Journal of New Theory
ISSN: 2149-1402

33 (2020) 85-94
Journal of New Theory
http://www.newtheory.org
Open Access

Similarity Measures of Pythagorean Neutrosophic Sets with Dependent Neutrosophic Components Between T and F

Jansi Rajan[1], Mohana Krishnaswamy[2]

Article History
Received: 08.01.2019
Accepted: 26.12.2020
Published: 31.12.2020
Original Article

Abstract − Clustering plays an important role in data mining, pattern recognition and machine learning. This paper proposes Pythagorean neutrosophic clustering methods based on similarity measures between Pythagorean neutrosophic sets with T and F are dependent neutrosophic components [PN-Set]. First, we define a generalized distance measure between PN-Sets and propose two distance-based similarity measures of PN-Sets. Then, we present a clustering algorithm based on the similarity measures of PN-Sets to cluster Pythagorean neutrosophic data. Finally, an illustrative example is given to demonstrate the application and effectiveness of the developed clustering methods.

Keywords − *Pythagorean neutrosophic Sets with T and F are dependent neutrosophic components, clustering algorithm, distance measure, similarity measure.*

3. Introduction

Fuzzy sets were firstly initiated by L.A. Zadeh [1] in 1965. Zadeh's idea of fuzzy set evolved as a new tool to deal with uncertainties in real-life problems and discussed only membership function. After the extensions of fuzzy set theory Atanassov [2] generalized this concept and introduced a new set called intuitionistic fuzzy set (IFS) in 1986, which can describe the non-membership grade of an imprecise event along with its membership grade under a restriction that the sum of both membership and non-membership grades does not exceed 1. IFS has its greatest use in practical multiple attribute decision-making problems. In some practical problems, the sum of membership and non-membership degree to which an alternative satisfying attribute provided by decision-maker (DM) may be bigger than 1.

Yager [3] was decided to introduce the new concept known as Pythagorean fuzzy sets. Pythagorean fuzzy sets have a limitation that their square sum is less than or equal to 1. IFS was failed to deal with indeterminate and inconsistent information which exist in beliefs system; therefore, Smarandache [4] in 1995 introduced a new concept known as neutrosophic set (NS) which generalizes fuzzy sets and intuitionistic fuzzy sets and so on. A neutrosophic set includes truth membership, falsity membership and indeterminacy membership.

In 2006, Smarandache introduced, for the first time, the degree of dependence (and consequently the degree of independence) between the components of the fuzzy set, and also between the components of the neutrosophic set. In 2016, the refined neutrosophic set was generalized to the degree of dependence or independence of subcomponents [5]. In neutrosophic set [5], if truth membership and falsity membership are

[1] mathematicsgasc@gmail.com (Corresponding Author); [2] riyaraju1116@gmail.com
[1,2] Department of Mathematics, Nirmala College for Women, Coimbatore, India

100% dependent and indeterminacy is 100% independent, that is $0 \leq u_A(x) + \zeta_A(x) + v_A(x) \leq 2$. Sometimes in real life, we face many problems which cannot be handled by using neutrosophic for example when $u_A(x) + \zeta_A(x) + v_A(x) > 2$. So, Pythagorean neutrosophic set with T and F are dependent neutrosophic components [PN-SET] of the condition is as their square sum does not exceed 2. Here, T and F are dependent neutrosophic components, and we make $u_A(x), v_A(x) as$ Pythagorean, then $\left(u_A(x)\right)^2 + \left(v_A(x)\right)^2 \leq 1$ with $u_A(x), v_A(x) \in [0,1]$. If $\zeta_A(x)$ is independent of them, then $0 \leq \zeta_A(x) \leq 1$. Then, $0 \leq \left(u_A(x)\right)^2 + \left(\zeta_A(x)\right)^2 + \left(v_A(x)\right)^2 \leq 2$, with $u_A(x), \zeta_A(x), v_A(x) \in [0,1]$.

Recently, Ye [6,7] presented the correlation coefficient of single-valued neutrosophic sets (SVNSsaa0 and the cross-entropy measure of SVNSs and applied them to single-valued neutrosophic decision-making problems. Then, Ye [8] proposed similarity measures between interval neutrosophic sets and their applications in multicriteria decision making. Xu [9] and Zhang [10] proposed a clustering algorithm. J. Ye [11] also introduced the clustering methods using Distance-based similarity measures of single-valued neutrosophic sets.

This paper proposes a Pythagorean neutrosophic clustering algorithm to deal with data represented by Pythagorean neutrosophic set with dependent neutrosophic components between T and F [PN-Set, in short]. We define a generalized distance measure between PN-Sets and propose two distance-based similarity measures of PN-Sets. Then, we present a clustering algorithm based on the similarity measures of PN-Sets to cluster Pythagorean neutrosophic data and gives an illustrative example.

4. Preliminaries

Definition 2.1 [2]

Let E be a universe. An intuitionistic fuzzy set A on E can be defined as follows:

$$A = \{< x, u_A(x), v_A(x) >: x \in E\}.$$

where $u_A: E \to [0,1]$ and $v_A: E \to [0,1]$ such that $0 \leq u_A(x) + v_A(x) \leq 1$ for any $x \in E$.

Here, $u_A(x)$ and $v_A(x)$ is the degree of membership and degree of non-membership of the element x, respectively.

Definition 2.2 [12,13]

Let X be a nonempty set, and I the unit interval [0,1]. A Pythagorean fuzzy set S is an object having the form $A = \{(x, u_A(x), v_A(x)): x \in X\}$ where the functions $u_A: X \to [0,1]$ and $v_A: X \to [0,1]$ denote respectively the degree of membership and degree of non-membership of each element $x \in X$ to the set P, and $0 \leq \left(u_A(x)\right)^2 + (v_A(x))^2 \leq 1$ for each $x \in X$.

Definition 2.3[4]

Let X be a nonempty set (universe). A neutrosophic set A on X is an object of the form:

$$\text{A}= \left\{\left(x, u_A(x), \zeta_A(x), v_A(x)\right): x \in X\right\}.$$

Where $u_A(x), \zeta_A(x), v_A(x) \in [0,1], 0 \leq u_A(x) + \zeta_A(x) + v_A(x) \leq 2, for\ all\ x\ in\ X.$ $u_A(x)$ is the degree of membership, $\zeta_A(x)$ is the degree of indeterminacy and $v_A(x)$ is the degree of non-membership. Here $u_A(x)\ and\ v_A(x)$ are dependent components and $\zeta_A(x)$ is an independent component.

Definition 2.4 [4]

Let X be a nonempty set, and I the unit interval [0,1]. A neutrosophic set A and B of the form

$$A = \{(x, u_A(x), \zeta_A(x), v_A(x)) : x \in X\} \text{ and } B = \{(x, u_B(x), \zeta_B(x), v_B(x)) : x \in X\}$$

Then,

$$A^C = \{(x, v_A(x), 1 - \zeta_A(x), u_A(x)) : x \in X\} \text{ or } A^C = \{(x, v_A(x), \zeta_A(x), u_A(x)) : x \in X\}$$

$$A \cup B = \{(x, \max(u_A(x), u_B(x)), \min(\zeta_A(x), \zeta_B(x)), \min(v_A(x), v_B(x))) : x \in X\}$$

$$A \cap B = \{(x, \min(u_A(x), u_B(x)), \max(\zeta_A(x), \zeta_B(x)), \max(v_A(x), v_B(x)) : x \in X\}$$

3. Distance-Based Similarity Measures between PN-Sets

Definition 3.1

Let X be a nonempty set (universe). A PN-Set M on X is an object of the form:

$$M = \{(x, u_M(x), \zeta_M(x), v_M(x)) : x \in X\},$$

Where $u_M(x), \zeta_M(x), v_M(x) \in [0,1], 0 \leq (u_M(x))^2 + (\zeta_M(x))^2 + (v_M(x))^2 \leq 2,$ for all $x \in X$. $u_M(x)$ is the degree of membership, $\zeta_M(x)$ is the degree of indeterminacy and $v_M(x)$ is the degree of non-membership. Here $u_M(x)\ and\ v_M(x)$ are dependent components and $\zeta_M(x)$ is an independent component.

Definition 3.2

Let X be a nonempty set and I the unit interval [0,1]. A PN-Sets M and N of the form

$$M = \{(x, u_M(x), \zeta_M(x), v_M(x)) : x \in X\} \text{ and } N = \{(x, u_N(x), \zeta_N(x), v_N(x)) : x \in X\}.$$

Then,

$$M^C = \{(x, v_M(x), 1 - \zeta_M(x), u_M(x)) : x \in X\} \text{or } M^C = \{(x, v_M(x), \zeta_M(x), u_M(x)) : x \in X\}$$

$$M \cup N = \{(x, \max(u_M(x), u_N(x)), \min(\zeta_M(x), \zeta_N(x)), \min(v_M(x), v_N(x))) : x \in X\}$$

$$M \cap N = \{(x, \min(u_M(x), u_N(x)), \max(\zeta_M(x), \zeta_N(x)), \max(v_M(x), v_N(x))) : x \in X\}$$

For two PN-Sets S and T in a universe of discourse $X = \{x_1, x_2, \ldots, x_n\}$, which are denoted by $M = \{(x_i, u_M(x_i), \zeta_M(x_i), v_M(x_i)) : x_i \in X\}$ and $N = \{(x_i, u_N(x_i), \zeta_N(x_i), v_N(x_i)) : x_i \in X\}$, where $u_M(x_i)$, $\zeta_M(x_i)$, $v_M(x_i)$, $u_N(x_i)$, $\zeta_N(x_i)$, $v_N(x_i) \in [0,1]$ for every $x_i \in X$. Let us consider the weight $w_i (i = 1,2, \ldots, n)$ of an element $x_i (i = 1,2, \ldots, n)$, with $w_i \geq 0$ $(i = 1,2, \ldots, n)$, and $\sum_{i=1}^{n} w_i = 1$. Then, we define the generalized PN weighted distance measure:

$$d_p(M, N) = \left\{\frac{1}{3}\sum_{i=1}^{n} w_i \left[\left|u_M^2(x_i) - u_N^2(x_i)\right|^p + \left|\zeta_M^2(x_i) - \zeta_N^2(x_i)\right|^p + \left|v_M^2(x_i) - v_N^2(x_i)\right|^p\right]\right\}^{\frac{1}{p}} \quad (1)$$

where $p > 0$. When $p = 1,2$, we can obtain the PN weighted Hamming distance, and the PN weighted Euclidean distance, respectively, as follows:

$$d_1(M,N) = \frac{1}{3}\sum_{i=1}^{n} w_i\left[\left|u_M^2(x_i) - u_N^2(x_i)\right| + \left|\zeta_M^2(x_i) - \zeta_N^2(x_i)\right| + \left|v_M^2(x_i) - v_N^2(x_i)\right|\right] \quad (2)$$

$$d_2(M,N) = \left\{\frac{1}{3}\sum_{i=1}^{n} w_i\left[\left|u_M^2(x_i) - u_N^2(x_i)\right|^2 + \left|\zeta_M^2(x_i) - \zeta_N^2(x_i)\right|^2 + \left|v_M^2(x_i) - v_N^2(x_i)\right|^2\right]\right\}^{\frac{1}{2}} \quad (3)$$

Therefore, Eqs. (2) and (3) are the special cases of (1). Then, for the distance measure, we have the following proposition.

Proposition 3.3. The above-defined distance $d_p(M,N)\ for\ p > 0$ satisfies the following properties:

(DP1) $0 \le d_p(M,N) \le 1$;

(DP2) $d_p(M,N) = 0$ if and only if M = N;

(DP3) $d_p(M,N) = d_p(N,M)$;

(DP4) If $M \subseteq N \subseteq O$, O is a PN-Set in X, then $d_p(M,O) \ge d_p(M,N)\ \ and\ \ d_p(M,O) \ge d_p(N,O)$.

Proof:

It is easy to see that $d_p(M,N)$ satisfies the properties (DP1)-DP43). Therefore, we only prove (DP4). Let M ⊆ $N \subseteq O$, then, $u_M(x_i) \le u_N(x_i) \le u_O(x_i), \zeta_M(x_i) \ge \zeta_N(x_i) \ge \zeta_O(x_i)\ and\ v_M(x_i) \ge v_N(x_i) \ge v_O(x_i)$ for every $x_i \in X$. Also, $u_M^2(x_i) \le u_N^2(x_i) \le u_O^2(x_i), \zeta_M^2(x_i) \ge \zeta_N^2(x_i) \ge \zeta_O^2(x_i)$, and $v_M^2(x_i) \le v_N^2(x_i) \le v_O^2(x_i)$, for every $x_i \in X$.

Then, we obtain the following relations:

$$\left|u_M^2(x_i) - u_N^2(x_i)\right|^p \le \left|u_M^2(x_i) - u_O^2(x_i)\right|^p, \left|u_N^2(x_i) - u_O^2(x_i)\right|^p \le \left|u_M^2(x_i) - u_O^2(x_i)\right|^p,$$

$$\left|\zeta_M^2(x_i) - \zeta_N^2(x_i)\right|^p \le \left|\zeta_M^2(x_i) - \zeta_O^2(x_i)\right|^p, \left|\zeta_N^2(x_i) - \zeta_O^2(x_i)\right|^p \le \left|\zeta_M^2(x_i) - \zeta_O^2(x_i)\right|^p,$$

$$\left|v_M^2(x_i) - v_N^2(x_i)\right|^p \le \left|v_M^2(x_i) - v_O^2(x_i)\right|^p, \left|v_N^2(x_i) - v_O^2(x_i)\right|^p \le \left|v_M^2(x_i) - v_O^2(x_i)\right|^p,$$

Hence,

$$\left|u_M^2(x_i) - u_N^2(x_i)\right|^p + \left|\zeta_M^2(x_i) - \zeta_N^2(x_i)\right|^p + \left|v_M^2(x_i) - v_N^2(x_i)\right|^p$$

$$\le \left|u_M^2(x_i) - u_O^2(x_i)\right|^p + \left|\zeta_M^2(x_i) - \zeta_O^2(x_i)\right|^p + \left|v_M^2(x_i) - v_O^2(x_i)\right|^p$$

$$\left|u_N^2(x_i) - u_O^2(x_i)\right|^p + \left|\zeta_N^2(x_i) - \zeta_O^2(x_i)\right|^p + \left|v_N^2(x_i) - v_O^2(x_i)\right|^p$$

$$\le \left|u_M^2(x_i) - u_O^2(x_i)\right|^p + \left|\zeta_M^2(x_i) - \zeta_O^2(x_i)\right|^p + \left|v_M^2(x_i) - v_O^2(x_i)\right|^p$$

Combining the above inequalities with the above-defined distance formula (1), we can obtain $d_p(M,O) \ge d_p(M,N)$ and $d_p(M,O) \ge d_p(N,O)$ for $p > 0$. Thus, the property (DP4) is satisfied.

This completes the proof.

Note that similarity and distance (dissimilarity) measures are complementary: when the first increases, the second decreases. Normalized distance measure and similarity measure are dual concepts.

Thus, $S(M,N) = 1 - d(M,N)$ and vice versa. The properties of distance measures below are complementary to those of similarity measure.

Proposition 3.4 Let A and B be two PN-Sets in a universe of discourse $X = \{x_1, x_2, \dots, x_n\}$; $S(A,B)$ is called a Pythagorean neutrosophic similarity measure, which should satisfy the following properties:

(SP1) $0 \le S(M,N) \le 1$;

(SP2) $S(M,N) = 0$ if and only if A = B;

(SP3) $S(M,N) = S(N,M)$;

(SP4) If $M \subseteq N \subseteq O$, C is a PN-Set in X, then $S(M,O) \ge S(M,N)$ and $S(M,O) \ge S(N,O)$.

Assume that there are two PN-sets $M = \{(x_i, u_M(x_i), \zeta_M(x_i), v_M(x_i)): x_i \in X\}$ and $N = \{(x_i, u_N(x_i), \zeta_N(x_i), v_N(x_i)): x_i \in X\}$ in a universe of discourse $X = \{x_1, x_2, \dots, x_n\}$. Thus, according to the relationship between the distance and the similarity measure, we can obtain the following PN similarity measure:

$$S_1(M,N) = 1 - d_p(M,N) = 1 - \left\{\frac{1}{3}\sum_{i=1}^{n} w_i[|u_M^2(x_i) - u_N^2(x_i)|^p + |\zeta_M^2(x_i) - \zeta_N^2(x_i)|^p + |v_M^2(x_i) - v_N^2(x_i)|^p]\right\}^{\frac{1}{p}} \quad (4)$$

Obviously, we can easily prove that $S_1(M,N)$ satisfies the properties (SP1) - (SP4) in Proposition 2 by the relationship between the distance and the similarity measure and the proof of Proposition 1, which is omitted here.

Furthermore, we can also propose another PN similarity measure:

$$S_2(M,N) = \frac{1 - d_p(M,N)}{1 + d_p(M,N)} = \frac{1 - \left\{\frac{1}{3}\sum_{i=1}^{n} w_i[|u_M^2(x_i) - u_N^2(x_i)|^p + |\zeta_M^2(x_i) - \zeta_N^2(x_i)|^p + |v_M^2(x_i) - v_N^2(x_i)|^p]\right\}^{\frac{1}{p}}}{1 + \left\{\frac{1}{3}\sum_{i=1}^{n} w_i[|u_M^2(x_i) - u_N^2(x_i)|^p + |\zeta_M^2(x_i) - \zeta_N^2(x_i)|^p + |v_M^2(x_i) - v_N^2(x_i)|^p]\right\}^{\frac{1}{p}}} \quad (5)$$

Then, the similarity measure $S_2(M,N)$ also satisfied the properties (SP1) - (SP4) in Proposition 2.

Proof:

It is easy to see that $S_2(M,N)$ satisfies the properties (SP1) - (SP3). Therefore, we only prove the property (SP4).

As we obtain $d_p(M,O) \ge d_p(M,N)$ and $d_p(M,O) \ge d_p(N,O)$ for p > 0 from the property (DP4) in Proposition 1, there are $1 - d_p(M,N) \ge 1 - d_p(M,O), 1 - d_p(N,O) \ge 1 - d_p(M,O), 1 + d_p(M,N) \le 1 + d_p(M,O)$ and $1 + d_p(N,O) \le 1 + d_p(M,O)$.Then, there are the following inequalities:

$$\frac{1 - d_p(M,N)}{1 + d_p(M,N)} \ge \frac{1 - d_p(M,O)}{1 + d_p(M,O)}$$

and

$$\frac{1 - d_p(N,O)}{1 + d_p(N,O)} \ge \frac{1 - d_p(M,O)}{1 + d_p(M,O)}$$

Then, there are $S(M,O) \le S(M,N) and\ S(M,O) \le S(N,O)$. Hence, the property (SP4) is satisfied.

This completes the proof.

Example 3.5:

Assume that we have the following three PN-Sets in a universe of distance $X = \{x_1, x_2\}$:

$$A = \{(x_1, 0.1, 0.9, 0.6), (x_1, 0.1, 0.9, 0.6)\}$$

$$B = \{(x_1, 0.7, 0.8, 0.4), (x_1, 0.4, 0.6, 0.7)\}$$

$$C = \{(x_1, 0.8, 0.1, 0.3), (x_1, 0.4, 0.3, 0.1)\}$$

Then, there are $A \subseteq B \subseteq C$, with $u_A(x_i) \leq u_B(x_i) \leq u_C(x_i), \zeta_A(x_i) \leq \zeta_A(x_i) \leq \zeta_B(x_i)\zeta_C(x_i)$ and $v_A(x_i) \leq v_B(x_i)v_C(x_i)$ for each x_i in X={x_1, x_2}, and the weight vector $w = (0.4,0.6)^T$.

By applying Eq. (4) (take p = 1), the similarity measures between the PN-Sets are as follows:

$$S_1(A,B) = 0.7427, S_1(B,C) = 0.7367, S_1(A,C) = 0.4793.$$

Hence, $S_1(A,C) \leq S_1(A,B)$ and $S_1(A,C) \leq S_1(B,C)$

By applying Eq. (5) for p = 1, the similarity measures between the PN-Sets are as follows:

$$S_2(A,B) = 0.5907, S_2(B,C) = 0.5832, S_2(A,C) = 0.3152.$$

Hence, $S_2(A,C) \leq S_2(A,B)$ and $S_2(A,C) \leq S_2(B,C)$.

4. Clustering Algorithm Based on the Similarity Measures of PN-Sets

In this section, we can apply the proposed similarity measures of PN-Sets to clustering analysis under a PN environment. Based on the intuitionistic fuzzy clustering algorithm proposed by Zhang [1] and Xu [9].

Definition 4.1 Assume that $A = (A_1, A_2, \dots, A_m)$ is a set of PN-Sets and $C = (S_{ij})_{m\times m}$ is a similarity matrix, where $S_{ij} = S_K(A_i, A_j)(k = 1,2)$ and $S_{ij} \in [0,1]$ for $i, j = 1,2, \dots, m$, with $S_{ii} = 1$ for $i = 1,2, \dots, m$, and $S_{ij} = S_{ji}$, for $i, j = 1,2, \dots, m$.

Definition 4.2 [9,10] Let $C = (S_{ij})_{m\times m}$ be a similarity matrix, if $C^2 = C_{\circ} C = (\bar{S}_{ij})_{m\times m}$, then C^2 is called a composition matrix of C, where $\bar{S}_{ij} = \max_k\{min(S_{ik}, S_{kj})\}$, for $i, j = 1,2, \dots, m$.

Definition 4.3 [9,10] Let $C = (S_{ij})_{m\times m}$ be a similarity matrix, if $C^2 \subseteq C$, i.e., $\bar{S}_{ij} \leq S_{ij}$ for $i, j = 1,2, \dots, m$, then C is called an equivalent similarity matrix.

Definition 4.4 [9,10] Let $C = (S_{ij})_{m\times m}$ be a similarity matrix. Then, after finite time compositions of C:

$$C \to C^2 \to C^4 \to \cdots \to C^{2^k} \to \cdots, \quad (6)$$

there must exist a positive integer k such that $C^{2^k} = C^{2^{(k+1)}}$, then C^{2^k} is also an equivalent similarity matrix.

Definition 4.5 [9,10] Let $C = (S_{ij})_{m\times m}$ be an equivalent similarity matrix. Then, $C_\lambda = (S_{ij}^\lambda)_{m\times m}$ is called the λ-cutting matrix of C, where

$$S_{ij}^\lambda = \begin{cases} 0, S_{ij} < \lambda; \\ 1, S_{ij} \geq \lambda \end{cases} \text{ for } i, j = 1,2, \dots, m, \quad (7)$$

and λ is the confidence level with $\lambda \in [0,1]$.

Assume that $A = \{A_1, A_2, \dots, A_m\}$ is a set of PN-Set, where $A_j = \left\{\left(x_i, u_{A_j}, \zeta_{A_j}, v_{A_j}\right) : x_i \in X\right\}$ $(j = 1,2, \dots, m)$ in a universe of discourse $X = \{x_1, x_2, \dots, x_n\}$ is a PN. Let w_i be the weight for each element x_i $(i = 1,2, \dots, n)$, with $w_i \in [0,1]$, and $\sum_{i=1}^n w_i = 1$. Then, we can give the algorithm of clustering PN-Sets as follows:

Step 1. By use of Eqs. (4) or (5), one can calculate the similarity measure degrees of PN-Sets, and then construct a similarity matrix $C = (S_{ij})_{m\times m}$, where $S_{ij} = S_k(A_i, A_j)$ $(k = 1,2)$ for $i, j = 1,2, \dots, m$.

Step 2. The process of building the composition matrices is repeated until it holds that

$$C \to C^2 \to C^4 \to \cdots \to C^{2^k} = C^{2^{(k+1)}}$$

which implies that C^{2^k} is an equivalent similarity matrix, which is denoted by $\bar{C} = (\bar{S}_{ij})_{m\times m}$.

Step 3. For the equivalent similarity matrix $\bar{C} = (\bar{S}_{ij})_{m\times m}$, we can construct a λ-cutting matrix $\bar{C}_\lambda = (\bar{S}_{ij}^{\lambda})_{m\times m}$ of $\bar{C}$ by Eq(7); if all the elements of the ith row or column in $\bar{C}_\lambda$ are the same as the corresponding elements of the ith row or column, we conceive object sets A_i *and* A_j are the same class.

5. Illustrative Example

A car market is going to classify five different cars of A_j $(j = 1,2,\dots,5)$. Every car has six evaluation attributes: (i) x_1, fuel consumption; (ii) x_2, price; (iii) x_3, coefficient friction; (iv) x_4, comfortable degree; (v) x_5, safety. The characteristics of each car under the six attributes are represented by the form of PN-SETs, and then the Pythagorean neutrosophic data are as follows:

$$A_1 = \{x_1,(0.5,0.9,0.8), x_2,(0.6,0.7,0.7), x_3,(0.4,0.2,0.5), x_4,(0.7,0.8,0.5), x_5,(0.1,0.6,0.3)\}$$

$$A_2 = \{x_1,(0.1,0.7,0.8), x_2,(0.6,0.9,0.8), x_3,(0.5,0.2,0.4), x_4,(0.3,0.5,0.1), x_5,(0.7,0.3,0.5)\}$$

$$A_3 = \{x_1,(0.1,0.7,0.8), x_2,(0.5,0.6,0.7), x_3,(0.2,0.8,0.6), x_4,(0.4,0.3,0.9), x_5,(0.9,0.1,0.4)\}$$

$$A_4 = \{x_1,(0.3,0.6,0.7), x_2,(0.8,0.7,0.6), x_3,(0.1,0.9,0.5), x_4,(0.4,0.6,0.2), x_5,(0.5,0.2,0.7)\}$$

$$A_5 = \{x_1,(0.4,0.6,0.7), x_2,(0.9,0.6,0.1), x_3,(0.8,0.9,0.6), x_4,(0.5,0.2,0.3), x_5,(0.6,0.2,0.5)\}$$

If the weight vector of the attributes, x_i $(i = 1,2,3,4,5,6)$ $is\ w = \left(\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5}\right)^T$, then we utilize the two Pythagorean neutrosophic similarity measures to classify the five different cars of A_j $(j = 1,2,3,4,5)$ by the Pythagorean neutrosophic clustering algorithms.

5.1 Clustering Analysis using Eq. (4)

Step 1. Utilize the similarity measure formula (4) (take p = 2) to calculate the similarity measure between each pair of PN-SETs A_i *and* A_j $(i,j = 1,2,3,4,5)$ and construct the following similarity matrix:

$$C = \begin{bmatrix} 1 & 0.7358 & 0.6287 & 0.6824 & 0.6128 \\ 0.7358 & 1 & 0.7596 & 0.7339 & 0.6574 \\ 0.6287 & 0.7596 & 1 & 0.7095 & 0.6995 \\ 0.6824 & 0.7339 & 0.7095 & 1 & 0.7742 \\ 0.6168 & 0.6574 & 0.6995 & 0.7742 & 1 \end{bmatrix}$$

Step 2. Obtain equivalent similarity matrices by limited time composition of C:

$$C^2 = \begin{bmatrix} 1 & 0.7358 & 0.7358 & 0.7339 & 0.6574 \\ 0.7358 & 1 & 0.7596 & 0.7339 & 0.6574 \\ 0.7358 & 0.7596 & 1 & 0.7339 & 0.6574 \\ 0.7339 & 0.7339 & 0.7339 & 1 & 0.6574 \\ 0.6574 & 0.6574 & 0.6574 & 0.6574 & 1 \end{bmatrix}$$

$$C^4 = \begin{bmatrix} 1 & 0.7358 & 0.7358 & 0.7339 & 0.6574 \\ 0.7358 & 1 & 0.7596 & 0.7339 & 0.6574 \\ 0.7358 & 0.7596 & 1 & 0.7339 & 0.6574 \\ 0.7339 & 0.7339 & 0.7339 & 1 & 0.6574 \\ 0.6574 & 0.6574 & 0.6574 & 0.6574 & 1 \end{bmatrix}$$

Obviously, $C^4 = C^2$ implies that C^2 is an equivalent similarity matrix, denoted by $\bar{C}$.

Step 3. When λ has different values, we can construct a λ-cutting matrix $\bar{C}_\lambda = (\bar{S}_{ij}^{\lambda})_{m\times m}$ of $\bar{C}$ by Eq.(7) and obtain different categories, which give the following discussion:

(i) If $0 \leq \lambda \leq 0.6574$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

then the cars are the same category:$\{A_1, A_2, A_3, A_4, A_5\}$.

(ii) If $0.6574 < \lambda \leq 0.7339$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into two categories: $\{A_1, A_2, A_3, A_4\}, \{A_5\}$.

(iii) If $0.7339 < \lambda \leq 0.7358$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into three categories: $\{A_1, A_2, A_3\}, \{A_4\}, \{A_5\}$.

(iv) If $0.7358 < \lambda \leq 0.7596$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into four categories: $\{A_1\}, \{A_2, A_3\}, \{A_4\}, \{A_5\}$.

(v) If $0.7596 < \lambda \leq 1$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into five categories: $\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}$.

5.2 Clustering Analysis Using Eq. (5)

Step 1. Utilize the similarity measure formula (5) (take p = 2) to calculate the similarity measure between each pair of PN-SETs A_i *and* A_j $(i, j = 1,2,3,4,5)$ and construct the following similarity matrix:

$$C = \begin{bmatrix} 1 & 0.5820 & 0.4585 & 0.5179 & 0.4459 \\ 0.5820 & 1 & 0.6124 & 0.5797 & 0.4896 \\ 0.4585 & 0.6124 & 1 & 0.5498 & 0.5379 \\ 0.5179 & 0.5797 & 0.5498 & 1 & 0.6316 \\ 0.4459 & 0.4896 & 0.5379 & 0.6316 & 1 \end{bmatrix}$$

Step 2. Obtain equivalent similarity matrices by limited time composition of C:

$$C^2 = \begin{bmatrix} 1 & 0.5820 & 0.5820 & 0.5797 & 0.4896 \\ 0.5820 & 1 & 0.6124 & 0.5797 & 0.4896 \\ 0.5820 & 0.6124 & 1 & 0.5797 & 0.4896 \\ 0.5797 & 0.5797 & 0.5797 & 1 & 0.4896 \\ 0.4896 & 0.4896 & 0.4896 & 0.4896 & 1 \end{bmatrix}$$

$$C^4 = \begin{bmatrix} 1 & 0.5820 & 0.5820 & 0.5797 & 0.4896 \\ 0.5820 & 1 & 0.6124 & 0.5797 & 0.4896 \\ 0.5820 & 0.6124 & 1 & 0.5797 & 0.4896 \\ 0.5797 & 0.5797 & 0.5797 & 1 & 0.4896 \\ 0.4896 & 0.4896 & 0.4896 & 0.4896 & 1 \end{bmatrix}$$

Obviously, $C^4 = C^2$ implies that C^2 is an equivalent similarity matrix, denoted by $\bar{C}$.

Step 3. When λ has different values, we can construct a λ-cutting matrix $\bar{C_\lambda} = \left(\bar{S}_{ij}^{\lambda}\right)_{m\times m}$ of $\bar{C}$ by Eq.(7) and obtain different categories, which give the following discussion:

(i) If $0 \leq \lambda \leq 0.4896$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

then the cars are the same category:$\{A_1, A_2, A_3, A_4, A_5\}$.

(ii) If $0.4896 < \lambda \leq 0.5797$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into two categories: $\{A_1, A_2, A_3, A_4\}, \{A_5\}$.

(iii) If $0.5797 < \lambda \leq 0.5820$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into three categories: $\{A_1, A_2, A_3\}, \{A_4\}, \{A_5\}$.

(iv) If $0.5820 < \lambda \leq 0.6124$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into four categories: $\{A_1\}, \{A_2, A_3\}, \{A_4\}, \{A_5\}$.

(v) If $0.6124 < \lambda \leq 1$

$$\bar{C_\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

then the cars can be divided into five categories: $\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}$.

Conclusion

This paper introduced a generalized PN weighted distance measure and presented two distance-based similarity measures in a PN setting. Then, a PN clustering algorithm was established based on the two similarity measures. Finally, an illustrative example was given to demonstrate the application and effectiveness of the PN clustering methods.

References

[1] L. A. Zadeh, Fuzzy Sets, Information and Control, 8(1965) 338- 353.

[2] K. Atanassov, Intuitionistic Fuzzy Sets, Fuzzy Sets and Systems, 20(1986) 87-96.

[3] R. R. Yager, A. M. Abbasov, Pythagorean Membership Grades, Complex Numbers and Decision Making, International Journal of Intelligent Systems, 28 (2013) 436-452.

[4] F. Smarandache, A Unifying Field in Logics: Neutrosophic Logic, Neutrosophy, Neutrosophic Set, Neutrosophic Probability; American Research Press: Rehoboth, DE, USA, 1999.

[5] F. Smarandache, Degree of Dependence and Independence of the (Sub)Components of Fuzzy Set and Neutrosophic set. Neutrosophic Sets Systems, 11(2016) 95–97.

[6] J. Ye, Single-valued Neutrosophic Cross-entropy for Multicriteria Decision-making Problems, Applied Mathematical Modelling, 38 (2014) 1170-1175.

[7] J. Ye, Multicriteria Decision-making Method Using the Correlation Coefficient Under Single-valued Neutrosophic Environment, International Journal of General Systems, 42(4) (2013) 386–394.

[8] J. Ye, Similarity Measure Between Interval Neutrosophic Sets and Their Applications in Multicriteria Decision Making, Journal of Intelligent and Fuzzy systems 26(2014) 165–172.

[9] Z. S. Xu, J. Chen, J. J. Wu, Clustering Algorithm for Intuitionistic Fuzzy Sets, Information Science, 19 (2008) 3775-3790.

[10] H. M. Zhang, Z. S. Xu, Q. Chen, Clustering Method of Intuitionistic Fuzzy Sets, Control Decision, 22 (2007) 882-888.

[11] J. Ye, Clustering Methods using Distance-Based Similarity Measures of Single-valued Neutrosophic Sets, Journal Intelligent Systems, 23 (2014) 379-389.

[12] X. Peng, Y. Yang, Some Results for Pythagorean Fuzzy Sets, International Journal of Intelligent Systems, 30 (2015) 1133-1160.

[13] R. R. Yager, Pythagorean Fuzzy Subsets, in: Proc Joint IFSA World Congress and NAFIPS Annual Meeting, Edmonton, Canada (2013) 57-61.