then showed that, under some conditions, their proposed weighted distance coincided with the normalized Hamming distance between two IFSs. In this paper, we use a different approach from that of Narukawa and Torra [32], and propose a new similarity measure between two IFSs based on the Sugeno integral with respect to fuzzy measures.

For any two IFSs $\tilde{A}$ and $\tilde{B}$ in IFS$(X)$, we define

\[
\tilde{C}_\mu(\tilde{A}, \tilde{B}) = \{(x, |\mu_\tilde{A}(x) - \mu_\tilde{B}(x)|) | x \in X\} \quad \text{and} \quad \tilde{C}_\nu(\tilde{A}, \tilde{B}) = \{(x, |\nu_\tilde{A}(x) - \nu_\tilde{B}(x)|) | x \in X\}.
\]

We now define two Sugeno integrals $E(\tilde{C}_\mu(\tilde{A}, \tilde{B}))$ and $E(\tilde{C}_\nu(\tilde{A}, \tilde{B}))$ as follows:

\[
E(\tilde{C}_\mu(\tilde{A}, \tilde{B})) = \frac{1}{n} \sum_{i=1}^{n} K_{x_i} \int_{C_{\mu_{\tilde{A}}}(\tilde{A}, \tilde{B})} \tilde{C}_\mu(\tilde{A}, \tilde{B}) \, dm,
\]

where $K_{x_i}$ is the number of $x \in C_{\mu_{\tilde{A}}}(\tilde{A}, \tilde{B})$ such that $\mu_{\mu_{\tilde{A}}}(x) = x_i$, and

\[
E(\tilde{C}_\nu(\tilde{A}, \tilde{B})) = \frac{1}{n} \sum_{i=1}^{n} K_{x_i} \int_{C_{\mu_{\tilde{A}}}(\tilde{A}, \tilde{B})} \tilde{C}_\nu(\tilde{A}, \tilde{B}) \, dm,
\]

where $K_{x_i}$ is the number of $x \in C_{\nu_{\tilde{A}}}(\tilde{A}, \tilde{B})$ such that $\mu_{\nu_{\tilde{A}}}(x) = x_i$. Intuitively, we can see the defined Sugeno integral $E(\tilde{C}_\mu(\tilde{A}, \tilde{B}))$ as the expected total difference between $\tilde{A}$ and $\tilde{B}$ based on their degrees of memberships $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$. On the other hand, the defined Sugeno integral $E(\tilde{C}_\nu(\tilde{A}, \tilde{B}))$ can be seen as the expected total difference between $\tilde{A}$ and $\tilde{B}$ based on their degrees of nonmemberships $\nu_{\tilde{A}}(x)$ and $\nu_{\tilde{B}}(x)$.

**Proposition 3.1.** The defined Sugeno integrals $E(\tilde{C}_\mu(\tilde{A}, \tilde{B}))$ and $E(\tilde{C}_\nu(\tilde{A}, \tilde{B}))$ satisfy the following properties:

\begin{align*}
(D1) & \quad 0 \leq E(\tilde{C}_\mu(\tilde{A}, \tilde{B})) < \infty \quad \text{and} \quad 0 \leq E(\tilde{C}_\nu(\tilde{A}, \tilde{B})) < \infty. \\
(D2) & \quad E(\tilde{C}_\mu(\tilde{A}, \tilde{B})) = 0 \quad \text{(or} \quad E(\tilde{C}_\nu(\tilde{A}, \tilde{B})) = 0 \text{)} \quad \text{if and only if} \quad \tilde{A} = \tilde{B}. \\
(D3) & \quad E(\tilde{C}_\mu(\tilde{A}, \tilde{B})) = E(\tilde{C}_\mu(\tilde{B}, \tilde{A})) \quad \text{and} \quad E(\tilde{C}_\nu(\tilde{A}, \tilde{B})) = E(\tilde{C}_\nu(\tilde{B}, \tilde{A})). \\
(D4) & \quad \text{Let} \quad A, B \quad \text{and} \quad C \quad \text{in} \quad \text{IFS}(X). \quad \text{If} \quad A \subseteq B \subseteq C \quad \text{then} \\
& \quad E(\tilde{C}_\mu(\tilde{A}, \tilde{B})) \leq E(\tilde{C}_\mu(\tilde{A}, \tilde{C})), \quad E(\tilde{C}_\mu(\tilde{B}, \tilde{C})) \leq E(\tilde{C}_\mu(\tilde{B}, \tilde{A})); \\
& \quad E(\tilde{C}_\nu(\tilde{A}, \tilde{B})) \leq E(\tilde{C}_\nu(\tilde{A}, \tilde{C})), \quad E(\tilde{C}_\nu(\tilde{B}, \tilde{C})) \leq E(\tilde{C}_\nu(\tilde{B}, \tilde{A})).
\end{align*}

**Proof.** It is easy to prove the properties (D1), (D2) and (D3). We therefore only prove D (4). Since $\tilde{A} \subseteq \tilde{B} \subseteq \tilde{C}$, $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x) \leq \mu_{\tilde{C}}(x)$ and $\nu_{\tilde{A}}(x) \geq \nu_{\tilde{B}}(x) \geq \nu_{\tilde{C}}(x)$, $\forall x \in X$. These imply that

\[
|\mu_{\tilde{A}}(x) - \mu_{\tilde{B}}(x)| \leq |\mu_{\tilde{A}}(x) - \mu_{\tilde{C}}(x)|, \quad \forall x \in X
\]

and
\[ |v_A^-(x) - v_B^-(x)| \leq |v_A^-(x) - v_C^-(x)|, \quad \forall x \in X. \]

Hence, we have

\[ \tilde{C}_\mu(A, B) \subseteq \tilde{C}_\mu(A, \tilde{C}) \quad \text{and} \quad \tilde{C}_\nu(A, B) \subseteq \tilde{C}_\nu(A, \tilde{C}). \]

We then obtain

\[ E(\tilde{C}_\mu(A, B)) = \frac{1}{n} \sum_{i=1}^{n} K_x \int_{\tilde{C}_\mu(A, B)} \tilde{C}_\mu(A, B) dm = \frac{1}{n} \sum_{i=1}^{n} K_x \sup_{\beta \in [0, 1]} \left[ \beta \wedge m(\tilde{C}_\mu, \tilde{C}_\mu) \cap \tilde{C}_\mu(\tilde{C}_\mu, B) \right] \]

\[ \leq \frac{1}{n} \sum_{i=1}^{n} K_x \sup_{\beta \in [0, 1]} \left[ \beta \wedge m(\tilde{C}_\mu, \tilde{C}_\mu) \cap \tilde{C}_\mu(A, C) \cap \tilde{C}_\mu(A, B) \right] = \frac{1}{n} \sum_{i=1}^{n} K_x \int_{\tilde{C}_\mu(A, C)} \tilde{C}_\mu(A, C) dm = E(\tilde{C}_\mu(A, \tilde{C})). \]

Similarly, we can also obtain \( E(\tilde{C}_\mu(B, \tilde{C})) \leq E(\tilde{C}_\mu(A, \tilde{C})), E(\tilde{C}_\nu(A, B)) \leq E(\tilde{C}_\nu(A, \tilde{C})), \text{and } E(\tilde{C}_\mu(B, \tilde{C})) \leq E(\tilde{C}_\mu(A, \tilde{C})). \]

Based on the defined Sugeno integrals \( E(\tilde{C}_\mu(A, B)) \) and \( E(\tilde{C}_\nu(A, B)) \), we propose a similarity measure of IFSs. For any \( \tilde{A} \) and \( \tilde{B} \) in \( IFS(X) \), we define the similarity measure \( S_{new}(\tilde{A}, \tilde{B}) \) as follows:

\[ S_{new}(\tilde{A}, \tilde{B}) = \exp \left( -\frac{1}{2} \left[ E(\tilde{C}_\mu(A, B)) + E(\tilde{C}_\nu(A, B)) \right] \right). \]  

(11)

It is easy to see that the proposed similarity measure \( S_{new}(\tilde{A}, \tilde{B}) \) satisfies the properties (P1)–(P4) of a similarity measure.

4. Numerical examples and comparisons

As follows, we give several examples to compare the proposed similarity with some existing similarity measures [11,22,23,28,29]. For convenience, we consider \( p = 1 \) and \( \omega_n = 1/3, i = 1, 2, 3 \) in similarity measures \( S_0, S_1, S_2, S_3 \) and \( S_{mod} \).

**Example 1.** Assume that there are three IFS patterns in \( X = \{x_1, x_2, x_3\} \). The three patterns are denoted as follows:

\( \tilde{A}_1 = \{(x_1, 0.3, 0.3), (x_2, 0.2, 0.2), (x_3, 0.1, 0.1)\}; \)

\( \tilde{A}_2 = \{(x_1, 0.2, 0.2), (x_2, 0.2, 0.2), (x_3, 0.2, 0.2)\}; \)

\( \tilde{A}_3 = \{(x_1, 0.4, 0.4), (x_2, 0.4, 0.4), (x_3, 0.4, 0.4)\}. \)

Assume that a sample \( \tilde{B} = \{(x_1, 0.3, 0.3), (x_2, 0.2, 0.2), (x_3, 0.1, 0.1)\} \) is given. By Eqs. (1)–(8), we have

\[ S_0'(\tilde{A}_1, \tilde{B}) = S_0'(\tilde{A}_2, \tilde{B}) = S_0'(\tilde{A}_3, \tilde{B}) = 1; \]

\[ S_0'(\tilde{A}_1, \tilde{B}) = 1, \quad S_0'(\tilde{A}_2, \tilde{B}) = 0.933, \quad S_0'(\tilde{A}_3, \tilde{B}) = 0.800; \]

\[ S_0'(\tilde{A}_1, \tilde{B}) = 1, \quad S_0'(\tilde{A}_2, \tilde{B}) = 0.967, \quad S_0'(\tilde{A}_3, \tilde{B}) = 0.900; \]

\[ S_0'(\tilde{A}_1, \tilde{B}) = 1, \quad S_0'(\tilde{A}_2, \tilde{B}) = 0.956, \quad S_0'(\tilde{A}_3, \tilde{B}) = 0.867; \]

\[ S_{mod}'(\tilde{A}_1, \tilde{B}) = 1, \quad S_{mod}'(\tilde{A}_2, \tilde{B}) = 0.933, \quad S_{mod}'(\tilde{A}_3, \tilde{B}) = 0.800; \]

\[ S_1(\tilde{A}_1, \tilde{B}) = 1, \quad S_1(\tilde{A}_2, \tilde{B}) = 0.933, \quad S_1(\tilde{A}_3, \tilde{B}) = 0.800; \]

\[ S_1(\tilde{A}_1, \tilde{B}) = 1, \quad S_1(\tilde{A}_2, \tilde{B}) = 0.933, \quad S_1(\tilde{A}_3, \tilde{B}) = 0.713; \]

\[ S_1(\tilde{A}_1, \tilde{B}) = 1, \quad S_1(\tilde{A}_2, \tilde{B}) = 0.875, \quad S_1(\tilde{A}_3, \tilde{B}) = 0.667. \]

On the other hand, we obtain

\[ \tilde{C}_{1,1} = \{(x_1, 0), (x_2, 0), (x_3, 0)\}; \quad \tilde{C}_{1,1} = \{(x_1, 0), (x_2, 0), (x_3, 0)\}; \]

\[ \tilde{C}_{2,2} = \{(x_1, 0.1), (x_2, 0), (x_3, 0.1)\}; \quad \tilde{C}_{2,2} = \{(x_1, 0.1), (x_2, 0), (x_3, 0.1)\}; \]

\[ \tilde{C}_{3,3} = \{(x_1, 0.1), (x_2, 0.2), (x_3, 0.3)\}; \quad \tilde{C}_{3,3} = \{(x_1, 0.1), (x_2, 0.2), (x_3, 0.3)\}. \]

Thus, we have the proposed similarity measure \( S_{new} \) by Eq. (11) with

\[ S_{new}(\tilde{A}_1, \tilde{B}) = 1, \quad S_{new}(\tilde{A}_2, \tilde{B}) = 0.936, \quad S_{new}(\tilde{A}_3, \tilde{B}) = 0.741. \]

Obviously, the data indicate that \( \tilde{A}_1 = \tilde{B} \). This means that sample \( \tilde{B} \) matches pattern \( \tilde{A}_1 \). The proposed similarity measures together with those of Liang and Shi [28], Mitchell [29] and Hung and Yang [22] show the correct classification according to the principle of maximum membership degree. However, Dengfeng and Chuntian’s [11] similarity measure \( S_0 \) does not classify this sample well because \( S_0'(\tilde{A}_1, \tilde{B}) = S_0'(\tilde{A}_2, \tilde{B}) = S_0'(\tilde{A}_3, \tilde{B}) = 1. \)
Example 2. Assume that there are two IFS patterns in $X = \{x_1, x_2, x_3\}$. The two patterns are denoted as follows:

\[ \tilde{A}_1 = \{(x_1, 0.2, 0.2), (x_2, 0.2, 0.2), (x_3, 0.2, 0.2)\}; \]
\[ \tilde{A}_2 = \{(x_1, 0.4, 0.4), (x_2, 0.4, 0.4), (x_3, 0.4, 0.4)\}. \]

Assume that a sample $\tilde{B} = \{(x_1, 0.3, 0.3), (x_2, 0.3, 0.3), (x_3, 0.1, 0.3)\}$ is given. By Eqs. (1)–(8), we have

\[ S^p_{\tilde{A}_1, \tilde{B}} = S^p_{\tilde{A}_2, \tilde{B}} = 0.967; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.900, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.867; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.933, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.900; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.900, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.867; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.900, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.833; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.849, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.757; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.818, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.714. \]

We obtain that

\[ \tilde{C}_{1, \mu} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.1)\}, \quad \tilde{C}_{1, v} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.1)\}; \]
\[ \tilde{C}_{2, \mu} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.3)\}, \quad \tilde{C}_{2, v} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.1)\}. \]

Thus, we have the proposed similarity measure $S_{\text{new}}$ by Eq. (11) with

\[ S_{\text{new}}(\tilde{A}_1, \tilde{B}) = 0.905, \quad S_{\text{new}}(\tilde{A}_2, \tilde{B}) = 0.819. \]

The data obviously show that sample $\tilde{B}$ is closer to $\tilde{A}_1$ than to $\tilde{A}_2$. According to the principle of the maximum degree of similarity between IFSs, sample $\tilde{B}$ matches pattern $\tilde{A}_1$ in the above calculation of similarity measures. However, we find that the similarity measures $S^p_{\tilde{A}}$ [11, 28] are not suitable to classify this sample.

Example 3. Assume that there are three IFS patterns in $X = \{x_1, x_2, x_3\}$. The three patterns are denoted as follows:

\[ \tilde{A}_1 = \{(x_1, 0.1, 0.1), (x_2, 0.5, 0.1), (x_3, 0.1, 0.9)\}; \]
\[ \tilde{A}_2 = \{(x_1, 0.5, 0.5), (x_2, 0.7, 0.3), (x_3, 0.0, 0.8)\}; \]
\[ \tilde{A}_3 = \{(x_1, 0.7, 0.2), (x_2, 0.1, 0.8), (x_3, 0.4, 0.4)\}. \]

Also assume that a sample $\tilde{B} = \{(x_1, 0.4, 0.4), (x_2, 0.6, 0.2), (x_3, 0.0, 0.8)\}$ is given. By Eqs. (1)–(8), we have

\[ S^p_{\tilde{A}_1, \tilde{B}} = S^p_{\tilde{A}_2, \tilde{B}} = 1, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.600; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.833, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.933, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.600; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.917, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.967, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.600; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.899, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.956, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.722; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.833, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.933, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.600; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.833, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.933, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.567; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.757, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.898, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.444; \]
\[ S^p_{\tilde{A}_1, \tilde{B}} = 0.714, \quad S^p_{\tilde{A}_2, \tilde{B}} = 0.875, \quad S^p_{\tilde{A}_3, \tilde{B}} = 0.395. \]

According to

\[ \tilde{C}_{1, \mu} = \{(x_1, 0.3), (x_2, 0.1), (x_3, 0.1)\}, \quad \tilde{C}_{1, v} = \{(x_1, 0.3), (x_2, 0.1), (x_3, 0.1)\}, \]
\[ \tilde{C}_{2, \mu} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.1)\}, \quad \tilde{C}_{2, v} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.1)\}, \]
\[ \tilde{C}_{3, \mu} = \{(x_1, 0.3), (x_2, 0.5), (x_3, 0.4)\}, \quad \tilde{C}_{3, v} = \{(x_1, 0.2), (x_2, 0.6), (x_3, 0.4)\}, \]

we have the proposed similarity measure $S_{\text{new}}$ by Eq. (11) with

\[ S_{\text{new}}(\tilde{A}_1, \tilde{B}) = 0.741, \quad S_{\text{new}}(\tilde{A}_2, \tilde{B}) = 0.936, \quad S_{\text{new}}(\tilde{A}_3, \tilde{B}) = 0.632. \]
The data show that sample $\tilde{B}$ is closer to $\tilde{A}_2$ than to $\tilde{A}_1$ and to $\tilde{A}_3$. From the above calculation of similarity measures, sample $\tilde{B}$ matches pattern $A_2$ according to the principle of maximum member degree of similarity between IFSs. However, the similarity measures $S^c_i$ [11] cannot be used effectively to classify this sample.

**Example 4.** Assume that there are two IFS patterns in $X = \{x_1, x_2, x_3\}$. The two patterns are denoted as follows:

$\tilde{A}_1 = \{(x_1, 0.1, 0.4), (x_2, 0.4, 0.3), (x_3, 0.3, 0.1)\}$;

$\tilde{A}_2 = \{(x_1, 0.3, 0.4), (x_2, 0.3, 0.4), (x_3, 0.1, 0.1)\}$.

Assume that a sample $\tilde{B} = \{(x_1, 0.2, 0.2), (x_2, 0.2, 0.2), (x_3, 0.2, 0.2)\}$ is given. Hence

$\tilde{C}_{1,1} = \{(x_1, 0.1), (x_2, 0.2), (x_3, 0.1)\}$, $\tilde{C}_{1,2} = \{(x_1, 0.2), (x_2, 0.1), (x_3, 0.1)\}$;

$\tilde{C}_{2,1} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.1)\}$, $\tilde{C}_{2,2} = \{(x_1, 0.2), (x_2, 0.2), (x_3, 0.1)\}$.

By Eq. (11), we have

$S_{new}(\tilde{A}_1, \tilde{B}) = 0.818$, $S_{new}(\tilde{A}_2, \tilde{B}) = 0.861$.

But,

$S^c_i(\tilde{A}_1, \tilde{B}) = S^c_i(\tilde{A}_2, \tilde{B}) = 0.867$;

$S_{i}(\tilde{A}_1, \tilde{B}) = S_{i}(\tilde{A}_2, \tilde{B}) = 0.833$;

$S_{hy1}(\tilde{A}_1, \tilde{B}) = S_{hy1}(\tilde{A}_2, \tilde{B}) = 0.653$;

$S_{hy2}(\tilde{A}_1, \tilde{B}) = S_{hy2}(\tilde{A}_2, \tilde{B}) = 0.641$.

The data show that the sample $\tilde{B}$ should be closer to $\tilde{A}_2$ than to $\tilde{A}_1$. These results illustrate that the similarity measures $S^c_i$ [28], $S_i$, $S_{hy1}$ and $S_{hy2}$ [23] cannot be used effectively to classify this sample. However, the proposed measure $S_{new}$ can correctly indicate that sample $\tilde{B}$ matches pattern $A_2$ according to the principle of maximum member degree of similarity between IFSs.

**Example 5.** Assume that there are two IFS patterns in $X = \{x_1, x_2, x_3\}$. The two patterns are denoted as follows:

$\tilde{A}_1 = \{(x_1, 0.3, 0.3), (x_2, 0.4, 0.1), (x_3, 0.4, 0.4)\}$;

$\tilde{A}_2 = \{(x_1, 0.4, 0.4), (x_2, 0.4, 0.2), (x_3, 0.4, 0.2)\}$.

Assume that a sample $\tilde{B} = \{(x_1, 0.2, 0.2), (x_2, 0.2, 0.2), (x_3, 0.2, 0.2)\}$ is given. Let us consider Mitchell’s [29] similarity measure $S^p_{mod}$ for any $p \geq 1$ with its calculations as follows:

$S^p_{mod}(\tilde{A}_1, \tilde{B}) = \frac{1}{2} (\rho_{\mu}(\tilde{A}_1, \tilde{B}) + \rho_{\nu}(\tilde{A}_1, \tilde{B}))$,

where

$\rho_{\mu}(\tilde{A}_1, \tilde{B}) = 1 - \frac{1}{\sqrt[3]{3}} \sqrt[3]{(0.1)^p + (0.2)^p + (0.2)^p}$

and

$\rho_{\nu}(\tilde{A}_1, \tilde{B}) = 1 - \frac{1}{\sqrt[3]{3}} \sqrt[3]{(0.1)^p + (0.1)^p + (0.2)^p}$.

Then we have

$S^p_{mod}(\tilde{A}_1, \tilde{B}) = 1 - \frac{1}{2 \times \sqrt[3]{3}} \left( \sqrt[3]{1 + (2)^{p-1}} + \sqrt[3]{2 + (2)^p} \right) \times 0.1$.

Similarly, we can find

$S^p_{mod}(\tilde{A}_2, \tilde{B}) = 1 - \frac{1}{2 \times \sqrt[3]{3}} \left( \sqrt[3]{3(2)^p + 2} \right) \times 0.1$.

Thus, we have $S^p_{mod}(\tilde{A}_1, \tilde{B}) > S^p_{mod}(\tilde{A}_2, \tilde{B})$. On the other hand, we can find that

$\tilde{C}_{1,1} = \{(x_1, 0.1), (x_2, 0.2), (x_3, 0.2)\}$, $\tilde{C}_{1,2} = \{(x_1, 0.1), (x_2, 0.1), (x_3, 0.2)\}$;

$\tilde{C}_{2,1} = \{(x_1, 0.2), (x_2, 0.2), (x_3, 0.2)\}$, $\tilde{C}_{2,2} = \{(x_1, 0.2), (x_2, 0), (x_3, 0)\}$.
By Eq. (11), we have
\[ S_{\text{new}}(\tilde{A}_1, \tilde{B}) = \exp(-0.9/6), \quad S_{\text{new}}(\tilde{A}_2, \tilde{B}) = \exp(-0.8/6). \]

Thus, we have \( S_{\text{new}}(\tilde{A}_2, \tilde{B}) > S_{\text{new}}(\tilde{A}_1, \tilde{B}) \). The data obviously show that the sample \( \tilde{B} \) should be closer to \( \tilde{A}_2 \) than to \( \tilde{A}_1 \). These results illustrate that the proposed similarity measure \( S_{\text{new}} \) can correctly indicate that sample \( \tilde{B} \) matches pattern \( A_2 \) according to the principle of maximum member degree of similarity between IFSs. However, Mitchell’s \([29]\) similarity measure \( S_{\text{mod}}^p \) cannot be used effectively to classify this sample because \( S_{\text{mod}}^p(\tilde{A}_1, \tilde{B}) > S_{\text{mod}}^p(\tilde{A}_2, \tilde{B}) \).

From the previous examples, we see that the proposed similarity measure \( S_{\text{new}} \) performs better than the existing similarity measures \([11,22,23,28,29]\). In the next section, we apply the proposed measure to pattern recognition.

5. Application to pattern recognition

Let \( O = \{O_1, \ldots, O_n\} \) be a set of \( n \) objects. The aim of cluster analysis is to cluster \( O \) into \( c \) clusters. In this section, we consider that each pair of objects in \( O \) is represented by the relation \( r_{ij} \). This means that \( r_{ij} \) represents the degree of similarity between the objects \( O_i \) and \( O_j \). Usually, \( r_{ij} \) satisfies the following well-known conditions:

(i) \( 0 \leq r_{ij} \leq 1 \) for \( 1 \leq i, j \leq n \).
(ii) (Reflexivity) \( r_{ii} = 1 \) for \( 1 \leq i \leq n \).
(iii) (Symmetry) \( r_{ij} = r_{ji} \) for \( 1 \leq i, j \leq n \).

For the sake of convenience, the notation \( R = [r_{ij}] \) is used. We should mention here that if the matrix \( R \) is obtained from the correlations of IFSs, clustering methods for IFSs based on max–min compositions of the correlation matrix \( R \) of IFSs had been proposed in \([41,46]\). In fact, these clustering methods are similar to Yang and Shih \([42]\). It has been shown in \([42]\) that the max–min compositions of the relation matrix \( R \) can ultimately reach a max–min transitivity so that its \( \lambda \)-cut can correspond to an equivalent class. Yang and Wu \([43]\) claimed that their proposed similarity-based clustering method (SCM) is robust to initials, different cluster shapes and cluster numbers. However, the SCM was originally constructed for a data set with feature vectors. In this paper, we use SCM for constructed similarity matrix \( R \). One reason for this is that the SCM for the similarity matrix \( R \) should be more robust than those in \([41,42,44,46]\). In addition, we need to modify the original SCM such that it is suitable for the relation matrix \( R \). Let us consider the relation data
\[
 r_i = (r_{i1}, r_{i2}, \ldots, r_{in}), \quad i = 1, 2, \ldots, n,
\]
that represent vectors of relational strengths with the other objects. Using Yang and Wu’s concept \([43]\), we use the following similarity measure \( S(r_i, z_k) \) between \( r_i \) and the \( k \)th cluster center \( z_k \) with
\[
 S(r_i, z_k) = \exp\left( -\frac{\|r_i - z_k\|^2}{\beta} \right),
\]
where \( \|r_i - z_k\| \) is the Euclidean norm and \( \beta \) is the normalized constant defined by
\[
 \beta = \frac{\sum_{i=1}^n \|r_i - \bar{r}\|^2}{n}, \quad \text{with} \quad \bar{r} = \frac{1}{n} \sum_{i=1}^n r_i.
\]

Then the total similarity measure is given by
\[
 J(z) = \sum_{k=1}^c \sum_{i=1}^n \left( \exp\left( -\frac{\|r_i - z_k\|^2}{\beta} \right) \right)^{\gamma},
\]
where \( z = (z_1, \ldots, z_c) \) and \( \gamma > 0 \).

Since the clustering result is influenced by \( \gamma \), Yang and Wu \([43]\) proposed a correlation comparison algorithm (CCA) to select \( \gamma \). Based on this algorithm, we consider the total similarity function \( J(r_i)_{\gamma_m} \) for each data point \( r_i \) with
\[
 J(r_i)_{\gamma_m} = \sum_{j=1}^n \left( \exp\left( -\frac{\|r_i - r_j\|^2}{\beta} \right) \right)^{\gamma_m}, \quad i = 1, \ldots, n,
\]
where \( \gamma_m = m, m = 1, 2, 3, \ldots \). The correlation between the values of \( J(r_i)_{\gamma_1} \) and \( J(r_i)_{\gamma_2} \) are calculated. That is, CCA is based on a correlation comparison procedure with “\( \gamma_1 = 1, \gamma_2 = 2 \)”, “\( \gamma_2 = 2, \gamma_3 = 3 \)”, “\( \gamma_3 = 3, \gamma_4 = 4 \)”, etc. The CCA can be summarized as follows:
CCA algorithm

S1. Set \( m = 1 \) and give a threshold \( \rho \).
S2. Calculate the correlation of the values of \( \tilde{J}(r_i)_{m=1} \) and \( \tilde{J}(r_i)_{m=1} \).
S3. If the correlation is greater than or equal to the threshold \( \rho \)
    THEN choose \( \gamma_m \) to be the estimate of \( \gamma \);
    ELSE \( m = m + 1 \) and GOTO S2.

Since Yang and Wu [43] suggested a threshold around 0.97 ~ 0.999, we use 0.99 as the threshold in this section. After the parameter \( \gamma \) is estimated using CCA, the next step is to find a \( z_i \) that maximizes the objective function \( J(z) \). Differentiating \( J(z) \) with respect to all \( z_k \), we obtain

\[
z_k = \frac{\sum_{i=1}^{n} r_i S_k^r}{\sum_{i=1}^{n} S_k^r}, \quad (14)
\]

where

\[
S_k = S(r_i, z_k) = \exp \left( -\frac{||r_i - z_k||^2}{\beta} \right), \quad (15)
\]

Note that \( z_k \) in Eq. (14) cannot be solved directly. However, they can be solved using the fixed-point iterative method. This forms the similarity clustering algorithm (SCA) as follows:

**SCA algorithm**

(S1) Fix any \( \epsilon > 0 \). Choose initial values, say \( z_k^{(0)} \), \( k = 1, \ldots, c \).
(S2) Calculate \( S_k^{(0)} \) using \( z_k^{(0)} \) and Eq. (15).
(S3) Update \( z_k^{(0)} \) by \( z_k^{(1)} \) using \( S_k^{(0)} \) and Eq. (14).
(S4) IF \( \max_k ||z_k^{(0)} - z_k^{(1)}|| < \epsilon \), stop. Otherwise, set \( z_k^{(1)} = z_k^{(0)} \) and GO TO (S2).

When the SCA algorithm is processed, the problem is to determine what kind of initialization can guarantee that all peaks (clusters) are found simultaneously. To solve this problem, we adopt Yang and Wu’s [43] suggestion to set all data points to be the initial centers (i.e., \( z^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)}) = (r_1, \ldots, r_n) \)). Because SCA is processed with \( z^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)}) = (r_1, \ldots, r_n) \), the final \( n \) cluster centers show the final states of all data points. This provides a method to classify the data points using their final states. For example, \( x_1^{(0)} \) and \( x_n^{(0)} \) converge at the same peak. Then \( z_k^{(0)} \) and \( z_n^{(0)} \) should belong to the same cluster and so should \( x_1 \) and \( x_n \). Therefore, when \( z^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)}) = (r_1, \ldots, r_n) \) is centralized to \( c^* \) clusters, the data set is also classified into those \( c^* \) clusters simultaneously. By processing the final states of \( z^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)}) = (r_1, \ldots, r_n) \) into the single linkage algorithm, the optimal cluster number \( c^* \) and the identified clusters are found simultaneously. Thus, the SCM procedure for a relational matrix can be summarized as follows.

**SCM for a relational matrix \( R \)**

**Step 1.** Estimate \( \gamma \) using CCA;
**Step 2.** Process SCA with \( z^{(0)} = (z_1^{(0)}, \ldots, z_n^{(0)}) = (r_1, \ldots, r_n) \).
**Step 3.** Process the single linkage algorithm with the final states of the data points.
**Step 4.** Find the optimal cluster number \( c^* \) according to the dendrogram.
**Step 5.** Identify these \( c^* \) clusters.

Next, we adopt the 15 IFS patterns used by [24]. The SCM and Xu et al. [41] methods are used to analyze then compare these patterns.

**Example 6.** The 15 IFS patterns from Hung and Yang [24] are used in this example. These 15 IFS patterns in \( X = \{x\} \) are as follows:

\[
\tilde{A}_1 = \{(x, 0.91, 0.08)\}, \quad \tilde{A}_2 = \{(x, 0.93, 0.07)\}, \quad \tilde{A}_3 = \{(x, 0.87, 0.12)\}, \quad \tilde{A}_4 = \{(x, 0.85, 0.14)\}, \quad \tilde{A}_5 = \{(x, 0.79, 0.20)\}, \quad \tilde{A}_6 = \{(x, 0.19, 0.80)\}, \quad \tilde{A}_7 = \{(x, 0.10, 0.82)\}, \quad \tilde{A}_8 = \{(x, 0.06, 0.90)\}, \quad \tilde{A}_9 = \{(x, 0.03, 0.82)\}, \quad \tilde{A}_{10} = \{(x, 0.07, 0.73)\}, \quad \tilde{A}_{11} = \{(x, 0.50, 0.50)\}, \quad \tilde{A}_{12} = \{(x, 0.45, 0.55)\}, \quad \tilde{A}_{13} = \{(x, 0.40, 0.50)\}, \quad \tilde{A}_{14} = \{(x, 0.42, 0.48)\}, \quad \tilde{A}_{15} = \{(x, 0.46, 0.46)\}.
\]
We use the proposed similarity measure of Eq. (11) to calculate the similarity degrees between these 15 IFS patterns. We obtain the relational matrix $R$ as follows:

$$
R = \begin{pmatrix}
1 & 0.985 & 0.961 & 0.942 & 0.887 & 0.461 & 0.434 & 0.445 & 0.475 & 0.660 & 0.628 & 0.628 & 0.641 & 0.660 \\
1 & 0.946 & 0.928 & 0.874 & 0.480 & 0.454 & 0.427 & 0.438 & 0.468 & 0.651 & 0.619 & 0.619 & 0.631 & 0.651 \\
0.980 & 0.923 & 0.508 & 0.480 & 0.452 & 0.463 & 0.494 & 0.687 & 0.654 & 0.654 & 0.667 & 0.687 \\
1 & 0.942 & 0.517 & 0.489 & 0.461 & 0.472 & 0.504 & 0.701 & 0.667 & 0.667 & 0.680 & 0.701 \\
1 & 0.549 & 0.489 & 0.502 & 0.535 & 0.744 & 0.708 & 0.708 & 0.722 & 0.745 \\
1 & 0.946 & 0.891 & 0.914 & 0.909 & 0.737 & 0.775 & 0.775 & 0.760 & 0.737 \\
1 & 0.942 & 0.966 & 0.942 & 0.698 & 0.733 & 0.733 & 0.719 & 0.698 \\
1 & 0.946 & 0.914 & 0.657 & 0.691 & 0.691 & 0.677 & 0.657 \\
1 & 0.937 & 0.674 & 0.708 & 0.708 & 0.694 & 0.674 \\
1 & 0.719 & 0.756 & 0.756 & 0.741 & 0.719 \\
1 & 0.951 & 0.951 & 0.951 & 0.961 \\
1 & 0.951 & 0.951 & 0.951 \\
1 & 0.980 & 0.951 \\
1 & 0.970 \\
1
\end{pmatrix}
$$

The hierarchical clustering tree of the final states of all data points obtained by the SCM for the relational matrix $R$ is shown in Fig. 1. Increase in the $y$-coordinate represents the distance between clusters. Fig. 1 shows that there are three well-separated clusters among the final states of the data points. Hence, the optimal cluster number is $c^* = 3$. The corresponding clustering results are also shown in Table 1.

Now, suppose we have another IFS pattern $\bar{B} = \{(x, 0.60, 0.38)\}$. What cluster does this pattern belong to? To answer this question, we use the proposed similarity measure $S_{\text{new}}$ to calculate the average similarity between $B$ and clusters $C_1, C_2, C_3$ as follows:

$$
S_{\text{new}}(C_1, \bar{B}) = \frac{1}{5} \sum_{i=1}^{5} S_{\text{new}}(\bar{A}_i, \bar{B}) = \frac{1}{5}(0.737 + 0.726 + 0.767 + 0.783 + 0.831) = 0.769;
$$

$$
S_{\text{new}}(C_2, \bar{B}) = \frac{1}{5} \sum_{i=6}^{10} S_{\text{new}}(\bar{A}_i, \bar{B}) = \frac{1}{5}(0.660 + 0.625 + 0.589 + 0.604 + 0.644) = 0.624;
$$

$$
S_{\text{new}}(C_3, \bar{B}) = \frac{1}{5} \sum_{i=11}^{15} S_{\text{new}}(\bar{A}_i, \bar{B}) = \frac{1}{5}(0.896 + 0.852 + 0.852 + 0.869 + 0.896) = 0.873.
$$

It is evident that pattern $\bar{B}$ matches cluster $C_3$ according to the principle of the maximum degree of similarities. Thus, we can recognize the given IFS pattern $\bar{B}$.

Xu et al. [41] recently proposed a clustering algorithm for IFSs. Next, we make comparisons between the proposed SCM and the method proposed by Xu et al. [41]. In [41], Xu et al. proposed a weighted association coefficient $C(\bar{A}, \bar{B})$ of two IFS $\bar{A}$ and $\bar{B}$ (i.e., Eq. (28) in [41]). We calculate $C(\bar{A}, \bar{B})$ for the 15 IFS patterns and obtain the association matrix $C$ as follows:

Fig. 1. Hierarchical clustering tree through the SCM for the relational matrix $R$. 

93–109
The clustering results for 15 IFSs patterns by the SCM.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>$\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4, \tilde{A}_5$</td>
</tr>
<tr>
<td>$C_2$</td>
<td>$\tilde{A}_6, \tilde{A}_7, \tilde{A}_8, \tilde{A}<em>9, \tilde{A}</em>{10}$</td>
</tr>
<tr>
<td>$C_3$</td>
<td>$\tilde{A}<em>{11}, \tilde{A}</em>{12}, \tilde{A}<em>{13}, \tilde{A}</em>{14}, \tilde{A}_{15}$</td>
</tr>
</tbody>
</table>

\[
C = \begin{pmatrix}
1 & 0.979 & 0.960 & 0.960 & 0.943 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 \\
1 & 0.940 & 0.920 & 0.861 & 0.268 & 0.173 & 0.137 & 0.098 & 0.134 & 0.575 & 0.525 & 0.468 & 0.488 & 0.529 \\
1 & 0.981 & 0.922 & 0.339 & 0.241 & 0.197 & 0.163 & 0.195 & 0.642 & 0.593 & 0.530 & 0.550 & 0.591 \\
1 & 0.943 & 0.369 & 0.270 & 0.218 & 0.191 & 0.221 & 0.667 & 0.619 & 0.554 & 0.573 & 0.615 \\
1 & 0.459 & 0.354 & 0.279 & 0.272 & 0.306 & 0.745 & 0.701 & 0.628 & 0.646 & 0.687 \\
1 & 0.981 & 0.898 & 0.953 & 0.886 & 0.732 & 0.777 & 0.705 & 0.687 & 0.675 \\
1 & 0.917 & 0.988 & 0.902 & 0.668 & 0.720 & 0.665 & 0.644 & 0.624 \\
1 & 0.915 & 0.821 & 0.589 & 0.640 & 0.586 & 0.565 & 0.546 \\
1 & 0.906 & 0.611 & 0.668 & 0.628 & 0.605 & 0.579 \\
1 & 0.692 & 0.749 & 0.715 & 0.692 & 0.665 \\
1 & 0.990 & 0.900 & 0.900 & 0.920 \\
1 & 0.901 & 0.897 & 0.911 \\
1 & 0.995 & 0.982 \\
1 & 0.982 \\
1 & 1
\end{pmatrix}
\]

Starting from the association matrix $C$, Xu et al. [41] used the max–min composition to obtain an equivalent association matrix. The final equivalent association matrix from $C$ based on the max–min composition is $C^{16}$ as follows:

\[
C^{16} = \begin{pmatrix}
1 & 0.979 & 0.960 & 0.960 & 0.943 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 \\
1 & 0.960 & 0.960 & 0.943 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 \\
1 & 0.981 & 0.943 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 \\
1 & 0.943 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 \\
1 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 & 0.745 \\
1 & 0.981 & 0.917 & 0.981 & 0.906 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 \\
1 & 0.917 & 0.988 & 0.906 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 \\
1 & 0.917 & 0.906 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 \\
1 & 0.906 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 \\
1 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 & 0.777 \\
1 & 0.990 & 0.920 & 0.920 & 0.920 \\
1 & 0.902 & 0.920 & 0.920 \\
1 & 0.995 & 0.982 \\
1 & 0.982 \\
1 & 1
\end{pmatrix}
\]

The hierarchical clustering tree for the equivalent association matrix $C^{16}$ obtained by the Xu et al. [41] method is shown in Fig. 2. Increase in the y-coordinate represents the distance between clusters. Fig. 2 also shows that there are three well-separated clusters for the 15 IFS patterns so that the corresponding clustering results from [41] are the same as those shown in Table 1. However, if we compare Fig. 1 with Fig. 2, we find that there are differences between two figures for the lower $\lambda$-cuts. We find that the lower $\lambda$-cuts in Fig. 1 merged. Thus, it is clearer because of the three well-separated clusters shown in the hierarchical clustering tree. This is because the SCM can automatically merge these closer data points and re-organizes the data points according to the structure of data. Moreover, the SCM mechanism can self-organize cluster numbers and it is also robust to different cluster volumes and noisy points. These properties have been previously discussed [43].

Next, we consider a practical experiment. We borrow the car data set from Xu et al. [41]. We then apply the SCM, and the methods proposed in [46, 41] to the car data set, and make comparisons between them.
Example 7. The real car data set from [41] is used in this example. Each data point has six attributes: \( x_1 \): fuel economy; \( x_2 \): aerodynamic degree; \( x_3 \): price; \( x_4 \): comfort; \( x_5 \): design; \( x_6 \): safety. The data set is shown in Table 2.

We use the proposed similarity measure in Eq. (11) to calculate the similarity degrees between these 10 car data. We obtain the relational matrix \( R \) as follows:

\[
R = \begin{pmatrix}
1 & 0.572 & 0.549 & 0.513 & 0.607 & 0.819 & 0.577 & 0.522 & 0.505 & 0.654 \\
1 & 0.577 & 0.577 & 0.861 & 0.705 & 0.549 & 0.607 \\
1 & 0.577 & 0.607 & 0.819 & 0.705 & 0.607 & 0.577 \\
1 & 0.522 & 0.670 & 0.549 & 0.518 & 0.905 & 0.497 \\
1 & 0.602 & 0.607 & 0.607 & 0.577 & 0.779 \\
1 & 0.607 & 0.577 & 0.650 & 0.607 \\
1 & 0.741 & 0.522 & 0.577 \\
1 & 0.472 & 0.670 \\
1 & 0.549 \\
1
\end{pmatrix}
\]

The hierarchical clustering tree of the final states of all data points obtained by the SCM for the relational matrix \( R \) is shown in Fig. 3. Increase in the \( y \)-coordinate represents the distance between clusters. Fig. 3 shows that there are four well-separated clusters with \( C_1 = \{4, 9\} \), \( C_2 = \{1, 6\} \), \( C_3 = \{5, 10\} \) and \( C_4 = \{2, 3, 7, 8\} \) in the final states of the data points, but three clusters are also reasonable.

Xu et al. [41] calculated the weighted association coefficient \( C(A, B) \) for the 10 car data set by using the weight \( \omega = (0.15, 0.10, 0.30, 0.20, 0.15, 0.10) \) and obtained the final equivalent association matrix \( \mathbf{C}^8 \) based on the max–min composition [41]. According to the hierarchical clustering tree (or \( \lambda \)-cut structure) from the equivalent association matrix \( \mathbf{C}^8 \), we find that the clustering results with two clusters \( C_1 = \{1, 6\} \) and \( C_2 = \{2, 3, 4, 5, 7, 8, 9, 10\} \) and five clusters \( C_1 = \{1, 6\} \), \( C_2 = \{2\} \), \( C_3 = \{3, 5, 7, 10\} \), \( C_4 = \{8\} \) and \( C_5 = \{4, 9\} \) are the two most probable cases.

Table 2
The car data set.

| \( x_1 \) | \( \mu_A(x_1) \) | \( v_A(x_1) \) | \( x_2 \) | \( \mu_A(x_2) \) | \( v_A(x_2) \) | \( x_3 \) | \( \mu_A(x_3) \) | \( v_A(x_3) \) | \( x_4 \) | \( \mu_A(x_4) \) | \( v_A(x_4) \) | \( x_5 \) | \( \mu_A(x_5) \) | \( v_A(x_5) \) | \( x_6 \) | \( \mu_A(x_6) \) | \( v_A(x_6) \) |
| \( A_1 \) | 0.30 | 0.40 | 0.20 | 0.70 | 0.40 | 0.50 | 0.80 | 0.10 | 0.40 | 0.50 | 0.20 | 0.70 |
| \( A_2 \) | 0.40 | 0.30 | 0.50 | 0.10 | 0.60 | 0.20 | 0.20 | 0.70 | 0.30 | 0.60 | 0.70 | 0.20 |
| \( A_3 \) | 0.40 | 0.20 | 0.60 | 0.10 | 0.80 | 0.10 | 0.20 | 0.60 | 0.30 | 0.70 | 0.50 | 0.20 |
| \( A_4 \) | 0.30 | 0.40 | 0.90 | 0.00 | 0.80 | 0.10 | 0.70 | 0.10 | 0.10 | 0.80 | 0.20 | 0.80 |
| \( A_5 \) | 0.80 | 0.10 | 0.70 | 0.20 | 0.70 | 0.00 | 0.40 | 0.10 | 0.80 | 0.20 | 0.40 | 0.60 |
| \( A_6 \) | 0.40 | 0.30 | 0.30 | 0.50 | 0.20 | 0.60 | 0.70 | 0.10 | 0.50 | 0.40 | 0.30 | 0.60 |
| \( A_7 \) | 0.60 | 0.40 | 0.40 | 0.20 | 0.70 | 0.20 | 0.30 | 0.60 | 0.30 | 0.70 | 0.60 | 0.10 |
| \( A_8 \) | 0.90 | 0.10 | 0.70 | 0.20 | 0.70 | 0.10 | 0.40 | 0.50 | 0.40 | 0.50 | 0.80 | 0.00 |
| \( A_9 \) | 0.40 | 0.40 | 1.00 | 0.00 | 0.90 | 0.10 | 0.60 | 0.20 | 0.20 | 0.70 | 0.10 | 0.80 |
| \( A_{10} \) | 0.90 | 0.10 | 0.80 | 0.00 | 0.60 | 0.30 | 0.50 | 0.20 | 0.80 | 0.10 | 0.60 | 0.40 |
According to [46], an equivalent similarity matrix $R^8$ is obtained. Based on the hierarchical clustering tree (or $k$-cut structure) from the equivalent similarity matrix $R^8$, we find that the clustering results with two clusters $C_1 = \{4, 9\}$ and $C_2 = \{1, 2, 3, 5, 6, 7, 8, 10\}$ and four clusters $C_1 = \{1, 6\}$, $C_2 = \{2, 3, 7, 8\}$, $C_3 = \{4, 9\}$ and $C_4 = \{5, 10\}$ are the two most probable cases.

The advantage to the method proposed by Xu et al. [41] is that attribute weights are considered, so that it can count information from different attributes and remain more flexible. However, its drawback is that attribute weights become parameters. This leads to problems in determining which weight is better and how to choose attribute weights. In fact, different weights affect the final results in [41]. Next, we demonstrate this phenomenon as follows. We calculate the weighted association coefficient $C(A, B)$ for the 10 car data using another weight $w = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$. The association matrix $C$ is obtained as follows:

\[
C = \begin{pmatrix}
1 & 0.637 & 0.651 & 0.689 & 0.650 & 0.908 & 0.682 & 0.564 & 0.655 & 0.602 \\
1 & 0.938 & 0.646 & 0.656 & 0.683 & 0.921 & 0.798 & 0.631 & 0.679 \\
1 & 0.714 & 0.700 & 0.639 & 0.932 & 0.810 & 0.712 & 0.685 \\
1 & 0.738 & 0.627 & 0.651 & 0.676 & 0.969 & 0.705 \\
1 & 0.665 & 0.679 & 0.825 & 0.757 & 0.929 \\
1 & 0.661 & 0.579 & 0.594 & 0.651 \\
1 & 0.851 & 0.652 & 0.705 \\
1 & 0.691 & 0.869 \\
1 & 0.730 \\
1
\end{pmatrix}
\]

The final equivalent association matrix from $C$ based on the max–min composition is $C^8$ as follows:

\[
C^8 = \begin{pmatrix}
1 & 0.689 & 0.689 & 0.689 & 0.689 & 0.908 & 0.689 & 0.689 & 0.689 & 0.689 \\
1 & 0.938 & 0.757 & 0.851 & 0.689 & 0.932 & 0.851 & 0.757 & 0.851 \\
1 & 0.757 & 0.851 & 0.689 & 0.932 & 0.851 & 0.757 & 0.851 \\
1 & 0.757 & 0.689 & 0.757 & 0.757 & 0.969 & 0.757 \\
1 & 0.689 & 0.851 & 0.869 & 0.757 & 0.929 \\
1 & 0.689 & 0.689 & 0.869 & 0.689 & 0.869 & 0.689 & 0.869 \\
1 & 0.851 & 0.757 & 0.851 \\
1 & 0.757 & 0.869 \\
1 & 0.757 \\
1
\end{pmatrix}
\]

The hierarchical clustering tree for the equivalent association matrix $C^8$ obtained by the method proposed by Xu et al. [41] is shown in Fig. 4. Increase in the $y$-coordinate represents the distance between clusters. Fig. 4 shows that there are three separated clusters for the car data set with $C_1 = \{4, 9\}$, $C_2 = \{2, 3, 5, 7, 8, 10\}$ and $C_3 = \{1, 6\}$.

Furthermore, the hierarchical clustering tree obtained by the method of Xu et al. [41] using the weight $w = (1/8, 1/8, 1/8, 1/8, 1/8, 2/8, 2/8)$ is shown in Fig. 5, where the three clusters $C_1 = \{4, 9\}$, $C_2 = \{2, 3, 5, 7, 8, 10\}$ and $C_3 = \{1, 6\}$ are the most probable cases. The hierarchical clustering tree obtained by the method proposed by Xu et al. [41] using the weight $w = (0.1, 0.1, 0.1, 0.1, 0.3, 0.3)$
is shown in Fig. 6, where the three clusters $C_1 = \{4, 9\}$, $C_2 = \{1, 6\}$ and $C_3 = \{2, 3, 7\}$ and the four clusters $C_1 = \{4, 9\}$, $C_2 = \{1, 6\}$, $C_4 = \{2, 3, 7\}$ and $C_5 = \{5, 10\}$ are the two most probable cases. The hierarchical clustering tree obtained by the method of Xu et al. [41] using the weight $w = (0.1, 0.1, 0.1, 0.1, 0.1, 0.5)$ is shown in Fig. 7, where the five clusters $C_1 = \{4, 9\}$, $C_2 = \{1, 6\}$, $C_3 = \{2, 3, 7\}$, $C_4 = \{8\}$ and $C_5 = \{5, 10\}$ are the most probable cases. The hierarchical clustering tree obtained by the Xu et al. [41] method using the weight $w = (0.1, 0.1, 0.1, 0.1, 0.1, 0.3, 0.3)$ is shown in Fig. 8, where the four clusters $C_1 = \{2, 3, 7, 8\}$, $C_2 = \{4, 9\}$, $C_3 = \{1, 6\}$ and $C_4 = \{5, 10\}$ are the most probable cases. As a whole, the attribute weights in the method proposed by Xu et al. [41] actually affect their final clustering results.
6. Conclusion

There are several similarity measures between IFSs proposed by researchers that can provide effective ways to deal with IFSs in some cases. However, in other cases, these measures may obtain unreasonable results. In this paper, we propose a new similarity measure of IFSs based on the Sugeno integral. By using examples, we show that the proposed similarity measure can deal with problems more reasonably than most existing similarity measures. At the same time, the proposed similarity measure can overcome the drawbacks of existing measures. In a practical example, we embed the proposed similarity measure with a robust clustering algorithm and then apply it in pattern recognition. The clustering methods for IFSs [41,46] are also discussed and compared. We find that the proposed SCM can automatically merge these closer data points and re-organize the data points according to the structure of data so that it can more easily demonstrate the optimal cluster number.

Acknowledgements

The authors thank the anonymous reviewers for their constructive comments to improve the presentation of the paper. This work was supported in part by the National Science Council of Taiwan under Miin-Shen Yang’s Grant: NSC-97-2118-M-033-002-MY2.

References


