

Research Article

Correlation Measures of Dual Hesitant Fuzzy Sets

Lei Wang, Mingfang Ni, and Lei Zhu

Institute of Communication Engineering, PLA University of Science and Technology, Nanjing, Jiangsu 210007, China

Correspondence should be addressed to Lei Wang; iponly@126.com

Received 23 June 2013; Revised 9 September 2013; Accepted 9 September 2013

Academic Editor: Toly Chen

Copyright © 2013 Lei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The dual hesitant fuzzy sets (DHFSs) were proposed by Zhu et al. (2012), which encompass fuzzy sets, intuitionistic fuzzy sets, hesitant fuzzy sets, and fuzzy multisets as special cases. Correlation measures analysis is an important research topic. In this paper, we define the correlation measures for dual hesitant fuzzy information and then discuss their properties in detail. One numerical example is provided to illustrate these correlation measures. Then we present a direct transfer algorithm with respect to the problem of complex operation of matrix synthesis when reconstructing an equivalent correlation matrix for clustering DHFSs. Furthermore, we prove that the direct transfer algorithm is equivalent to transfer closure algorithm, but its asymptotic time complexity and space complexity are superior to the latter. Another real world example, that is, diamond evaluation and classification, is employed to show the effectiveness of the association coefficient and the algorithm for clustering DHFSs.

1. Introduction

Correlation indicates how well two variables move together in a linear fashion. In other words, correlation reflects a linear relationship between two variables. It is an important measure in data analysis, in particular in decision making, medical diagnosis, pattern recognition, and other real world problems [1–7]. Zadeh [8] introduced the concept of fuzzy sets (FSs) whose basic component is only a membership function with the nonmembership function being one minus the membership function. In fuzzy environments, Hung and Wu [9] used the concept of “expected value” to define the correlation coefficient of fuzzy numbers, which lies in $[-1, 1]$. Hong [10] considered the computational aspect of the T_ω -based extension principle when the principle is applied to a correlation coefficient of L - R fuzzy numbers and gave the exact solution of a fuzzy correlation coefficient without programming or the aid of computer resources. Atanassov [11, 12] gave a generalized form of fuzzy set, called intuitionistic fuzzy set (IFS), which is characterized by a membership function and a non-membership function. In intuitionistic fuzzy environments, Gerstenkorn and Mańko [13] defined a function measuring the correlation of IFSs and introduced a coefficient of such a correlation. Bustince and Burillo [14] introduced the concepts of correlation and correlation coefficient of interval-valued intuitionistic fuzzy

sets (IVIFSs) [12]. Hung [15] and Mitchell [16] derived the correlation coefficient of intuitionistic fuzzy sets from a statistical viewpoint by interpreting an intuitionistic fuzzy set as an ensemble of ordinary fuzzy sets. Hung and Wu [17] proposed a method to calculate the correlation coefficient of intuitionistic fuzzy sets by means of “centroid.” Xu [18] gave a detailed survey on association analysis of intuitionistic fuzzy sets and pointed out that most existing methods deriving association coefficients cannot guarantee that the association coefficient of any two intuitionistic fuzzy sets equals one if and only if these two intuitionistic fuzzy sets are the same. Szmidt and Kacprzyk [5] discussed a concept of correlation for data represented as intuitionistic fuzzy set adopting the concepts from statistics and proposed a formula for measuring the correlation coefficient (lying in $[-1, 1]$) of intuitionistic fuzzy sets. Robinson and Amirtharaj [19] defined the correlation coefficient of interval vague sets lying in the interval $[0, 1]$ and proposed a new method for computing the correlation coefficient of interval vague sets lying in the interval $[-1, 1]$ using α -cuts over the vague degrees through statistical confidence intervals which is presented by an example. Instead of using point-based membership as in fuzzy sets, interval-based membership is used in a vague set. In [20], Robinson and Amirtharaj presented a detailed comparison between vague sets and intuitionistic fuzzy sets and defined the correlation coefficient of vague sets through

simple examples. Hesitant fuzzy sets (HFSs) were originally introduced by Torra [21, 22]. In hesitant fuzzy environments, Chen et al. [23] derived some correlation coefficient formulas for HFSs and applied them to two real world examples by using clustering analysis under hesitant fuzzy environments. Xu and Xia [24] defined the correlation measures for hesitant fuzzy information and then discussed their properties in detail.

Recently, Dubois and Prade introduced the definition of dual hesitant fuzzy set. Dual hesitant fuzzy set can reflect human's hesitance more objectively than the other classical extensions of fuzzy set (intuitionistic fuzzy set, type-2 fuzzy set (T-2FS) [25], hesitant fuzzy set, etc.). The motivation to propose the DHFSs is that when people make a decision, they are usually hesitant and irresolute for one thing or another which makes it difficult to reach a final agreement. They further indicated that DHFSs can better deal with the situations that permit both the membership and the nonmembership of an element to a given set having a few different values, which can arise in a group decision making problem. For example, in the organization, some decision makers discuss the membership degree 0.6 and the non-membership 0.3 of an alternative A that satisfies a criterion x . Some possibly assign (0.8, 0.2), while the others assign (0.7, 0.2). No consistency is reached among these decision makers. Accordingly, the difficulty of establishing a common membership degree and a non-membership degree is not because we have a margin of error (intuitionistic fuzzy set) or some possibility distribution values (type-2 fuzzy set), but because we have a set of possible values (hesitant fuzzy set). For such a case, the satisfactory degrees can be represented by a dual hesitant fuzzy element $\{(0.6, 0.8, 0.7), (0.3, 0.2)\}$, which is obviously different from intuitionistic fuzzy number (0.8, 0.2) or (0.7, 0.2) and hesitant fuzzy number $\{0.6, 0.8, 0.7\}$. The aforementioned measures, however, cannot be used to deal with the correlation measures of dual hesitant fuzzy information. Thus, it is very necessary to develop some theories for dual hesitant fuzzy sets. However, little has been done about this issue. In this paper, we mainly discuss the correlation measures of dual hesitant fuzzy information. To do this, the remainder of the paper is organized as follows. Section 2 presents some basic concepts related to DHFSs, HFSs, and IFSs. In Section 3, we propose some correlation measures of dual hesitant fuzzy elements, obtain several important conclusions, and given an example to illustrate the correlation measures. In Section 4, we propose a direct transfer clustering algorithm based on DHFSs and then use a numerical example to illustrate our algorithm. Finally, Section 5 concludes the paper with some remarks and presents future challenges.

2. Preliminaries

2.1. DHFSs, HFSs, and IFSs

Definition 1 (see [26]). Let X be a fixed set then a dual hesitant fuzzy set (DHFS) D on X is described as;

$$D = \{\langle x, h(x), g(x) \rangle \mid x \in X\} \quad (1)$$

in which $h(x)$ and $g(x)$ are two sets of some values in $[0, 1]$, denoting the possible membership degrees and non-membership degrees of the element $x \in X$ to the set D , respectively, with the conditions

$$0 \leq \gamma, \quad \eta \leq 1, \quad 0 \leq \gamma^+ + \eta^+ \leq 1, \quad (2)$$

where $\gamma \in h(x)$, $\eta \in g(x)$, $\gamma^+ \in h^+(x) = \cup_{\gamma \in h(x)} \max\{\gamma\}$, and $\eta^+ \in g^+(x) = \cup_{\eta \in g(x)} \max\{\eta\}$ for all $x \in X$. For convenience, the pair $d_E(x) = (h_E(x), g_E(x))$ is called a dual hesitant fuzzy element (DHFE), denoted by $d = (h, g)$, with the conditions $\gamma \in h(x)$, $\eta \in g(x)$, $\gamma^+ \in h^+(x) = \cup_{\gamma \in h(x)} \max\{\gamma\}$, and $\eta^+ \in g^+(x) = \cup_{\eta \in g(x)} \max\{\eta\}$, $0 \leq \gamma, \eta \leq 1$ and $0 \leq \gamma^+ + \eta^+ \leq 1$.

Definition 2 (see [21, 22]). Let X be a fixed set; a hesitant fuzzy set (HFS) A on X is in terms of a function that when applied to X returns a subset of $[0, 1]$, which can be represented as the following mathematical symbol:

$$A = \{\langle x, h_A(x) \rangle \mid x \in X\}, \quad (3)$$

where $h_A(x)$ is a set of values in $[0, 1]$, denoting the possible membership degrees of the element $x \in X$ to the set A . For convenience, we call $h_A(x)$ a hesitant fuzzy element (HFE). We use $\langle x, h_A \rangle$ for all $x \in X$ to represent HFSs.

Definition 3 (see [11, 12]). Let X be a fixed set, an intuitionistic fuzzy set (IFS) A on X is an object having the form

$$A = \{\langle x, \mu_A(x), \nu_A(x) \rangle \mid x \in X\} \quad (4)$$

which is characterized by a membership function μ_A and a non-membership function ν_A , where $\mu_A : X \rightarrow [0, 1]$ and $\nu_A : X \rightarrow [0, 1]$, with the condition $0 \leq \mu_A(x) + \nu_A(x) \leq 1$, for all $x \in X$. We use $\langle x, \mu_A, \nu_A \rangle$ for all $x \in X$ to represent IFSs considered in the rest of the paper without explicitly mentioning it. Furthermore, $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$ is called a hesitancy degree or an intuitionistic index of x in A . In the special case $\pi(x) = 0$, that is, $\mu_A(x) + \nu_A(x) = 1$, the IFS A reduces to an FS.

2.2. Correlation Coefficients of HFSs and IFSs. Many approaches [4, 13, 17, 20, 21] have been introduced to compute the correlation coefficients of IFSs. Let $X = \{x_1, x_2, \dots, x_n\}$ be a discrete universe of discourse, for any two A and B on X .

The correlation of the IFSs A and B is defined as [13]

$$C_{\text{IFS}_1}(A, B) = \sum_{i=1}^n (\mu_A(x_i) \mu_B(x_i) + \nu_A(x_i) \nu_B(x_i)), \quad (5)$$

Then, the correlation coefficient of the IFSs A and B is defined as

$$\begin{aligned} \rho_{\text{IFS}_1}(A, B) &= \frac{C_{\text{IFS}_1}(A, B)}{(C_{\text{IFS}_1}(A, A) \cdot C_{\text{IFS}_1}(B, B))^{1/2}} \\ &= \left(\sum_{i=1}^n (u_A(x_i) u_B(x_i) + v_A(x_i) v_B(x_i)) \right) \\ &\quad \times \left(\left(\left(\sum_{i=1}^n (u_A^2(x_i) + v_A^2(x_i)) \right) \right. \right. \\ &\quad \left. \left. \cdot \left(\sum_{i=1}^n (u_B^2(x_i) + v_B^2(x_i)) \right) \right) \right)^{1/2} \right)^{-1}. \end{aligned} \tag{6}$$

In [23], Chen et al. defined the correlation and correlation coefficient for HFSs as follows, respectively:

$$\begin{aligned} C_{\text{HFS}_1}(A, B) &= \sum_{i=1}^n \left(\frac{1}{l_i} \sum_{j=1}^{l_i} h_{A_{\sigma(j)}}(x_i) h_{B_{\sigma(j)}}(x_i) \right), \\ \rho_{\text{HFS}_1}(A, B) &= \frac{C_{\text{HFS}_1}(A, B)}{(C_{\text{HFS}_1}(A, A) \cdot C_{\text{HFS}_1}(B, B))^{1/2}} \\ &= \left(\sum_{i=1}^n \left(\frac{1}{l_i} \sum_{j=1}^{l_i} h_{A_{\sigma(j)}}(x_i) h_{B_{\sigma(j)}}(x_i) \right) \right) \\ &\quad \times \left(\left(\left(\sum_{i=1}^n \left(\frac{1}{l_i} \sum_{j=1}^{l_i} h_{A_{\sigma(j)}}^2(x_i) \right) \right) \right. \right. \\ &\quad \left. \left. \cdot \left(\sum_{i=1}^n \left(\frac{1}{l_i} \sum_{j=1}^{l_i} h_{B_{\sigma(j)}}^2(x_i) \right) \right) \right) \right)^{1/2} \right)^{-1}, \end{aligned} \tag{7}$$

where $l_i = \max\{l(h_A(x_i)), l(h_B(x_i))\}$ for each x_i in X , and $l(h_A(x_i))$ and $l(h_B(x_i))$ represent the number of values in $h_A(x_i)$ and $h_B(x_i)$, respectively. We will talk about l_i in detail in the next section.

3. Correlation Measures of DHFEs

In this section, we first introduce the concept of correlation and correlation coefficient for DHFSs and then propose several correlation coefficient formulas and discuss their properties.

We arrange the elements in $d_E(x) = (h_E(x), g_E(x))$ in decreasing order and let $\gamma_E^{\sigma(i)}(x)$ be the i th largest value in $h_E(x)$ and $\eta_E^{\sigma(j)}(x)$ the j th largest value in $g_E(x)$. Let $l_h(d_E(x_i))$ the number of values in $h_E(x_i)$ and $l_g(d_E(x_i))$ be the number of values in $g_E(x_i)$. For convenience, $l(d(x_i)) = (l_h(d(x_i)), l_g(d(x_i)))$. In most cases, for two DHFSs A and B , $l(d_A(x_i)) \neq l(d_B(x_i))$; that is, $l_h(d_A(x_i)) \neq l_h(d_B(x_i))$,

$l_g(d_A(x_i)) \neq l_g(d_B(x_i))$. To operate correctly, we should extend the shorter one until both of them have the same length when we compare them. In [24, 27], Xu and Xia extended the shorter one by adding different values in hesitant fuzzy environments. Similarly, Torra [21] also applied this ideal to derive some correlation coefficient formulas for HFSs. In fact, we can extend the shorter one by adding any value in it. The selection of this value mainly depends on the decision makers' risk preferences. Optimists anticipate desirable outcomes and may add the maximum value, while pessimists expect unfavorable outcomes and may add the minimum value. The same situation can also be found in many existing references [13, 14].

We define several correlation coefficients for DHFEs.

Definition 4. For two DHFSs A and B on X , the correlation of A and B , denoted as $C_{\text{DHFS}_1}(A, B)$, is defined by

$$\begin{aligned} C_{\text{DHFS}_1}(A, B) &= \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A_{\sigma(j)}}(x_i) \gamma_{B_{\sigma(j)}}(x_i) \right. \\ &\quad \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A_{\sigma(k)}}(x_i) \eta_{B_{\sigma(k)}}(x_i) \right). \end{aligned} \tag{8}$$

Definition 5. For two DHFSs A and B on X , the correlation coefficient of A and B , denoted as $\rho_{\text{DHFS}_1}(A, B)$, is defined by:

$$\begin{aligned} \rho_{\text{DHFS}_1}(A, B) &= \frac{C_{\text{DHFS}_1}(A, B)}{(C_{\text{DHFS}_1}(A, A) \cdot C_{\text{DHFS}_1}(B, B))^{1/2}} \\ &= \left(\sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A_{\sigma(j)}}(x_i) \gamma_{B_{\sigma(j)}}(x_i) \right. \right. \\ &\quad \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A_{\sigma(k)}}(x_i) \eta_{B_{\sigma(k)}}(x_i) \right) \right) \\ &\quad \times \left(\left(\left(\sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A_{\sigma(j)}}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A_{\sigma(k)}}^2(x_i) \right) \right) \right. \right. \\ &\quad \left. \left. \cdot \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B_{\sigma(j)}}^2(x_i) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B_{\sigma(k)}}^2(x_i) \right) \right) \right)^{1/2} \right)^{-1}. \end{aligned} \tag{9}$$

Definition 6. For two DHFSs A and B on X , the correlation coefficient of A and B , denoted as $\rho_{\text{DHFS}_2}(A, B)$, is defined by

$$\begin{aligned} \rho_{\text{DHFS}_2}(A, B) &= \frac{C_{\text{DHFS}_1}(A, B)}{\max\{C_{\text{DHFS}_1}(A, A), C_{\text{DHFS}_1}(B, B)\}} \\ &= \left(\sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}(x_i) \gamma_{B\sigma(j)}(x_i) \right. \right. \\ &\quad \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}(x_i) \eta_{B\sigma(k)}(x_i) \right) \right) \\ &\times \left(\max \left\{ \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}^2(x_i) \right), \right. \right. \\ &\quad \left. \left. \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B\sigma(j)}^2(x_i) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B\sigma(k)}^2(x_i) \right) \right\} \right)^{-1}. \end{aligned} \quad (10)$$

Theorem 7. For two DHFSs A and B , the correlation coefficient of A and B , denoted as $\rho_{\text{DHFS}_i}(A, B)$ ($i = 1, 2$), should satisfy the following properties:

- (1) $0 \leq \rho_{\text{DHFS}_i}(A, B) \leq 1$;
- (2) $A = B \Rightarrow \rho_{\text{DHFS}_i}(A, B) = 1$;
- (3) $\rho_{\text{DHFS}_i}(A, B) = \rho_{\text{DHFS}_i}(B, A)$; $i = 1, 2$.

Proof. (1) The inequality $0 \leq \rho_{\text{DHFS}_1}(A, B)$ and $0 \leq \rho_{\text{DHFS}_2}(A, B)$ is obvious. Below let us prove that $\rho_{\text{DHFS}_1}(A, B) \leq 1$, $\rho_{\text{DHFS}_2}(A, B) \leq 1$:

$$\begin{aligned} C_{\text{DHFS}_1}(A, B) &= \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}(x_i) \gamma_{B\sigma(j)}(x_i) \right. \\ &\quad \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}(x_i) \eta_{B\sigma(k)}(x_i) \right) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^{l_{h(i)}} \frac{\gamma_{A\sigma(j)}(x_i) \gamma_{B\sigma(j)}(x_i)}{l_{h(i)}} \right. \\ &\quad \left. + \sum_{k=1}^{l_{g(i)}} \frac{\eta_{A\sigma(k)}(x_i) \eta_{B\sigma(k)}(x_i)}{l_{g(i)}} \right) \end{aligned}$$

$$\begin{aligned} &= \left(\sum_{j=1}^{l_{h(1)}} \frac{\gamma_{A\sigma(j)}(x_1)}{\sqrt{l_{h(1)}}} \cdot \frac{\gamma_{B\sigma(j)}(x_1)}{\sqrt{l_{h(1)}}} \right. \\ &\quad \left. + \sum_{j=1}^{l_{h(2)}} \frac{\gamma_{A\sigma(j)}(x_2)}{\sqrt{l_{h(2)}}} \cdot \frac{\gamma_{B\sigma(j)}(x_2)}{\sqrt{l_{h(2)}}} \right. \\ &\quad \left. + \dots + \sum_{j=1}^{l_{h(n)}} \frac{\gamma_{A\sigma(j)}(x_n)}{\sqrt{l_{h(n)}}} \cdot \frac{\gamma_{B\sigma(j)}(x_n)}{\sqrt{l_{h(n)}}} \right) \\ &\quad + \left(\sum_{k=1}^{l_{g(1)}} \frac{\eta_{A\sigma(k)}(x_1)}{\sqrt{l_{g(1)}}} \cdot \frac{\eta_{B\sigma(k)}(x_1)}{\sqrt{l_{g(1)}}} \right. \\ &\quad \left. + \sum_{k=1}^{l_{g(2)}} \frac{\eta_{A\sigma(k)}(x_2)}{\sqrt{l_{g(2)}}} \cdot \frac{\eta_{B\sigma(k)}(x_2)}{\sqrt{l_{g(2)}}} \right. \\ &\quad \left. + \dots + \sum_{k=1}^{l_{g(n)}} \frac{\eta_{A\sigma(k)}(x_n)}{\sqrt{l_{g(n)}}} \cdot \frac{\eta_{B\sigma(k)}(x_n)}{\sqrt{l_{g(n)}}} \right). \end{aligned} \quad (11)$$

Using the Cauchy-Schwarz inequality

$$\begin{aligned} &(x_1 y_1 + x_2 y_2 + \dots + x_n y_n)^2 \\ &\leq (x_1^2 + x_2^2 + \dots + x_n^2) \cdot (y_1^2 + y_2^2 + \dots + y_n^2), \end{aligned} \quad (12)$$

where $(x_1, x_2, \dots, x_n) \in R^n$, $(y_1, y_2, \dots, y_n) \in R^n$, we obtain

$$\begin{aligned} C_{\text{DHFS}_1}(A, B)^2 &\leq \left(\sum_{j=1}^{l_{h(1)}} \frac{\gamma_{A\sigma(j)}^2(x_1)}{l_{h(1)}} + \sum_{j=1}^{l_{h(2)}} \frac{\gamma_{A\sigma(j)}^2(x_2)}{l_{h(2)}} \right. \\ &\quad \left. + \dots + \sum_{j=1}^{l_{h(n)}} \frac{\gamma_{A\sigma(j)}^2(x_n)}{l_{h(n)}} + \sum_{k=1}^{l_{g(1)}} \frac{\eta_{A\sigma(k)}^2(x_1)}{l_{g(1)}} \right. \\ &\quad \left. + \sum_{k=1}^{l_{g(2)}} \frac{\eta_{A\sigma(k)}^2(x_2)}{l_{g(2)}} + \dots + \sum_{k=1}^{l_{g(n)}} \frac{\eta_{A\sigma(k)}^2(x_n)}{l_{g(n)}} \right) \\ &\quad \cdot \left(\sum_{j=1}^{l_{h(1)}} \frac{\gamma_{B\sigma(j)}^2(x_1)}{l_{h(1)}} + \sum_{j=1}^{l_{h(2)}} \frac{\gamma_{B\sigma(j)}^2(x_2)}{l_{h(2)}} \right. \\ &\quad \left. + \dots + \sum_{j=1}^{l_{h(n)}} \frac{\gamma_{B\sigma(j)}^2(x_n)}{l_{h(n)}} + \sum_{k=1}^{l_{g(1)}} \frac{\eta_{B\sigma(k)}^2(x_1)}{l_{g(1)}} \right. \\ &\quad \left. + \sum_{k=1}^{l_{g(2)}} \frac{\eta_{B\sigma(k)}^2(x_2)}{l_{g(2)}} + \dots + \sum_{k=1}^{l_{g(n)}} \frac{\eta_{B\sigma(k)}^2(x_n)}{l_{g(n)}} \right) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}^2(x_i) \right) \\
 &\quad \cdot \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B\sigma(k)}^2(x_i) \right) \\
 &= C_{DHFS_1}(A, A) \cdot C_{DHFS_1}(B, B).
 \end{aligned} \tag{13}$$

Therefore,

$$C_{DHFS_1}(A, B) \leq (C_{DHFS_1}(A, A))^{1/2} \cdot (C_{DHFS_1}(B, B))^{1/2}. \tag{14}$$

So, $0 \leq \rho_{DHFS_1}(A, B) \leq 1$.

In fact, we have

$$\begin{aligned}
 &(x_1^2 + x_2^2 + \dots + x_n^2) \cdot (y_1^2 + y_2^2 + \dots + y_n^2) \\
 &\leq (\max((x_1^2 + x_2^2 + \dots + x_n^2), (y_1^2 + y_2^2 + \dots + y_n^2)))^2, \\
 &\sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}^2(x_i) \right) \\
 &\quad \cdot \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B\sigma(k)}^2(x_i) \right) \\
 &\leq \left(\max \left\{ \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}^2(x_i) \right), \right. \right. \\
 &\quad \left. \left. \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B\sigma(k)}^2(x_i) \right) \right\} \right)^2.
 \end{aligned} \tag{15}$$

Then

$$\begin{aligned}
 &(C_{DHFS_1}(A, A) \cdot C_{DHFS_1}(B, B))^{1/2} \\
 &\leq \max \{C_{DHFS_1}(A, A), C_{DHFS_1}(B, B)\}.
 \end{aligned} \tag{16}$$

We also obtain $0 \leq \rho_{DHFS_2}(A, B) \leq 1$.

(2) and (3) are straightforward. \square

Moreover, from the proof of Theorem 7, we have Theorem 8 easily.

Theorem 8. For two DHFSs A and B on X , then $\rho_{DHFS_1}(A, B) \geq \rho_{DHFS_2}(A, B)$.

However, from Theorem 7, we notice that all the above correlation coefficients cannot guarantee that the correlation coefficient of any two DHFSs equals one if and only if these two DHFSs are the same. Thus, how to derive the correlation coefficients of the DHFSs satisfying this desirable property is an interesting research topic. To solve this issue, in what follows, we develop a new method to calculate the correlation coefficient of the DHFSs A and B .

Definition 9. For two DHFSs A and B on X , the correlation coefficient of A and B , denoted as $\rho_{DHFS_3}(A, B)$, is defined by

$$\begin{aligned}
 &\rho_{DHFS_3}(A, B) \\
 &= \left(\frac{1}{2n} \sum_{i=1}^n \left(\frac{\Delta\gamma_{\min}^\lambda + \Delta\gamma_{\max}^\lambda}{\Delta\gamma_i^\lambda + \Delta\gamma_{\max}^\lambda} + \frac{\Delta\eta_{\min}^\lambda + \Delta\eta_{\max}^\lambda}{\Delta\eta_i^\lambda + \Delta\eta_{\max}^\lambda} \right) \right)^{1/\lambda},
 \end{aligned} \tag{17}$$

where

$$\begin{aligned}
 \Delta\gamma_i^\lambda &= \frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \left| \gamma_{A\sigma(j)}(x_i) - \gamma_{B\sigma(j)}(x_i) \right|^\lambda, \\
 \Delta\gamma_{\min}^\lambda &= \min_i \{ \Delta\gamma_i^\lambda \}, \quad \Delta\gamma_{\max}^\lambda = \max_i \{ \Delta\gamma_i^\lambda \}, \\
 \Delta\eta_i^\lambda &= \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \left| \eta_{A\sigma(k)}(x_i) - \eta_{B\sigma(k)}(x_i) \right|^\lambda, \\
 \Delta\eta_{\min}^\lambda &= \min_i \{ \Delta\eta_i^\lambda \}, \quad \Delta\eta_{\max}^\lambda = \max_i \{ \Delta\eta_i^\lambda \}, \\
 &\lambda > 0.
 \end{aligned} \tag{18}$$

Equation (17) is motivated by the generalized idea provided by Xu [18]. Obviously, the greater the value of $\rho_{DHFS_3}(A, B)$, the closer A to B . By Definition 9, we have Theorem 10.

Theorem 10. The correlation coefficient $\rho_{DHFS_3}(A, B)$ satisfies the following properties:

- (1) $0 \leq \rho_{DHFS_3}(A, B) \leq 1$;
- (2) $A = B \Leftrightarrow \rho_{DHFS_3}(A, B) = 1$;
- (3) $\rho_{DHFS_3}(A, B) = \rho_{DHFS_3}(B, A)$.

Proof. (1) The inequality $0 \leq \rho_{DHFS_3}(A, B)$ is obvious. Below let us prove that $\rho_{DHFS_3}(A, B) \leq 1$:

$$\begin{aligned}
 &\frac{\Delta\gamma_{\min}^\lambda + \Delta\gamma_{\max}^\lambda}{\Delta\gamma_i^\lambda + \Delta\gamma_{\max}^\lambda} + \frac{\Delta\eta_{\min}^\lambda + \Delta\eta_{\max}^\lambda}{\Delta\eta_i^\lambda + \Delta\eta_{\max}^\lambda} \quad \text{for } i = 1, 2, \dots, n \\
 &= \frac{(\Delta\gamma_{\min}^\lambda / \Delta\gamma_{\max}^\lambda) + 1}{(\Delta\gamma_i^\lambda / \Delta\gamma_{\max}^\lambda) + 1} + \frac{(\Delta\eta_{\min}^\lambda / \Delta\eta_{\max}^\lambda) + 1}{(\Delta\eta_i^\lambda / \Delta\eta_{\max}^\lambda) + 1} \\
 &\leq 2.
 \end{aligned} \tag{19}$$

We obtain

$$\begin{aligned}
 &\rho_{DHFS_3}(A, B) \\
 &= \frac{1}{2n} \sum_{i=1}^n \left(\frac{\Delta\gamma_{\min}^\lambda + \Delta\gamma_{\max}^\lambda}{\Delta\gamma_i^\lambda + \Delta\gamma_{\max}^\lambda} + \frac{\Delta\eta_{\min}^\lambda + \Delta\eta_{\max}^\lambda}{\Delta\eta_i^\lambda + \Delta\eta_{\max}^\lambda} \right) \\
 &\leq \frac{1}{2n} \cdot 2n \\
 &= 1.
 \end{aligned} \tag{20}$$

(2) and (3) are obvious. \square

Usually, in practical applications, the weight of each element $x_i \in X$ should be taken into account, and, so, we present the following weighted correlation coefficient. Assume that the weight of the element $x_i \in X$ is w_i ($i = 1, 2, \dots, n$) with $w_i \in [0, 1]$ and $\sum_{i=1}^n w_i = 1$; then we extend the correlation coefficient formulas given:

$$\begin{aligned} \rho_{\text{DHFS-}w_1}(A, B) &= \left(\sum_{i=1}^n w_i \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}(x_i) \gamma_{B\sigma(j)}(x_i) \right. \right. \\ &\quad \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}(x_i) \eta_{B\sigma(k)}(x_i) \right) \right) \\ &\quad \times \left(\left(\sum_{i=1}^n w_i \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}^2(x_i) + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}^2(x_i) \right) \right. \right. \\ &\quad \left. \left. \cdot \sum_{i=1}^n w_i \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B\sigma(j)}^2(x_i) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B\sigma(k)}^2(x_i) \right) \right) \right)^{1/2} \right)^{-1}, \end{aligned} \quad (21)$$

$$\begin{aligned} \rho_{\text{DHFS-}w_2}(A, B) &= \left(\sum_{i=1}^n w_i \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}(x_i) \gamma_{B\sigma(j)}(x_i) \right. \right. \\ &\quad \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}(x_i) \eta_{B\sigma(k)}(x_i) \right) \right) \\ &\quad \times \left(\max \left\{ w_i \sum_{i=1}^n \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{A\sigma(j)}^2(x_i) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{A\sigma(k)}^2(x_i) \right), \right. \right. \\ &\quad \left. \left. \sum_{i=1}^n w_i \left(\frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \gamma_{B\sigma(j)}^2(x_i) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \eta_{B\sigma(k)}^2(x_i) \right) \right\} \right)^{-1}, \end{aligned} \quad (22)$$

$$\begin{aligned} \rho_{\text{DHFS-}w_3}(A, B) &= \left(\frac{1}{2} \sum_{i=1}^n w_i \left(\frac{\Delta\gamma_{\min}^\lambda + \Delta\gamma_{\max}^\lambda}{\Delta\gamma_i^\lambda + \Delta\gamma_{\max}^\lambda} + \frac{\Delta\eta_{\min}^\lambda + \Delta\eta_{\max}^\lambda}{\Delta\eta_i^\lambda + \Delta\eta_{\max}^\lambda} \right) \right)^{1/\lambda}, \end{aligned} \quad (23)$$

where

$$\begin{aligned} \Delta\gamma_i^\lambda &= \frac{1}{l_{h(i)}} \sum_{j=1}^{l_{h(i)}} \left| \gamma_{A\sigma(j)}(x_i) - \gamma_{B\sigma(j)}(x_i) \right|^\lambda, \\ \Delta\gamma_{\min}^\lambda &= \min_i \{ \Delta\gamma_i^\lambda \}, \quad \Delta\gamma_{\max}^\lambda = \max_i \{ \Delta\gamma_i^\lambda \}, \\ \Delta\eta_i^\lambda &= \frac{1}{l_{g(i)}} \sum_{k=1}^{l_{g(i)}} \left| \eta_{A\sigma(k)}(x_i) - \eta_{B\sigma(k)}(x_i) \right|^\lambda, \\ \Delta\eta_{\min}^\lambda &= \min_i \{ \Delta\eta_i^\lambda \}, \quad \Delta\eta_{\max}^\lambda = \max_i \{ \Delta\eta_i^\lambda \}, \\ &\quad \lambda > 0. \end{aligned} \quad (24)$$

Note that all these formulas satisfy the properties in Theorem 7.

In what follows, we use a medical diagnosis problem in [28, 29] to illustrate the developed correlation coefficient formulas. Actually, this is also a pattern recognition problem.

Example 11. To make a proper diagnosis $Q = \{Q_1$ (viral fever), Q_2 (malaria), Q_3 (typhoid), Q_4 (stomach problem), and Q_5 (chest problem)} for a patient with the given values of the symptoms, $S = \{S_1$ (temperature), S_2 (headache), S_3 (cough), S_4 (stomach pain), and S_5 (chest pain)}, Xu [18] considered all possible diagnoses and symptoms as HFEs. Utilizing DHFSs can take much more information into account; the more values we obtain from patients, the greater epistemic certainty we have. So, in this paper, we use DHFEs to deal with such cases; each symptom is described by a DHFE, which is described by two sets (γ_{ij}) and (η_{ij}) . (γ_{ij}) indicates the degree that symptoms characteristic S_i satisfies the considered diagnoses G_j and (η_{ij}) indicates the degree that the symptoms characteristic S_i does not satisfy the considered diagnoses G_j . The data are given in Table 1. The set of patients is $P = \{\text{Al, Bob, Joe, Ted}\}$. The symptoms which can be also described by DHFEs are given in Table 2. We need to seek a diagnosis for each patient.

We utilize the correlation coefficient ρ_{DHFS_1} to derive a diagnosis for each patient. All the results for the considered patients are listed in Table 3. From the arguments in Table 3, we can find that Ted suffers from viral fever, Al and Joe from malaria, and Bob from stomach problem.

If we utilize the correlation coefficient formulas ρ_{DHFS_2} and ρ_{DHFS_3} to derive a diagnosis, then the results are listed in Tables 4 and 5, respectively.

From Tables 3–5 we know that the results obtained by different correlation coefficient formulas are different. That is because these correlation coefficient formulas are based on different linear relationships.

4. Clustering Method Based on Direct Transfer Algorithm for HFSs

Based on clustering algorithms for IFSS [30, 31], and HFSs [23], and the correlation coefficient formulas developed previously for DHFSs, in what follows, we propose a direct

TABLE 1: Symptoms characteristic of the considered diagnoses.

	Temperature	Headache	Cough	Stomach pain	Chest pain
Viral fever	{{(0.6, 0.4, 0.3), (0.2, 0.0)}}	{{(0.7, 0.5, 0.3, 0.2), (0.3, 0.1)}}	{{(0.5, 0.3), (0.5, 0.4, 0.2)}}	{{(0.5, 0.4, 0.3, 0.2, 0.1), (0.5, 0.3)}}	{{(0.5, 0.4, 0.2, 0.1), (0.5, 0.4, 0.3)}}
Malaria	{{(0.9, 0.8, 0.7), (0.1, 0.0)}}	{{(0.5, 0.3, 0.2, 0.1), (0.4, 0.3)}}	{{(0.2, 0.1), (0.7, 0.6, 0.5)}}	{{(0.6, 0.5, 0.3, 0.2, 0.1), (0.3, 0.2)}}	{{(0.4, 0.3, 0.2, 0.1), (0.6, 0.5, 0.4)}}
Typhoid	{{(0.6, 0.3, 0.1), (0.3, 0.2)}}	{{(0.9, 0.8, 0.7, 0.6), (0.1, 0.0)}}	{{(0.5, 0.3), (0.5, 0.4, 0.3)}}	{{(0.5, 0.4, 0.3, 0.2, 0.1), (0.5, 0.4)}}	{{(0.6, 0.4, 0.3, 0.2), (0.4, 0.3, 0.2)}}
Stomach problem	{{(0.5, 0.4, 0.2), (0.5, 0.3)}}	{{(0.4, 0.3, 0.2, 0.1), (0.4, 0.3)}}	{{(0.4, 0.3), (0.6, 0.5, 0.4)}}	{{(0.9, 0.8, 0.7, 0.6, 0.5), (0.1, 0.0)}}	{{(0.5, 0.4, 0.2, 0.1), (0.5, 0.4, 0.3)}}
Chest problem	{{(0.3, 0.2, 0.1), (0.7, 0.6)}}	{{(0.5, 0.3, 0.2, 0.1), (0.5, 0.3)}}	{{(0.3, 0.2), (0.6, 0.4, 0.3)}}	{{(0.7, 0.6, 0.5, 0.3, 0.2), (0.2, 0.1)}}	{{(0.8, 0.7, 0.6, 0.5), (0.2, 0.1, 0.0)}}

TABLE 2: Symptoms characteristic of the considered patients.

	Temperature	Headache	Cough	Stomach pain	Chest pain
Al	{{(0.9, 0.7, 0.5), (0.1, 0.0)}}	{{(0.4, 0.3, 0.2, 0.1), (0.5, 0.4)}}	{{(0.4, 0.3), (0.5, 0.4, 0.2)}}	{{(0.6, 0.5, 0.4, 0.2, 0.1), (0.3, 0.2)}}	{{(0.4, 0.3, 0.2, 0.1), (0.5, 0.4, 0.3)}}
Bob	{{(0.5, 0.4, 0.2), (0.5, 0.3)}}	{{(0.5, 0.4, 0.3, 0.1), (0.4, 0.3)}}	{{(0.2, 0.1), (0.7, 0.6, 0.5)}}	{{(0.9, 0.8, 0.6, 0.5, 0.4), (0.1, 0.0)}}	{{(0.5, 0.4, 0.3, 0.2), (0.5, 0.4, 0.3)}}
Joe	{{(0.9, 0.7, 0.6), (0.1, 0.0)}}	{{(0.7, 0.4, 0.3, 0.1), (0.2, 0.1)}}	{{(0.3, 0.2), (0.5, 0.4, 0.3)}}	{{(0.6, 0.4, 0.3, 0.2, 0.1), (0.4, 0.3)}}	{{(0.6, 0.3, 0.2, 0.1), (0.4, 0.3, 0.2)}}
Ted	{{(0.8, 0.7, 0.5), (0.2, 0.1)}}	{{(0.6, 0.5, 0.4, 0.2), (0.4, 0.3)}}	{{(0.5, 0.3), (0.5, 0.4, 0.3)}}	{{(0.6, 0.4, 0.3, 0.2, 0.1), (0.4, 0.3)}}	{{(0.5, 0.4, 0.2, 0.1), (0.5, 0.4, 0.3)}}

TABLE 3: Values of ρ_{DHFS_i} for each patient to the considered set of possible diagnoses.

	Viral fever	Malaria	Typhoid	Stomach problem	Chest problem
Al	0.9257	0.9620	0.7957	0.8680	0.7110
Bob	0.8380	0.8791	0.8041	0.9922	0.9035
Joe	0.9427	0.9521	0.8757	0.9329	0.7026
Ted	0.9718	0.9472	0.8890	0.8790	0.7644

transfer algorithm to clustering analysis with respect to the problem of complex operation of matrix synthesis when reconstructing analogical relation to equivalence relation clustering under hesitant fuzzy environments. Before doing this, some concepts are introduced firstly.

Definition 12. Let A_j ($j = 1, 2, \dots, m$) be m DHFs; then $C = (\rho_{ij})_{m \times m}$ is called an association matrix, where $\rho_{ij} = \rho(A_i, A_j)$ is the association coefficient of A_i and A_j , which has the following properties:

- (1) $0 \leq \rho_{ij} \leq 1$, for all $i, j = 1, 2, \dots, m$;
- (2) $\rho_{ij} = 1$ if and only if $A_i = A_j$;
- (3) $\rho_{ij} = \rho_{ji}$, for all $i, j = 1, 2, \dots, m$.

Definition 13 (see [23, 30]). Let $C = (\rho_{ij})_{m \times m}$ be an association matrix; if $C^2 = C \circ C = (\bar{\rho}_{ij})_{m \times m}$, then C^2 is called a composition matrix of C , where $\bar{\rho}_{ij} = \max\{\min\{\rho_{ik}, \rho_{kj}\}\}$, for all $i, j = 1, 2, \dots, m$.

Based on Definition 13, we have the following theorem.

Theorem 14 (see [23, 30]). Let $C = (\rho_{ij})_{m \times m}$ be an association matrix; then the composition matrix C^2 is also an association matrix.

Theorem 15 (see [23, 30]). Let $C = (\rho_{ij})_{m \times m}$ be an association matrix; then, for any nonnegative integer k , the composition matrix $C^{2^{k+1}}$ derived from $C^{2^{k+1}} = C^{2^k} \circ C^{2^k}$ is also an association matrix.

Definition 16 (see [23, 30]). Let $C = (\rho_{ij})_{m \times m}$ be an association matrix, if $C^2 \subseteq C$, that is,

$$\max_k \{ \min\{\rho_{ik}, \rho_{kj}\} \} \leq \rho_{ij}, \quad \forall i, j = 1, 2, \dots, m, \quad (25)$$

then C is called an equivalent association matrix.

By the transitivity principle of equivalent matrix, we can easily prove the following theorem.

Theorem 17 (see [23, 30, 32]). Let $C = (\rho_{ij})_{m \times m}$ be an association matrix; then, after the finite times of compositions: $C \rightarrow C^2 \rightarrow C^4 \rightarrow \dots \rightarrow C^{2^k} \rightarrow \dots$, there must exist a positive integer k such that $C^{2^k} = C^{2^{k+1}}$, and C^{2^k} is also an equivalent association matrix.

Definition 18 (see [23, 30, 31]). Let $C = (\rho_{ij})_{m \times m}$ be an equivalent correlation matrix. Then we call $C_\lambda = (\lambda \rho_{ij})_{m \times m}$ the λ -cutting matrix of C , where

$$\lambda \rho_{ij} = \begin{cases} 0 & \text{if } \rho_{ij} < \lambda, \\ 1 & \text{if } \rho_{ij} \geq \lambda, \end{cases} \quad i, j = 1, 2, \dots, m \quad (26)$$

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

TABLE 4: Values of ρ_{DHFS_2} for each patient to the considered set of possible diagnoses.

	Viral fever	Malaria	Typhoid	Stomach problem	Chest problem
Al	0.8718	0.8888	0.7571	0.8286	0.7591
Bob	0.7586	0.8451	0.7960	0.9852	0.8889
Joe	0.9075	0.8607	0.8152	0.8712	0.6500
Ted	0.8764	0.9140	0.8835	0.8678	0.7550

TABLE 5: Values of ρ_{DHFS_3} for each patient to the considered set of possible diagnoses.

	Viral fever	Malaria	Typhoid	Stomach problem	Chest problem
Al	0.8085	0.7739	0.7900	0.7515	0.8026
Bob	0.7480	0.7714	0.6824	0.7547	0.8006
Joe	0.7925	0.7887	0.7548	0.6878	0.8100
Ted	0.8063	0.7244	0.7516	0.7386	0.8230

Next, a traditional transfer closure algorithm is given as follows.

Step 1. Let $\{A_1, A_2, \dots, A_m\}$ be a set of DHFSs in $X = \{x_1, x_2, \dots, x_n\}$. We can calculate the correlation coefficients of the DHFSs and then construct a correlation matrix $C = (\rho_{ij})_{m \times m}$, where $\rho_{ij} = \rho(A_i, A_j)$.

Step 2. Check whether $C = (\rho_{ij})_{m \times m}$ is an equivalent correlation matrix; that is, check whether it satisfies $C^2 \subseteq C$, where

$$C^2 = C \circ C = (\bar{\rho}_{ij})_{m \times m}, \quad \bar{\rho}_{ij} = \max_k \left\{ \min \{ \rho_{ik}, \rho_{kj} \} \right\}, \quad (27)$$

$$i, j = 1, 2, \dots, m.$$

If it does not hold, we construct the equivalent correlation matrix $C^{2^k} : C \rightarrow C^2 \rightarrow C^4 \rightarrow \dots \rightarrow C^{2^k} \rightarrow \dots$, until $C^{2^k} = C^{2^{k+1}}$.

Step 3. For a confidence level λ , we construct a λ -cutting matrix $C_\lambda = (\lambda \rho_{ij})_{m \times m}$ through Definition 18 in order to classify the DHFSs A_j ($j = 1, 2, \dots, m$). If all elements of the i th line (column) are the same as the corresponding elements of the j th line (column) in C_λ , then the DHFSs A_i and A_j are of the same type. By means of this principle, we can classify all these m A_j ($j = 1, 2, \dots, m$).

By analyzing the aforementioned transfer closure algorithm, this algorithm has one drawback such as complex operation of matrix synthesis when reconstructing the equivalent correlation matrix. In this paper, we have the following theorem of the correlation coefficients in dual hesitant fuzzy environment.

Theorem 19. For all $x, y \in A$, for the confidence level λ , if $\exists x_1, x_2, x_3, \dots, x_p$, when $\rho(x, x_1) \geq \lambda$, $\rho(x_1, x_2) \geq \lambda$, $\rho(x_2, x_3) \geq \lambda, \dots, \rho(x_p, y) \geq \lambda$, then x, x_1, x_2, \dots, x_p and y are of the same type.

Proof. we are motivated by the generalized idea based on the transitivity principle of ordinary equivalent relation R : for all $x, y \in A$ (here, A is an ordinary set, not a fuzzy set), $\exists x_1, x_2, x_3, \dots, x_p$, when $(x, x_1) \in R$, $(x_1, x_2) \in R, \dots, (x_k, x_{k+1}) \in R, \dots, (x_p, y) \in R$, we can have $(x, y) \in R$.

And from Definition 18, we can see that the λ -cutting matrix of C is an ordinary correlation matrix, which completes the proof of Theorem 19. \square

From the above theoretical analysis, we propose a direct transfer algorithm for clustering DHFSs as follows.

Step 1. Let $A = \{A_1, A_2, \dots, A_m\}$ be a set of DHFSs in $X = \{x_1, x_2, \dots, x_n\}$. We can calculate the correlation coefficients of the DHFSs and then construct a correlation matrix $C = (\rho_{ij})_{m \times m}$, where $\rho_{ij} = \rho(A_i, A_j)$.

Step 2. By setting the threshold to the confidence level λ , we can construct a λ -cutting matrix $C_\lambda = (\lambda \rho_{ij})_{m \times m}$. If $\rho_{ij} = 1$, this means that the DHFSs A_i and A_j are of the same type. By means of this principle, we can classify all these m A_j ($j = 1, 2, \dots, m$).

We can see that the transfer closure algorithm must construct the equivalent correlation matrix $C \rightarrow C^2 \rightarrow C^4 \rightarrow \dots \rightarrow C^{2^k} \rightarrow \dots$, until $C^{2^k} = C^{2^{k+1}}$ and then construct a λ -cutting matrix $C_\lambda = (\lambda \rho_{ij})_{m \times m}$ through Definition 18 in order to classify the DHFSs A_j ($j = 1, 2, \dots, m$). Simply, the transfer algorithm only constructs a λ -cutting matrix $C_\lambda = (\lambda \rho_{ij})_{m \times m}$ by setting the threshold to the confidence level λ and then classifies the DHFSs A_j ($j = 1, 2, \dots, m$) directly. In what follows, we will talk about the relationship between the transfer closure algorithm and the direct transfer algorithm.

Theorem 20. The clustering results are the same by the transfer closure algorithm and the direct transfer algorithm, at the same confidence level.

Proof. (1) For a confidence level λ , for all $A_i, A_j \in A$, $\exists x_1, x_2, x_3, \dots, x_p$, if $\rho(A_i, x_1) \geq \lambda$, $\rho(x_1, x_2) \geq \lambda$, $\rho(x_2, x_3) \geq \lambda, \dots, \rho(x_p, A_j) \geq \lambda$, then A_i and A_j are of the same type by the direct transfer algorithm.

Assume that we construct the equivalent correlation matrix C^{2^k} when we employ the transfer closure algorithm. We must prove that $\rho_{C^{2^k}}(A_i, A_j) \geq \lambda$. Consider

$$\begin{aligned} & \rho_{C^2}(A_i, A_j) \\ &= \max_{A_q \in A} \left\{ \min \left\{ \rho_C(A_i, A_q), \rho_C(A_q, A_j) \right\} \right\}, \\ & \quad i, j, q = 1, 2, \dots, m \\ &= \max_{A_q \in A, A_i \neq A_j} \left\{ \max \left\{ \min \left\{ \rho_C(A_i, A_q), \rho_C(A_q, A_j) \right\} \right\}, \right. \\ & \quad \left. \rho_C(A_i, A_j) \right\} \\ &\geq \rho_C(A_i, A_j). \end{aligned} \quad (28)$$

So $C^2 \supseteq C$, and, for the same reason, we have $C^{2^k} \supseteq C^{2^{k-1}} \supseteq C^{2^{k-2}}, \dots, \supseteq C^2 \supseteq C$. Consider

$$\begin{aligned} & \rho_{C^{2^k}}(A_i, A_j) \\ &= \max_{A_q \in A} \left\{ \min \left\{ \rho_{C^{2^k}}(A_i, A_q), \rho_{C^{2^k}}(A_q, A_j) \right\} \right\}, \\ & \quad i, j, q = 1, 2, \dots, m \\ &= \max_{A_q \in A, A_q \neq A_p} \left\{ \max \left\{ \min \left\{ \rho_{C^{2^k}}(A_i, A_p), \rho_{C^{2^k}}(A_p, A_j) \right\}, \right. \right. \\ & \quad \left. \left. \min \left\{ \rho_{C^{2^k}}(A_i, A_{x_p}), \right. \right. \right. \\ & \quad \left. \left. \left. \rho_{C^{2^k}}(A_{x_p}, A_j) \right\} \right\} \right\} \\ &\geq \min \left\{ \rho_{C^{2^k}}(A_i, A_{x_p}), \rho_{C^{2^k}}(A_{x_p}, A_j) \right\} \\ &\geq \min \left\{ \rho_{C^{2^k}}(A_i, A_{x_{p-1}}), \right. \\ & \quad \left. \rho_{C^{2^k}}(A_{x_{p-1}}, A_{x_p}), \rho_{C^{2^k}}(A_{x_p}, A_j) \right\} \\ &\geq \min \left\{ \rho_{C^{2^k}}(A_i, A_{x_1}), \rho_{C^{2^k}}(A_{x_1}, A_{x_2}) \right. \\ & \quad \left. , \dots, \rho_{C^{2^k}}(A_i, A_{x_{p-1}}), \rho_{C^{2^k}}(A_{x_{p-1}}, A_{x_p}), \right. \\ & \quad \left. \rho_{C^{2^k}}(A_{x_p}, A_j) \right\} \\ &\geq \min \left\{ \rho_C(A_i, A_{x_1}), \rho_C(A_{x_1}, A_{x_2}) \right. \\ & \quad \left. , \dots, \rho_C(A_i, A_{x_{p-1}}), \right. \\ & \quad \left. \rho_C(A_{x_{p-1}}, A_{x_p}), \rho_C(A_{x_p}, A_j) \right\} \\ &\geq \lambda. \end{aligned} \quad (29)$$

For a confidence level λ , when we get that A_i and A_j are of the same type using the direct transfer algorithm, we can also have the same clustering results by the transfer closure algorithm.

(2) For a confidence level λ , for all $A_i, A_j \in A$, \exists the equivalent correlation matrix C^{2^k} , $\rho_{C^{2^k}}(A_i, A_j) \geq \lambda$, then, A_i and A_j are of the same type by the transfer closure algorithm.

Let

$$\begin{aligned} & \rho_{C^{2^k}}(A_i, A_j) \\ &= \max_{A_q \in A} \left\{ \min \left\{ \rho_{C^{2^{k-1}}}(A_i, A_q), \rho_{C^{2^{k-1}}}(A_q, A_j) \right\} \right\}. \end{aligned} \quad (30)$$

Then $\exists x_1$, $\rho_{C^{2^k}}(A_i, A_j) = \min \left\{ \rho_{C^{2^{k-1}}}(A_i, A_{x_1}), \rho_{C^{2^{k-1}}}(A_{x_1}, A_j) \right\} \geq \lambda$, $\rho_{C^{2^{k-1}}}(A_i, A_{x_1}) \geq \lambda$, $\rho_{C^{2^{k-1}}}(A_{x_1}, A_j) \geq \lambda$.

So A_i and A_j are the same type in $C^{2^{k-1}}$ by the direct transfer algorithm.

For the same reason, $\exists x_2$, $\rho_{C^{2^{k-2}}}(A_i, A_{x_2}) \geq \lambda$, $\rho_{C^{2^{k-2}}}(A_{x_2}, A_{x_1}) \geq \lambda$. $\exists x_3$, $\rho_{C^{2^{k-2}}}(A_{x_1}, A_{x_3}) \geq \lambda$, $\rho_{C^{2^{k-2}}}(A_{x_3}, A_j) \geq \lambda$.

A_i and A_j are the same type in $C^{2^{k-2}}$ by the direct transfer algorithm.

So, $\exists x_1, x_2, x_3, \dots, x_{2^k}$, $\rho_C(A_i, A_{x_1}) \geq \lambda$, $\rho_C(A_{x_1}, A_{x_2}) \geq \lambda$, $\rho_C(A_{x_2}, A_{x_3}) \geq \lambda, \dots, \rho_C(A_{x_{2^k}}, A_j) \geq \lambda$.

A_i and A_j are the same type in C by the direct transfer algorithm.

For a confidence level λ , when we get A_i and A_j are of the same type using the transfer closure algorithm, we can also have the same clustering results by the direct transfer algorithm, which completes the proof. \square

We assume $A = \{A_1, A_2, \dots, A_m\}$ to be a set of DHFSs, and we construct the equivalent correlation matrix $C^{2^k} : C \rightarrow C^2 \rightarrow C^4 \rightarrow \dots \rightarrow C^{2^k} \rightarrow \dots$, until $C^{2^k} = C^{2^{k+1}}$ and then construct a λ -cutting matrix $C_\lambda = (\lambda \rho_{ij})_{m \times m}$ for the transfer closure algorithm. Consequently, the running time of the transfer closure algorithm is $T_{\text{tca}} = O(km^3 + km^2)$; by the same arguments, the direct transfer algorithm requires $T_{\text{dta}} = O(m^2)$ time on the same example. And we have established $S_{\text{tca}} = O(m^2)$ space bound at least for the step of constructing the equivalent correlation matrix based on the transfer closure algorithm, while, for the transfer algorithm, it constructs a λ -cutting matrix $C_\lambda = (\lambda \rho_{ij})_{m \times m}$ by setting the threshold to the confidence level λ and needs $S_{\text{tca}} = O(m)$ space bound. We can see that the computational complexity of both two algorithms ranges depends on the number of m , and the direct transfer algorithm exhibits better behavior.

Below, we conduct experiments in order to demonstrate the effectiveness of the proposed clustering algorithm for DHFSs.

Example 21. Every diamond is a miracle of time and place and chance. Like snowflakes, no two are exactly alike. Every consumer shopping for diamonds is faced with endless diamond combinations. In addition to different diamond combinations, prices are also influenced by market supply and demand conditions, fashion trends, and so forth. While

TABLE 6: Diamond data set.

	“D” color	“FL” clarity	“3 excellent” cut	“1 carat” weight
A_1	{{(0.5, 0.4, 0.3); (0.4, 0.2)}}	{{(0.6, 0.5); (0.3, 0.2, 0.1)}}	{{(0.6, 0.4, 0.3); (0.4, 0.2, 0.1)}}	{{(0.6); (0.4)}}
A_2	{{(0.8, 0.7, 0.6); (0.2, 0.1)}}	{{(0.7, 0.6); (0.3, 0.2, 0.1)}}	{{(0.7, 0.6, 0.5); (0.3, 0.2, 0.1)}}	{{(0.7); (0.2)}}
A_3	{{(0.9, 0.8, 0.7); (0.1, 0.0)}}	{{(0.8, 0.7); (0.2, 0.1, 0.0)}}	{{(0.8, 0.7, 0.6); (0.2, 0.1, 0.0)}}	{{(0.9); (0.1)}}
A_4	{{(0.4, 0.3, 0.1); (0.6, 0.5)}}	{{(0.6, 0.5); (0.4, 0.2, 0.1)}}	{{(0.6, 0.5, 0.4); (0.3, 0.2, 0.1)}}	{{(0.3); (0.6)}}
A_5	{{(0.6, 0.5, 0.4); (0.3, 0.2)}}	{{(0.3, 0.2); (0.6, 0.5, 0.4)}}	{{(0.3, 0.2, 0.1); (0.6, 0.5, 0.4)}}	{{(0.1); (0.8)}}
A_6	{{(0.6, 0.5, 0.4); (0.4, 0.2)}}	{{(0.7, 0.6); (0.3, 0.2, 0.1)}}	{{(0.2, 0.1, 0.0); (0.7, 0.2, 0.1)}}	{{(0.8); (0.1)}}
A_7	{{(0.8, 0.6, 0.5); (0.2, 0.1)}}	{{(0.6, 0.5); (0.3, 0.2, 0.1)}}	{{(0.4, 0.3, 0.2); (0.5, 0.4, 0.3)}}	{{(0.5); (0.4)}}
A_8	{{(0.7, 0.6, 0.5); (0.2, 0.0)}}	{{(0.4, 0.3); (0.6, 0.5, 0.4)}}	{{(0.6, 0.5, 0.4); (0.4, 0.3, 0.2)}}	{{(0.8); (0.2)}}
A_9	{{(0.4, 0.3, 0.2); (0.6, 0.5)}}	{{(0.4, 0.3); (0.6, 0.5, 0.4)}}	{{(0.2, 0.1, 0.0); (0.8, 0.6, 0.5)}}	{{(0.2); (0.6)}}
A_{10}	{{(0.2, 0.1, 0.0); (0.7, 0.6)}}	{{(0.8, 0.6); (0.2, 0.1, 0.0)}}	{{(0.6, 0.5, 0.3); (0.4, 0.2, 0.1)}}	{{(0.7); (0.3)}}

consumers’ tastes and budgets change, most seek to find a fair price for the diamond they choose. Until the middle of the twentieth century, there was no agreed upon standard by which diamonds could be judged. No matter how beautiful a diamond may look you simply cannot see its true quality. GIA created the first and now globally accepted standard for describing diamonds: color, clarity, cut, and carat weight. Concerning color, the less color in the stone there is, the more desirable and valuable it is. Grades run from “D” to “X.” Clarity measures the amount, size, and placement of internal “inclusions”, and external “blemishes.” Grades run from “Flawless” to “included.” Cut does not refer to a diamond’s shape but to the proportion and arrangement of its facets and the quality of workmanship. Grades range from “excellent” to “poor.” Carat refers to a diamond’s weight. Generally speaking, the higher the carat weight, the more

expensive the stone. Two diamonds of equal carat weight, however, can have very different quality and price when the other three Cs are considered. We choose a “perfect” diamond whose 4C is “D” color, “FL” clarity, “3 excellent” cut, and “1carat” weight. For the convenience of analysis, the weight vector of these attributes is $w = (0.25, 0.25, 0.25, 0.25)$. Here, there are ten diamonds. In order to better make the assessment, several evaluation organizations are requested. The normalized evaluation diamond data, represented by DHFSs, are displayed in Table 6.

Now we utilize the direct transfer algorithm to cluster the ten diamonds, which involves the following steps.

Step 1. Utilize (21) to calculate the association coefficients, and then construct an association matrix:

$$C = \begin{pmatrix} 1.0000 & 0.9495 & 0.9010 & 0.9227 & 0.7571 & 0.8270 & 0.9542 & 0.9123 & 0.7778 & 0.9241 \\ 0.9495 & 1.0000 & 0.9853 & 0.8053 & 0.6436 & 0.8948 & 0.9457 & 0.9404 & 0.6226 & 0.8260 \\ 0.9010 & 0.9853 & 1.0000 & 0.7164 & 0.5146 & 0.8697 & 0.8904 & 0.9076 & 0.4867 & 0.7811 \\ 0.9227 & 0.8053 & 0.7164 & 1.0000 & 0.8174 & 0.7353 & 0.8490 & 0.7463 & 0.8484 & 0.9025 \\ 0.7571 & 0.6436 & 0.5146 & 0.8174 & 1.0000 & 0.6003 & 0.8240 & 0.6997 & 0.9316 & 0.5822 \\ 0.8270 & 0.8948 & 0.8697 & 0.7353 & 0.6003 & 1.0000 & 0.9048 & 0.8750 & 0.6948 & 0.8540 \\ 0.9542 & 0.9457 & 0.8904 & 0.8490 & 0.8240 & 0.9048 & 1.0000 & 0.9105 & 0.7993 & 0.8018 \\ 0.9123 & 0.9404 & 0.9076 & 0.7463 & 0.6997 & 0.8750 & 0.9105 & 1.0000 & 0.6826 & 0.7612 \\ 0.7778 & 0.6226 & 0.4867 & 0.8484 & 0.9316 & 0.6948 & 0.7993 & 0.6826 & 1.0000 & 0.7206 \\ 0.9241 & 0.8260 & 0.7811 & 0.9025 & 0.5822 & 0.8540 & 0.8018 & 0.7612 & 0.7206 & 1.0000 \end{pmatrix} \tag{31}$$

Step 2. We give a detailed sensitivity analysis with respect to the confidence level, and, by (26), we get all the possible classifications of the ten diamonds; see Table 7 and Figure 1.

From the above numerical analysis, under the group setting, the experts’ evaluation information usually does not reach an agreement for the objects that need to be classified. Example 21 clearly shows that the clustering algorithm based on DHFSs provides a proper way to resolve this issue.

In the following, a comparison is made among the method proposed in this paper, Chen et al.’s method [23], and Zhao et al.’s method [31] in Table 8.

Through Table 8, it is worthy of pointing out that the clustering results of the direct transfer clustering method proposed in this paper are exactly the same with those of Chen et al.’s transfer clustering method and Zhao et al.’s Boole method, but our method does not need to use the transitive closure technique to calculate the equivalent matrix of the association matrix and thus requires much less computational effort than Chen et al.’s method. The computational complexity of Chen et al.’s method and Zhao et al.’s method has, relatively, high computational complexity, which indeed motivates our clustering method proposed in this paper. Furthermore, from Example 21 we can see that

TABLE 7: The clustering result of 10 diamonds.

Class	Confidence level	Dual hesitant fuzzy clustering algorithm
10	$0.9853 < \lambda \leq 1$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
9	$0.9542 < \lambda \leq 0.9853$	$\{A_2, A_3\}, \{A_1\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
8	$0.9495 < \lambda \leq 0.9542$	$\{A_2, A_3\}, \{A_1, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
7	$0.9404 < \lambda \leq 0.9495$	$\{A_1, A_2, A_3, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
6	$0.9316 < \lambda \leq 0.9404$	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_9\}, \{A_{10}\}$
5	$0.9241 < \lambda \leq 0.9316$	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}, \{A_{10}\}$
4	$0.9227 < \lambda \leq 0.9241$	$\{A_1, A_2, A_3, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}$
3	$0.9123 < \lambda \leq 0.9227$	$\{A_1, A_2, A_3, A_4, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_6\}$
2	$0.8484 < \lambda \leq 0.9123$	$\{A_1, A_2, A_3, A_4, A_6, A_7, A_8, A_{10}\}, \{A_5, A_9\}$
1	$0 < \lambda \leq 0.8484$	$\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$

TABLE 8: Comparisons of the derived results.

Classes	The results derived by the direct transfer algorithm method	The results derived by Chen et al.'s transfer algorithm method	The results derived by Zhao et al.'s Boole method
10	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_1\}, \{A_2\}, \{A_3\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
9	$\{A_2, A_3\}, \{A_1\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_2, A_3\}, \{A_1\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_2, A_3\}, \{A_1\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_7\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
8	$\{A_2, A_3\}, \{A_1, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_2, A_3\}, \{A_1, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_2, A_3\}, \{A_1, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
7	$\{A_1, A_2, A_3, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_1, A_2, A_3, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$	$\{A_1, A_2, A_3, A_7\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_8\}, \{A_9\}, \{A_{10}\}$
6	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_9\}, \{A_{10}\}$	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_9\}, \{A_{10}\}$	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_4\}, \{A_5\}, \{A_6\}, \{A_9\}, \{A_{10}\}$
5	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}, \{A_{10}\}$	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}, \{A_{10}\}$	$\{A_1, A_2, A_3, A_7, A_8\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}, \{A_{10}\}$
4	$\{A_1, A_2, A_3, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}$	$\{A_1, A_2, A_3, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}$	$\{A_1, A_2, A_3, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_4\}, \{A_6\}$
3	$\{A_1, A_2, A_3, A_4, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_6\}$	$\{A_1, A_2, A_3, A_4, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_6\}$	$\{A_1, A_2, A_3, A_4, A_7, A_8, A_{10}\}, \{A_5, A_9\}, \{A_6\}$
2	$\{A_1, A_2, A_3, A_4, A_6, A_7, A_8, A_{10}\}, \{A_5, A_9\}$	$\{A_1, A_2, A_3, A_4, A_6, A_7, A_8, A_{10}\}, \{A_5, A_9\}$	$\{A_1, A_2, A_3, A_4, A_6, A_7, A_8, A_{10}\}, \{A_5, A_9\}$
1	$\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$	$\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$	$\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}\}$

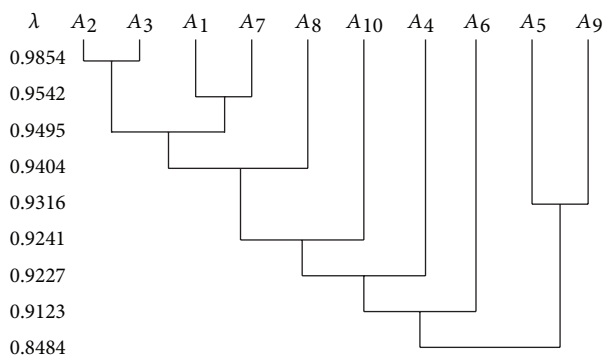


FIGURE 1: The clustering result of 10 diamonds.

the clustering results have much to do with the threshold; the smaller the confidence level is, the more detailed the clustering will be.

5. Conclusions

Dual hesitant fuzzy set, as an extension of fuzzy set, can describe the situation that people have hesitancy when they make a decision more objectively than other extensions of fuzzy set (interval-valued fuzzy set, intuitionistic fuzzy set, type-2 fuzzy set, and fuzzy multiset). In this paper, the correlation coefficients for DHFSs have been studied. Their properties have been discussed, and the differences and correlations among them have been investigated in detail. We have made the clustering analysis under dual hesitant fuzzy environments with one typical real world example. To further extend the application range of the present clustering algorithm, in particular for the case that needs to assign weights for different experts, it will be necessary to generalize the original definition of DHFSs.

Given that DHFSs are a suitable technique of denoting uncertain information that is widely encountered in daily life and the latent applications of our algorithm in the field of data

mining, information retrieval and pattern recognition, and so forth, may be the directions for future research.

Acknowledgments

The authors are very grateful to the anonymous reviewers for their insightful and constructive comments and suggestions that have led to an improved version of this paper. This work is supported by the National Nature Science Foundation of China (no. 70971136).

References

- [1] Z. Xu, "Choquet integrals of weighted intuitionistic fuzzy information," *Information Sciences*, vol. 180, no. 5, pp. 726–736, 2010.
- [2] P. Bonizzoni, G. D. Vedova, R. Dondi, and T. Jiang, "Correlation clustering and consensus clustering," in *Algorithms and Computation*, vol. 3827 of *Lecture Notes in Computer Science*, pp. 226–235, 2005.
- [3] H. P. Kriegel, P. Kröger, E. Schubert, and A. Zimek, "A general framework for increasing the robustness of PCA-based correlation clustering algorithms," in *Scientific and Statistical Database Management*, vol. 5069 of *Lecture Notes in Computer Science*, pp. 418–435, 2008.
- [4] D. G. Park, Y. C. Kwun, J. H. Park, and I. Y. Park, "Correlation coefficient of interval-valued intuitionistic fuzzy sets and its application to multiple attribute group decision making problems," *Mathematical and Computer Modelling*, vol. 50, no. 9–10, pp. 1279–1293, 2009.
- [5] E. Szmids and J. Kacprzyk, "Correlation of intuitionistic fuzzy sets," in *Computational Intelligence for Knowledge-Based Systems Design*, vol. 6178 of *Lecture Notes in Computer Science*, pp. 169–177, 2010.
- [6] G. W. Wei, H. J. Wang, and R. Lin, "Application of correlation coefficient to interval-valued intuitionistic fuzzy multiple attribute decision-making with incomplete weight information," *Knowledge and Information Systems*, vol. 26, no. 2, pp. 337–349, 2011.
- [7] J. Ye, "Multicriteria fuzzy decision-making method using entropy weights-based correlation coefficients of interval-valued intuitionistic fuzzy sets," *Applied Mathematical Modelling*, vol. 34, no. 12, pp. 3864–3870, 2010.
- [8] L. A. Zadeh, "Fuzzy sets," *Information and Computation*, vol. 8, pp. 338–353, 1965.
- [9] W. L. Hung and J. W. Wu, "A note on the correlation of fuzzy numbers by expected interval," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 4, pp. 517–523, 2001.
- [10] D. H. Hong, "Fuzzy measures for a correlation coefficient of fuzzy numbers under T_W (the weakest t-norm)-based fuzzy arithmetic operations," *Information Sciences*, vol. 176, no. 2, pp. 150–160, 2006.
- [11] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.
- [12] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 31, no. 3, pp. 343–349, 1989.
- [13] T. Gerstenkorn and J. Mańko, "Correlation of intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 44, no. 1, pp. 39–43, 1991.
- [14] H. Bustince and P. Burillo, "Correlation of interval-valued intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 74, no. 2, pp. 237–244, 1995.
- [15] W. L. Hung, "Using statistical viewpoint in developing correlation of intuitionistic fuzzy sets," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, no. 4, pp. 509–516, 2001.
- [16] H. B. Mitchell, "A correlation coefficient for intuitionistic fuzzy sets," *International Journal of Intelligent Systems*, vol. 19, no. 5, pp. 483–490, 2004.
- [17] W. L. Hung and J. W. Wu, "Correlation of intuitionistic fuzzy sets by centroid method," *Information Sciences*, vol. 144, no. 1–4, pp. 219–225, 2002.
- [18] Z. Xu, "On correlation measures of intuitionistic fuzzy sets," *Intelligent Data Engineering and Automated Learning*, vol. 4224, pp. 16–24, 2006.
- [19] P. J. Robinson and E. C. H. Amirtharaj, "Vague correlation coefficient of interval vague sets," *International Journal of Fuzzy System Applications*, vol. 2, no. 1, pp. 18–34, 2012.
- [20] P. J. Robinson and E. C. H. Amirtharaj, "A short primer on the correlation coefficient of vague sets," *International Journal of Fuzzy System Applications*, vol. 1, no. 2, pp. 55–69, 2011.
- [21] V. Torra, "Hesitant fuzzy sets," *International Journal of Intelligent Systems*, vol. 25, no. 6, pp. 529–539, 2010.
- [22] V. Torra and Y. Narukawa, "On hesitant fuzzy sets and decision," in *Proceedings of the 18th IEEE International Conference on Fuzzy Systems*, pp. 1378–1382, Jeju Island, Korea, August 2009.
- [23] N. Chen, Z. Xu, and M. Xia, "Correlation coefficients of hesitant fuzzy sets and their applications to clustering analysis," *Applied Mathematical Modelling*, vol. 37, no. 4, pp. 2197–2211, 2013.
- [24] Z. Xu and M. Xia, "On distance and correlation measures of hesitant fuzzy information," *International Journal of Intelligent Systems*, vol. 26, no. 5, pp. 410–425, 2011.
- [25] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, vol. 144, Academic Press, New York, NY, USA, 1980.
- [26] B. Zhu, Z. Xu, and M. Xia, "Dual hesitant fuzzy sets," *Journal of Applied Mathematics*, vol. 2012, Article ID 879629, 13 pages, 2012.
- [27] Z. Xu and M. Xia, "Distance and similarity measures for hesitant fuzzy sets," *Information Sciences*, vol. 181, no. 11, pp. 2128–2138, 2011.
- [28] E. Szmids and J. Kacprzyk, "A similarity measure for intuitionistic fuzzy sets and its application in supporting medical diagnostic reasoning," in *Artificial Intelligence and Soft Computing*, vol. 3070 of *Lecture Notes in Computer Science*, pp. 388–393, 2004.
- [29] I. K. Vlachos and G. D. Sergiadis, "Intuitionistic fuzzy information—applications to pattern recognition," *Pattern Recognition Letters*, vol. 28, no. 2, pp. 197–206, 2007.
- [30] Z. Xu, J. Chen, and J. Wu, "Clustering algorithm for intuitionistic fuzzy sets," *Information Sciences*, vol. 178, no. 19, pp. 3775–3790, 2008.
- [31] H. Zhao, Z. Xu, and Z. Wang, "Intuitionistic fuzzy clustering algorithm based on boolean matrix and association measure," *International Journal of Information Technology & Decision Making*, vol. 12, no. 1, pp. 95–118, 2013.
- [32] P. Z. Wang, *Fuzzy Set Theory and Applications*, Shanghai Scientific and Technical Publishers, Shanghai, China, 1983.