



# On Similarity Measures Between Pythagorean Fuzzy Sets Derived from Overlap and Grouping Functions

Zihang Jia<sup>1</sup> · Junsheng Qiao<sup>2</sup> · Minghao Chen<sup>1</sup>

Received: 10 July 2022 / Revised: 13 January 2023 / Accepted: 9 March 2023 / Published online: 25 April 2023  
© The Author(s) under exclusive licence to Taiwan Fuzzy Systems Association 2023

**Abstract** In 2013, Yager introduced the theory of Pythagorean fuzzy set (PFS) to generalize the application range of fuzzy sets. Compared with fuzzy set and intuitionistic fuzzy set (IFS), PFS is more capable to handle uncertainty in the real world. Similarity measure is an important topic in the theory of PFS and has been applied in various fields, such as classification, clustering, decision making, and so on. Meanwhile, overlap and grouping functions, as two novel continuous binary aggregation functions, have been discussed in the literature for applications in image processing, decision making, classification and so on. The main purpose of this paper is to construct similarity measures between PFSs based on overlap and grouping functions. Firstly, on the basis of overlap and grouping functions, we introduce two novel construction methods of similarity measure between PFSs, called *O*-similarity measure and *G*-similarity measure, respectively. Secondly, we use some numerical examples to illustrate characteristics of *O*-similarity measure in detail. Finally, we show the applications of *O*-similarity measure in classification and clustering to demonstrate effectiveness and superiority of

the proposed methods in the environment of expert assessments and data sets. The results of experiments show that *O*-similarity measure has better performance than some existing similarity measures under some conditions.

**Keywords** Pythagorean fuzzy set · Similarity measure · Classification · Clustering · Overlap function · Grouping function

## 1 Introduction

The theory of fuzzy set is proposed by Zadeh [1] in 1965 to describe uncertainty of phenomenon. After that, in 1986, Atanassov proposed the theory of IFS [2]. The IFS is characterized by the membership function and the non-membership function. Therefore, compared with the fuzzy set only characterized by the membership function, the IFS can depict the uncertainty of data more detailedly.

In 2013, Yager [3, 4] introduced the Pythagorean theorem into the theory of IFS and proposed the theory of PFS. As the extension of IFS, PFS has greater representation space for expressing the uncertainty of phenomenon [5]. Therefore, in the past decade, the theory of PFS was developing rapidly. Especially in applications, the theory of PFS has been widely applied in decision making [4, 6–8], attribute reduction [9], self-driving vehicle [10], conflict analysis [11], medical waste treatment [12], medical diagnosis [13], and so on.

Similarity measure is an important topic in the theory of fuzzy set and has been applied in various fields [5, 7, 14]. Ever since PFSs' appearance, many authors have paid attention to the similarity measure between PFSs. Peng et al. [14] proposed 12 similarity measures between PFSs which are derived from distance measures between PFSs. Zeng

✉ Minghao Chen  
chenmh130264@aliyun.com

Zihang Jia  
zihangjia@outlook.com

Junsheng Qiao  
jsqiao@nwnu.edu.cn; jsqiao@whu.edu.cn

<sup>1</sup> School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning, People's Republic of China

<sup>2</sup> College of Mathematics and Statistics, Northwest Normal University, Lanzhou 730070, Gansu, People's Republic of China

et al. [7] proposed some similarity measures between PFSs which take into account five parameters of PFSs. Zhang et al. [15] proposed four similarity measures between PFSs based on exponential function. Firozja et al. [16] proposed a construction method of similarity measure between PFSs based on continuous triangular conorms (t-conorms, for short).

Bustince et al. [17, 18] pointed out that the associativity property of triangular norms (t-norms, for short) and t-conorms is not required for many applications, such as decision making and classification problems. In 2010, Bustince et al. [17] proposed the definition of overlap functions that are not necessarily associative binary aggregation functions. After that, in 2012, Bustince et al. [18] proposed the definition of grouping functions. Since then, overlap and grouping functions have been successfully applied in numerous fields, such as classification [19, 20], decision making [21, 22], image processing [23], fuzzy community detection [18], and so on.

It is worth noting that the associativity property of t-conorm does not work in the construction method of similarity measure between PFSs given by [16]. Meanwhile, overlap and grouping function, as two novel binary continuous aggregation functions, are not necessarily associative and widely applied in numerous application problems. In addition, there are some application areas which similarity measure between PFSs and overlap and grouping functions work together, see, e.g., decision making [7, 21, 22] and classification [5, 19, 20]. Therefore, the purpose of this article is to introduce two novel methods to construct similarity measures between PFSs based on overlap functions and grouping functions. In this way, we can not only construct more useful similarity measures between PFSs from the mathematical point of view, but also provide more potential applications of similarity measures between PFSs in actual classification and decision making problems from the application viewpoint.

The rest of this work is organized as follows. In Sect. 2, we review some concepts used in this paper. In Sect. 3, we introduce two novel construction methods of similarity measure between PFSs based on overlap and grouping functions. In Sect. 4, we provide some numerical examples to illustrate the superiority and reasonability of proposed methods. In Sect. 5, we apply the proposed methods to classification and clustering problems. In Sect. 6, this article is summarized.

## 2 Preliminaries

### 2.1 Fuzzy Sets, Intuitionistic Fuzzy Sets and Pythagorean Fuzzy Sets

Let  $U$  and  $V$  be two nonempty universes and  $I$  be the unit interval  $[0, 1]$ . The notation  $\text{Map}(U, V)$  denotes the family

of all mappings from  $U$  to  $V$ . For each  $A \in \text{Map}(U, I)$ ,  $A$  is called a fuzzy set on  $U$  [1].

Let  $I_1^2 = \{(x_i, y_i) \in I \times I \mid 0 \leq x_i + y_i \leq 1\}$  and  $I_p^2 = \{(x_p, y_p) \in I \times I \mid 0 \leq x_p^2 + y_p^2 \leq 1\}$  where  $I_1^2$  and  $I_p^2$  denote the set of all intuitionistic fuzzy numbers (IFNs) [24] and Pythagorean fuzzy numbers (PFNs) [6], respectively. Actually, the set  $I_1^2$  is a subset of the set  $I_p^2$  (see Fig. 1).

Let  $p_1 = (\alpha_1, \beta_1)$  and  $p_2 = (\alpha_2, \beta_2)$  be two PFNs. Denote that

$$p_1^C = (\beta_1, \alpha_1);$$

$$p_1 \leq p_2 \iff \alpha_1 \leq \alpha_2 \text{ and } \beta_1 \geq \beta_2;$$

$$p_1 = p_2 \iff \alpha_1 = \alpha_2 \text{ and } \beta_1 = \beta_2.$$

**Definition 2.1** (See [2]) Let  $U$  be a universe. For each  $A \in \text{Map}(U, I_1^2)$ ,  $A$  is called an IFS on  $U$  where  $A(u) = (\mu_A(u), \nu_A(u))$  for each  $u \in U$ ,  $\mu_A(u)$  and  $\nu_A(u)$  denote membership degree and non-membership degree, respectively, of  $u$  in the IFS  $A$ . For each  $u \in U$ , the hesitancy degree of  $u$  denoted by  $\pi_A(u)$  is defined as  $\pi_A(u) = 1 - \mu_A(u) - \nu_A(u)$ .

**Definition 2.2** (See [3]) Let  $U$  be a universe. For each  $A \in \text{Map}(U, I_p^2)$ ,  $A$  is called a PFS on  $U$  where  $A(u) = (\mu_A(u), \nu_A(u))$  for each  $u \in U$ ,  $\mu_A(u)$  and  $\nu_A(u)$  denote membership degree and non-membership degree, respectively, of  $u$  in the PFS  $A$ . For each  $u \in U$ , the hesitancy degree of  $u$  denoted by  $\pi_A(u)$  is defined as  $\pi_A(u) = \sqrt{1 - (\mu_A(u))^2 - (\nu_A(u))^2}$ .

**Remark 2.1** It has been pointed out that the set  $I_1^2$  is a subset of the set  $I_p^2$ . Therefore, from Definitions 2.1 and 2.2, one has that IFS is a special case of PFS.

Let  $A, B \in \text{Map}(U, I_p^2)$  where  $A(u) = (\mu_A(u), \nu_A(u))$  and  $B(u) = (\mu_B(u), \nu_B(u))$  are two PFNs for each  $u \in U$ . Denote that

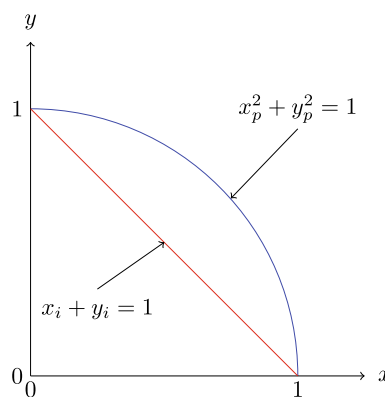


Fig. 1 The comparison between PFNs and IFNs

$$A^C(u) = (A(u))^C;$$

$$A \subseteq B \iff A(u) \leq B(u) \text{ for each } u \in U;$$

$$A = B \iff A(u) = B(u) \text{ for each } u \in U.$$

### 2.2 Fuzzy Negations, t-Norms, t-Conorms, Overlap and Grouping Functions

**Definition 2.3** (See [25]) A non-increasing mapping  $N : I \rightarrow I$  is called a fuzzy negation if  $N(0) = 1$  and  $N(1) = 0$ . Further, a fuzzy negation  $N$  is called strong or involutive if  $N(N(x)) = x$  for each  $x \in I$ .

For each  $x \in I$ ,  $N_S(x) = 1 - x$ ,  $N_{1,\lambda}(x) = \frac{1-x}{1+\lambda x}$  ( $\lambda \in ]-1, +\infty[$ ) and  $N_{2,\varphi}(x) = (1 - x^\varphi)^{\frac{1}{\varphi}}$  ( $\varphi \in ]0, +\infty[$ ) are three types strong fuzzy negations [26] (see Fig. 2).

**Lemma 2.1** Let  $N$  be a strong fuzzy negation. Then the following statements hold.

- (1)  $N(x) = 1$  if and only if  $x = 0$ ;
- (2)  $N(x) = 0$  if and only if  $x = 1$ ;

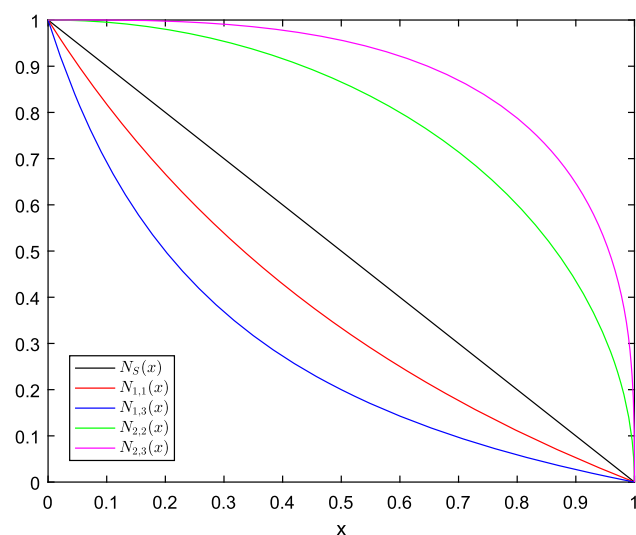
**Proof**

- (1) Let  $N(x) = 1$ . Since  $N(1) = 0$  and  $N(N(x)) = x$  for each  $x \in I$ , it follows that  $N(x) = 1 \iff N(N(x)) = N(1) \iff x = 0$ .

Therefore, one concludes that  $N(x) = 1$  if and only if  $x = 0$ .

- (2) It can be proved in a similar way as for statement (1). □

**Definition 2.4** (See [25]) A mapping  $T : I \times I \rightarrow I$  (resp.  $S : I \times I \rightarrow I$ ) is called a t-norm (resp. t-conorm) if it is



**Fig. 2** The curves of some strong fuzzy negations

commutative, is associative, is increasing and has 1 (resp. 0) as the neutral element.

Further, a t-norm  $T$  is called continuous if it is continuous in both arguments at the same time. A t-norm  $T$  is called positive if  $T(x, y) = 0$ ; then either  $x = 0$  or  $y = 0$ .

**Definition 2.5** (See [17]) A binary function  $O : I \times I \rightarrow I$  is called an overlap function if, for any  $x, y \in I$ , it satisfies the following conditions:

- (O1)  $O$  is commutative;
- (O2)  $O(x, y) = 0$  if and only if  $xy = 0$ ;
- (O3)  $O(x, y) = 1$  if and only if  $xy = 1$ ;
- (O4)  $O$  is increasing;
- (O5)  $O$  is continuous.

We now list some commonly used overlap functions from [27–30] as follows.

**Example 2.1**

- (1) Any continuous and positive t-norm is an overlap function.
- (2) For any  $p > 0$ , the function  $O_p : I \times I \rightarrow I$  given, for each  $x, y \in I$ , by

$$O_p(x, y) = x^p y^p$$

is an overlap function.

- (3) The function  $O_{DB} : I \times I \rightarrow I$  given, for each  $x, y \in I$ , by

$$O_{DB}(x, y) = \begin{cases} \frac{2xy}{x+y} & \text{if } x+y \neq 0, \\ 0 & \text{if } x+y = 0, \end{cases}$$

is an overlap function.

- (4) The function  $O_+ : I \times I \rightarrow I$  given, for each  $x, y \in I$ , by

$$O_+(x, y) = \frac{2xy}{1+xy}$$

is an overlap function.

- (5) The function  $O_{Mid} : I \times I \rightarrow I$  given, for each  $x, y \in I$ , by

$$O_{Mid}(x, y) = xy \frac{x+y}{2}$$

is an overlap function.

- (6) The function  $O_{mM} : I \times I \rightarrow I$  given, for each  $x, y \in I$ , by

$$O_{mM}(x, y) = \min\{x, y\} \max\{x^2, y^2\}$$

is an overlap function.

**Definition 2.6** (See [18]) A binary function  $G : I \times I \rightarrow I$  is called a grouping function if, for any  $x, y \in I$ , it satisfies the following conditions:

- (G1)  $G$  is commutative;
- (G2)  $G(x, y) = 0$  if and only if  $x = y = 0$ ;
- (G3)  $G(x, y) = 1$  if and only if  $x = 1$  or  $y = 1$ ;
- (G4)  $G$  is increasing;
- (G5)  $G$  is continuous.

**Lemma 2.2** (See [18]) *Let  $N : I \rightarrow I$  be a strong fuzzy negation. Then the following statements are equivalent.*

- (1) A binary function  $G : I \times I \rightarrow I$  is a grouping function.
- (2) There exists an overlap function  $O$  such that, for each  $x, y \in I$ ,  

$$G(x, y) = N(O(N(x), N(y))).$$

The grouping function  $G$  (resp. overlap function  $O$ ) given by Lemma 2.2 is called the  $N$ -dual grouping function of  $O$  (resp. overlap function of  $G$ ).

**Lemma 2.3** (See [18]) *Let  $G$  be a grouping function. If  $G$  is associative, then  $G$  becomes a  $t$ -conorm.*

### 3 Two Novel Construction Methods of Similarity Measure Between PFSs

To begin with, we review the concept of similarity measure between PFSs.

**Definition 3.1** (See [7]) A mapping  $\mathbb{S} : \text{Map}(U, I_P^2) \times \text{Map}(U, I_P^2) \rightarrow I$  is called a similarity measure between PFSs if, for each  $A, B, C \in \text{Map}(U, I_P^2)$ , it satisfies the following conditions:

- (SP1)  $0 \leq \mathbb{S}(A, B) \leq 1$ ;
- (SP2)  $\mathbb{S}(A, B) = \mathbb{S}(B, A)$ ;
- (SP3)  $\mathbb{S}(A, B) = 1$  if and only if  $A = B$ ;
- (SP4) If  $A \subseteq B \subseteq C$ , then  $\mathbb{S}(A, C) \leq \min\{\mathbb{S}(A, B), \mathbb{S}(B, C)\}$ .

#### 3.1 Similarity Measure Derived from Overlap Functions

In this subsection, we propose a novel method to construct similarity measure between PFSs based on overlap functions. To begin with, we propose the similarity measure between PFNs based on overlap functions as follows.

**Definition 3.2** Let  $k \in [1, +\infty[$ ,  $N_a$  and  $N_b$  be two strong fuzzy negations and  $O$  be an overlap function. The mapping  $\mathbb{S}_P^O : I_P^2 \times I_P^2 \rightarrow I$  is defined as

$$\mathbb{S}_P^O(p_1, p_2) = \sqrt[k]{O(N_a(|\alpha_1 - \alpha_2|^k), N_b(|\beta_1 - \beta_2|^k))} \quad (1)$$

where  $p_1 = (\alpha_1, \beta_1), p_2 = (\alpha_2, \beta_2) \in I_P^2$ .

**Theorem 3.1** *Let  $\mathbb{S}_P^O$  be the mapping given by Eq. (1). Then, for each  $p_1 = (\alpha_1, \beta_1), p_2 = (\alpha_2, \beta_2), p_3 = (\alpha_3, \beta_3) \in I_P^2$ , the following statements hold.*

- (1)  $0 \leq \mathbb{S}_P^O(p_1, p_2) \leq 1$ ;
- (2)  $\mathbb{S}_P^O(p_1, p_2) = \mathbb{S}_P^O(p_2, p_1)$ ;
- (3)  $\mathbb{S}_P^O(p_1, p_2) = 1$  if and only if  $p_1 = p_2$ ;
- (4) If  $p_1 \leq p_2 \leq p_3$ , then  $\mathbb{S}_P^O(p_1, p_3) \leq \min\{\mathbb{S}_P^O(p_1, p_2), \mathbb{S}_P^O(p_2, p_3)\}$ .

**Proof**

- (1) Since  $\alpha_1, \alpha_2, \beta_1, \beta_2 \in I$ , one has that  $|\alpha_1 - \alpha_2|, |\beta_1 - \beta_2| \in I$ . Further, the unit interval  $I$  is closed under  $k$ -power operation for each  $k \in ]0, +\infty[$ . In addition, from Definitions 2.3 and 2.5, since strong fuzzy negations  $N_a$  and  $N_b$  are two unary operations on  $I$  and overlap function  $O$  is a binary operation on  $I$ , one obtains that  $0 \leq \mathbb{S}_P^O(p_1, p_2) \leq 1$ .
- (2) From Eq. (1), it follows that

$$\begin{aligned} \mathbb{S}_P^O(p_1, p_2) &= \sqrt[k]{O(N_a(|\alpha_1 - \alpha_2|^k), N_b(|\beta_1 - \beta_2|^k))} \\ &= \sqrt[k]{O(N_a(|\alpha_2 - \alpha_1|^k), N_b(|\beta_2 - \beta_1|^k))} \\ &= \mathbb{S}_P^O(p_2, p_1). \end{aligned}$$

Therefore, one concludes that

- (3) From Eq. (1), it follows that

$$\begin{aligned} \mathbb{S}_P^O(p_1, p_2) = 1 &\iff \sqrt[k]{O(N_a(|\alpha_1 - \alpha_2|^k), N_b(|\beta_1 - \beta_2|^k))} = 1 \\ &\iff O(N_a(|\alpha_1 - \alpha_2|^k), N_b(|\beta_1 - \beta_2|^k)) = 1. \end{aligned}$$

Then, by condition (O3) of Definition 2.5, one has that

$$\mathbb{S}_P^O(p_1, p_2) = 1 \iff N_a(|\alpha_1 - \alpha_2|^k) = N_b(|\beta_1 - \beta_2|^k) = 1.$$

Further, from statement (1) of Lemma 2.1, one has that strong fuzzy negation  $N(x) = 1$  if and only if  $x = 0$ . Thus, since  $N_a$  and  $N_b$  are two strong fuzzy negations, it follows that

$$\begin{aligned} \mathbb{S}_P^O(p_1, p_2) = 1 &\iff |\alpha_1 - \alpha_2|^k = |\beta_1 - \beta_2|^k = 0 \\ &\iff p_1 = p_2. \end{aligned}$$

Therefore, one concludes that  $\mathbb{S}_P^O(p_1, p_2) = 1$  if and only if  $p_1 = p_2$ .

- (4) Let  $p_1 = (\alpha_1, \beta_1), p_2 = (\alpha_2, \beta_2), p_3 = (\alpha_3, \beta_3) \in I_P^2$  with  $p_1 \leq p_2 \leq p_3$ . Then, since  $k \in [1, +\infty[$ , it follows that

$$|\alpha_1 - \alpha_3|^k \geq \max\{|\alpha_1 - \alpha_2|^k, |\alpha_2 - \alpha_3|^k\}$$

and

$$|\beta_1 - \beta_3|^k \geq \max\{|\beta_1 - \beta_2|^k, |\beta_2 - \beta_3|^k\}.$$

Further, by Definition 2.3 and condition (O4) of Definition 2.5, since  $O$  is increasing and  $N_a, N_b$  are non-increasing, it follows that

$$\begin{aligned} \mathbb{S}_p^O(p_1, p_3) &= \sqrt[k]{O(N_a(|\alpha_1 - \alpha_3|^k), N_b(|\beta_1 - \beta_3|^k))} \\ &\leq \sqrt[k]{O(N_a(|\alpha_1 - \alpha_2|^k), N_b(|\beta_1 - \beta_2|^k))} \\ &= \mathbb{S}_p^O(p_1, p_2) \end{aligned}$$

and

$$\begin{aligned} \mathbb{S}_p^O(p_1, p_3) &= \sqrt[k]{O(N_a(|\alpha_1 - \alpha_3|^k), N_b(|\beta_1 - \beta_3|^k))} \\ &\leq \sqrt[k]{O(N_a(|\alpha_2 - \alpha_3|^k), N_b(|\beta_2 - \beta_3|^k))} \\ &= \mathbb{S}_p^O(p_2, p_3). \end{aligned}$$

Therefore, one concludes that  $\mathbb{S}_p^O(p_1, p_3) \leq \min\{\mathbb{S}_p^O(p_1, p_2), \mathbb{S}_p^O(p_2, p_3)\}$ . □

For convenience, the mapping  $\mathbb{S}_p^O$  defined in Definition 3.2 is always said to be the  $O$ -similarity measure between PFNs.

**Definition 3.3** Let  $U = \{u_1, u_2, \dots, u_n\}$  be a nonempty finite universe,  $\mathbb{S}_p^O$  be the  $O$ -similarity measure between PFNs given by Eq. (1) and  $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in ]0, 1]^n$  with  $\sum_{i=1}^n \omega_i = 1$ . The mapping  $\mathbb{S}^O : \text{Map}(U, I_p^2) \times \text{Map}(U, I_p^2) \rightarrow I$  is defined as

$$\mathbb{S}^O(A, B) = \sum_{i=1}^n \omega_i \mathbb{S}_p^O(A(u_i), B(u_i)) \tag{2}$$

where  $A, B \in \text{Map}(U, I_p^2)$ .

In the following,  $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in ]0, 1]^n$  is always said to be the weight vector of universe.

**Theorem 3.2** Let  $U = \{u_1, u_2, \dots, u_n\}$  be a nonempty finite universe and  $\mathbb{S}^O$  be the mapping given by Eq. (2). Then, for each  $A, B, C \in \text{Map}(U, I_p^2)$ , the following statements hold.

- (1)  $0 \leq \mathbb{S}^O(A, B) \leq 1$ ;
- (2)  $\mathbb{S}^O(A, B) = \mathbb{S}^O(B, A)$ ;
- (3)  $\mathbb{S}^O(A, B) = 1$  if and only if  $A = B$ ;
- (4) If  $A \subseteq B \subseteq C$ , then  $\mathbb{S}^O(A, C) \leq \min\{\mathbb{S}^O(A, B), \mathbb{S}^O(B, C)\}$ .

**Proof** Statements (1) and (2) can be immediately derived from Eqs. (1) and (2) and statements (1) and (2) of Theorem 3.1. Therefore, we only verify statements (3) and (4) as follows.

(C) Since  $\omega_i \in ]0, 1]$ , from Eqs. (1) and (2), it follows that

$$\begin{aligned} \mathbb{S}^O(A, B) = 1 &\iff \sum_{i=1}^n \omega_i \mathbb{S}_p^O(A(u_i), B(u_i)) = 1 \\ &\iff \mathbb{S}_p^O(A(u_i), B(u_i)) = 1 \quad \text{for } i = 1, 2, \dots, n. \end{aligned}$$

Further, from statement (3) of Theorem 3.1, it follows that

$$\mathbb{S}^O(A, B) = 1 \iff A(u_i) = B(u_i),$$

for each  $u_i \in U$ . Therefore, one concludes that  $\mathbb{S}^O(A, B) = 1$  if and only if  $A = B$ .

(D) Let  $A, B, C \in \text{Map}(U, I_p^2)$  with  $A \subseteq B \subseteq C$ . Then, from statement (4) of Theorem 3.1, one obtains that

$$\mathbb{S}_p^O(A(u_i), C(u_i)) \leq \min\{\mathbb{S}_p^O(A(u_i), B(u_i)), \mathbb{S}_p^O(B(u_i), C(u_i))\}$$

for each  $u_i \in U$ . Thus, for each  $u_i \in U$ , it follows that

$$\sum_{i=1}^n \omega_i \mathbb{S}_p^O(A(u_i), C(u_i)) \leq \sum_{i=1}^n \omega_i \mathbb{S}_p^O(A(u_i), B(u_i))$$

and

$$\sum_{i=1}^n \omega_i \mathbb{S}_p^O(A(u_i), C(u_i)) \leq \sum_{i=1}^n \omega_i \mathbb{S}_p^O(B(u_i), C(u_i)),$$

respectively. Therefore, one concludes that  $\mathbb{S}^O(A, C) \leq \min\{\mathbb{S}^O(A, B), \mathbb{S}^O(B, C)\}$ . □

From Theorem 3.2, one has that the mapping  $\mathbb{S}^O$  given by Definition 3.3 is a similarity measure between PFSs. The mapping  $\mathbb{S}^O$  is always said to be  $O$ -similarity measure between PFSs.

**Remark 3.1** In Definition 3.3,  $\omega_i \neq 0$  for each  $i \in \{1, 2, \dots, n\}$  is important and necessary. Consider the following case. Let  $U = \{u_1, u_2\}$ ,  $\omega = (0, 1)$ ,  $\mathbb{S}^O$  be the  $O$ -similarity measure between PFSs and  $A, B \in \text{Map}(U, I_p^2)$  where  $A = \frac{(0.5, 0.5)}{u_1} + \frac{(0.5, 0.5)}{u_2}$  and  $B = \frac{(1, 0)}{u_1} + \frac{(0.5, 0.5)}{u_2}$ . Then, by Eqs. (1) and (2), it follows that

$$\mathbb{S}^O(A, B) = \mathbb{S}_p^O(A(u_2), B(u_2)) = 1.$$

However,  $A$  is not equivalent to  $B$ . Therefore, under this case,  $\mathbb{S}^O$  not satisfies condition (SP3) of Definition 3.1, that is,  $\mathbb{S}^O$  is not a similarity measure between PFSs.

### 3.2 Similarity Measure Derived from Grouping Functions

**Definition 3.4** Let  $k \in [1, +\infty[$ ,  $N$  be a strong fuzzy negation and  $G$  be a grouping function. The mapping  $\mathbb{S}_p^G : I_p^2 \times I_p^2 \rightarrow I$  is defined as

$$\mathbb{S}_p^G(p_1, p_2) = \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right)\right)}, \tag{3}$$

where  $p_1 = (\alpha_1, \beta_1), p_2 = (\alpha_2, \beta_2) \in I_p^2$ .

**Theorem 3.3** Let  $\mathbb{S}_p^G$  be the mapping given by Eq. (3). For each  $p_1 = (\alpha_1, \beta_1), p_2 = (\alpha_2, \beta_2), p_3 = (\alpha_3, \beta_3) \in I_p^2$ , the following statements hold.

- (1)  $0 \leq \mathbb{S}_p^G(p_1, p_2) \leq 1$ ;
- (2)  $\mathbb{S}_p^G(p_1, p_2) = \mathbb{S}_p^G(p_2, p_1)$ ;
- (3)  $\mathbb{S}_p^G(p_1, p_2) = 1$  if and only if  $p_1 = p_2$ ;
- (4) If  $p_1 \leq p_2 \leq p_3$ , then  $\mathbb{S}_p^G(p_1, p_3) \leq \min\{\mathbb{S}_p^G(p_1, p_2), \mathbb{S}_p^G(p_2, p_3)\}$ .

**Proof**

- (1) From Definitions 2.3 and 2.6, it can be proved in a similar way as for statement (1) of Theorem 3.1.
- (2) From Eq. (3), it follows that

$$\begin{aligned} \mathbb{S}_p^G(p_1, p_2) &= \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right)\right)} \\ &= \sqrt[k]{N\left(G\left(|\alpha_2 - \alpha_1|^k, |\beta_2 - \beta_1|^k\right)\right)} \\ &= \mathbb{S}_p^G(p_2, p_1). \end{aligned}$$

Therefore, one concludes that

$$\mathbb{S}_p^G(p_1, p_2) = \mathbb{S}_p^G(p_2, p_1).$$

- (3) From Eq. (3), it follows that

$$\begin{aligned} \mathbb{S}_p^G(p_1, p_2) = 1 &\iff \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right)\right)} = 1 \\ &\iff N\left(G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right)\right) = 1. \end{aligned}$$

Further, from statement (1) of Lemma 2.1, one has that strong fuzzy negation  $N(x) = 1$  if and only if  $x = 0$ . Therefore, from condition (G2) of Definition 2.6, since  $N$  is a strong fuzzy negation, it follows that

$$\begin{aligned} \mathbb{S}_p^G(p_1, p_2) = 1 &\iff G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right) = 0 \\ &\iff |\alpha_1 - \alpha_2|^k = |\beta_1 - \beta_2|^k = 0 \\ &\iff p_1 = p_2. \end{aligned}$$

Therefore, one concludes that  $\mathbb{S}_p^G(p_1, p_2) = 1$  if and only if  $p_1 = p_2$ .

- (4) Let  $p_1 = (\alpha_1, \beta_1), p_2 = (\alpha_2, \beta_2), p_3 = (\alpha_3, \beta_3) \in I_p^2$  with  $p_1 \leq p_2 \leq p_3$ . Then, since  $k \in [1, +\infty[$ , it follows that

$$|\alpha_1 - \alpha_3|^k \geq \max\left\{|\alpha_1 - \alpha_2|^k, |\alpha_2 - \alpha_3|^k\right\}$$

and

$$|\beta_1 - \beta_3|^k \geq \max\left\{|\beta_1 - \beta_2|^k, |\beta_2 - \beta_3|^k\right\}.$$

Further, from Definition 2.3 and condition (G4) of Definition 2.6, since  $G$  is increasing and  $N$  is non-increasing, it follows that

$$\begin{aligned} \mathbb{S}_p^G(p_1, p_3) &= \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_3|^k, |\beta_1 - \beta_3|^k\right)\right)} \\ &\leq \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right)\right)} \\ &= \mathbb{S}_p^G(p_1, p_2) \end{aligned}$$

and

$$\begin{aligned} \mathbb{S}_p^G(p_1, p_3) &= \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_3|^k, |\beta_1 - \beta_3|^k\right)\right)} \\ &\leq \sqrt[k]{N\left(G\left(|\alpha_2 - \alpha_3|^k, |\beta_2 - \beta_3|^k\right)\right)} \\ &= \mathbb{S}_p^G(p_2, p_3). \end{aligned}$$

Therefore, one concludes that  $\mathbb{S}_p^G(p_1, p_3) \leq \min\{\mathbb{S}_p^G(p_1, p_2), \mathbb{S}_p^G(p_2, p_3)\}$ . □

For convenience, the mapping  $\mathbb{S}_p^G$  defined in Definition 3.4 is always said to be the  $G$ -similarity measure between PFNs.

**Remark 3.2**

- (1) In Definition 3.4, if we take  $N = N_s, k = 2$  and  $G$  as an associative grouping function, then, by Lemma 2.3, it follows that

$$\mathbb{S}_p^G(p_1, p_2) = \sqrt{1 - S\left(|\alpha_1 - \alpha_2|^2, |\beta_1 - \beta_2|^2\right)}, \tag{4}$$

that is, the  $G$ -similarity measure  $\mathbb{S}_p^G(p_1, p_2)$  becomes the  $s$ -similarity measure between PFNs given by Definition 3.1 of [16].

- (2) In Definitions 3.2 and 3.4, if we take  $N_a = N_b = N$  and  $O$  as  $N$ -dual overlap function of grouping function  $G$ , then, by Lemma 2.2, it follows that



$$\begin{aligned} \mathbb{S}_p^O(p_1, p_2) &= \sqrt[k]{O(N(|\alpha_1 - \alpha_2|^k), N(|\beta_1 - \beta_2|^k))} \\ &= \sqrt[k]{N\left(G\left(|\alpha_1 - \alpha_2|^k, |\beta_1 - \beta_2|^k\right)\right)} \\ &= \mathbb{S}_p^G(p_1, p_2). \end{aligned}$$

Thus, we establish the contact between  $O$ -similarity measure  $\mathbb{S}_p^O$  and  $G$ -similarity measure  $\mathbb{S}_p^G$  by Lemma 2.2. Therefore, in the following, we only discuss  $O$ -similarity measure.

**Definition 3.5** Let  $U = \{u_1, u_2, \dots, u_n\}$  be a nonempty finite universe,  $\mathbb{S}_p^G$  be the  $G$ -similarity measure between PFNs given by Eq. (3) and  $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in [0, 1]^n$  with  $\sum_{i=1}^n \omega_i = 1$ . The mapping  $\mathbb{S}^G : \text{Map}(U, I_p^2) \times \text{Map}(U, I_p^2) \rightarrow I$  is defined as

$$\mathbb{S}^G(A, B) = \sum_{i=1}^n \omega_i \mathbb{S}_p^G(A(u_i), B(u_i)), \tag{5}$$

where  $A, B \in \text{Map}(U, I_p^2)$ .

**Theorem 3.4** Let  $U = \{u_1, u_2, \dots, u_n\}$  be a nonempty finite universe and  $\mathbb{S}^G$  be the mapping given by Eq. (5). For each  $A, B, C \in \text{Map}(U, I_p^2)$ , the following statements hold.

- (1)  $0 \leq \mathbb{S}^G(A, B) \leq 1$ ;
- (2)  $\mathbb{S}^G(A, B) = \mathbb{S}^G(B, A)$ ;
- (3)  $\mathbb{S}^G(A, B) = 1$  if and only if  $A = B$ ;
- (4) If  $A \subseteq B \subseteq C$ , then  $\mathbb{S}^G(A, C) \leq \min\{\mathbb{S}^G(A, B), \mathbb{S}^G(B, C)\}$ .

**Proof** From Definitions 3.4 and 3.5 and Theorem 3.3, it can be proved in a similarity way as for Theorem 3.2.  $\square$

From Theorem 3.4, one has that the mapping  $\mathbb{S}^G$  given by Definition 3.5 is a similarity measure between PFNs. The mapping  $\mathbb{S}^G$  is always said to be  $G$ -similarity measure between PFNs.

### 4 Numerical Examples

In this section, to verify the properties of  $O$ -similarity between PFNs, we expound some numerical examples.

**Example 4.1** Let  $U = \{u\}$  be a universe,  $\alpha, \beta \in [0, 1]$  with  $0 \leq \alpha^2 + \beta^2 \leq 1$ , and  $A, B \in \text{Map}(U, I_p^2)$  where  $A = \frac{(\alpha, \beta)}{u}$  and  $B = \frac{(\beta, \alpha)}{u}$ .

If we take overlap function  $O$  as  $O_+$ , strong fuzzy negations  $N_a$  and  $N_b$  as  $N_s$  and  $k = 2$ , the degree of similarity between PFNs  $A$  and  $B$  calculated by  $\mathbb{S}^{O_+}$  are

shown in Fig. 3. We now show some analyses to this example from Fig. 3 as follows.

- (1)  $\mathbb{S}^{O_+}(A, B)$  gets the maximum value 1 if and only if  $A = B$  ( $\alpha = \beta$ );
- (2) If  $\alpha = 0, \beta = 1$  or  $\alpha = 1, \beta = 0$ , then  $\mathbb{S}^{O_+}(A, B)$  gets the minimum value 0;
- (3) For each  $\alpha, \beta \in [0, 1]$  with  $0 \leq \alpha^2 + \beta^2 \leq 1$ ,  $\mathbb{S}^{O_+}(A, B) \in [0, 1]$ ;
- (4)  $\mathbb{S}^{O_+}(A, B) = \mathbb{S}^{O_+}(B, A)$ , that is,  $\mathbb{S}^{O_+}$  is commutative.

Thus, Example 4.1 shows that the  $O$ -similarity measure  $\mathbb{S}^{O_+}$  satisfies conditions (SP1)–(SP3) of Definition 3.1.

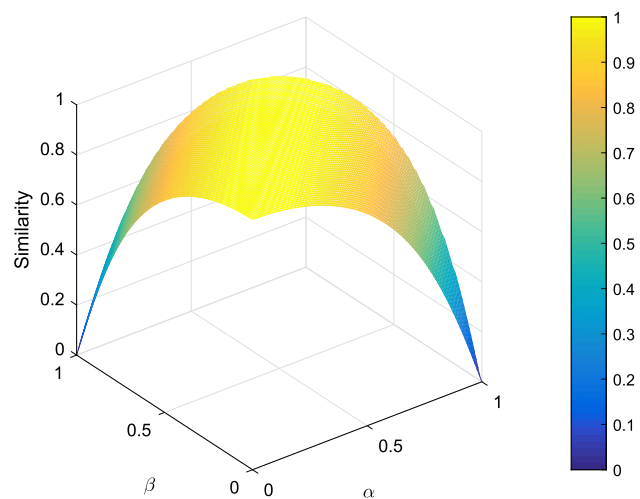
**Example 4.2** Let  $U = \{u\}$  be a universe and  $A, B, C \in \text{Map}(U, I_p^2)$  where  $A = \frac{(0.1, 0.9)}{u}$ ,  $B = \frac{(0.4, 0.8)}{u}$  and  $C = \frac{(0.6, 0.6)}{u}$ , respectively.

It is obvious that  $A \subseteq B \subseteq C$ . If we take overlap function  $O$  as  $O_{DB}$ , strong fuzzy negation  $N_a$  as  $N_s$ , strong fuzzy negation  $N_b$  as  $N_{1,2}$  and  $k = 1$ , then one has that  $\mathbb{S}^{O_{DB}}(A, C) = 0.4667$ ,  $\mathbb{S}^{O_{DB}}(A, B) = 0.7241$  and  $\mathbb{S}^{O_{DB}}(B, C) = 0.6667$ , respectively. Thus, one obtains that  $\mathbb{S}^{O_{DB}}(A, C) \leq \min\{\mathbb{S}^{O_{DB}}(A, B), \mathbb{S}^{O_{DB}}(B, C)\}$ .

Thus, Example 4.2 shows that the  $O$ -similarity measure  $\mathbb{S}^{O_{DB}}$  satisfies condition (SP4) of Definition 3.1.

Further, we show the superiority and reasonability of  $O$ -similarity measure as follows. Let  $U = \{u_1, u_2, \dots, u_n\}$  be a universe and  $A, B \in \text{Map}(U, I_p^2)$  where  $A(u_i) = (\mu_A(u_i), \nu_A(u_i))$  and  $B(u_i) = (\mu_B(u_i), \nu_B(u_i))$  for each  $u_i \in U$ . Some existing similarity measure between  $A$  and  $B$  are listed in Table 1.

**Example 4.3** This numerical example is adopted from [36] which includes six cases. All PFNs in this example are



**Fig. 3** The similarity measure of PFNs given in Example 4.1

on universe  $U = \{u\}$ . The degree of similarity between PFSs calculated by existing methods are shown in Table 2.

Then, we construct some  $O$ -similarity measures to calculate the degree of similarity of these cases. In this example, we always take  $k = 2$  and  $N_a$  as  $N_s$ . If we determine the strong fuzzy negation  $N_b$ , then the corresponding  $O$ -similarity measure is denoted  $\mathbb{S}_{N_b}^O$ . Then the degree of similarity between PFSs calculated by different  $O$ -similarity measures are shown in Table 3.

In the end, we show some analyses based on Tables 2 and 3.

- (1) For case 1, the degree of similarity between PFSs  $A$  and  $B$  calculated by  $\mathbb{S}_C$  and  $\mathbb{S}_P^2$  are 1. However,  $A$  is not equivalent to  $B$ . Thus, the  $O$ -similarity measures used in this example are better than  $\mathbb{S}_C$  and  $\mathbb{S}_P^2$  under case 1.
- (2) For cases 1 and 2, the results calculated by  $\mathbb{S}_{HK}$  are same. However, cases 1 and 2 are different. Meanwhile, the degree of similarity of cases 1 and 2 cannot be calculated by  $\mathbb{S}_Y$  due to “the division by zero problem”. Thus, the  $O$ -similarity measures used in this example are better than  $\mathbb{S}_{HK}$  and  $\mathbb{S}_Y$  under cases 1 and 2.

- (3) For cases 3 and 4, the similarity measures  $\mathbb{S}_{HY1}$ ,  $\mathbb{S}_{HY2}$ ,  $\mathbb{S}_{HY3}$ ,  $\mathbb{S}_C$ ,  $\mathbb{S}_L$ ,  $\mathbb{S}_{HK}$ ,  $\mathbb{S}_{N_s}^{O_1}$ ,  $\mathbb{S}_{N_s}^{O_2}$ ,  $\mathbb{S}_{N_s}^{O_{0.5}}$ ,  $\mathbb{S}_{N_s}^{O_+}$ ,  $\mathbb{S}_{N_s}^{O_{DB}}$ ,  $\mathbb{S}_{N_s}^{O_{MM}}$  and  $\mathbb{S}_{N_s}^{O_{Mid}}$  cannot distinguish these two different cases. Thus, these similarity measures need to improved. It is worth noting that the other  $O$ -similarity measures used in this example can distinguish these two different cases. That is why two strong fuzzy negations  $N_a$  and  $N_b$  are considered in Definition 3.2.
- (4) For cases 5 and 6,  $O$ -similarity measures cannot distinguish these two two cases. Thus, under these cases, some existing similarity measures are better than  $O$ -similarity measures.

**Example 4.4** In this example, we consider five special cases. In this example, all PFSs are on universe  $U = \{u\}$ . Further, we always take  $k = 1$  and strong fuzzy negations  $N_a = N_b = N_s$ . Then the degree of similarity between PFSs calculated by different similarity measures are shown in Table 4. We now show some analyses based on Table 4 as follows:

For cases 1 and 2, the similarity measures  $\mathbb{S}_P^1$ ,  $\mathbb{S}_P^2$ ,  $\mathbb{S}_P^3$ ,  $\mathbb{S}_P^4$ ,  $\mathbb{S}_P^5$ ,  $\mathbb{S}_Z^1$ ,  $\mathbb{S}_Z^2$ ,  $\mathbb{S}_Z^3$  and  $\mathbb{S}_P^4$  cannot distinguish these two

**Table 1** Some existing similarity measures between PFSs

Author	Similarity measure between PFSs
Hung and Yang [31]	$\mathbb{S}_{HY1}(A, B) = 1 - \frac{1}{n} \sum_{i=1}^n \max\{ \mu_A(u_i) - \mu_B(u_i) ,  v_A(u_i) - v_B(u_i) \}$ $\mathbb{S}_{HY2}(A, B) = \frac{\exp(\mathbb{S}_{HY1}(A, B) - 1) - \exp(-1)}{1 - \exp(-1)}$ $\mathbb{S}_{HY3}(A, B) = \frac{\mathbb{S}_{HY1}(A, B)}{2 - \mathbb{S}_{HY1}(A, B)}$
Ye [32]	$\mathbb{S}_Y(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{\mu_A(u_i)\mu_B(u_i) + v_A(u_i)v_B(u_i)}{\sqrt{\mu_A^2(u_i) + v_A^2(u_i)}\sqrt{\mu_B^2(u_i) + v_B^2(u_i)}}$
Chen [33]	$\mathbb{S}_C(A, B) = 1 - \frac{1}{2n} \sum_{i=1}^n  \mu_A(u_i) - v_A(u_i) - (\mu_B(u_i) - v_B(u_i)) $
Li et al. [34]	$\mathbb{S}_L(A, B) = 1 - \sqrt{\frac{\sum_{i=1}^n ((\mu_A(u_i) - \mu_B(u_i))^2 + (v_A(u_i) - v_B(u_i))^2)}{2n}}$
Hong and Kim [35]	$\mathbb{S}_{HK}(A, B) = 1 - \frac{1}{2n} \sum_{i=1}^n ( \mu_A(u_i) - \mu_B(u_i)  +  v_A(u_i) - v_B(u_i) )$
Peng et al. [14]	$\mathbb{S}_P^1(A, B) = 1 - \frac{1}{2n} \sum_{i=1}^n ( \mu_A^2(u_i) - \mu_B^2(u_i)  +  v_A^2(u_i) - v_B^2(u_i)  +  \pi_A^2(u_i) - \pi_B^2(u_i) )$ $\mathbb{S}_P^2(A, B) = 1 - \frac{1}{2n} \sum_{i=1}^n  \mu_A^2(u_i) - \mu_B^2(u_i) - (v_A^2(u_i) - v_B^2(u_i)) $ $\mathbb{S}_P^3(A, B) = 1 - \frac{1}{n} \sum_{i=1}^n \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) \}$ $\mathbb{S}_P^4(A, B) = \frac{1}{n} \sum_{i=1}^n \frac{1 - \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) \}}{1 + \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) \}}$ $\mathbb{S}_P^5(A, B) = \frac{\sum_{i=1}^n (1 - \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) \})}{\sum_{i=1}^n (1 + \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) \})}$
Zhang et al. [15]	$\mathbb{S}_Z^1(A, B) = \frac{1}{n} \sum_{i=1}^n \left[ 2^{1 - \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) ,  \pi_A^2(u_i) - \pi_B^2(u_i) \}} - 1 \right]$ $\mathbb{S}_Z^2(A, B) = \frac{1}{n} \sum_{i=1}^n \left[ 2^{1 - \frac{1}{2}( \mu_A^2(u_i) - \mu_B^2(u_i)  +  v_A^2(u_i) - v_B^2(u_i)  +  \pi_A^2(u_i) - \pi_B^2(u_i) )} - 1 \right]$ $\mathbb{S}_Z^3(A, B) = \frac{1}{n} \sum_{i=1}^n \left[ 2^{1 - \max\{ \mu_A^2(u_i) - \mu_B^2(u_i) ,  v_A^2(u_i) - v_B^2(u_i) \}} - 1 \right]$ $\mathbb{S}_Z^4(A, B) = \frac{1}{n} \sum_{i=1}^n \left[ 2^{1 - \frac{1}{2}( \mu_A^2(u_i) - \mu_B^2(u_i)  +  v_A^2(u_i) - v_B^2(u_i) )} - 1 \right]$



**Table 2** The degree of similarity between PFSs calculated by existing methods in Example 4.3

Case	1	2	3	4	5	6	Counterintuitive case
	$A(u) = (0.5, 0.5)$ $B(u) = (0, 0)$	$A(u) = (0.6, 0.4)$ $B(u) = (0, 0)$	$A(u) = (0, 0.87)$ $B(u) = (0.28, 0.55)$	$A(u) = (0.6, 0.27)$ $B(u) = (0.28, 0.55)$	$A(u) = (0.125, 0.075)$ $B(u) = (0.175, 0.025)$	$A(u) = (0.5, 0.45)$ $B(u) = (0.55, 0.4)$	
$S_{HY1}$ [31]	0.5000	0.4000	<b>0.6800</b>	<b>0.6800</b>	<b>0.9500</b>	<b>0.9500</b>	Cases 3, 4, 5 and 6
$S_{HY2}$ [31]	0.3775	0.2862	<b>0.5668</b>	<b>0.5668</b>	<b>0.9228</b>	<b>0.9228</b>	Cases 3, 4, 5 and 6
$S_{HY3}$ [31]	0.3333	0.2500	<b>0.5152</b>	<b>0.5152</b>	<b>0.9048</b>	<b>0.9048</b>	Cases 3, 4, 5 and 6
$S_Y$ [32]	<i>NaN</i>	<i>NaN</i>	0.8912	0.7794	0.9216	0.9946	Cases 1 and 2
$S_C$ [33]	<b>1.0000</b>	0.9000	<b>0.7000</b>	<b>0.7000</b>	<b>0.9500</b>	<b>0.9500</b>	Cases 1, 3, 4, 5 and 6
$S_L$ [34]	0.5000	0.4901	<b>0.6993</b>	<b>0.6993</b>	<b>0.9500</b>	<b>0.9500</b>	Cases 3, 4, 5 and 6
$S_{HK}$ [35]	<b>0.5000</b>	<b>0.5000</b>	<b>0.7000</b>	<b>0.7000</b>	<b>0.9500</b>	<b>0.9500</b>	Cases 1, 2, 3, 4, 5 and 6
$S_P^1$ [14]	0.5000	0.4800	0.5456	0.7184	0.9850	0.9475	-
$S_P^2$ [14]	<b>1.0000</b>	0.9000	0.7336	0.7444	0.9900	0.9525	Case 1
$S_P^3$ [14]	0.7500	0.6400	0.5456	0.7184	0.9850	0.9475	-
$S_P^4$ [14]	0.6000	0.4706	0.3751	0.5605	0.9704	0.9002	-
$S_P^5$ [14]	0.6000	0.4706	0.3751	0.5605	0.9704	0.9002	-
$S_Z^1$ [15]	0.4142	0.3947	0.4596	0.6454	0.9793	0.9285	-
$S_Z^2$ [15]	0.4142	0.3947	0.4596	0.6454	0.9793	0.9285	-
$S_Z^3$ [15]	0.6818	0.5583	0.4596	0.6454	0.9793	0.9285	-
$S_Z^4$ [15]	0.6818	0.6702	0.6628	0.6753	0.9862	0.9352	-

“Bolditalic” denotes unreasonable results. “NaN” denotes it cannot compute the degree of similarity due to “the division by zero problem”

**Table 3** The degree of similarity between PFSs calculated by  $O$ -similarity measures in Example 4.3

Case	1	2	3	4	5	6	Counterintuitive case	
$S_{N_1}^{O_1}$	0.7500	$A(u) = (0.5, 0.5)$ $B(u) = (0, 0)$	$A(u) = (0.6, 0.4)$ $B(u) = (0, 0)$	$A(u) = (0, 0.87)$ $B(u) = (0.28, 0.55)$	$A(u) = (0.6, 0.27)$ $B(u) = (0.28, 0.55)$	$A(u) = (0.125, 0.075)$ $B(u) = (0.175, 0.025)$	$A(u) = (0.5, 0.45)$ $B(u) = (0.55, 0.4)$	Cases 3, 4, 5 and 6
$S_{N_{1,1}}^{O_1}$	0.6708	0.6808	0.8662	0.8758	<b>0.9963</b>	<b>0.9963</b>	<b>0.9975</b>	Cases 5 and 6
$S_{N_2}^{O_2}$	0.5625	0.5376	<b>0.8272</b>	<b>0.8272</b>	<b>0.9950</b>	<b>0.9950</b>	<b>0.9950</b>	Cases 3, 4, 5 and 6
$S_{N_{1,2}}^{O_2}$	0.3750	0.4073	0.6866	0.7151	<b>0.9901</b>	<b>0.9901</b>	<b>0.9901</b>	Cases 5 and 6
$S_{N_1}^{O_{0.5}}$	0.8660	0.8563	<b>0.9537</b>	<b>0.9537</b>	<b>0.9987</b>	<b>0.9987</b>	<b>0.9987</b>	Cases 3, 4, 5 and 6
$S_{N_{1,3}}^{O_{0.5}}$	0.7530	0.7763	0.8919	0.9046	<b>0.9969</b>	<b>0.9969</b>	<b>0.9969</b>	Cases 5 and 6
$S_{N_1}^{O_+}$	0.8485	0.8362	<b>0.9515</b>	<b>0.9515</b>	<b>0.9987</b>	<b>0.9987</b>	<b>0.9987</b>	Cases 3, 4, 5 and 6
$S_{N_{2,2}}^{O_+}$	0.9173	0.8800	0.9780	0.9719	<b>0.9994</b>	<b>0.9994</b>	<b>0.9994</b>	Cases 5 and 6
$S_{N_2}^{O_{0.8}}$	0.8660	0.8523	<b>0.9536</b>	<b>0.9536</b>	<b>0.9987</b>	<b>0.9987</b>	<b>0.9987</b>	Cases 3, 4, 5 and 6
$S_{N_{2,2}}^{O_{0.8}}$	0.9194	0.8812	0.9781	0.9719	<b>0.9994</b>	<b>0.9994</b>	<b>0.9994</b>	Cases 5 and 6
$S_{N_1}^{O_{0.01}}$	0.6495	0.6720	<b>0.8731</b>	<b>0.8731</b>	<b>0.9963</b>	<b>0.9963</b>	<b>0.9963</b>	Cases 3, 4, 5 and 6
$S_{N_{2,3}}^{O_{0.01}}$	0.8615	0.7989	0.9597	0.9473	<b>0.9987</b>	<b>0.9987</b>	<b>0.9987</b>	Cases 5 and 6
$S_{N_1}^{O_{0.01}^{id}}$	0.6495	0.6307	<b>0.8674</b>	<b>0.8674</b>	<b>0.9963</b>	<b>0.9963</b>	<b>0.9963</b>	Cases 3, 4, 5 and 6
$S_{N_{2,4}}^{O_{0.01}^{id}}$	0.8095	0.7243	0.9410	0.9228	<b>0.9981</b>	<b>0.9981</b>	<b>0.9981</b>	Cases 5 and 6

“**Bolditalic**” denotes unreasonable results

**Table 4** The degree of similarity between PFSs in Example 4.4

Case	1 $A(u) = (0.5, 0.1)$ $B(u) = (0.5, 0.3)$	2 $A(u) = (0.6, 0.3)$ $B(u) = (0.6, \frac{\sqrt{17}}{10})$	3 $A(u) = (0.1, 0.5)$ $B(u) = (0.2, 0.7)$	4 $A(u) = (0.4, 0.7)$ $B(u) = (\frac{\sqrt{19}}{10}, \frac{\sqrt{73}}{10})$	5 $A(u) = (0.4, 0.7)$ $B(u) = (\frac{\sqrt{13}}{10}, 0.5)$	Counterintuitive case
$\mathbb{S}_P^1$ [14]	<b>0.9200</b>	<b>0.9200</b>	<b>0.7300</b>	<b>0.7300</b>	<b>0.7300</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_P^2$ [14]	<b>0.9600</b>	<b>0.9600</b>	<b>0.8950</b>	<b>0.8950</b>	<b>0.8950</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_P^3$ [14]	<b>0.9200</b>	<b>0.9200</b>	<b>0.7600</b>	<b>0.7600</b>	<b>0.7600</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_P^4$ [14]	<b>0.8519</b>	<b>0.8519</b>	<b>0.6129</b>	<b>0.6129</b>	<b>0.6129</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_P^5$ [14]	<b>0.8519</b>	<b>0.8519</b>	<b>0.6129</b>	<b>0.6129</b>	<b>0.6129</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_Z^1$ [15]	<b>0.8921</b>	<b>0.8921</b>	<b>0.6586</b>	<b>0.6586</b>	<b>0.6586</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_Z^2$ [15]	<b>0.8921</b>	<b>0.8921</b>	<b>0.6586</b>	<b>0.6586</b>	<b>0.6586</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_Z^3$ [15]	<b>0.8921</b>	<b>0.8921</b>	<b>0.6935</b>	<b>0.6935</b>	<b>0.6935</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}_Z^4$ [15]	<b>0.9453</b>	<b>0.9453</b>	<b>0.8213</b>	<b>0.8213</b>	<b>0.8213</b>	Cases 1, 2, 3, 4 and 5
$\mathbb{S}^{O_1}$	0.8000	0.8877	0.7200	0.8153	0.7684	–
$\mathbb{S}^{O_2}$	0.6400	0.7880	0.5184	0.6646	0.5905	–
$\mathbb{S}^{O_{0.5}}$	0.8944	0.9422	0.8485	0.9029	0.8766	–
$\mathbb{S}^{O_+}$	0.8889	0.9405	0.8372	0.8982	0.8691	–
$\mathbb{S}^{O_{DB}}$	0.8889	0.9405	0.8471	0.9010	0.8730	–
$\mathbb{S}^{O_{MM}}$	0.8000	0.8877	0.6480	0.7860	0.7381	–
$\mathbb{S}^{O_{Mid}}$	0.7200	0.8378	0.6120	0.7377	0.6764	–

“Bolditalic” denotes unreasonable results

different cases, although some of them consider the degree of hesitancy. For cases 3, 4 and 5, the same problem arises with these similarity measures. Therefore,  $O$ -similarity measures are better than these similarity measures under these cases.

### 5 Applications of $O$ -Similarity Measure in Classification and Clustering Problems

In this section, we design a series of experiments to verify the performance of  $O$ -similarity measure. These experiments are implemented on MATLAB R2016a, which is installed on a private computer with 64-bit Windows 10 and Intel Core i5-8300H CPU at 2.30 GHz, 16.0 GB RAM.

#### 5.1 Application in Classification on the Basis of a Classical Example

Let  $U = \{u_1, u_2, u_3, u_4\}$  be a universe and  $P_i$  be known target with class label  $C_i \in \text{Map}(U, I_p^2)$  ( $i = 1, 2, 4$ ) where

$$C_1 = \frac{(0.3, 0.3)}{u_1} + \frac{(0.4, 0.4)}{u_2} + \frac{(0.4, 0.4)}{u_3} + \frac{(0.4, 0.4)}{u_4},$$

$$C_2 = \frac{(0.5, 0.5)}{u_1} + \frac{(0.1, 0.1)}{u_2} + \frac{(0.5, 0.5)}{u_3} + \frac{(0.1, 0.1)}{u_4} \text{ and}$$

$$C_3 = \frac{(0.5, 0.4)}{u_1} + \frac{(0.4, 0.5)}{u_2} + \frac{(0.3, 0.3)}{u_3} + \frac{(0.2, 0.2)}{u_4},$$

respectively. The unknown target  $Q$  is given by

$$Q = \frac{(0.4, 0.4)}{u_1} + \frac{(0.5, 0.5)}{u_2} + \frac{(0.2, 0.2)}{u_3} + \frac{(0.3, 0.3)}{u_4}.$$

The goal is to identify which class  $Q$  belongs to. This example is also used in [14]. In this subsection, we always take strong fuzzy negations  $N_a$  and  $N_b$  as  $N_s$ ,  $\omega = (0.25, 0.25, 0.25, 0.25)$  and  $k = 1$ . Then the classified result of different similarity measures are shown in Table 5.

From Table 5, we show some analyses for this example as follows. Most similarity measures can determine the classified results and have the same classified result. There are only three similarity measures cannot determine the classified results, such as  $\mathbb{S}_Y$ ,  $\mathbb{C}$  and  $\mathbb{S}_P^2$ . This result shows that the  $O$ -similarity measure is reasonable and effective, as well as being better superior to some similarity measures.

**Table 5** The classified results of different similarity measures in Sect. 5.1

	$\mathbb{S}^O(C_1, Q)$	$\mathbb{S}^O(C_2, Q)$	$\mathbb{S}^O(C_3, Q)$	Target
$\mathbb{S}_{HY1}$ [31]	0.8750	0.7500	<b>0.9000</b>	$P_3$
$\mathbb{S}_{HY2}$ [31]	0.8141	0.6501	<b>0.8495</b>	$P_3$
$\mathbb{S}_{HY3}$ [31]	0.7778	0.6000	<b>0.8182</b>	$P_3$
$\mathbb{S}_Y$ [32]	<b>1.0000</b>	<b>1.0000</b>	0.9969	<b>Cannot be determined</b>
$\mathbb{S}_C$ [33]	<b>1.0000</b>	<b>1.0000</b>	0.9750	<b>Cannot be determined</b>
$\mathbb{S}_L$ [34]	0.8677	0.7261	<b>0.9134</b>	$P_3$
$\mathbb{S}_{HK}$ [35]	0.8750	0.7500	<b>0.9250</b>	$P_3$
$\mathbb{S}_P^1$ [14]	0.8250	0.6900	<b>0.9050</b>	$P_3$
$\mathbb{S}_P^2$ [14]	<b>1.0000</b>	<b>1.0000</b>	0.9775	<b>Cannot be determined</b>
$\mathbb{S}_P^3$ [14]	0.9125	0.8450	<b>0.9300</b>	$P_3$
$\mathbb{S}_P^4$ [14]	0.8397	0.7381	<b>0.8698</b>	$P_3$
$\mathbb{S}_P^5$ [14]	0.8391	0.7316	<b>0.8692</b>	$P_3$
$\mathbb{S}_Z^1$ [15]	0.7722	0.6211	<b>0.8726</b>	$P_3$
$\mathbb{S}_Z^2$ [15]	0.7722	0.6211	<b>0.8726</b>	$P_3$
$\mathbb{S}_Z^3$ [15]	0.8825	0.7984	<b>0.9055</b>	$P_3$
$\mathbb{S}_Z^4$ [15]	0.8825	0.7984	<b>0.9352</b>	$P_3$
$\mathbb{S}^{O_{0.5}}$	0.9912	0.9613	<b>0.9962</b>	$P_3$
$\mathbb{S}^{O_1}$	0.9825	0.9250	<b>0.9925</b>	$P_3$
$\mathbb{S}^{O_2}$	0.9655	0.8588	<b>0.9850</b>	$P_3$
$\mathbb{S}^{O_+}$	0.9911	0.9589	<b>0.9962</b>	$P_3$
$\mathbb{S}^{O_{DB}}$	0.9912	0.9613	<b>0.9962</b>	$P_3$
$\mathbb{S}^{O_{mM}}$	0.9739	0.8909	<b>0.9900</b>	$P_3$
$\mathbb{S}^{O_{Mid}}$	0.9739	0.8909	<b>0.9888</b>	$P_3$

In table 5, the bold content in columns 2-4 represents the highest similarity. The bold content in column 5 indicates unreasonable results

### 5.2 Application in Classification on the Basis of Six Datasets from the UCI Database

In this section, we show the effectiveness of the  $O$ -similarity measures by repeating 2-Fold Cross Validation in the environment of six datasets. The six datasets are downloaded from the UCI machine learning data repository (<http://archive.ics.uci.edu/ml>). The detailed information of the datasets is described in Table 6.

We use the method given by [5] to generate PFSs for six datasets. Then we apply various existing similarity measure and  $O$ -similarity measures to the classification method given by [5] (see Algorithm 1) under fuzzy environment and repeat 100 times 2-Fold Cross Validation for each experiments. In this subsection, we always take strong fuzzy negations  $N_a$  and  $N_b$  as  $N_s$ ,  $k = 1$  and  $\omega_i = \frac{1}{n}$  for  $i = 1, 2, \dots, n$  whenever the universe  $U = \{u_1, u_2, \dots, u_n\}$ . The classification results are shown in Table 7.

**Algorithm 1** [5] The classification method in the fuzzy framework

**Input:** The set of known target  $T = \{T_1, T_2, \dots, T_m\}$ , the set of unknown target  $T_U = \{T_{u1}, T_{u2}, \dots, T_{un}\}$  and a similarity measure  $\mathbb{S}$ .

**Output:** Classification results.

- 1: **for**  $i=1$  to  $n$  **do**
- 2:     **for**  $j=1$  to  $m$  **do**
- 3:         Measure the degree of similarity between  $T_{ui}$  and  $T_j$  by  $\mathbb{S}$ ;
- 4:     **end for**
- 5:     Select the maximum  $\mathbb{S}(T_{ui}, T_\alpha)$ ;
- 6:     Classify the unknown target  $T_{ui}$  to known target  $T_\alpha$ ;
- 7: **end for**

We now show some analyses on the basis of Table 7 as follows:

- (1) For datasets “Wine”, “Blood” and “Banknote”, there is always one  $O$ -similarity measure which has the highest accuracy of classification in compared with some existing similarity measures.
- (2) For datasets “Seed”, “Haberman” and “Wireless”, although the highest accuracy of classification is always given by  $\mathbb{S}_Y$ , there is always one  $O$ -similarity

measure  $\mathbb{S}^O$  such that the accuracy of classification given by  $\mathbb{S}^O$  is second only to  $\mathbb{S}_Y$ .

Therefore, on the basis of the above analyses, the performance of the  $O$ -similarity measures is superior to other methods under some conditions.

### 5.3 Application in Clustering on the Basis of a Classical Example

In this subsection, we use the example given by [14] to illustrate the application of  $O$ -similarity measure in clustering. To facilitate understanding, we briefly introduce the dataset as follows.

Software project [14]: The dataset contains ten software projects  $P_i$  ( $i = 1, 2, \dots, 10$ ) and five criteria such as economic feasibility  $u_1$ , technological feasibility  $u_2$ , staff feasibility  $u_3$ , period feasibility  $u_4$  and legal feasibility  $u_5$ . The weight vector of five criteria is given by  $\omega = (0.1, 0.15, 0.2, 0.3, 0.25)$ . The data of evaluation information given by experts are represented by the PFSs. The details of this dataset are shown Table 8.

**Table 6** Basic information of experimental datasets

Dataset	Data source	Objects	Attribute	Decision class
Wine	UCI	178	13	3
Seeds	UCI	210	7	3
Haberman	UCI	306	3	2
Blood	UCI	748	5	2
Banknote	UCI	1372	5	4
Wireless	UCI	2000	7	4

**Table 7** The classified accuracy of different similarity measures in subsection 5.2

	Wine	Seed	Haberman	Blood	Banknote	Wireless
$\mathbb{S}_{HY1}$ [31]	0.7003	0.8844	0.7271	0.4771	0.7852	0.9554
$\mathbb{S}_{HY2}$ [31]	0.7003	0.8844	0.7271	0.4771	0.7852	0.9554
$\mathbb{S}_{HY3}$ [31]	0.7003	0.8844	0.7271	0.4771	0.7852	0.9554
$\mathbb{S}_Y$ [32]	0.6525	<b>0.8854</b>	<b>0.7274</b>	0.4431	0.7850	<b>0.9586</b>
$\mathbb{S}_C$ [33]	0.7056	0.8844	0.7271	0.4771	0.7852	0.9554
$\mathbb{S}_L$ [34]	0.6994	0.8844	0.7271	0.4591	0.7852	0.9554
$\mathbb{S}_{HK}$ [35]	0.7129	0.8844	0.7271	0.4771	0.7852	0.9554
$\mathbb{S}_P^1$ [14]	0.7107	0.8842	0.7271	0.4763	0.7852	0.9509
$\mathbb{S}_P^2$ [14]	0.7009	0.8842	0.7271	0.4771	0.7852	0.9509
$\mathbb{S}_P^3$ [14]	0.7052	0.8842	0.7270	0.5003	0.7852	0.9509
$\mathbb{S}_P^4$ [14]	0.7019	0.8840	0.7270	0.5047	0.7853	0.9500
$\mathbb{S}_P^5$ [14]	0.7052	0.8842	0.7270	0.5003	0.7852	0.9509
$\mathbb{S}_Z^1$ [15]	0.7199	0.8839	0.7270	0.5050	0.7853	0.9503
$\mathbb{S}_Z^2$ [15]	0.7101	0.8839	0.7270	0.4883	0.7853	0.9503
$\mathbb{S}_Z^3$ [15]	0.7041	0.8839	0.7270	0.5050	0.7853	0.9503
$\mathbb{S}_Z^4$ [15]	0.7129	0.8842	0.7271	0.4860	0.7852	0.9509
$\mathbb{S}^{O_{0.5}}$	0.6947	0.8845	0.7271	0.4771	0.7852	0.9566
$\mathbb{S}^{O_1}$	0.7167	0.8844	0.7270	0.5029	0.7852	0.9554
$\mathbb{S}^{O_2}$	0.7140	0.8843	0.7269	0.5000	<b>0.7856</b>	0.9524
$\mathbb{S}^{O_+}$	0.7133	0.8845	0.7271	0.4998	0.7851	0.9562
$\mathbb{S}^{O_{DB}}$	0.6925	0.8846	0.7271	0.4771	0.7852	0.9568
$\mathbb{S}^{O_{mM}}$	<b>0.7206</b>	0.8845	0.7270	<b>0.5094</b>	0.7853	0.9558
$\mathbb{S}^{O_{Mid}}$	0.7197	0.8845	0.7270	<b>0.5094</b>	0.7853	0.9547

“Bolditalic” denotes the highest accuracy of classification for a dataset

In this subsection, we always take overlap function  $O$  as  $O_+$ , strong fuzzy negations  $N_a$  and  $N_b$  as  $N_s$  and  $k = 1$ . According to classification algorithm given by [14], the clustering results of the ten software projects calculated by  $\mathbb{S}^{O_+}$  are shown in Table 9. It is worth noting that clustering result with seven clusters  $\{P_1\}$ ,  $\{P_2, P_5\}$ ,  $\{P_3\}$ ,  $\{P_4\}$ ,  $\{P_6\}$ ,  $\{P_7\}$ ,  $\{P_8, P_9, P_{10}\}$  is supported by [14]. In addition, we obtain the reasonable results only through three iterations. The number of iterations is equal to Peng’s work [14] in the clustering algorithm.

### 5.4 Application in Clustering on the Basis of Datasets from the UCI Database

In this subsection, we show the effective of the  $O$ -similarity measures in clustering problems. We apply different similarity measures to six datasets which are used in Sect. 5.2. We repeat 100 times 2-Fold Cross Validation for each dataset. We use the method given by [5] to generate PFSs for datasets and use the classical  $k$ -means algorithm to cluster the generated PFSs. In this subsection, we always take strong fuzzy negations  $N_a$  as  $N_s$ , strong fuzzy negation

$N_b$  as  $N_{1,1}$ ,  $k = 1$  and  $\omega_i = \frac{1}{n}$  for  $i = 1, 2, \dots, n$  whenever the universe  $U = \{u_1, u_2, \dots, u_n\}$ . We analyze the performance of similarity measures listed in Table 1 and  $O$ -similarity measures by purity of clustering results. The experiments results are shown in Table 10.

We now show some analyses on the basis of Table 10 as follows:

- (1) For datasets “Wine”, “Banknote” and “Wireless”, there ia always one  $O$ -similarity measure which has the highest purity of clustering in compared with some existing similarity measures.
- (2) For dataset “Seed”, although the highest purity of clustering is given by  $\mathbb{S}_Y$ , the second highest purity of clustering is given by  $\mathbb{S}^{O_{0.5}}$  and  $\mathbb{S}^{O_{DB}}$ . Meanwhile, the difference between the two is only 0.0093.
- (3) For dataset “Haberman”, the highest purity of clustering is given by  $\mathbb{S}_p^1$ . The highest purity of clustering calculated by  $O$ -similarity measures is given by  $\mathbb{S}^{O_2}$ . The difference between the two is only 0.0008.
- (4) For dataset “Blood”, the highest purity of clustering is given by  $\mathbb{S}_Y$  and  $\mathbb{S}_p^1$ . The highest purity of clustering calculated by  $O$ -similarity measures is given by  $\mathbb{S}^{O_{0.5}}$  and  $\mathbb{S}^{O_{DB}}$ . The difference between the two is only 0.0002.

Therefore, on the basis of the above analyses, the performance of the  $O$ -similarity measures is superior to others methods under some conditions.

## 6 Conclusion

We mainly study the construction methods of similarity measure between PFSs based on overlap and grouping functions. To be precise, we achieved the followings:

- (1) We proposed two construction methods of similarity measures between PFSs, called  $O$ -similarity measure

**Table 8** The evaluation information of ten software projects [14]

	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$
$P_1$	(0.5,0.7)	(0.4,0.3)	(0.5,0.4)	(0.6,0.6)	(0.7,0.6)
$P_2$	(0.6,0.7)	(0.6,0.5)	(0.5,0.6)	(0.7,0.6)	(0.6,0.5)
$P_3$	(0.4,0.3)	(0.5,0.5)	(0.6,0.4)	(0.8,0.4)	(0.4,0.4)
$P_4$	(0.5,0.3)	(0.6,0.4)	(0.4,0.5)	(0.5,0.6)	(0.5,0.4)
$P_5$	(0.6,0.7)	(0.6,0.6)	(0.6,0.4)	(0.7,0.5)	(0.7,0.3)
$P_6$	(0.7,0.4)	(0.6,0.3)	(0.5,0.8)	(0.8,0.5)	(0.6,0.2)
$P_7$	(0.6,0.5)	(0.5,0.6)	(0.7,0.6)	(0.7,0.6)	(0.5,0.1)
$P_8$	(0.6,0.7)	(0.4,0.3)	(0.8,0.5)	(0.6,0.4)	(0.7,0.3)
$P_9$	(0.5,0.7)	(0.6,0.3)	(0.7,0.5)	(0.4,0.4)	(0.7,0.3)
$P_{10}$	(0.4,0.7)	(0.6,0.3)	(0.4,0.5)	(0.5,0.4)	(0.8,0.3)

**Table 9** The clustering result of ten software projects in Sect. 5.3

Confident Level	Clusters
$0.9342 < \lambda \leq 1$	$\{P_1\}, \{P_2\}, \{P_3\}, \{P_4\}, \{P_5\}, \{P_6\}, \{P_7\}, \{P_8\}, \{P_9\}, \{P_{10}\}$
$0.9305 < \lambda \leq 0.9342$	$\{P_1\}, \{P_2\}, \{P_3\}, \{P_4\}, \{P_5\}, \{P_6\}, \{P_7\}, \{P_8, P_9\}, \{P_{10}\}$
$0.9031 < \lambda \leq 0.9305$	$\{P_1\}, \{P_2\}, \{P_3\}, \{P_4\}, \{P_5\}, \{P_6\}, \{P_7\}, \{P_8, P_9, P_{10}\}$
$0.8976 < \lambda \leq 0.9031$	$\{P_1\}, \{P_2, P_5\}, \{P_3\}, \{P_4\}, \{P_6\}, \{P_7\}, \{P_8, P_9, P_{10}\}$
$0.8936 < \lambda \leq 0.8976$	$\{P_1, P_2, P_5\}, \{P_3\}, \{P_4\}, \{P_6\}, \{P_7\}, \{P_8, P_9, P_{10}\}$
$0.8817 < \lambda \leq 0.8936$	$\{P_1, P_2, P_5, P_8, P_9, P_{10}\}, \{P_3\}, \{P_4\}, \{P_6\}, \{P_7\}$
$0.8778 < \lambda \leq 0.8817$	$\{P_1, P_2, P_4, P_5, P_8, P_9, P_{10}\}, \{P_3\}, \{P_6\}, \{P_7\}$
$0.8628 < \lambda \leq 0.8778$	$\{P_1, P_2, P_4, P_5, P_7, P_8, P_9, P_{10}\}, \{P_3\}, \{P_6\}$
$0.8609 < \lambda \leq 0.8628$	$\{P_1, P_2, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}\}, \{P_3\}$
$0 \leq \lambda \leq 0.8609$	$\{P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}\}$



**Table 10** The purity of clustering of different similarity measures in Sect. 5.4

	Wine	Seed	Haberman	Blood	Banknote	Wireless
$\mathbb{S}_{\text{HY1}}$ [31]	0.8319	0.8390	0.7372	0.7621	0.7371	0.8577
$\mathbb{S}_{\text{HY2}}$ [31]	0.8319	0.8390	0.7372	0.7621	0.7371	0.8577
$\mathbb{S}_{\text{HY3}}$ [31]	0.8319	0.8390	0.7372	0.7621	0.7371	0.8577
$\mathbb{S}_{\text{Y}}$ [32]	0.6797	<b>0.8485</b>	0.7371	<b>0.7623</b>	0.7309	0.8613
$\mathbb{S}_{\text{C}}$ [33]	0.8319	0.8390	0.7372	0.7621	0.7371	0.8577
$\mathbb{S}_{\text{L}}$ [34]	0.7527	0.8384	0.7371	0.7621	0.7130	0.8576
$\mathbb{S}_{\text{HK}}$ [35]	0.8319	0.8390	0.7372	0.7621	0.7371	0.8577
$\mathbb{S}_{\text{P}}^1$ [14]	0.8088	0.8351	<b>0.7383</b>	<b>0.7623</b>	0.7351	0.8469
$\mathbb{S}_{\text{P}}^2$ [14]	0.8400	0.8348	0.7372	0.7621	0.7371	0.8558
$\mathbb{S}_{\text{P}}^3$ [14]	0.8075	0.8351	0.7378	0.7622	0.7351	0.8470
$\mathbb{S}_{\text{P}}^4$ [14]	0.8378	0.8359	0.7375	0.7621	0.7475	0.8600
$\mathbb{S}_{\text{P}}^5$ [14]	0.8075	0.8351	0.7378	0.7622	0.7351	0.8470
$\mathbb{S}_{\text{Z}}^1$ [15]	0.8288	0.8309	0.7378	0.7622	0.7390	0.8513
$\mathbb{S}_{\text{Z}}^2$ [15]	0.8095	0.8309	0.7378	0.7622	0.7390	0.8514
$\mathbb{S}_{\text{Z}}^3$ [15]	0.8445	0.8309	0.7378	0.7621	0.7390	0.8514
$\mathbb{S}_{\text{Z}}^4$ [15]	0.8349	0.8347	0.7371	0.7621	0.7444	0.8537
$\mathbb{S}^{O_{0.5}}$	0.8506	0.8392	0.7371	0.7621	0.7416	0.8577
$\mathbb{S}^{O_1}$	0.8534	0.8338	0.7371	0.7620	0.7447	0.8641
$\mathbb{S}^{O_2}$	<b>0.8614</b>	0.8382	0.7375	0.7620	0.7414	0.8657
$\mathbb{S}^{O_+}$	0.8408	0.8388	0.7370	0.7620	0.7381	0.8601
$\mathbb{S}^{O_{\text{DB}}}$	0.8485	0.8392	0.7371	0.7621	0.7432	0.8578
$\mathbb{S}^{O_{\text{mM}}}$	0.8578	0.8361	0.7373	0.7620	<b>0.7489</b>	<b>0.8662</b>
$\mathbb{S}^{O_{\text{Mid}}}$	0.8557	0.8360	0.7374	0.7620	0.7391	0.8661

“Bolditalic” denotes the highest purity of clustering for a dataset

and  $G$ -similarity measure, based on overlap and grouping functions. The connection between  $O$ -similarity measure and  $G$ -similarity measure was established. We also established the connection between  $G$ -similarity measure and  $s$ -similarity measure derived from  $t$ -conorms.

- (2) We used four numerical examples to illustrate the effectiveness, reliability and deficiencies of  $O$ -similarity measure by contrast with other existing similarity measures.
- (3) We applied some  $O$ -similarity measures to classification and clustering problems. On the one hand, we showed the application of  $O$ -similarity measure in classification and clustering by using classical examples. On the other hand, we used six datasets from UCI database to illustrate the effectiveness and reliability by comparing with some existing similarity measures. The experimental results showed that  $O$ -similarity measure had good performance in data-driven environments.

In the further work, on the one hand, we will research the construction methods of similarity measures between PFSs

based on other aggregation functions, such as uninorms [37]. On the other hand, we will study the differences among the  $O$ -similarity measures constructed through different overlap functions, strong fuzzy negations and parameter  $k$ .

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (11771111 and 62166037) and the Science and Technology Program of Gansu Province (20JR10RA101).

## References

1. Zadeh, L.A.: Fuzzy sets. *Inf. Control* **8**, 338–353 (1965)
2. Atanassov, K.T.: Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **20**, 87–96 (1986)
3. Yager, R.R.: Pythagorean fuzzy subsets. In: 2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013, pp. 57–61 (2013)
4. Yager, R.R.: Pythagorean membership grades in multicriteria decision making. *IEEE Trans. Fuzzy Syst.* **22**, 958–965 (2014)
5. Gao, X., Deng, Y.: Generating method of Pythagorean fuzzy sets from the negation of probability. *Eng. Appl. Artif. Intell.* **105**, 104403 (2021)

6. Zhang, X., Xu, Z.: Extension of TOPSIS to multiple criteria decision making with Pythagorean fuzzy sets. *Int. J. Intell. Syst.* **29**, 1061–1078 (2014)
7. Zeng, W., Li, D., Yin, Q.: Distance and similarity measures of Pythagorean fuzzy sets and their applications to multiple criteria group decision making. *Int. J. Intell. Syst.* **33**, 2236–2254 (2018)
8. Xu, T.-T., Zhang, H., Li, B.-Q.: Pythagorean fuzzy entropy and its application in multiple-criteria decision-making. *Int. J. Fuzzy Syst.* **22**, 1552–1564 (2020)
9. Zhang, S.-P., Sun, P., Mi, J.-S., Feng, T.: Belief function of Pythagorean fuzzy rough approximation space and its applications. *Int. J. Approx. Reason.* **119**, 58–80 (2020)
10. Bakioglu, G., Atahan, A.O.: AHP integrated TOPSIS and VIKOR methods with Pythagorean fuzzy sets to prioritize risks in self-driving vehicles. *Appl. Soft Comput.* **99**, 106948 (2021)
11. Lang, G., Miao, D., Fujita, H.: Three-way group conflict analysis based on Pythagorean fuzzy set theory. *IEEE Trans. Fuzzy Syst.* **28**, 447–461 (2020)
12. Liu, P., Rani, P., Mishra, A.R.: A novel Pythagorean fuzzy combined compromise solution framework for the assessment of medical waste treatment technology. *J Clean Prod* **292**, 126047 (2021)
13. Nan, T., Zhang, H., He, Y.: Pythagorean fuzzy full implication triple I method and its application in medical diagnosis. *Int. J. Fuzzy Syst.* (2022). <https://doi.org/10.1007/s40815-022-01261-8>
14. Peng, X., Yuan, H., Yang, Y.: Pythagorean fuzzy information measures and their applications. *Int. J. Intell. Syst.* **32**, 991–1029 (2017)
15. Zhang, Q., Hu, J., Feng, J., Liu, A., Li, Y.: New similarity measures of Pythagorean fuzzy sets and their applications. *IEEE Access* **7**, 138192–138202 (2019)
16. Firozja, M.A., Agheli, B., Jamkhaneh, E.B.: A new similarity measure for Pythagorean fuzzy sets. *Complex Intell. Syst.* **6**, 67–74 (2020)
17. Bustince, H., Fernandez, J., Mesiar, R., Montero, J., Orduna, R.: Overlap functions. *Nonlinear Anal. Theory Methods Appl.* **72**, 1488–1499 (2010)
18. Bustince, H., Pagola, M., Mesiar, R., Hullermeier, E., Herrera, F.: Grouping, overlap, and generalized bientropic functions for fuzzy modeling of pairwise comparisons. *IEEE Trans. Fuzzy Syst.* **20**, 405–415 (2012)
19. Elkano, M., Galar, M., Sanz, J., Bustince, H.: Fuzzy rule-based classification systems for multi-class problems using binary decomposition strategies: on the influence of n-dimensional overlap functions in the fuzzy reasoning method. *Inf. Sci.* **332**, 94–114 (2016)
20. Dimuro, G.P., Fernández, J., Bedregal, B., Mesiar, R., Sanz, J.A., Lucca, G., Bustince, H.: The state-of-art of the generalizations of the Choquet integral: from aggregation and pre-aggregation to ordered directionally monotone functions. *Inf. Fusion* **57**, 27–43 (2020)
21. Elkano, M., Galar, M., Sanz, J.A., Schiavo, P.F., Pereira, S., Dimuro, G.P., Borges, E.N., Bustince, H.: Consensus via penalty functions for decision making in ensembles in fuzzy rule-based classification systems. *Appl. Soft Comput.* **67**, 728–740 (2018)
22. Jia, Z., Qiao, J.: On decision evaluation functions in three-way decision spaces derived from overlap and grouping function. *Soft Comput.* **24**, 15159–15178 (2020)
23. Jurio, A., Bustince, H., Pagola, M., Pradera, A., Yager, R.R.: Some properties of overlap and grouping functions and their application to image thresholding. *Fuzzy Sets Syst.* **229**, 69–90 (2013)
24. Xu, Z., Yager, R.R.: Some geometric aggregation operators based on intuitionistic fuzzy sets. *Int. J. Gen. Syst.* **35**, 417–433 (2006)
25. Klement, E.P., Mesiar, R., Pap, E.: *Triangular Norms*. Kluwer, Dordrecht (2000)
26. Hu, B.Q.: Three-way decisions space and three-way decisions. *Inf. Sci.* **281**, 21–52 (2014)
27. Bedregal, B., Dimuro, G.P., Bustince, H., Barrenechea, E.: New results on overlap and grouping functions. *Inf. Sci.* **249**, 148–170 (2013)
28. Dimuro, G.P., Bedregal, B., Bustince, H., Asiáin, M.J., Mesiar, R.: On additive generators of overlap functions. *Fuzzy Sets Syst.* **287**, 76–96 (2016)
29. Qiao, J., Hu, B.Q.: On multiplicative generators of overlap and grouping functions. *Fuzzy Sets Syst.* **332**, 1–24 (2018)
30. Qiao, J., Hu, B.Q.: On interval additive generators of interval overlap functions and interval grouping functions. *Fuzzy Sets Syst.* **323**, 19–55 (2017)
31. Hung, W.-L., Yang, M.-S.: Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. *Pattern Recognit. Lett.* **25**, 1603–1611 (2004)
32. Ye, J.: Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Math. Comput. Model.* **53**, 91–97 (2011)
33. Chen, S.-M.: Similarity measures between vague sets and between elements. *IEEE Trans. Syst. Man Cybern. B* **27**, 153–158 (1997)
34. Li, Y., Olson, D.L., Qin, Z.: Similarity measures between intuitionistic fuzzy (vague) sets: a comparative analysis. *Pattern Recognit. Lett.* **28**, 278–285 (2007)
35. Hong, D.H., Kim, C.: A note on similarity measures between vague sets and between elements. *Inf. Sci.* **115**, 83–96 (1999)
36. Chen, S.-M., Chang, C.-H.: A novel similarity measure between Atanassov's intuitionistic fuzzy sets based on transformation techniques with applications to pattern recognition. *Inf. Sci.* **291**, 96–114 (2015)
37. Yager, R.R., Rybalov, A.: Uninorm aggregation operators. *Fuzzy Sets Syst.* **80**, 111–120 (1996)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



**Zihang Jia** received the B.S. Degree in Information and Computing Sciences from Northwest Normal University, Lanzhou, China, in 2022. He is currently working toward the M.S. Degree at the School of Mathematical Sciences of Dalian University of Technology, Dalian, China. His current research interests include Aggregation Functions and Three-way Decisions.



**Junsheng Qiao** received the B.S. Degree in Mathematics and Applied Mathematics from Northwest Normal University, Lanzhou, China, in 2012, the M.S. Degree in Fundamental Mathematics in Sichuan University, Chengdu, China, in 2015 and the Ph.D. Degree in Computational Mathematics in Wuhan University, Wuhan, China, in 2018. He is currently an Associate Professor with the College of Mathematics and Statistics, Northwest Normal

University, China. He is also a Member of the Associate Editors of Computational and Applied Mathematics, a Member of the Associate Editors of Journal of Intelligent & Fuzzy Systems, a Member of the Associate Editorial Board of Journal of Fuzzy Logic and Modeling in Engineering and a member of Granular Computing and Knowledge Discovery Professional Committee of Chinese Society for Artificial

Intelligence. His current research interests include Fuzzy Logics, Aggregation Functions, Rough Sets and Three-way Decisions.



**Minghao Chen** received the B.S. Degree in Mathematics from Harbin Institute of Technology, China, in 1983, the M.S. Degree in Mathematics from Harbin Institute of Technology, China, in 1989, the Ph.D. Degree in Mathematics from Harbin Institute of Technology, China, in 2007. He is currently a Professor with School of Mathematical Sciences of Dalian University of Technology, Dalian, China. His main fields of research are Uncertain

Dynamical System, Fuzzy Optimization, and Rough Set Theory and Its Application.