

New Distance Measures on Dual Hesitant Fuzzy Sets and Their Application in Pattern Recognition

Li Xin^a, Zhang Xiaohong^{*b}

College of Arts and Sciences, Shanghai Maritime University, Shanghai 201306, China
^a14749275@qq.com, ^bzxhonghz@263.net,

Keywords: Dual hesitant fuzzy sets, Distance Measure, Hesitance degree, Pattern recognition.

Abstract: The concept of dual hesitant fuzzy sets (DHFSs), which was first introduced as a new extension of fuzzy sets and hesitant fuzzy sets in 2012, is a useful tool to deal with the vagueness and ambiguity in many practical problems under hesitant fuzzy environment. Normally, we use the definition of distance to describe the relationship of two DHFSs. However, considering that the existing distance measures of DHFSs still have some major shortcomings, so in this paper, we firstly introduce a new concept –hesitance degree of each dual hesitant fuzzy element (DHFE) to these existing distance measures and then develop several novel distance measures in which both the values and the numbers of values of DHFE are taken into account. The properties of these new distance measures are discussed. Finally, we apply our proposed distance measures of DHFSs in pattern recognition making to illustrate their validity and applicability.

1. Introduction

The concept of fuzzy set was introduced by Zadeh in 1965 as an extension of the classical notion of sets (see [1, 2]). In classical set theory, the membership of elements in a set is assessed in binary terms according to a bivalent condition –an element either belongs to or does not belong to the set, by contract, fuzzy sets theory permits the gradual assessment of the membership of elements in a set, this is described with the aid of a membership function valued in the real unit interval $[0, 1]$. Since its original definition, several extensions have been proposed for fuzzy sets, among them, we can underline, for their relevance in this paper, hesitant fuzzy sets (HFSs) and dual hesitance fuzzy sets (DHFSs). The HFSs define the membership of an element and the membership degree may be a set of possible values rather than inter-values $[0, 1]$. On the basis of HFSs, Zhu and Xu [3] introduced the definition of DHFSs which uses the membership hesitancy function and the non-membership hesitancy function to support a more exemplary and flexible access to assign values for each element in the domain, it is a very useful tool to deal with vagueness and ambiguity in the pattern recognition problems under hesitant fuzzy environment.

Considering how to describe the relationship of two given fuzzy sets, researchers proposed the concept of distance measures which are used for estimating the degree of distance between two fuzzy sets. The most widely used distance measures are the Hamming distance, Euclidean distance and Hausdorff metric. Based on these researchers, Su and Xu have made some significant extensions for

these measures to deal with DHFSs. Now, these distance measures have been extensively applied in some fields such as decision making, pattern recognition, machine learning and market prediction and so on.

However, in the process of practical application, we found that these existing distance measures of DHFSs still have some notable short comings, the most obvious is that they can only cover the divergence of the values but fail to consider the numbers of value of two given DHFSs. However, the main characteristic of DHFSs is that they can describe the hesitant situations flexibly, such a hesitation is depicted by a number of values of DHFSs being greater than just one. Hence, it's very necessary to take into account both the difference of the values and that of the numbers when we study the difference between the DHFSs. In this paper, we referred to [4, 5], we propose some new distance measures for DHFSs in this paper by taking into account the hesitance of the hesitant fuzzy sets and investigate their application in practical pattern recognition.

Now, we firstly review the definition of DHFSs and the properties of their distance, and then list several frequently-used distance measures of DHFSs.

Definition 1.1. Let X be a fixed set, then a dual hesitant fuzzy set (DHFSs) H on X is described as:

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

in which $h(x)$ and $g(x)$ are two sets of value in $[0, 1]$ denoting the possible membership degrees and non-membership degrees of the element $x \in H$ to the set, respectively.

Definition 1.2. Let A and B be two DHFSs on $X = \{x_1, x_2, \dots, x_n\}$, then the distance between A and B denoted as $d(A, B)$, which satisfy the following properties:

- (1) $0 \leq d(A, B) \leq 1$;
- (2) $d(A, B) = 0$ only if $A = B$;
- (3) $d(A, B) = d(B, A)$.

Definition 1.3. Let elements in $d_E(x) = (h_E(x), g_E(x))$ in decreasing order, and let $\gamma_E^{\delta(i)}$ be the i th largest value in $h_E(x)$ and $\eta_E^{\delta(j)}$ be the j th largest value in $g_E(x)$. Let $l_h(d_E(x_i))$ and $l_g(d_E(x_i))$ be the number of values in $h_E(x_i)$ and $g_E(x_i)$, respectively. But in most case, $l_h(d_E(x_i)) \neq l_g(d_E(x_i))$. To operate correctly, we should extend the shorter one making both of them have the length by adding different values.

On the basis of aboving definitions, we can refer several existing distance measures for DHFSs now in [3].

Definition 1.4. we define a dual hesitant normalized Hamming distances at first :

$$d_1(A, B) = \sum_{i=1}^n \left\{ \frac{1}{n} \left\{ \frac{1}{l_{x_i}} \left\{ \sum_{j=1}^{\#h_{x_i}} |\gamma_A^{\delta(j)}(x_i) - \gamma_B^{\delta(j)}(x_i)|^\lambda + \sum_{k=1}^{\#g_{x_i}} |\eta_A^{\delta(k)}(x_i) - \eta_B^{\delta(k)}(x_i)|^\lambda \right\} \right\} \right\}^{\frac{1}{\lambda}} \quad (2)$$

in which $\lambda > 0$. $l_{x_i} = (\#h_{x_i}) + (\#g_{x_i})$ and $\#h$ and $\#g$ are the numbers of the elements in h and g respectively.

Definition 1.5. If we take the weight of each element into account, the following weighted distance measures for DHFSs can be attained ($\omega_i \in [0, 1]$, $\sum_{i=1}^n \omega_i = 1$ and $\lambda > 0$):

$$d_2(A, B) = \sum_{i=1}^n \left\{ \omega_i \left\{ \frac{1}{l_{x_i}} \left\{ \sum_{j=1}^{\#h_{x_i}} |\gamma_A^{\delta(j)}(x_i) - \gamma_B^{\delta(j)}(x_i)|^\lambda + \sum_{k=1}^{\#g_{x_i}} |\eta_A^{\delta(k)}(x_i) - \eta_B^{\delta(k)}(x_i)|^\lambda \right\} \right\} \right\}^{\frac{1}{\lambda}} \quad (3)$$

2. New distance measures for DHFSs with their application to pattern recognition

In this section, we will propose a simple but convincing example to reveal the disadvantage of the above-mentioned distance measures at first. And then by introducing the definition of hesitance degree, we can get our modified distance measures. At last, we utilize the proposed distance measures to a practical pattern recognition example to prove their validity and superiority.

Example 2.1. Let $X=\{x_i\}$, assume that there exist two patterns which are presented by DHFSs A and B, $A=\{\{0.69, 0.75\}, \{0.37, 0.76\}\}, \{\{0.21, 0.53, 0.74\}, \{0.96\}\}, \{\{0.22, 0.31, 0.60\}, \{0.61\}\}, \{\{0.31, 0.67\}, \{0.44, 0.69\}\}, B=\{\{0.34, 0.35\}\}$. Now, there is a sample to be recognized which is represented by another DHFS $H= \{\{0.34, 0.53\}, \{0.21, 0.54\}\}, \{\{0.16, 0.31, 0.52\}, \{0.53\}\}, \{\{0.09, 0.15, 0.39\}, \{0.42\}\}, \{\{0.23, 0.49\}, \{0.29, 0.51\}\}$.

Firstly, we should analyze this question by our above-mentioned knowledge, it is obviously that the difference of the membership values between A and H as well as that between B and H are almost the same, but the structure of H and that of A is almost totally uniform, which is quite different from that of B. what else, the number of values of H is the same as that of A, but different from that of B to a great extent. As we stressed before, a hesitation is depicted by a number of values of DHFSs being greater than just one, so the number of values is equally important. So through discussion, we think it is easy to understand that H should belong to the pattern A.

However, by applying the existing distance measure equation (2), we can obtain $d(A, H)=0.0584$ and $d(B, H)=0.0366$, so we get the result that H should belong to the pattern B, it is obviously contrast our analysis. The error is because the existing distance measures can only cover the divergence of the values but fail to consider the numbers of value of two given DHFSs. As we all known one small false step will make a great difference, so it's very necessary to take into account both the difference of the values and that of the numbers when we study the difference between the DHFSs.

In the following, we propose some new distance measures between DHFSs by taking into account the hesitance extent of each DHFE, which can overcome the above-mentioned shortcoming. Before that, we first introduce a new concept as follows:

Definition 2.1. Let A be a DHFSs on $X=\{x_1, x_2, \dots, x_n\}$, $f_A(x_i)$ and $g_A(x_i)$ are the membership function and non-membership function of A. $l(f_A(x_i))$ and $l(g_A(x_i))$ are the length of $f_A(x_i)$ and $g_A(x_i)$, respectively.

$$\mu(h_A(x_i)) = 1 - \frac{1}{2} \left[\frac{1}{l(f_A(x_i))} + \frac{1}{l(g_A(x_i))} \right] \quad (4)$$

$$\mu(h_A) = \frac{1}{n} \sum_{i=1}^n \mu(h_A(x_i)) \quad (5)$$

We call $\mu(h_A(x_i))$ the hesitance degree of $h_A(x_i)$, $\mu(h_A)$ the hesitance degree of h_A . The value of $\mu(h_A)$ reflects the degree of hesitance for a decisionmaker to determine the membership for h_A . In the following, we present some new distance measures which include the value of $\mu(h_A)$.

Definition 2.2. Let h_A and h_B be a DHFS on $X=\{x_1, x_2, \dots, x_n\}$. Then the normalized Hamming distance with hesitance degree between $h_A(X_i)$ and $h_B(X_i)$ can be redefined as ($\lambda > 0$):

$$d_{in} = \left\{ \sum_{i=1}^n \left\{ \frac{1}{2n} \left[|\mu(h_A(x_i)) - \mu(h_B(x_i))|^\lambda + \frac{1}{l_{x_i}} \left\{ \sum_{j=1}^{\#h_{x_i}} |\gamma_A^{\delta(j)}(x_i) - \gamma_B^{\delta(j)}(x_i)|^\lambda + \sum_{k=1}^{\#g_{x_i}} |\eta_A^{\delta(k)}(x_i) - \eta_B^{\delta(k)}(x_i)|^\lambda \right\} \right] \right\} \right\} \quad (6)$$

Definition 2.3. A generalized dual hesitant weighted distance with hesitance degree between $h_A(x_i)$ and $h_B(x_i)$ is given as ($\lambda > 0$):

$$d_{2n} = \left\{ \sum_{i=1}^n \left\{ \omega_i \left\{ \left| \mu(h_A(x_i)) - \mu(h_B(x_i)) \right|^\lambda + \frac{1}{I_{x_i}} \left\{ \sum_{j=1}^{\#h_{x_i}} \left| \gamma_A^{\delta(j)}(x_i) - \gamma_B^{\delta(j)}(x_i) \right|^\lambda + \sum_{k=1}^{\#g_{x_i}} \left| \eta_A^{\delta(k)}(x_i) - \eta_B^{\delta(k)}(x_i) \right|^\lambda \right\} \right\} \right\} \right\} \quad (7)$$

Now we need to prove (6) satisfy the conditions of Definition 1.2.

(1) It is obviously that $0 \leq d(A, B) \leq 1$, because all the values of DHFSs are obtained in the interval $[0,1]$;

(2) Necessity: if $d_{in}(A, B) = 0$, then

$$|\mu(h_A(x_i)) - \mu(h_B(x_i))| = 0, |\gamma_A^{\delta(j)}(x_i) - \gamma_B^{\delta(j)}(x_i)| = 0, |\eta_A^{\delta(k)}(x_i) - \eta_B^{\delta(k)}(x_i)| = 0.$$

We can know $h_A(x) = h_B(x)$, so $A=B$.

Sufficiency: if $A=B$, then

$$|\mu(h_A(x_i)) - \mu(h_B(x_i))| = 0, |\gamma_A^{\delta(j)}(x_i) - \gamma_B^{\delta(j)}(x_i)| = 0, |\eta_A^{\delta(k)}(x_i) - \eta_B^{\delta(k)}(x_i)| = 0.$$

We can know $d_{in}(A, B) = 0$.

(3) Obviously, $d_{in}(A, B) = d_{in}(B, A)$.

We can prove equation (7) satisfy the conditions of Definition 1.2. as well.

Now, we reconsider Example 3.1 by applying the above equation (6), we can obtain that $d(A, H)=0.0292$ and $d(B, H)=0.0704$, so this result is accord with our analysis.

3. The application in pattern recognition

To validate the proposed distance measures in practical application, we present another example in this section: an available example quoting from [6].

The problem of building materials recognition is common in pattern recognition. Let each of metal materials be related to four attribute indices $G_j(j= 1, 2, 3, 4)$, let the weight vector of the attributes $G_j(j= 1, 2, 3, 4)$ be $\omega = (0.40, 0.22, 0.18, 0.20)^T$, all data of other metal material be expressed in Table 1. In order to recognize which pattern a new metal material $B = \{\{0.8, 0.2\}, \{0.8, 0.2\}, \{0.5, 0.2\}, \{0.7, 0.3\}\}$. By applying the above-mentioned weighted equation, we can obtain table 2-4.

Table 1. DHFSs for building materials

A \ G	G ₁	G ₂	G ₃	G ₄
A ₁	{{0.5,0.6}{0.3}}	{{0.2}{0.7,0.8}}	{{0.3,0.4}{0.5,0.6}}	{{0.5,0.6,0.7}{0.3}}
A ₂	{{0.8}{0.2}}	{{0.6,0.7,0.8}{0.2}}	{{0.1,0.2}{0.3}}	{{0.2}{0.6,0.7,0.8}}
A ₃	{{0.7,0.8}{0.2}}	{{0.2,0.3,0.4}{0.5}}	{{0.4,0.5}{0.2}}	{{0.2,0.4}{0.5,0.6}}
A ₄	{{0.3,0.4}{0.6}}	{{0.4,0.5}{0.3,0.4}}	{{0.3,0.4}{0.6}}	{{0.4,0.5}{0.5}}
A ₅	{{0.7}{0.3}}	{{0.4,0.5}{0.3,0.4}}	{{0.3}{0.5,0.6,0.7}}	{{0.5}{0.4,0.5}}

Table 2. Distances among Ai and B calculated by equation (3)

λ \ A	A ₁	A ₂	A ₃	A ₄	A ₅	Ranking
$\lambda = 1$	0.1185	0.0680	0.0685	0.1500	0.0990	A ₄ > A ₁ > A ₅ > A ₃ > A ₂
$\lambda = 2$	0.0876	0.0496	0.0487	0.1091	0.0723	A ₄ > A ₁ > A ₅ > A ₂ > A ₃
$\lambda = 4$	0.0687	0.0433	0.0414	0.0948	0.0637	A ₄ > A ₁ > A ₅ > A ₂ > A ₃
$\lambda = 6$	0.0895	0.0415	0.0394	0.0919	0.0617	A ₄ > A ₁ > A ₅ > A ₂ > A ₃

Table 3. Hesitance degree of A_i and B

A \ G	G_1	G_2	G_3	G_4	Hesitance degree of A_i and B
A_1	0.2500	0.2500	0.5000	0.3334	0.3334
A_2	0.5000	0.3334	0.2500	0.3334	0.3542
A_3	0.2500	0.3334	0.2500	0.5000	0.3334
A_4	0.2500	0.5000	0.2500	0.2500	0.3125
A_5	0.5000	0.5000	0.3334	0.2500	0.3959

Table 4. Distances among A_i and B calculated by equation (7)

	λ \ A	A_1	A_2	A_3	A_4	A_5	Ranking
$\alpha = 0.1$ $\beta = 0.9$	$\lambda = 1$	0.4810	0.3535	0.3243	0.5255	0.4200	$A_4 > A_1 > A_5 > A_2 > A_3$
	$\lambda = 2$	0.2288	0.1900	0.1586	0.1987	0.1795	$A_1 > A_4 > A_2 > A_5 > A_3$
	$\lambda = 4$	0.1663	0.1389	0.1080	0.1248	0.1228	$A_1 > A_2 > A_4 > A_5 > A_3$
	$\lambda = 6$	0.1530	0.1274	0.0970	0.1098	0.1146	$A_1 > A_2 > A_5 > A_4 > A_3$
$\alpha = 0.3$ $\beta = 0.7$	$\lambda = 1$	0.4430	0.3535	0.3229	0.4765	0.4200	$A_4 > A_1 > A_5 > A_2 > A_3$
	$\lambda = 2$	0.2015	0.1830	0.1480	0.1776	0.1813	$A_1 > A_2 > A_5 > A_4 > A_3$
	$\lambda = 4$	0.1330	0.1147	0.0880	0.1008	0.1048	$A_1 > A_2 > A_5 > A_4 > A_3$
	$\lambda = 6$	0.1120	0.1006	0.0763	0.0862	0.0913	$A_1 > A_2 > A_5 > A_4 > A_3$
$\alpha = 0.5$ $\beta = 0.5$	$\lambda = 1$	0.4058	0.3675	0.3217	0.4275	0.4225	$A_1 > A_4 > A_5 > A_2 > A_3$
	$\lambda = 2$	0.1742	0.1758	0.1374	0.1565	0.1831	$A_5 > A_1 > A_2 > A_4 > A_3$
	$\lambda = 4$	0.0996	0.0905	0.0678	0.0768	0.0868	$A_1 > A_2 > A_5 > A_4 > A_3$
	$\lambda = 6$	0.0864	0.0738	0.0555	0.0626	0.0681	$A_1 > A_2 > A_5 > A_4 > A_3$
$\alpha = 0.7$ $\beta = 0.3$	$\lambda = 1$	0.3670	0.3745	0.3201	0.3785	0.4200	$A_5 > A_4 > A_2 > A_1 > A_3$
	$\lambda = 2$	0.1469	0.1687	0.1267	0.1354	0.1849	$A_5 > A_2 > A_1 > A_4 > A_3$
	$\lambda = 4$	0.0662	0.0575	0.0476	0.0528	0.0688	$A_5 > A_1 > A_2 > A_4 > A_3$
	$\lambda = 6$	0.0531	0.0470	0.0347	0.0390	0.0449	$A_1 > A_2 > A_5 > A_4 > A_3$
$\alpha = 0.9$ $\beta = 0.1$	$\lambda = 1$	0.3290	0.3815	0.3187	0.3295	0.4200	$A_5 > A_2 > A_1 > A_4 > A_3$
	$\lambda = 2$	0.1196	0.1616	0.1061	0.1143	0.1867	$A_5 > A_2 > A_1 > A_4 > A_3$
	$\lambda = 4$	0.0329	0.0421	0.0275	0.0288	0.0508	$A_5 > A_2 > A_1 > A_4 > A_3$
	$\lambda = 6$	0.0198	0.0202	0.0140	0.0154	0.0216	$A_5 > A_2 > A_1 > A_4 > A_3$

It is clear in the table 2, when the values of λ are different, the optimum metal material is different (A_2 or A_3), so the traditional equation only consider the membership values but cannot take into account the hesitance degree of each hesitant fuzzy element. In the table 3, it is obviously that the hesitance degree of A_3 and B is relatively smaller than that of A_2 and B. We also can obtain that no matter how much the value of is, the minimal distance is the distance among A_3 and B in tables 4. Based on the minimum distance principle, it is easy to get the conclusion that A_3 is the optimum metal material.

4. Summary

This paper present a new definition of DHFSs based on the original definition by introducing the concept of hesitance degree and investigated their application. We also apply our proposed new distance measures of DHFSs in pattern recognition. Compared to the existing definitions, the proposed definition has a better distinction to some degree. We also look forward to make some further development about DHFSs.

Acknowledgements

This research was financially supported by the National Science Foundation of China (Grant No. 61573240, 61473239).

References

- [1] L.A.Zadeh: Fuzzy Sets, Information and control 8 (1965), p. 338-353.
- [2] Xiaohong Zhang, Daowu Pei, Jianhua Dai: *Fuzzy Mathematics and Rough Set Theory* (Tsinghua University press, Beijing 2013)
- [3] Bin Zhu, Zeshui Xu, Meimei Xia: Dual hesitant fuzzy sets, J. Appl. Math. 2012 (2012). p.1-13.
- [4] Zhan Su, Zeshui Xu, Haifeng Liu, Shousheng Liu, Distance and similarity measures for dual hesitant fuzzy sets and their applications in pattern recognition, Journal of Intelligent & Fuzzy Systems 29 (2015) p.731-745.
- [5] Qiuyan Zhan, Xiaohong Zhang, Zhenyu Chen: Modified definitions of distance and similarity measures of interval-valued hesitant fuzzy sets, Fuzzy Systems and Mathematics 30 (2016), in press.
- [6] Deqing Li, Wenyi Zeng, Junhong Li: New distance and similarity measures on hesitant fuzzy sets and their applications in multiple criteria decision making, Engineering Applications of Artificial Intelligence 40 (2015), p. 11–16.