

# A Chi-square Distance-based Similarity Measure of Single-valued Neutrosophic Set and Applications

H.P. Ren, S.X. Xiao, H. Zhou

## Haiping Ren

Teaching Department of Basic Subjects  
Jiangxi University of Science and Technology, 330013 Nanchang, China  
9520060004@jxust.edu.cn

## Shixiao Xiao\*

Chengyi University College  
Jimei University, 361021 Xiamen, China  
\*Corresponding author: xiaoshixiao@jmu.edu.cn

## Hui Zhou

School of Mathematics and Computer Science  
Yichun University, 336000 Yichun, China  
huihui7978@126.com

**Abstract:** The aim of this paper is to propose a new similarity measure of single-valued neutrosophic sets (SVNSs). The idea of the construction of the new similarity measure comes from Chi-square distance measure, which is an important measure in the applications of image analysis and statistical inference. Numerical examples are provided to show the superiority of the proposed similarity measure comparing with the existing similarity measures of SVNSs. A weighted similarity is also put forward based on the proposed similarity. Some examples are given to show the effectiveness and practicality of the proposed similarity in pattern recognition, medical diagnosis and multi-attribute decision making problems under single-valued neutrosophic environment.

**Keywords:** Chi-square distance measure, similarity measure, multi-attribute decision making, single-valued neutrosophic set.

## 1 Introduction

Since fuzzy set was first proposed by Zadeh and has achieved a great success in various fields due to its capability of handling uncertainty [4, 6, 9, 14, 21, 29]. Over the last decades, some extended fuzzy sets have been introduced by researchers, such as intuitionistic fuzzy set [1], vague set [10], interval-valued intuitionistic fuzzy set [2] and hesitant fuzzy set [28]. As an extension of Zadeh's fuzzy set, intuitionistic fuzzy set can better describe the situation when decision making process exists decision makers' hesitation than Zadeh's fuzzy set by adding a non-membership degree parameter [32]. In recent years, intuitionistic fuzzy set has received a lot attention and been applied to many fields, such as management decision, pattern recognition and medical diagnosis [5, 7, 18, 20, 27]. In practice, indeterminate and inconsistent information may occur, then intuitionistic fuzzy set cannot deal with these situations well because it only contains the true membership degree and the false membership degree (non-membership degree). For example, when an authority wants to choose the best candidate, ten experts are invited to take part in the decision. For one candidate he gained 10 votes from the experts. There are 3 votes "yes", 2 votes "no", 2 "gave up" and 3 "undecided". In this case, intuitionistic fuzzy set cannot describe it well. To overcome this shortcoming, Smarandache [26] introduced a concept of neutrosophic set, which is an extension of intuitionistic fuzzy set from philosophical point of

view. Neutrosophic set is defined as a set containing the degree of truth, indeterminacy, and falsity. For afore-mentioned example, the vote result can be expressed by a neutrosophic set. However, the original neutrosophic set is difficult to apply in practical problems. To applied it easily in science and engineering fields, Wang et al. [30] introduced the concept of single-valued neutrosophic set (SVNS), which is a subclass of Smarandache's neutrosophic set. Because SVNS is easy to express, it has been a useful mathematical tool for handling various practical problems involving imprecise, indeterminacy, and inconsistent data [11, 15, 19, 23, 25].

In recent years, the study and applications of information measures of fuzzy sets have received a lot attention. Similarity measure is one of the most important measurement tools for comparing the degree of similarity between two objects. Since Li and Chen [17] introduced the definition of the similarity measure between two intuitionistic fuzzy sets. Since then intuitionistic fuzzy similarity measures have received great attention. From a different point of view, many similarity measures are proposed and applied to solve various practical problems of MADM, pattern recognition and medical diagnosis, etc. [12, 13, 16, 22, 31, 34]. As an extension of intuitionistic fuzzy set, some similarity measures of neutrosophic set is developed from those of intuitionistic fuzzy sets, and some new similarity measures are also proposed, but the references are still rare ([35]). Based on vector similarity functions, some similarity measures between simplified neutrosophic sets are put forwards, such as similarity measures based on Jaccard, Dice, and cosine functions [33–35].

We find that the existing similarity measures have shortcomings, and the detail analysis can be found in Example 1. Then the aim of this paper is to develop a new similarity measure of SVNSs based on Chi-square distance measure, which is an important measure in statistical theory. We will show the advantage of the proposed similarity measure with existing similarity measures of SVNSs through comparison with some numerical examples. Three examples are provided to demonstrate the effectiveness and practicality of the proposed similarity in the application of pattern recognition, medical diagnosis and multi-attribute decision making.

The remains of this article are organized as follows: Section 2 will recall some basic concepts and properties of SVNSs and similarity measure. Section 3 introduces a new similarity measure between SVNSs based on ordinary Chi-square distance measure, and put forward a weighted similarity for further applications. Section 4 develops the applications of the proposed similarity measure with some examples. Finally, conclusions are provided in Section 5.

## 2 Preliminary knowledge

In this section, some basic concepts and properties of SVNSs and similarity measure are presented. Smarandache [26] originally introduced a concept of neutrosophic set from philosophical point of view.

**Definition 1** Let  $X$  be a universal set. A set is called a neutrosophic set, if it is characterized by three parameters: truth-membership function  $T_A(x)$ , indeterminacy-membership function  $I_A(x)$  and falsity-membership function  $F_A(x)$ . That is  $A$  has the following form:

$$A = \{ \langle x, T_A(x), I_A(x), F_A(x) \rangle \mid x \in X \}.$$

Here  $T_A(x), I_A(x), F_A(x) : X \rightarrow ]^{-}0, 1^{+}[$ ,  $]^{-}0, 1^{+}[$  is non-standard interval, and they satisfy

$$^{-}0 \leq \sup T_A(x) + \sup I_A(x) + \sup F_A(x) \leq 3^{+}$$

A neutrosophic is defined as a set containing the degree of truth, indeterminacy, and falsity. However, the original neutrosophic set is difficult to apply in practical problems. To apply it

easily in science and engineering fields, Wang et al. [30] introduced single-valued neutrosophic set (SVNS), which is a subclass of Smarandache's neutrosophic set.

**Definition 2** Let  $X$  be a universal set. A set  $A$  is called a SVNS, if it is characterized by three parameters: truth-membership function  $T_A(x)$ , indeterminacy-membership function  $I_A(x)$  and falsity-membership function  $F_A(x)$ . That is,  $A$  has the following form:

$$A = \{ \langle x, T_A(x), I_A(x), F_A(x) \rangle \mid x \in X \}.$$

Here  $T_A(x), I_A(x), F_A(x) : X \rightarrow [0, 1]$ , and they satisfy  $0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3$ . For convenience, when  $X = \{x\}$ , we briefly denote the element  $\langle x, T_A(x), I_A(x), F_A(x) \rangle$  of  $A$  by  $\langle T_A, I_A, F_A \rangle$ . Element  $\langle T_A, I_A, F_A \rangle$  is often named as a single-valued neutrosophic value (SVNV).

**Definition 3** [30] Let  $X$  be a universal set, and  $A = \{ \langle x, T_A(x), I_A(x), F_A(x) \rangle \mid x \in X \}$  and  $B = \{ \langle x, T_B(x), I_B(x), F_B(x) \rangle \mid x \in X \}$  are two SVNSs in  $X$ , then

(i) The complement of a SVNS  $A$  is

$$A^c = \{ \langle x, F_A(x), 1 - I_A(x), T_A(x) \rangle \mid x \in X \}$$

(ii)  $A \subseteq B$  if and only if

$$T_A(x) \leq T_B(x), I_A(x) \geq I_B(x), F_A(x) \geq F_B(x),$$

for all  $x$  in  $X$ . (iii)  $A = B$  if and only if  $A \subseteq B$  and  $B \subseteq A$ .

In the following discussion, we always use  $SVNSs(X)$  to denote the set of all SVNSs in  $X$ . Definition 4 will introduce the definition of a similarity measure between two SVNSs  $A$  and  $B$ .

**Definition 4** [35] Let  $A$  and  $B$  be two SVNSs, and  $S$  is a mapping  $S : SVNSs(X) \times SVNSs(X) \rightarrow [0, 1]$ . We call  $S(A, B)$  the similarity measure between  $A$  and  $B$  if it satisfies the following properties:

- (i)  $0 \leq S(A, B) \leq 1$ ;
- (ii)  $S(A, B) = 1$  if and only if  $A = B$ ;
- (iii)  $S(A, B) = S(B, A)$ ;
- (iv) If  $A \subseteq B \subseteq C$ , then  $S(A, C) \leq \min\{S(A, B), S(B, C)\}$ .

### 3 A new Chi-square distance-based similarity

This section contains two subsections. The first subsection will propose a new similarity measure between two SVNSs based on Chi-square distance measure. The second subsection will compare the proposed similarity measure with existing similarity measures of SVNSs.

#### 3.1 A new proposed similarity based on Chi-square distance

This section will propose a new similarity measure between two SVNSs based on Chi-square distance measure. The name of the Chi-square distance measure is derived from Pearson's Chi-squared test statistic  $\chi^2(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \frac{(x_i - y_i)^2}{x_i}$ , which is used to compare two discrete probability distributions. However, as a distance measure, the function  $d(x, y)$  should be symmetric for two objects  $x$  and  $y$ . Then Chi-square distance measure of two real vectors  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  is proposed as the following formula [24]:

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \frac{(x_i - y_i)^2}{x_i + y_i} \quad (1)$$

Chi-square distance measure is one of most important distance measure used in face recognition [24]. To avoid the fact that the denominator is zero, the numerical value is meaningless. For later use, we proposed a revised version of Chi-square distance measure formula as follows:

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \frac{(x_i - y_i)^2}{2 + x_i + y_i} \quad (2)$$

Note that the constant 2 can be changed as any other positive number. Let  $X = \{x_1, x_2, \dots, x_n\}$  be a universe set. Then for two given SVNNSs  $A = \{ \langle x_i, T_A(x_i), I_A(x_i), F_A(x_i) \rangle \mid x_i \in X \}$  and  $B = \{ \langle x_i, T_B(x_i), I_B(x_i), F_B(x_i) \rangle \mid x_i \in X \}$ , the new neutrosophic fuzzy information measure based on Chi-square distance  $S \triangleq S(A, B)$  is constructed as follows:

$$S = 1 - \frac{1}{2n} \sum_{i=1}^n \left[ \frac{(T_A(x_i) - T_B(x_i))^2}{2 + T_A(x_i) + T_B(x_i)} + \frac{(I_A(x_i) - I_B(x_i))^2}{2 + I_A(x_i) + I_B(x_i)} + \frac{(F_A(x_i) - F_B(x_i))^2}{2 + F_A(x_i) + F_B(x_i)} + |m_A(x_i) - m_B(x_i)| \right] \quad (3)$$

Where  $m_j(x_i) = \frac{1 + T_j(x_i) - F_j(x_i)}{2}$ ,  $j = 1, n$ .

To prove the information measure (3) is a valid similarity measure, we need the following lemma which can be easily proved by straightforward calculation.

**Lemma 1** Let  $a, b, c$  be three non-negative real numbers, and  $0 \leq a \leq b \leq c$ . Then

- (i)  $\frac{(a-c)^2}{2+a+c} \geq \frac{(a-b)^2}{2+a+b}$
- (ii)  $\frac{(a-c)^2}{2+a+c} \geq \frac{(b-c)^2}{2+b+c}$

**Theorem 1** Let  $X = \{x_1, x_2, \dots, x_n\}$  be a universe set.  $A = \{ \langle x_i, T_A(x_i), I_A(x_i), F_A(x_i) \rangle \mid x_i \in X \}$  and  $B = \{ \langle x_i, T_B(x_i), I_B(x_i), F_B(x_i) \rangle \mid x_i \in X \}$  are two SVNNSs. Then information measure  $S(A, B)$  given by (2) is a valid similarity measure between SVNNSs  $A$  and  $B$ . That is,  $S(A, B)$  satisfies the properties (i)-(iv) of Definition 4.

**Proof** (i) It is obvious that  $0 \leq S(A, B) \leq 1$ .

(ii) When  $A = B$ , i.e.  $T_A(x_i) = T_B(x_i), I_A(x_i) = I_B(x_i), F_A(x_i) = F_B(x_i)$ , for all  $x_i$  in  $X$ . Then

$$m_A(x_i) = \frac{1 + T_A(x_i) - F_A(x_i)}{2} = \frac{1 + T_B(x_i) - F_B(x_i)}{2} = m_B(x_i)$$

Hence we have  $S(A, B) = 1$ .

(iii) The result is obvious.

(iv) If  $A \subseteq B \subseteq C$ , i.e.

$$T_A(x_i) \leq T_B(x_i) \leq T_C(x_i), I_A(x_i) \geq I_B(x_i) \geq I_C(x_i), F_A(x_i) \geq F_B(x_i) \geq F_C(x_i),$$

Then  $0 \leq m_A(x_i) \leq m_B(x_i) \leq m_C(x_i)$ , for all  $x_i$  in  $X$ .

Consequently, we can get

$$|m_A(x_i) - m_C(x_i)| \geq |m_A(x_i) - m_B(x_i)|, |m_A(x_i) - m_C(x_i)| \geq |m_B(x_i) - m_C(x_i)|$$

Thus by (i) of Lemma 1, we have

$$\begin{aligned} \frac{(T_A(x_i) - T_C(x_i))^2}{2 + T_A(x_i) + T_C(x_i)} &\geq \frac{(T_A(x_i) - T_B(x_i))^2}{2 + T_A(x_i) + T_B(x_i)}, \\ \frac{(I_A(x_i) - I_C(x_i))^2}{2 + I_A(x_i) + I_C(x_i)} &\geq \frac{(I_A(x_i) - I_B(x_i))^2}{2 + I_A(x_i) + I_B(x_i)}, \end{aligned}$$

$$\frac{(F_A(x_i) - F_C(x_i))^2}{2 + F_A(x_i) + F_C(x_i)} \geq \frac{(F_A(x_i) - F_B(x_i))^2}{2 + F_A(x_i) + F_B(x_i)}.$$

Then we can easily conclude that  $S(A, C) \leq S(A, B)$ . By (ii) of Lemma 2, we have

$$\frac{(I_A(x_i) - I_C(x_i))^2}{2 + I_A(x_i) + I_C(x_i)} \geq \frac{(I_B(x_i) - I_C(x_i))^2}{2 + I_B(x_i) + I_C(x_i)},$$

$$\frac{(F_A(x_i) - F_C(x_i))^2}{2 + F_A(x_i) + F_C(x_i)} \geq \frac{(F_B(x_i) - F_C(x_i))^2}{2 + F_B(x_i) + F_C(x_i)}.$$

$$\frac{(T_A(x_i) - T_C(x_i))^2}{2 + T_A(x_i) + T_C(x_i)} \geq \frac{(T_B(x_i) - T_C(x_i))^2}{2 + T_B(x_i) + T_C(x_i)},$$

Then we can easily conclude that  $S(A, C) \leq S(B, C)$ . Hence  $S(A, C) \leq \min\{S(A, B), S(B, C)\}$ . This completes the proof of Theorem 1. If we consider the important degree of  $x_i \in X = \{x_1, x_2, \dots, x_n\}$ , then we can establish a weighted similarity measure  $S_W \triangleq S_W(A, B)$  between SNNSs  $A$  and  $B$  as follows:

$$S_W = 1 - \frac{1}{2} \sum_{i=1}^n w_i \left[ \frac{(T_A(x_i) - T_B(x_i))^2}{2 + T_A(x_i) + T_B(x_i)} + \frac{(I_A(x_i) - I_B(x_i))^2}{2 + I_A(x_i) + I_B(x_i)} + \frac{(F_A(x_i) - F_B(x_i))^2}{2 + F_A(x_i) + F_B(x_i)} + |m_A(x_i) - m_B(x_i)| \right] \quad (4)$$

where  $w_i$  ( $i = 1, 2, \dots, n$ ) is the important degree of the element  $x_i$ , they satisfy  $w_i \in [0, 1]$  and  $\sum_{i=1}^n w_i = 1$ . If we set  $w_i = \frac{1}{n}$  ( $i = 1, 2, \dots, n$ ), then  $S_W(A, B) = S(A, B)$ . Similar to the proof process of  $S_R(A, B)$  in Theorem 1, we can easily prove that the weighted similarity measure  $S_W(A, B)$  is also a valid similarity between two SNVSs  $A$  and  $B$ . That is  $S_W(A, B)$  satisfies the properties (i)-(iv) of Definition 4.

### 3.2 Comparison of various similarity measures

To demonstrate the validness and performance of the new proposed similarity measure, some numerical examples are used to compare it with existing similarity measures: Jaccard similarity  $S_J(A, B)$ , Dice similarity  $S_D(A, B)$ , Cosine Similarity  $S_C(A, B)$ , Improved cosine similarity  $C_1(A, B)$  and  $C_2(A, B)$ , Tangent function-based similarity  $T_1(A, B)$ ,  $T_2(A, B)$ , and Cotangent function-based similarity  $CoT_1(A, B)$ ,  $CoT_2(A, B)$ . These similarity measures are given as follows ([33], [34], [35]):

$$S_J(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{S_{J1}}{S_{J2}}, \quad (5)$$

where  $S_{J1} = T_A(x_i)T_B(x_i) + I_A(x_i)I_B(x_i) + F_A(x_i)F_B(x_i)$  and  $S_{J2} = (T_A^2(x_i) + I_A^2(x_i) + F_A^2(x_i)) + (T_B^2(x_i) + I_B^2(x_i) + F_B^2(x_i)) - S_{J1}$

$$S_D(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{2(T_A(x_i)T_B(x_i) + I_A(x_i)I_B(x_i) + F_A(x_i)F_B(x_i))}{(T_A^2(x_i) + I_A^2(x_i) + F_A^2(x_i)) + (T_B^2(x_i) + I_B^2(x_i) + F_B^2(x_i))}, \quad (6)$$

$$S_C(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{T_A(x_i)T_B(x_i) + I_A(x_i)I_B(x_i) + F_A(x_i)F_B(x_i)}{\sqrt{T_A^2(x_i) + I_A^2(x_i) + F_A^2(x_i)} \sqrt{T_B^2(x_i) + I_B^2(x_i) + F_B^2(x_i)}}, \quad (7)$$

$$C_1(A, B) = \frac{1}{n} \sum_{i=1}^n \cos \left[ \frac{\pi}{2} \max (|T_A(x_i) - T_B(x_i)|, |I_A(x_i) - I_B(x_i)|, |F_A(x_i) - F_B(x_i)|) \right], \quad (8)$$

$$C_2(A, B) = \frac{1}{n} \sum_{i=1}^n \cos \left[ \frac{\pi}{6} (|T_A(x_i) - T_B(x_i)| + |I_A(x_i) - I_B(x_i)| + |F_A(x_i) - F_B(x_i)|) \right], \quad (9)$$

$$T_1(A, B) = 1 - \frac{1}{n} \sum_{i=1}^n \tan \left[ \frac{\pi}{4} \max (|T_A(x_i) - T_B(x_i)|, |I_A(x_i) - I_B(x_i)|, |F_A(x_i) - F_B(x_i)|) \right], \quad (10)$$

$$T_2(A, B) = 1 - \frac{1}{n} \sum_{i=1}^n \tan \left[ \frac{\pi}{12} (|T_A(x_i) - T_B(x_i)| + |I_A(x_i) - I_B(x_i)| + |F_A(x_i) - F_B(x_i)|) \right], \quad (11)$$

$$CoT_1(A, B) = \frac{1}{n} \sum_{i=1}^n \cot \left[ \frac{\pi}{4} \max (|T_A(x_i) - T_B(x_i)|, |I_A(x_i) - I_B(x_i)|, |F_A(x_i) - F_B(x_i)|) \right], \quad (12)$$

$$CoT_2(A, B) = \frac{1}{n} \sum_{i=1}^n \cot \left[ \frac{\pi}{4} + \frac{\pi}{12} (|T_A(x_i) - T_B(x_i)| + |I_A(x_i) - I_B(x_i)| + |F_A(x_i) - F_B(x_i)|) \right], \quad (13)$$

**Example 1** Suppose that  $X = \{x\}$ , we consider pattern recognition problems with six pairs of SVNPs shown in Table 1. The calculated numerical values of these 9 existing similarity measures and proposed similarity measure are shown in Table 1.

Table 1: Values of the different similarity measures under different pairs of  $(A, B)$

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
<b>A</b>	<0.3,0.3,0.4>	<0.3,0.3,0.4>	<1,0,0>	<0.4,0.2,0.6>	<0.4,0.4,0.2>	<0.4,0.4,0.2>
<b>B</b>	<0.4,0.3,0.4>	<0.4,0.3,0.3>	<0,1,1>	<0.2,0.2,0.3>	<0.5,0.2,0.3>	<0.5,0.3,0.2>
$S_J(A, B)$	0.9737	0.9429	0	0	0.8500	0.9474
$S_D(A, B)$	0.9867	0.9706	0	0	0.9189	0.9730
$S_C(A, B)$	0.9910	0.9706	0	Null	0.9193	0.9733
$C_1(A, B)$	<b>0.9877</b>	<b>0.9877</b>	0	0	0.9511	0.9877
$C_2(A, B)$	0.9986	0.9945	0	<b>0.8660</b>	0.9781	0.9945
$T_1(A, B)$	<b>0.9213</b>	<b>0.9213</b>	0	0	0.8416	0.9213
$T_2(A, B)$	0.9738	0.9476	0	<b>0.7321</b>	0.8949	0.9476
$CoT_1(A, B)$	0.8541	0.8541	0	0	0.7265	0.8541
$CoT_2(A, B)$	0.9490	0.9004	0	<b>0.5774</b>	0.8098	0.9004
$S(A, B)$	0.9713	0.9463	0	<b>0.5833</b>	<b>0.9886</b>	<b>0.9714</b>

From Table 1, we can see that the similarity measures  $C_1(A, B)$  and  $T_1(A, B)$  cannot carry out the recognition between Case 1 and Case 2. For Case 4, there are only four reasonable similarity

measures  $C_2(A, B)$ ,  $T_2(A, B)$ ,  $CoT_2(A, B)$  and  $S(A, B)$  intuitively consistent. An interesting counter-intuitive case occurs when three SVNNSs  $A = \langle 0.4, 0.4, 0.2 \rangle$ ,  $B = \langle 0.5, 0.2, 0.3 \rangle$  and  $C = \langle 0.5, 0.3, 0.2 \rangle$ . They can be written in forms of intuitionistic fuzzy values as:  $A = \langle 0.4, 0.2 \rangle$ ,  $B = \langle 0.5, 0.3 \rangle$  and  $C = \langle 0.5, 0.2 \rangle$ , respectively. In this case, Boran and Akay [3] pointed out that it is expected that the similarity degree between  $A$  and  $B$  should not be less than the similarity degree between  $A$  and  $C$  since they are ordered as  $A \leq B \leq C$  according to score function and accuracy function. However, the similarity degree between  $A$  and  $C$  is greater than the similarity degree between  $A$  and  $B$  when the existing similarity measures are used (The results can be found in Table 1), which does not seem to be reasonable. Table 1 shows that our proposed similarity measure is in agreement with this analysis. According to the above analysis, the proposed similarity measure is the most reasonable similarity measure.

## 4 Applications

In the following discussion, we will give two examples in pattern recognition and medical diagnosis to demonstrate the effectiveness and practicability of the proposed similarity measure.

**Example 2** Assume that there are two patterns in  $X = \{x_1, x_2\}$ . The two patterns are expressed by SVNNSs, which are shown as follows:

$$A_1 = \{\langle x_1, 0.2, 0.0, 0.2 \rangle, \langle x_2, 0.2, 0.0, 0.2 \rangle, \langle x_3, 0.2, 0.0, 0.2 \rangle\},$$

$$A_2 = \{\langle x_1, 0.4, 0.0, 0.4 \rangle, \langle x_2, 0.4, 0.0, 0.4 \rangle, \langle x_3, 0.4, 0.0, 0.4 \rangle\}.$$

Assume that there is an object

$$B = \{\langle x_1, 0.3, 0.0, 0.3 \rangle, \langle x_2, 0.3, 0.0, 0.3 \rangle, \langle x_3, 0.2, 0.0, 0.3 \rangle\}$$

Our task is to classify the object  $B$  in  $A_1$  or  $A_2$ . According to the recognition principle of maximum similarity measure between SVNNSs, the process of assigning the object  $B$  to  $A_1$  or  $A_2$  is described by

$$k = \arg \max_{1 \leq i \leq 2} \{S_R(A_i, B)\} \quad (14)$$

By Eq. (3), we can get the similarity measures between  $A_1, A_2$  with  $B$ :  $S(A_1, B) = 0.9700$ ,  $S(A_2, B) = 0.9415$ . Then the pattern  $B$  is classified in  $A_1$  according to the recognition rule given by Eq. (14). This result is consistent with our intuition.

**Example 3** We consider the following pattern recognition problem: There are three patterns  $A_1, A_2$  and  $A_3$ , which are represented by SVNNSs in universe set  $X = \{x_1, x_2, x_3\}$ , as follows:

$$A_1 = \{\langle x_1, 1.0, 0.2, 0.0 \rangle, \langle x_2, 0.8, 0.3, 0.0 \rangle, \langle x_3, 0.7, 0.1, 0.1 \rangle\}$$

$$A_2 = \{\langle x_1, 0.8, 0.1, 0.1 \rangle, \langle x_2, 1.0, 0.1, 0.2 \rangle, \langle x_3, 0.9, 0.2, 0.1 \rangle\}$$

$$A_3 = \{\langle x_1, 0.6, 0.3, 0.2 \rangle, \langle x_2, 0.8, 0.2, 0.3 \rangle, \langle x_3, 0.6, 0.3, 0.2 \rangle\}$$

Given an unknown pattern  $B$ , which is represented by the SVNNS:

$$B = \{\langle x_1, 0.5, 0.3, 0.2 \rangle, \langle x_2, 0.6, 0.3, 0.2 \rangle, \langle x_3, 0.8, 0.2, 0.1 \rangle\}$$

Our task is to classify the pattern  $B$  in one of the classes  $A_1, A_2$  and  $A_3$ . According to the recognition principle of maximum similarity measure between SVNNSs, the process of assigning the pattern  $B$  to  $A_k$  ( $k = 1, 2, 3$ ) is described by

$$k = \arg \max_{1 \leq i \leq 3} \{S_R(A_i, B)\} \quad (15)$$

By Eq.(3), we can get the similarity measures between  $B$  with  $A_i$  ( $i = 1, 2, 3$ ) :

$$S(A_1, B) = 0.9115, S(A_2, B) = 0.8813, S(A_3, B) = 0.9345.$$

Then the pattern  $B$  is classified in  $A_3$  according to the recognition rule given by Eq. (15). Some medical diagnosis problems are very complex. Physicians need to use modern medical technologies to obtain a lot of information available to physicians for the help of decision, but the information is often incomplete, indeterminate and inconsistent. The SVNSSs proposed by Wang et al. [30] can be better choice to express this kind of information than Zadeh's fuzzy sets and intuitionistic fuzzy sets. Now in Example 4 we will utilize the proposed similarity measure to solve a class of medical diagnosis problems.

**Example 4** The medical diagnosis problem is adapted from De et al. [8]. Let  $Q = Q_1$  (Viral fever),  $Q_2$  (Malaria),  $Q_3$  (Typhoid),  $Q_4$  (Stomach problem),  $Q_5$  (Chest problem) be a set of diagnoses (diseases) and  $S = s_1$  (Temperature),  $s_2$  (Headache),  $s_3$  (Stomach pain),  $s_4$  (Cough),  $s_5$  (Chest pain) be a set of symptoms. Each diagnosis  $Q_i$  ( $i = 1, 2, 3, 4, 5$ ) can be represented by SVNSSs as follows:

$$Q_1 = \{ \langle s_1, 0.4, 0.6, 0.0 \rangle, \langle s_2, 0.3, 0.2, 0.5 \rangle, \langle s_3, 0.1, 0.2, 0.7 \rangle, \langle s_4, 0.4, 0.3, 0.3 \rangle, \langle s_5, 0.1, 0.2, 0.7 \rangle \}$$

$$Q_2 = \{ \langle s_1, 0.7, 0.3, 0.0 \rangle, \langle s_2, 0.2, 0.2, 0.6 \rangle, \langle s_3, 0.0, 0.1, 0.9 \rangle, \langle s_4, 0.7, 0.3, 0.0 \rangle, \langle s_5, 0.1, 0.1, 0.8 \rangle \}$$

$$Q_3 = \{ \langle s_1, 0.3, 0.4, 0.3 \rangle, \langle s_2, 0.6, 0.3, 0.1 \rangle, \langle s_3, 0.2, 0.1, 0.7 \rangle, \langle s_4, 0.2, 0.2, 0.6 \rangle, \langle s_5, 0.1, 0.0, 0.9 \rangle \}$$

$$Q_4 = \{ \langle s_1, 0.1, 0.2, 0.7 \rangle, \langle s_2, 0.2, 0.4, 0.4 \rangle, \langle s_3, 0.8, 0.2, 0.0 \rangle, \langle s_4, 0.2, 0.1, 0.7 \rangle, \langle s_5, 0.2, 0.1, 0.7 \rangle \}$$

$$Q_5 = \{ \langle s_1, 0.1, 0.1, 0.8 \rangle, \langle s_2, 0.0, 0.2, 0.8 \rangle, \langle s_3, 0.2, 0.0, 0.8 \rangle, \langle s_4, 0.2, 0.0, 0.8 \rangle, \langle s_5, 0.8, 0.1, 0.1 \rangle \}$$

Suppose there are two patients  $P_1$  and  $P_2$ , with respect to all the symptoms, can be represented by the following SVNSSs:

$$P_1 = \{ \langle s_1, 0.8, 0.1, 0.1 \rangle, \langle s_2, 0.6, 0.3, 0.1 \rangle, \langle s_3, 0.2, 0.0, 0.8 \rangle, \langle s_4, 0.6, 0.3, 0.1 \rangle, \langle s_5, 0.1, 0.3, 0.6 \rangle \}$$

$$P_2 = \{ \langle s_1, 0.0, 0.2, 0.8 \rangle, \langle s_2, 0.4, 0.4, 0.2 \rangle, \langle s_3, 0.6, 0.3, 0.1 \rangle, \langle s_4, 0.1, 0.7, 0.2 \rangle, \langle s_5, 0.1, 0.8, 0.1 \rangle \}$$

Our aim is to determine the patients  $P_1$  and  $P_2$  belong to which diagnosis of  $Q_j$  ( $j = 1, 2, 3, 4, 5$ ), respectively. Because the medical diagnosis problem is actually a pattern recognition problem, then we can use the diagnosis rule as follows: If  $k = \arg \max_{1 \leq j \leq 5} \{S_R(Q_j, P_i)\}$ , then we assign the patient  $P_1$  and  $P_2$  to the diagnosis  $Q_k$ .

$$S(Q_1, P_1) = 0.9160, S(Q_2, P_1) = 0.9360, \\ S(Q_3, P_1) = 0.9000, S(Q_4, P_1) = 0.7220$$

and  $S(Q_5, P_1) = 0.6640$ ,

$$S(Q_1, P_2) = 0.8103, S(Q_2, P_2) = 0.7259, \\ S(Q_3, P_2) = 0.8043, S(Q_4, P_2) = 0.8435$$

and  $S(Q_4, P_2) = 0.7553$  Then, By the above diagnosis rule, we can assign the patient  $P_1$  to the diagnosis  $Q_2$  (Malaria), and  $P_2$  to the diagnosis  $Q_4$  (Stomach problem). This result is in agreement with the one obtained in De et al. [8].

**Example 5** (Multi-attribute decision making) We consider a MADM problem adopted from Ye [33]. A manufacturing company wants to select the best global supplier from a set of four suppliers  $A = \{A_1, A_2, A_3, A_4\}$  whose core competencies are evaluated according to the four attributes  $O = \{O_1, O_2, O_3, O_4\}$ :  $o_1$  (the level of technology innovation),  $o_2$  (the control ability of flow),  $o_3$  (the ability of management),  $o_4$  (the level of service). The attributes are all benefit attributes. The weight vector for the four attributes determined by decision maker is

$$\mathbf{W} = (w_1, w_2, w_3, w_4)^T = (0.30, 0.25, 0.25, 0.20)^T$$

Suppose that the evaluation value of the alternative  $A_i (i = 1, 2, 3, 4)$  with respect to  $o_j (j = 1, 2, 3, 4)$  is a SNVN  $a_{ij} = \langle T_{ij}, I_{ij}, F_{ij} \rangle$ , which is obtained from a questionnaire of a domain expert. For example, when we ask the opinion of an expert about an alternative  $A_1$  with respect to an attribute  $o_1$ , he/she may say that the possibility in which the good statement is 0.5 and the poor statement is 0.3 and the degree in which he/she is not sure is 0.1. For the neutrosophic notation, it can be expressed as  $a_{11} = \langle T_{11}, I_{11}, F_{11} \rangle$ . The evaluation values are listed in Table 2.

Table 2: Evaluation values of each alternative with respect to each attribute

Alternatives	$o_1$	$o_2$	$o_3$	$o_4$
$A_1$	$\langle 0.75, 0.2, 0.3 \rangle$	$\langle 0.7, 0.2, 0.3 \rangle$	$\langle 0.65, 0.2, 0.25 \rangle$	$\langle 0.75, 0.2, 0.1 \rangle$
$A_2$	$\langle 0.8, 0.1, 0.2 \rangle$	$\langle 0.75, 0.2, 0.1 \rangle$	$\langle 0.75, 0.2, 0.1 \rangle$	$\langle 0.85, 0.1, 0.2 \rangle$
$A_3$	$\langle 0.7, 0.2, 0.2 \rangle$	$\langle 0.78, 0.2, 0.1 \rangle$	$\langle 0.85, 0.15, 0.1 \rangle$	$\langle 0.76, 0.2, 0.2 \rangle$
$A_4$	$\langle 0.8, 0.2, 0.1 \rangle$	$\langle 0.85, 0.2, 0.2 \rangle$	$\langle 0.7, 0.2, 0.2 \rangle$	$\langle 0.86, 0.1, 0.2 \rangle$

Now, we will propose a decision making method based on the proposed similarity measure to solve this problem and the detail steps is given as follows:

**Step 1** Determine the ideal solution  $A^*$  as follows:

$$A^* = (\langle T_j^*, I_j^*, F_j^* \rangle)_{1 \times 4} = (\langle \max_{1 \leq i \leq 4} (T_{ij}), \min_{1 \leq i \leq 4} (I_{ij}), \max_{1 \leq i \leq 4} (F_{ij}) \rangle)_{1 \times 4}$$

**Step 2** According to Eq. (4), calculate similarity measures between each alternative  $A_i (i = 1, 2, 3, 4)$  and the ideal solution  $A^*$  as follows:

$$S(A_1, A^*) = 0.9859, S(A_2, A^*) = 0.9955, S(A_3, A^*) = 0.9919$$

and  $S(A_4, A^*) = 0.9942$ .

**Step 3** According to the similarity measure values, the ranking order of the four suppliers is  $A_2 \succ A_4 \succ A_3 \succ A_1$ . Hence, the best supplier is  $A_2$ , which is in agreement with the result obtained by using weighted projection similarity and weighted Dice similarity methods (Ye [33]).

## 5 Conclusion

Neutrosophic sets are suitable to model the indeterminate and inconsistent information occurred in many practical problems. In this paper, we have proposed a new Chi-square distance-based similarity measure of SVNss. The new proposed similarity measure is a valid similarity measure and it can also overcome the counter-intuitive cases of the existing similarity measures by using some numerical examples. We have given the applications of the proposed similarity measure in pattern recognition and medical diagnosis. Furthermore, a multi-attribute decision making method is proposed through an example in which attribute values are expressed with SVNss.

As a prospect, the MADM method proposed in this paper could be applied to other MADM problems, such as the risk evaluation, credit evaluation. In the future work, we shall extend the proposed similarity to clustering analysis and image processing.

## Funding

This paper was supported by the National Natural Science Foundation of China (No.71661012) and Foundation of Jiangxi Educational Committee (No. GJJ170496).

## Author contributions

The authors contributed equally to this work.

## Conflict of interest

The authors declare no conflict of interest.

## Bibliography

- [1] Atanassov, K. T.(1986); Intuitionistic fuzzy sets, *Fuzzy Sets and Systems*, 20(1), 87–96, 1986.
- [2] Atanassov, K.; Gargov, G. (1989); Interval-valued intuitionistic fuzzy sets, *Fuzzy Sets and Systems*, 31(3), 343–349, 1989.
- [3] Boran, F. E; Akay, D. (2014); A biparametric similarity measure on intuitionistic fuzzy sets with applications to pattern recognition, *Information Sciences*, 255, 45–57, 2014.
- [4] Bozic, M.; Ducic, N.; Djordjevic, G.; Slavkovic, R. (2017); Optimization of Whег Robot Running with Simulation of Neuro-Fuzzy Control, *International Journal of Simulation Modelling*, 16(1), 19–30, 2017.
- [5] Chaira, T.; Panwar, A. (2014); An Atanassov’s intuitionistic fuzzy kernel clustering for medical image segmentation, *International Journal of Computational Intelligence Systems*, 7(2), 360–370, 2014.
- [6] Dalman H.; Gazel N.; Sivri M. (2016); A fuzzy set-based approach to multi-objective multi-item solid transportation problem under uncertainty, *International Journal of Fuzzy Systems*, 18(4), 716–729, 2016.
- [7] Das, S.; Guha, D.; Dutta, B. (2016); Medical diagnosis with the aid of using fuzzy logic and intuitionistic fuzzy logic, *Applied Intelligence*, 45(3), 850–867, 2016.
- [8] De S. K.; Biswas, R.; Roy, A. R. (2001); An application of intuitionistic fuzzy sets in medical diagnosis, *Fuzzy Sets & Systems*, 117(2), 209–213, 2001.
- [9] Dubois, D.; Prade, H.; Esteva, F. (2015); Fuzzy set modelling in case-based reasoning, *International Journal of Intelligent Systems*, 13(4), 345–373, 2015.
- [10] Gau, W. L.; Buehrer, D. J.(1993); Vague sets, *IEEE Transactions on Systems Man & Cybernetics*, 23(2), 610–614, 1993.
- [11] Huang, H. (2016); New distance measure of single-valued neutrosophic sets and its application, *International Journal of Intelligent Systems*, 31(10), 1021–1032, 2016.

- [12] Hung, W. L.; Yang, M. S. (2007); Similarity measures of intuitionistic fuzzy sets based on  $L_p$  metric, *International Journal of Approximate Reasoning*, 46(1), 120–136, 2007.
- [13] Hwang, C. M.; Yang, M. S. (2013); New construction for similarity measures between intuitionistic fuzzy sets based on lower, upper and middle fuzzy sets, *International Journal of Fuzzy Systems*, 15(3), 371–378, 2013.
- [14] Kar, A. K. (2015); A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network, *Journal of Computational Science*, 6, 23–33, 2015
- [15] Karaaslan, F. (2017); Correlation coefficients of single valued neutrosophic refined soft sets and their applications in clustering analysis, *Neural Computing & Applications*, 28(9), 2781–2793, 2017.
- [16] Le, H. S.; Phong, P. H. (2016); On the performance evaluation of intuitionistic vector similarity measures for medical diagnosis, *Journal of Intelligent & Fuzzy Systems*, 31(3), 1–12, 2016.
- [17] Li D. F.; Cheng C. T. (2012); New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions, *Pattern Recognition Letters*, 23(1), 221–225, 2002.
- [18] Li D. F.; Ren H. P. (2015); Multi-attribute decision making method considering the amount and reliability of intuitionistic fuzzy information, *Journal of Intelligent & Fuzzy Systems*, 28(4), 1877–1883, 2015.
- [19] Liu, P. D. (2016); The aggregation operators based on archimedean t-conorm and t-norm for single-valued neutrosophic numbers and their application to decision making, *International Journal of Fuzzy Systems*, 18(5), 1–15, 2016.
- [20] Meng, F.; Chen, X. (2016); Entropy and similarity measure of Atanassov's intuitionistic fuzzy sets and their application to pattern recognition based on fuzzy measures, *Pattern Analysis and Applications*, 19(1), 11–20, 2016.
- [21] Mousavi, M.; Yap, H. J.; Musa, S. N.; Dawal, S. Z. M. (2017); A Fuzzy Hybrid GA-PSO Algorithm for Multi-Objective AGV Scheduling in FMS, *International Journal of Simulation Modelling*, 16(1), 58–71, 2017.
- [22] Nguyen, H. (2016); A novel similarity/dissimilarity measure for intuitionistic fuzzy sets and its application in pattern recognition, *Expert Systems with Applications*, 45, 97–107, 2016
- [23] Peng, J. J.; Wang, J. Q.; Wang, J.; Zhang, H. Y.; Chen, X. H. (2016); Simplified neutrosophic sets and their applications in multi-criteria group decision-making problems, *International Journal of Systems Science*, 47(10), 2342–2358, 2016.
- [24] Perlibakas, V. (2004); Distance measures for PCA-based face recognition, *Pattern Recognition Letters*, 25(6), 711–724, 2004.
- [25] Pramanik, S.; Pramanik, S.; Giri, B. C. (2015); TOPSIS method for multi-attribute group decision-making under single-valued neutrosophic environment, *Neural Computing & Applications*, 27(3), 727–737, 2015.
- [26] Smarandache, F. (1999); A Unifying Field in Logics. Neutrosophy: Neutrosophic Probability, Set and Logic, *American Research Press*, 1999.

- [27] Tavana, M.; Zareinejad, M.; Caprio, D. D.; Kaviani M. A. (2016); An integrated intuitionistic fuzzy AHP and SWOT method for outsourcing reverse logistics, *Applied Soft Computing*, 40, 544–557, 2016.
- [28] Torra V. (2010); Hesitant fuzzy sets, *International Journal of Intelligent Systems*, 25(6), 529–539, 2010.
- [29] Vasavi, C.; Kumar G. S.; Murty, M. S. N.(2016); Generalized differentiability and integrability for fuzzy set-valued functions on time scales, *Soft Computing*, 20(3), 1093–1104, 2016.
- [30] Wang, H. B.; Smarandache, F.; Zhang Y. Q.; Sunderraman R. (2010); Single valued neutrosophic sets, *Multispace and Multistructure*, 4, 410–413, 2010.
- [31] Xu, Z. S.; Chen, J. (2008); An overview of distance and similarity measures of intuitionistic fuzzy sets, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 16(4), 529–555, 2008.
- [32] Xu Z. S.; Zhao N. (2016); Information fusion for intuitionistic fuzzy decision making: an overview, *Information Fusion*, 28, 10–23, 2016.
- [33] Ye, J.(2015); The generalized Dice measures for multiple attribute decision making under simplified neutrosophic environments, *Intelligent & Fuzzy Systems*, 31(1), 663–671, 2015.
- [34] Ye J.; Fu, J. (2016); Multi-period medical diagnosis method using a single valued neutrosophic similarity measure based on tangent function, *Comput Methods Programs Biomed*, 123, 142–149, 2016.
- [35] Ye, J. (2017); Single-valued neutrosophic similarity measures based on cotangent function and their application in the fault diagnosis of steam turbine, *Soft Computing*, 21(3), 1–9, 2017.