International Journal of Computational Intelligence Systems, Vol. 7, No. 2 (April 2014), 327-343

An Integrated Intuitionistic Fuzzy Similarity Measures for Medical Problems

Kuo-Chen Hung*
Department of Computer Science and Information Management, Hungkuang University, No. 1018, Sec. 6, Taiwan Boulevard, Shalu District, Taichung City 43302, Taiwan

Pei-Kuang Wang
Department of Logistics Management, National Defense University, No. 70, Sec. 2, Jhongyang N. Rd., Beitou District, Taipei City 112, Taiwan

**Corresponding author, E-mail: kuochen.hung@msa.hinet.net*

Received 24 February 2012

Accepted 6 November 2013

Abstract

The purpose of this paper is to develop an integrated similarity measures model based on intuitionistic fuzzy sets. This integrated model has improved two similarity measures methods: (1) Ye (*Mathematical and Computer Modelling*, 53, 91-97 (2011)) presented a novel cosine similarity measures method for handling pattern recognition problems based on intuitionistic fuzzy sets. However, in some cases, Ye's method can not give sufficient information to discriminate a sample between two patterns. Therefore, we provide an improved method for the similarity measure. (2) Hung et al. (*Computer-Aided Design*, 40, 447-454 (2008)) provided a new score function to measure the degree of suitability of each alternative. In this paper, we extend their method to modify the hesitation parameter with rate operations as a defuzzfication function for each characteristic, and then the defuzzy results as a parameter input new similarity measure method based on the Minkowski distance conception. Finally, the proposed similarity measures model is applied to two medical diagnosis problems to demonstrate the usefulness of this study. Furthermore, in order to make computing and ranking results easier, a computer-based interface system is also developed, and this system may help to make a decision more efficiently.

Keywords: Intuitionistic fuzzy sets; similarity measure; pattern recognition; medical diagnosis

1. Introduction

The concept of fuzzy sets was introduced by Zadeh[1] in 1965. Moreover, Atanassov[2] extended fuzzy sets to present intuitionistic fuzzy sets (IFSs) and this approach has been successfully applied in fields such as pattern recognition, machine learning, decision making and image processing. Among those applications, an important issue is the similarity measure method. In recent years, many different similarity measures methods between IFSs have been addressed. Chen[3,4] proposed some similarity functions to measure the degree of similarity between fuzzy sets and vague sets. Hong and Kim[5] provided an improved similarity measure method based on Chen's[4] method. Szmidt and Kacprzyk[6,7] introduced the Hamming distance between IFSs and proposed a similarity measure between IFSs based on the distance and its application in group decision making. Li and Cheng[8] also proposed similarity measures of IFSs and applied similarity measures to pattern recognition. Liang and Shi[9] proposed several similarity measures to differentiate different IFSs and discussed the relationships between these measures. Mitchell[10] a statistical viewpoint to modify Li and Cheng's[8] measures method. Hung and Yang[11] proposed another method to calculate the distance between IFSs based on the Hausdorff distance and used it to generate several similarity measures between IFSs that are suitable for use in linguistic variables. Grzegorzewski[12] had presented a new similarity measure method based on the Hausdorff metric; Liu[13] proposed some similarity measures between IFSs and between elements and applied the measure methods in pattern recognition. Wang and Xi[14] presented a hybrid method with Hamming distance and Hausdorff distance for similarity measure; Zhang and Fu[15] presented a similarity measure on three kinds of fuzzy sets. Xu[16] provided some similarity measures of intuitionistic fuzzy sets and their applications to multiple attribute decision making. Park et al.[17] addressed some similarity measures based on the Minkowski distance. Also, Hung and Yang[18] proposed a method to calculate the degree of similarity between IFSs, in which the proposed similarity measures are induced by L_p metric.

Xu and Xia[19] proposed a variety of distance measures including Hamming, Euclidean and Hausdorff distances for hesitant similarity measures, and the proposed method was applied to socio-economic decision making problem. Yusoff et al.[20] concentrated on Zhang and Fu's method[15] to develop a new similarity measure method for IFSs and some examples were given to validate these similarity measures. Mukherjee and Basu[21] based on matching function to solve a class of intuitionistic fuzzy assignment problem.

Herein, Ye[22] had also proposed a novel similarity measure function to solve some problem of pattern recognition and medical diagnosis. However, in some cases, the proposed measures of Ye[22] did not give sufficient information about the samples. The details of Ye's[22] method will be depicted in Section 3.1. Some recent development on similarity measures development based on IFSs are summarized in Table 1.

In this paper, we improve and extend some existing similarity measures to develop an integrated model to measure the degree of similarity between a sample and patterns. Unlike existing similarity measures, the proposed model can overcome some drawbacks of counter-intuition. Finally, we illustrate two applications in the context of colorectal cancer diagnosis and medical diagnosis by measuring similarity between IFSs.

The rest of this paper is organized as follows: In Section 2, the intuitionistic fuzzy sets are briefly depicted, and the cosine similarity measure function is discussed. In Section 3, two enhanced similarity measure methods are addressed. In Section 4, a proposed integrated similarity measure model is presented. Then the proposed model to handle medical pattern recognition problems is presented in Section 5. In Section 6, a computer based interface system is constructed. Finally, conclusions are drawn in Section 7.

Table 1. Some developments of similarity measure technology.

Author(s)	Year	Technology	Applied problem	Compared method(s)
Chen	1995	Hamming distance	Pattern recognition problem	None
Chen	1997	Hamming distance	Behavior analysis in an organization	None
Hong and Kim	1999	Modify Chen's (1995,1997) methods	Behavior analysis in an organization	Chen (1995, 1997)
Szmidt and Kacprzyk	2000	Hamming distance and Euclidean distance	Pattern recognition problem	None
Li and Cheng	2002	Euclidean distance	Pattern recognition problem	None
Liang and Shi	2003	Euclidean distance	Pattern recognition problem	Li and Cheng (2002)
Mitchell	2003	Modified Dengfeng–Chuntian (2002) method	Pattern recognition problem	None
Hung and Yang	2004	Hausdorff distance	Pattern recognition problem; Compound linguistic variables application	De et al. (2000)[23]; Li and Cheng (2002); Liang and Shi (2003); Mitchell (2003)
Grzegorzewski	2004	Hausdroff Hamming distance; Hausdorf Euclidean distance	Pattern recognition problem	None
Liu	2005	Euclidean distance	Pattern recognition problem	Hong and Kim (1999); Li and Cheng (2002)

Table 1. (Continued)

Wang and Xi	2005	Hybrid Hamming distance and Hausdorff distance	Pattern recognition problem; Minerals clustering	Liang and Shi (2003)
Zhang and Fu	2006	Hamming distance	Colorectal cancer diagnosis	None
Xu	2007	Hybrid distance ratio	Air-conditioning system selecting	None
Hung and Yang	2007	L_p metric measure	Material pattern recognition; Minerals clustering	Liang and Shi (2003); Mitchell (2003); Hung and Yang (2004); Wang and Xin (2005); Yang and Wu (2004)[24]
Park et al.	2007	Minkowski distance	Pattern recognition problem	Li and Cheng (2002); Liang and Shi (2003)
Xu and Xia	2011	Hesitant distance; Hesitant Hausdorff distance; Hesitant Hamming distance; Hesitant Euclidean distance; Hybrid hesitant Hamming distance	Socio-economic decision making	None
Yusoff et al.	2011	Hamming distance	Colorectal cancer diagnosis	Zhang and Fu (2006)
Ye	2011	Cosine similarity measure	Pattern recognition; Medical diagnosis	Li et al. (2007)[25]
Mukherjee and Basu	2012	Hybrid distance ratio	Assignment problem	None

2. Theoretical Background

2.1. *Intuitionistic fuzzy sets*

Fuzzy sets theory, proposed by Zadeh[1] in 1965, has been successfully applied in various fields. In this theory, the membership of an element to a fuzzy set is a single value between zero and one. But in reality, it may not always be certain that the degree of nonmembership of an element to a fuzzy set is just equal to 1 minus the degree of membership, i.e. there may be some hesitation degree. Thus, as a generalization of fuzzy sets, the concept of IFSs was introduced by Atanassov[2] in 1986. Gau and Buehrer[26] researched vague sets (VSs). On the basis of that research, Bustince and Burillo[27] pointed out that the notion of vague sets is the same as that of IFSs introduced by Atanassov.

The IFSs is as an extension of fuzzy sets. An IFSs A in a fixed set X is an objective with the expression

$$A = \left\{ \left\langle x, \mu_A(x), \nu_A(x) \right\rangle \middle| x \in X \right\} \tag{1}$$

Where the functions $\mu_A : X \to [0,1]$ and $\nu_A : X \to [0,1]$ denote the degree of membership and the degree of nonmembership of the element $x \in X$, respectively. For every $x \in X$,

$$0 \le \mu_A(x) + \nu_A(x) \le 1 . \tag{2}$$

When $\mu_A(x) + \nu_A(x) = 1$, for every $x \in X$, then the IFSs will degenerate to a fuzzy set. Hence, we can consider a fuzzy set with its membership function $\mu_A(x)$, having the IFSs expression as

$$A = \left\{ \left\langle x, \mu_A(x), 1 - \mu_A(x) \right\rangle \middle| x \in X \right\} \tag{3}$$

under the condition of $\nu_A(x) = 1 - \mu_A(x)$, for every $x \in X$.

We call the hesitation degree of an element $x \in X$ in A by the following expression:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x). \tag{4}$$

From Eq. (4) it is evident that

$$0 \le \pi_A(x) \le 1 \quad \text{for all} \quad x \in X. \tag{5}$$

Therefore, to describe an intuitionistic fuzzy set completely, we need at least two functions from the triplet: (1) membership function; (2) non-membership function; and (3) hesitancy degree[28,29].

2.2. *Cosine similarity measures for IFSs*

Cosine similarity measures (Bhattacharya[30], Salton and McGill[31]) are defined as the inner product of two vectors divided by the product of their lengths. This is nothing but the cosine of the angle between the vector representations of the two fuzzy sets.

Assume that $A = (\mu_A(x_1), \mu_A(x_2), ..., \mu_A(x_n))$ and $B = (\mu_B(x_1), \mu_B(x_2), ..., \mu_B(x_n))$ are two fuzzy sets in the universe of discourse $X = \{x_1, x_2, . . . , x_n\}$. A cosine similarity measure based on Bhattacharya's distance[30] between $\mu_A(x_i)$ and $\mu_B(x_i)$ can be defined as follows[31]:

$$C_F(A,B) = \frac{\sum_{i=1}^{n} \mu_A(x_i)\mu_B(x_i)}{\sqrt{\sum_{i=1}^{n} \mu_A^2(x_i)}\sqrt{\sum_{i=1}^{n} \mu_B^2(x_i)}} \tag{6}$$

The cosine of the angle between the vectors is between 0 and 1.

Furthermore, based on Eq. (6), Ye[22] addressed a new cosine similarity measure for two IFSs shown as the following,

Let two IFSs numbers $A = \{\langle x_i, \mu_A(x_i), \nu_A(x_i)\rangle | x_i \in X\}$ and $B = \{\langle x_i, \mu_B(x_i), \nu_B(x_i)\rangle | x_i \in X\}$, where the universe of disclose $X = \{x_i, i = 1, 2, \ldots, n\}$ perform similarity measure, then the extended cosine similarity measure is the following,

$$C_{IFS}(A,B) = \frac{1}{n}\sum_{i=1}^{n} \frac{\mu_A(x_i)\mu_B(x_i) + \nu_A(x_i)\nu_B(x_i)}{\sqrt{\mu_A^2(x_i) + \nu_A^2(x_i)}\sqrt{\mu_B^2(x_i) + \nu_B^2(x_i)}} \tag{7}$$

Therefore, the cosine similarity measure between IFSs A and B also satisfies the following properties[32]:

(P1) $0 \le C_{IFS}(A, B) \le 1$;

(P2) $C_{IFS}(A, B) = C_{IFS}(B, A)$;

(P3) $C_{IFS}(A, B) = 1$ if $A = B$, i.e. $\mu_A(x_i) = \mu_B(x_i)$ and $\nu_A(x_i) = \nu_B(x_i)$ for $i = 1, 2, . . . , n$.

For the relative proofs, please can refer to Ye[22]. However, Mitchell 10 think that (P3) stands no restriction on the number of ways $C_{IFS}(A, B)$ may equal one, as long as $C_{IFS}(A, B) = 1$ when $A = B$. Therefore, Mitchell[10] sugges using only strong similarity measures (P3') with "$C_{IFS}(A, B) = 1$ if and only if $A = B$".

Moreover, with the weights considered, Ye[22] had also presented a weighted cosine similarity measure shown as follows,

$$W_{IFS}(A,B) = \sum_{i=1}^{n} w_i \cdot \frac{\mu_A(x_i)\mu_B(x_i) + \nu_A(x_i)\nu_B(x_i)}{\sqrt{\mu_A^2(x_i) + \nu_A^2(x_i)}\sqrt{\mu_B^2(x_i) + \nu_B^2(x_i)}}, \tag{8}$$

where $w_i \in [0,1], i$ = 1, 2, …, n, and $\sum_{i=1}^{n} w_i = 1$. If the weights are $w_i = 1/n, i = 1, 2, ..., n$, then there is $W_{IFS}(A,B) = C_{IFS}(A,B)$.

3. Some Enhanced Similarity Measure Methods

3.1. *Improved method 1: New cosine similarity measure between IFSs*

Although Ye's[22] method has presented a novel cosine similarity measures for IFSs based on the angle, this method of similarity measurement is different from the traditional viewpoint, based on "Distance". However, since the method only considers the included angle between vectors consisting of parameters μ and ν, it will cause unsolvable problem in some cases. Now, we will illustrate counter example to describe the unsolvable problem.

Counter example: Let A^*, B and C be three patterns, and the set IFS values are represented by IFSs: A^* = (0.1, 0.2), B = (0.2, 0.4) and C = (0.25, 0.5). We can draw this as Fig. 1. Therefore, we can obtain the final evaluations using Eq. (7) by Ye's method. The results are

$C_{IFS}(A^*, B) = 1$, $C_{IFS}(B, C) = 1$ and $C_{IFS}(A^*, C) = 1$.

Hence, based on Ye's approach, A^*, B and C have same the degree of similarity, because they have same the included angle (0° angle). According to the strong property (P3') then it follows that $A^* = B = C$. However, it can be intuitively observe in Fig. 1, that the IFSs A^*, B and C are different, so how can their similarity measures can be one?

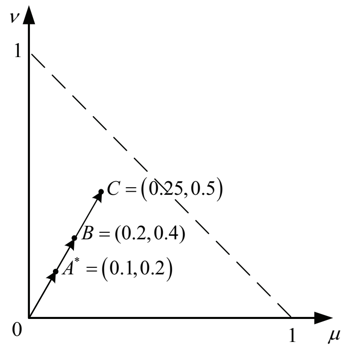

Fig. 1. A counter example using Ye's cosine similarity measures based on IFSs.

Based on the above counter example, Ye's[22] method still has some disadvantages in dealing with similarity measures. In order to overcome the above problem, an improved cosine similarity measure method is proposed.

The main cause of the unsolvable problem is that Ye's[22] method only considers the included angle between vectors consisting of membership and non-membership to overlook the hesitancy. This disadvantage provides us with an improved direction as follows:

Atanassov[2] proposed the conception of the intuitional fuzzy set that considered the value of hesitation. For satisfying the condition of $\mu(x)+\nu(x)+\pi(x)=1$ in the IFSs, vectors are altered from two-dimensional space to three-dimensional space (Fig. 2). Therefore, all the vectors in the IFSs are independent line in three dimensions. There are no two dependent vectors on the first quatrain on the plane $x+y+z=1$ to avoid the situation of the counter example.

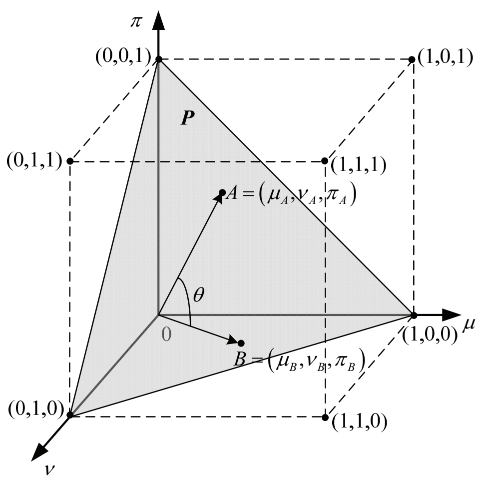

Fig. 2. The conception of an improved cosine similarity measure approach.

Based on the-above mentioned direction, an improved cosine similarity measure method is proposed as follows;

$$C_{IFS}^{new}(A,B)$$
$$=\frac{1}{n}\sum_{i=1}^{n}\frac{\mu_A(x_i)\mu_B(x_i)+\nu_A(x_i)\nu_B(x_i)+\pi_A(x_i)\pi_B(x_i)}{\sqrt{\mu_A^2(x_i)+\nu_A^2(x_i)+\pi_A^2(x_i)}\sqrt{\mu_B^2(x_i)+\nu_B^2(x_i)+\pi_B^2(x_i)}} \quad (9)$$

We will prove that our proposed cosine similarity measure will satisfy the following properties:

(P1) $0\le C_{IFS}^{new}(A, B)\le 1$;
(P2) $C_{IFS}^{new}(A, B) = C_{IFS}^{new}(B, A)$;
(P3) $C_{IFS}^{new}(A, B) = 1$ if and only if $A = B$, $\mu_A(x_i)=\mu_B(x_i)$ and $\nu_A(x_i)=\nu_B(x_i)$, for $\forall x_i\in X$.

Proof.

(P1) From the Cauchy-Schwarz inequality and the non-negative components, we know that

$$0\le \frac{\mu_A(x_i)\mu_B(x_i)+\nu_A(x_i)\nu_B(x_i)+\pi_A(x_i)\pi_B(x_i)}{\sqrt{\mu_A^2(x_i)+\nu_A^2(x_i)+\pi_A^2(x_i)}\sqrt{\mu_B^2(x_i)+\nu_B^2(x_i)+\pi_B^2(x_i)}} \le 1 \quad (10)$$

thus $0\le C_{IFS}^{new}(A, B)\le \frac{1}{n}\sum_{i=1}^{n}1=1$.

(P2) Because the C_{IFS}^{new} is symmetric with respect to IFSs A and B, thus (P2) is hold.

(P3) When $A=B$, there are $\mu_A(x_i)=\mu_B(x_i)$, $\nu_A(x_i)=\nu_B(x_i)$ and $\pi_A(x_i)=\pi_B(x_i)$ for i = 1,2,…, n. This implies that $C_{IFS}^{new}(A,B)=1$.

If $C_{IFS}^{new}(A,B)=1$, by Eqs. (9) and (10), we derive that

$$\frac{\mu_A(x_i)\mu_B(x_i)+\nu_A(x_i)\nu_B(x_i)+\pi_A(x_i)\pi_B(x_i)}{\sqrt{\mu_A^2(x_i)+\nu_A^2(x_i)+\pi_A^2(x_i)}\sqrt{\mu_B^2(x_i)+\nu_B^2(x_i)+\pi_B^2(x_i)}}=1, \quad (11)$$

for $i=1,\cdots,n$. If we set that $U_i=(\mu_A(x_i),\nu_A(x_i),\pi_A(x_i))$ and $V_i=(\mu_B(x_i),\nu_B(x_i),\pi_B(x_i))$ then Eq. (11) indicates that

$$\langle U_i,V_i\rangle=\sqrt{\langle U_i,U_i\rangle}\sqrt{\langle V_i,V_i\rangle}. \quad (12)$$

By Cauchy-Schwarz inequality, we have U_i and V_i parallel, such that there is a non-zero constant, say α_i, with

$$\frac{\mu_A(x_i)}{\mu_B(x_i)}=\frac{\nu_A(x_i)}{\nu_B(x_i)}=\frac{\pi_A(x_i)}{\pi_B(x_i)}=\alpha_i, \quad (13)$$

owing to $\mu_A(x_i)+\nu_A(x_i)+\pi_A(x_i)=1$ and $\mu_B(x_i)+\nu_B(x_i)+\pi_B(x_i)=1$ for $i=1,\cdots,n$ implying that $\alpha_i=1$ and $U_i=V_i$, for $i=1,\cdots,n$. Therefore, we have finished the proofs.

If we consider the axiom for similarity measure of Li and Cheng[8] and Mitchell[10], a well-defined similarity measure satisfies the following four properties:

(P1) $0\le C_{IFS}^{new}(A,B)\le 1$;
(P2) $C_{IFS}^{new}(A,B)=C_{IFS}^{new}(B,A)$;

(P3) $C_{IFS}^{new}(A,B)=1$ if and only if $A=B$;
(P4) If $A \subseteq B \subseteq C$, A, B, $C \in \text{IFSs}(X)$, then $C_{IFS}^{new}(A,C) \le C_{IFS}^{new}(A,B)$ and $C_{IFS}^{new}(A,C) \le C_{IFS}^{new}(B,C)$.

Proof.
We have proved (P1)-(P3) in the previous discussion. The detailed proof of (P4) is provided in the Appendix.

Based on Eq. (9), if we also consider the weights of x_i, a weighted cosine similarity measure between IFSs A and B is proposed as follows;

$$W_{IFS}^{new}(A,B)$$
$$=\sum_{i=1}^{n} w_i \cdot \frac{\mu_A(x_i)\mu_B(x_i)+\nu_A(x_i)\nu_B(x_i)+\pi_A(x_i)\pi_B(x_i)}{\sqrt{\mu_A^2(x_i)+\nu_A^2(x_i)+\pi_A^2(x_i)}\sqrt{\mu_B^2(x_i)+\nu_B^2(x_i)+\pi_B^2(x_i)}} \tag{14}$$

where $w_i \in [0,1]$, $i = 1, 2, \ldots, n$, and $\sum_{i=1}^{n} w_i = 1$.
If we take $w_i = 1/n$, $i = 1, 2, \ldots, n$, then there is $W_{IFS}^{new}(A,B)=C_{IFS}^{new}(A,B)$. Obviously, the weighted cosine similarity measure of two A and B also satisfies the properties (P1-P3') of Mitchell[10] and the properties (P1-P4) of Li and Cheng[8] and Mitchell[10].

Now, let we recall the above-mentioned counter example in section 3.1. Our proposed cosine similarity measures method, Eq. (9), was applied, and the calculated results are show in Table 2. Through the counter example, we can obtain that $C_{IFS}^{new}(A^*,B)=0.862$ and $C_{IFS}^{new}(A^*,C)=0.667$. Thus we know that $C_{IFS}^{new}(A^*,B) \succ C_{IFS}^{new}(A^*,C)$. However, Ye's method can not distinguish them. Therefore, we can observe that our proposed method is more distinguishable than Ye's[22] method.

Table 2. A comparison between Ye's method and our method in the counter example.

	Counter example	
	A^* = (0.1, 0.2), B = (0.2, 0.4), C = (0.25, 0.5)	
Ye's method	$C_{IFS}(A^*,B)=1$	$C_{IFS}(A^*,C)=1$
Our method	$C_{IFS}^{new}(A^*,B)=0.862$	$C_{IFS}^{new}(A^*,C)=0.667$

Although our proposed similarity measure method improves Ye's[22], our method is similar to other existing similarity methods. Unsolved problems still exist with our proposed method. Namely, we still can find some cases to imply identical ranking results. For example, using our proposed method, when $A^*=(1,0,0)$, $B=(0,1,0)$, $C=(0,0,1)$, then $C_{IFS}^{new}(A^*,B)=0$ and $C_{IFS}^{new}(A^*,C)=0$ implying that $C_{IFS}^{new}(A^*,B)=C_{IFS}^{new}(A^*,C)$. We still can not to distinguish which one (B or C) is more similar to A^*. Therefore, we need another similarity measure method to resolve the problem of identical ranking between two IFSs.

3.2. *Improved method 2: An extension of Hung et al. (2008) between IFSs*

In order to handle the MCDM problem, Hung et al.[33] had proposed a score function, K, shown as follows:

$$A=\left\{\left\langle x_i,\mu_A(x_i),\nu_A(x_i)\right\rangle \middle| x_i \in X\right\}$$

Definition 1. Let A be a IFS, $A=\left\{\left\langle x,\mu_A(x),\nu_A(x)\right\rangle \middle| x \in X\right\}$, where $\mu_A \in [0, 1]$, $\nu_A \in [0, 1]$, $\mu_A+\nu_A \le 1$. The unknown degree, or hesitancy degree, of A is denoted by π_A , and is defined by $\pi_A=1-\mu_A-\nu_A$, and $0 \le \pi_A \le 1$. Then a score function is provided as follows,

$$K(A)=\mu_A+\lambda\pi_A , \tag{15}$$

where $K(A) \in [0, 1]$, and $\lambda \in [0, 1]$. The parameter λ has been considered to express the percentage of hesitancy degree for pro. When $\lambda = 0$, it shows that the decision-maker is the most pessimistic, because it can not obtain anything from the part of hesitancy degree. When $\lambda = 0.5$, it shows that the decision-maker is fair and can obtain a half of the hesitancy degree. When $\lambda = 1$, it shows that the decision-maker is in the most optimistic situation, and can get complete, support for the hesitancy degree.

In this paper, we have modified the parameter λ based on Hung et al.'s[33] method. The value of λ should depend on the decision maker preference. We improve the parameter λ by modifying as a rate, $\mu_A/(\mu_A+\nu_A)$. Therefore, we can extend Hung et al.'s[33] method as a similarity measure, shown as the following:

$$K(A)=\mu_A+(\mu_A/(\mu_A+\nu_A))\pi_A . \tag{16}$$

Then Eq. (16) can be extended as a new similarity measure, shown as the following:

Definition 2. Let two IFS,

$A=\{\langle x_i,\mu_A(x_i),\nu_A(x_i)\rangle | x_i\in X\}$ and

$B=\{\langle x_i,\mu_B(x_i),\nu_B(x_i)\rangle | x_i\in X\}$,

where the universe of disclose $X=\{x_1, x_2, \ldots, x_n\}$; then we define a new similarity measure based on the normalized Minkowski distance as,

$$H_{IFS}^{new}(A,B)=1-\frac{1}{n^{1/p}}\sqrt[p]{\sum_{i=1}^{n}\left|K(A(x_i))-K(B(x_i))\right|^p}, \quad (17)$$

where $H_{IFS}^{new}(A,B)\in[0,1]$, p norm is a parameter with $1\le p<\infty$.

4. A Proposed Integrated Similarity Measures Model

In this paper, an integrated similarity measure model for measuring the degree of similarity between patterns and a sample with IFSs value has developed. Three similarity measure methods have been adopted in three different stages. (1) The first stage is the improved cosine similarity measure method, as Eq. (14); (2) the second stage is an extension of Hung et al.'s[33] method, as Eq. (17); (3) the third stage adopts the sum of degree through the pro viewpoint if necessary. A flowchart for the proposed integrated similarity measure model is presented in Fig. 3.

The details of the procedures and pseudo code of the algorithm are shown as follows:

Step 1: *Parameter input and intuitionistic fuzzy value input for each sample*

The worker needs to input some parameters, p norm and respective weight w_i, i = 1, 2, …, n. The relative intuitionistic fuzzy values also need to be input for each characteristic in the sample.

Step 2: *Performing the first similarity measure method*

In this step is performed the proposed first similarity measure method, $W_{IFS}^{new}(A,B)$, where the notation B denotes the pattern and A denotes the unknown sample. Namely, it starts Eq. (14) to calculate the similarity degree between each pattern and the unknown sample.

Step 3: *Clustering representation*

According to the results of performing Step 2, we can order all patterns and the sample based on their degree of similarity. The greatest similarity degree denotes that higher similarity exists between pattern and sample. Therefore, we know that the sample, S, belongs to the pattern, P. If the similarity degrees of all patterns can be distinguished, then all steps are stopped; otherwise, the next step will be executed.

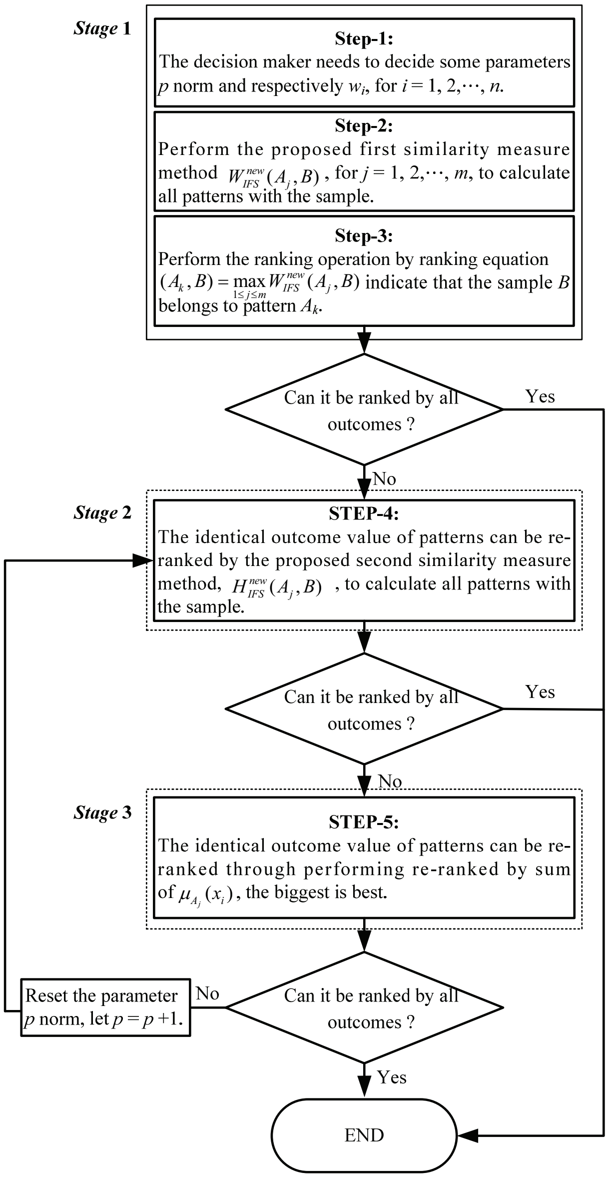

Fig. 3. The flowchart of our proposed integrated similarity measure method.

Step 4: *Re-ranking by the second similarity measure method*

If the outcomes are identical to those in Step 3, then Step 4 can begin to re-calculate the similarity degree focus on those identical outcomes. That is, the second similarity measure method, $H_{IFS}^{new}(A,B)$ has be started, where P denotes the pattern and S denotes the sample. Then it performs re-rank operations based on the calculated results of the second similarity measure method. If the similarity degrees of all patterns can be distinguished, then all steps are stopped; otherwise, the next step will be executed.

Step 5: *Re-ranking based on the degree of sum of pro*

Step 5 is similar to Step 4. If the outcomes in Step 4 are identical, then Step 5 will begin to re-calculate the similarity degree based on the method of the sum of pro. Then the re-rank operations will be performed again. If the similarity degrees of all patterns can be distinguished, then all steps are stopped; otherwise, the next step will be executed.

Step 6: *Reset the parameter p norm*

If the identical outcomes still cannot be distinguished, then the parameter p norm will be changed by adding one, and Step 4 will be executed to begin the next loop.

Algorithm (Pseudo code)

```
% Parameters setting
φ = 1, Input_W_Matrix;   % Setting initial parameters and
    weights for each criteria
Output_Set=[ ];   % Initial matrix is empty
Result_Sum =0;   % Initial value is zero
For j=1:m   % The total patterns are m.
   For i =1:n   % The total criteria are n for each pattern.
      Result= W_IFS^new(A_j,B) ;   % Stage 1: Performing first
similarity measure method, B is a sample and A is a set of patterns
      Result_Sum = Result_Sum + Result;
   End
   Output_Set=[ Output_Set, Result_Sum];
End
[Sort_Matrix_1, Sort_Index]=dsort(Output_Set); % The ranking
results by decrease (Showing relative similarity degree and their
address)
If Sort_Matrix_1[1,1] = Sort_Matrix_1[1,2] then   % Denotes the
existence of identical outcomes with the biggest value
      flag = 1;
Else
      flag = 0;   % Denotes the nonexistence identical outcomes,
thus stop all steps are stopped
      Show_Final_Ranking;
End
WHILE flag = 1 Do % Because of  existing identical outcomes,
thus it needs to execute "While_Do_Loop"
      For j=1:k   % The numbers of identical outcome are j = 1, 2,
..., k.
         For i =1:n
            Pattern_Output=K_function(P_j);   % Calculating by
equation (14)
            Sample_Output=K_function(S);   % Calculating by
equation (14)
         End
         Outcome= H_IFS^new(A_j,B) ;   % Stage 2: Performing second
similarity measure method
         Identical_Set = [Identical_Set, Outcome];
      End
      Sort_Matrix_2=dsort(Identical_Set);
      If Sort_Matrix_2[1,1] ≠ Sort_Matrix_2[1,2] then      %
Denotes the nonexistence of identical outcomes
            flag = 0;   % It will stop all steps
      Else
         For j=1:z   % The numbers of identical outcome, j = 1, 2,
..., z.
            For i =1:n
               Pro_Sum_Set = Pro_Sum_Set + μ_Pj(x_i) ;   %
Stage 3: Performing third similarity measure by sum of pro
            End
         End
         Sort_Matrix_3=dsort(Pro_Sum_Set);
         If Sort_Matrix_3[1,1] ≠ Sort_Matrix_3[1,2] then   %
Denotes the nonexistence identical outcomes
            flag = 0;   % All steps will be stopped
         Else
               p = p+1;   % Reset parameter φ norm, then go
to new WHILE_LOOP
         End if
      End if
End WHILE   % End of while loop
```

In order to describe the proposed similarity measure model is reasonable, a numerical example has been illustrated.

Numerical example: Let A_1, A_2 and A_3 be three patterns, and their intuitionistic fuzzy values are

$A_1 = \{\langle x_1, 0.2, 0.6\rangle, \langle x_2, 0.1, 0.7\rangle, \langle x_3, 0.0, 0.6\rangle\}$,

$A_2 = \{\langle x_1, 0.2, 0.6\rangle, \langle x_2, 0.0, 0.6\rangle, \langle x_3, 0.2, 0.8\rangle\}$ and

$A_3 = \{\langle x_1, 0.1, 0.5\rangle, \langle x_2, 0.2, 0.7\rangle, \langle x_3, 0.2, 0.8\rangle\}$.

Then we apply the procedures of our new model to handle this pattern recognition problem shown as the following:

Step 1: *Parameters input and intuitionistic fuzzy value input for each sample*

Let $p = 2$, $w_j = 1/3$, for $j = 1,2,3$, and input the sample, $B = \{\langle x_1, 0.15, 0.55\rangle, \langle x_2, 0.1, 0.65\rangle, \langle x_3, 0.15, 0.7\rangle\}$.

Step 2: *Performing the first similarity measure method*

By Eq. (14), we can obtain that

$W_{IFS}^{new}(A_1, B) = 0.963$ and

$W_{IFS}^{new}(A_2, B) = 0.975 = W_{IFS}^{new}(A_3, B)$.

Step 3: *Clustering representation*

Ranking all patterns with sample, we obtain that

$W_{IFS}^{new}(A_2,B)=W_{IFS}^{new}(A_3,B)>W_{IFS}^{new}(A_1,B)$.

The pattern A_2 and A_3 have the same outcome and have the biggest similarity degree; thus we can not distinguish whether sample *B* belongs to A_2 or A_3. Namely, the first similarity measure of our model can not help the decision maker to decide the pattern for the sample. Therefore, we need to star the second similarity measure in Step 4.

Step 4: *Second similarity measure method*

We try our second similarity measure to derive the identical outcomes, namely to perform Eq. (17) and obtain that $W_{IFS}^{new}(A_2,B)=0.919$ and $W_{IFS}^{new}(A_3,B)=0.940$. Namely, that $W_{IFS}^{new}(A_3,B)\succ W_{IFS}^{new}(A_2,B)$. Therefore, we know that sample *B* should be belong to pattern A_3, and all steps are stopped.

In this case, it needs two similarity measure methods to measure the degree of similarity.

5. Applications and Discussions

Pattern recognition plays an important part in human cognition. Humans are able to identify patterns that appear in many types of data, recognize instances of these patterns, and draw relevant conclusions. In this section, two applications for medical pattern recognition have been illustrated to describe the usefulness of the proposed method. Some issues from the applications and our proposed method are raised in section 5.3.

5.1. *Application 1—Medical diagnosis*

We will illustrate an application using the proposed approach for medical pattern recognition problem. Let us consider the same example as Szmidt and Kacprzyk[34,35].

They consist of a set of diseases D = {d_1(Viral fever), d_2(Malaria), d_3(Typhoid), d_4(Stomach problem), d_5(Chest pain)}, a set of symptoms S = {s_1(Temperature), s_2(Headache), s_3(Stomach pain), s_4(Cough), s_5(Chest pain)} and a set of patients P = { p_1(Al), p_2(Bob), p_3(Joe), p_4(Ted)}. Then the symptoms for each patient are given with IFSs values as follows;

$$p_1=\{\langle s_1,(0.8,0.1)\rangle,\langle s_2,(0.6,0.1)\rangle,\langle s_3,(0.2,0.8)\rangle,\langle s_4,(0.6,0.1)\rangle,\langle s_5,(0.1,0.6)\rangle\},$$

$$p_2=\{\langle s_1,(0.0,0.8)\rangle,\langle s_2,(0.4,0.4)\rangle,\langle s_3,(0.6,0.1)\rangle,\langle s_4,(0.1,0.7)\rangle,\langle x_3,(0.1,0.8)\rangle\},$$

$$p_3=\{\langle s_1,(0.8,0.1)\rangle,\langle s_2,(0.8,0.1)\rangle,\langle s_3,(0.0,0.6)\rangle,\langle s_4,(0.2,0.7)\rangle,\langle x_3,(0.0,0.5)\rangle\},$$

$$p_4=\{\langle s_1,(0.6,0.1)\rangle,\langle s_2,(0.5,0.4)\rangle,\langle s_3,(0.3,0.4)\rangle,\langle s_4,(0.7,0.2)\rangle,\langle x_3,(0.3,0.4)\rangle\}.$$

The characteristic symptoms for the diseases considered are given as follows,

$$d_1=\{\langle s_1,(0.4,0.0)\rangle,\langle s_2,(0.3,0.5)\rangle,\langle s_3,(0.1,0.7)\rangle,\langle s_4,(0.4,0.3)\rangle,\langle s_5,(0.1,0.7)\rangle\},$$

$$d_2=\{\langle s_1,(0.7,0.0)\rangle,\langle s_2,(0.2,0.6)\rangle,\langle s_3,(0.0,0.9)\rangle,\langle s_4,(0.7,0.0)\rangle,\langle x_3,(0.1,0.8)\rangle\},$$

$$d_3=\{\langle s_1,(0.3,0.3)\rangle,\langle s_2,(0.6,0.1)\rangle,\langle s_3,(0.2,0.7)\rangle,\langle s_4,(0.2,0.6)\rangle,\langle x_3,(0.1,0.9)\rangle\},$$

$$d_4=\{\langle s_1,(0.1,0.7)\rangle,\langle s_2,(0.2,0.4)\rangle,\langle s_3,(0.8,0.0)\rangle,\langle s_4,(0.2,0.7)\rangle,\langle x_3,(0.2,0.7)\rangle\},$$

$$d_5=\{\langle s_1,(0.1,0.8)\rangle,\langle s_2,(0.0,0.8)\rangle,\langle s_3,(0.2,0.8)\rangle,\langle s_4,(0.2,0.8)\rangle,\langle x_3,(0.8,0.1)\rangle\}.$$

In order to find a proper diagnosis, we calculate for each patient $p_i\in P$, where $i\in\{1,\ldots,4\}$, and the symmetric discrimination information measure for IFSs $C_{IFS}^{new}(s(p_i),d_k)$ between patient symptoms and the set of symptoms that are characteristic for each diseases is $d_k\in D$, with $k\in\{1,\ldots,5\}$. Based on Eq. (9), we assign to the *i*th patient the diagnosis whose symptoms have the highest symmetric discrimination information measured from the patient's symptoms. The diagnosed results for the considered patients are given as Table 3.

According to Table 3, the highest similarity score for each patient p_i from possible disease D points out a solution. As before, we know that Al may suffers from malaria; Bob may suffers from stomach problem; Joe may suffers from typhoid; and Ted may suffers from viral fever. We obtained the same results, i.e. the same quality of diagnosis for each patient when looking for the solution by applying the normalized Euclidean distance of Szmidt and Kacprzyk[34,35].

In addition, a comparison between Ye's[22] method, Szmidt and Kacprzyks' method[34,35] and our method for diagnostic results is shown as Table 4. From the Table

4, we know that our proposed method is the same as original study of medical diagnosis by Szmidt and Kacprzyk[34,35]. These final results of the work by Szmidt and Kacprzyk[34,35] have been verified by doctors. Different methods have been applied. Szmidt and Kacprzyk[34,35] used distance as a reference, but our method is based on computing the included angle. Therefore, our proposed method is an alternative method for the medical diagnosis application.

Table 3. The diagnostic results by the proposed approach.

Patients\Diseases	d_1	d_2	d_3	d_4	d_5	Ranking
p_1	0.8386	$\boxed{0.8765}$	0.8147	0.5185	0.4348	$d_2 \succ d_1 \succ d_3 \succ d_4 \succ d_5$
p_2	0.6536	0.4856	0.7839	$\boxed{0.9629}$	0.6570	$d_4 \succ d_3 \succ d_5 \succ d_1 \succ d_2$
p_3	0.7648	0.6749	$\boxed{0.8307}$	0.5363	0.4757	$d_3 \succ d_1 \succ d_2 \succ d_4 \succ d_5$
p_4	$\boxed{0.8751}$	0.8590	0.7706	0.6354	0.5769	$d_1 \succ d_2 \succ d_3 \succ d_4 \succ d_5$

Table 4. A comparison for different similarity measures approaches.

Patients\Diseases	Viral fever	Malaria	Typhoid	Stomach problem	Chest problem
Al	M^{Y}	M^{SK}, M^{NEW}	None	None	None
Bob	None	None	None	M^{SK}, M^{Y}, M^{NEW}	None
Joe	None	None	M^{SK}, M^{Y}, M^{NEW}	None	None
Ted	M^{SK}, M^{Y}, M^{NEW}	None	None	None	None

Szmidt and Kacprzyk[6,35]: As method 1 (M^{SK}); Ye[22]: As method 2 (M^{Y}); Our method: As method 3 (M^{NEW})

5.2. *Example 2—Cancer pattern recognition*

Colorectal cancer occurs frequently in developed countries. About 50% of patients eventually die from the local recurrence and/or distant metastasis within 5 years after curative resection[36]. Therefore, it is important to detect or predict a recurrent or metastasis tumor in the follow-up so that the appropriate therapy can be prescribed to increase the chance of survival.

Colorectal cancer forms initially in the mucosa lining of the bowel. In most cases, the first step in the formation of colorectal cancer is the appearance of polyps. When the abnormal cells within the polyps begin to spread and invade normal tissue, polyps become cancer growths. If no proper treatment is adopted, then the cancer can spread beyond the skin and the underlying tissues of the bowel wall, and it eventually may spread to the distant sites, such as the liver. To describe the condition of the patient, the following four possible outcomes are used: well, recurrence, metastasis and bad (recurrence and metastasis simultaneously). The main treatment for colorectal cancer is the surgical removal of the tumor, and the survival of a patient with colorectal cancer is dependent on four fundamental factors:

(1) The biology of that individual's malignancy,
(2) The immune response to the tumor,
(3) The time in the cancer patient's life history when the diagnosis is made,
(4) The adequacy of the treatment.

The measures of similarity between the IFSs can be used to measure the importance of a feature in a given classification task. Here we illustrate this problem in the context of colorectal cancer diagnosis as quoted by Zhang and Fu[15] and Yusoff et al.[20] to test their similarity measures. This sample shows the association between the key prognostic factors and the outcomes of the patients who are undergoing the follow-up program of colorectal cancer. A patient in a is in a follow-up program, may fall into any of the following states: metastasis, recurrence, bad, and well. If the state of a particular patient can be correctly determined, then the state information can be utilized to choose an appropriate treatment. A physician can subjectively judge the belongingness of each patient to the output classes.

Let C be an attribute set of a patient and the main 5 characters respectively are (C_1): the change of habit and character of stool; (C_2): bellyache; (C_3): ictus sileus; (C_4): chronic sileus and (C_5): anemia. As these characters usually are language variables, for each character, IFSs function established by fuzzy method or probability method, and obtains their character values. Consider a colorectal cancer sample whose 5 characters are quantified as $C = \{\langle C_1,(0.3,0.5)\rangle, \langle C_2,(0.4,0.4)\rangle, \langle C_3,(0.6,0.2)\rangle, \langle C_4,(0.5,0.1)\rangle$, and $\langle C_5,(0.9,0.0)\rangle \}$, and A_1, A_2, A_3 and A_4 are the character sets of the patterns, denoting metastasis, recurrence, bad (Metastasis and recurrence simultaneously) and well, as shown in Table 5.

Using Eq. (9) to compute the similarity measure between the sample C and the patterns (A_1, A_2, A_3 and A_4), we can obtain that

$C_{IFS}^{new}(A_1,C) = 0.9238$, $C_{IFS}^{new}(A_2,C) = 0.9102$, $C_{IFS}^{new}(A_3,C) = 0.5330$ and $C_{IFS}^{new}(A_4,C) = 0.7921$, respectively.

Hence,

$$C_{IFS}^{new}(A_1,C) > C_{IFS}^{new}(A_2,C) > C_{IFS}^{new}(A_4,C) > C_{IFS}^{new}(A_3,C).$$

Thus the ranking order of similarity measure is generated as follows:

$$C_{IFS}^{new}(A_1,C) \succ C_{IFS}^{new}(A_2,C) \succ C_{IFS}^{new}(A_4,C) \succ C_{IFS}^{new}(A_3,C).$$

Since the sample C with the pattern A_1 has the biggest similarity degree, we know that the sample C is similarity A_1 (Metastasis pattern), unlike other patterns. The sample C may suffers from metastasis.

For comparison, the final results of Zhang and Fu[15], Yusoff et al.[20] and our proposed method are shown as Table 6. According to Table 6, we know that our proposed method is the same as type II of Zhang and Fu[15] and Yusoff et al.[20]. Therefore, our proposed method can serve as an alternative method for medical similarity measures based on IFSs.

Table 5. Character sets of the patterns.

Patterns	Characters				
	C_1	C_2	C_3	C_4	C_5
A_1	<0.4, 0.4>	<0.3, 0.3>	<0.5, 0.1>	<0.5, 0.2>	<0.6, 0.2>
A_2	<0.2, 0.6>	<0.3, 0.5>	<0.2, 0.3>	<0.7, 0.1>	<0.8, 0.0>
A_3	<0.1, 0.9>	<0.0, 1.0>	<0.2, 0.7>	<0.1, 0.8>	<0.2, 0.8>
A_4	<0.8, 0.2>	<0.9, 0.0>	<1.0, 0.0>	<0.7, 0.2>	<0.6, 0.4>

Table 6. A comparison of the similarity degrees of different similarity measures between all characters.

	Similarity measures				
	(A_1 vs. C)	(A_2 vs. C)	(A_3 vs. C)	(A_4 vs. C)	Ranking
Zhang and Fu (I) (2006)	0.880 ❶	0.880 ❶	0.490 ❸	0.670 ❷	$(A_1 = A_2) \succ A_4 \succ A_3$
Zhang and Fu (II) (2006)	0884 ❶	0.870 ❷	0.449 ❹	0.671 ❸	$A_1 \succ A_2 \succ A_4 \succ A_3$
Yusoff et al. (2011)	0.850 ❶	0.834 ❷	0.340 ❹	0.644 ❸	$A_1 \succ A_2 \succ A_4 \succ A_3$
Our method	0.9238 ❶	0.9102 ❷	0.5330 ❹	0.7921 ❸	$A_1 \succ A_2 \succ A_4 \succ A_3$

5.3. *Discussion*

From the above two applications for medical pattern recognition, we provide the following descriptions:

(1) In the practical case of medical diagnosis study, the proposed method can provide a useful way to help doctors perform preliminary diagnosis. The proposed method differs from previous methods for medical diagnosis decision-making due to the fact that the proposed method considers the degree of hesitation, unlike other existing approaches. In the future, the proposed method may be a merit of the preliminary diagnosis models for solving the medical diagnosis problem using IFSs.

(2) If the difference between two adjacent ranking scores is very close, then observing Table 3, the suffered outcomes of Ted and Al, the scores of Viral Fever and Malaria are close. Besides, from Table 6, sample C has both A_1 (Metastasis pattern) and A_2 (Recurrence pattern), of which the scores are close in application 2. If it appears in the first and second ranks, doctors need to make an advanced diagnosis with their expert knowledge to evaluate the values of the two IFSs in the uncertain environment to achieve the optimal judgment. Therefore, the purpose of advanced medical diagnosis is to ensure the quality of medical diagnosis.

6. Computer-base Interface

As more and more decisions in real organizational settings are made, applying IFSs into medical diagnosis analysis to deal with imprecision, uncertainty and fuzziness in decision-making may become a popular research topic in the current uncertain environment. Applying IFSs to support doctors can provide a useful way to help the decision analyzer make his/her decisions efficiently.

In this paper, in order to make computing and ranking the results much easier and to increase the recruiting productivity, we have developed an information system called the intuitionistic fuzzy sets medical diagnosis system (IFSMDS). The design architecture of the system is shown in Fig. 4, and the home page is presented in Fig. 5. This prototype system was developed with Visual Basic 6 and ACCESS on N-tier client server architecture. On the IFSMDS, doctors need to key in the patient's name and symptom data. The intuitive values of each patient on each relation between symptoms and diseases are shown as illustrated in Fig. 6. The system can calculate the assessment value of each patient vs. diseases on each symptom. The diagnosed results are shown as Fig. 7. The greatest similarity degree indicates that the patient is more likely to suffer from the corresponding disease.

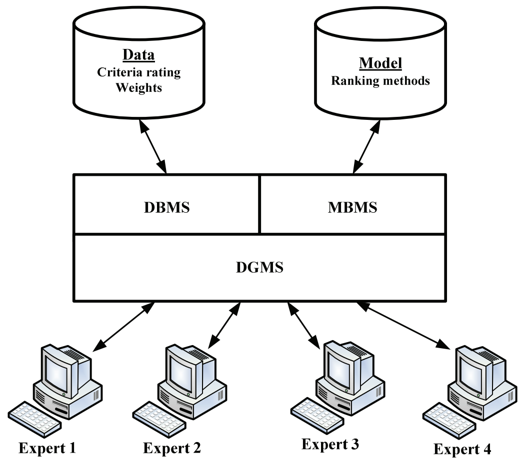

Fig. 4. The architecture of the IFSMDS.

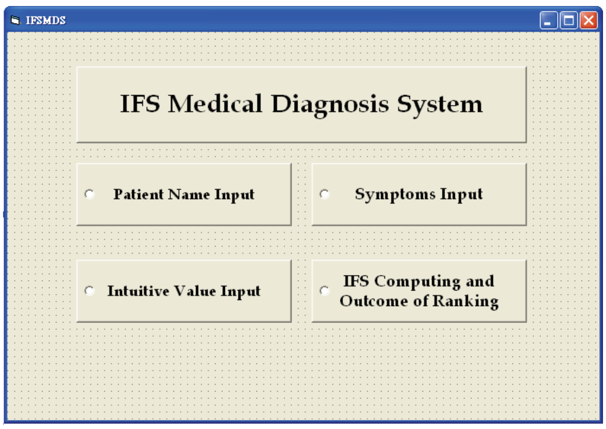

Fig. 5. The functional interface of the IFSMDS.

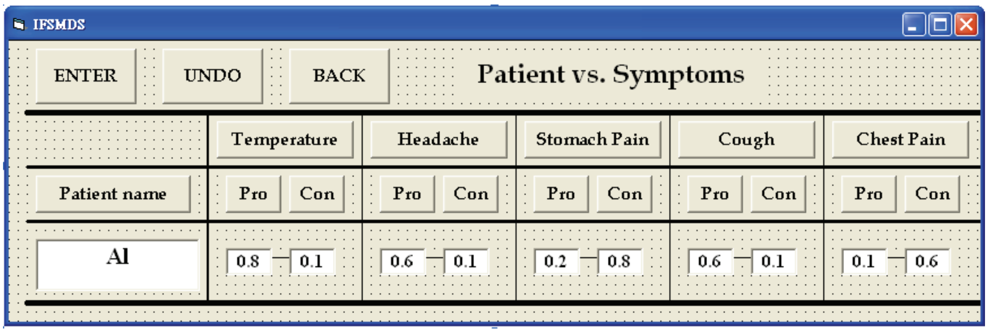

Fig. 6. Input intuitionistic diagnosed value on each symptom for patient.

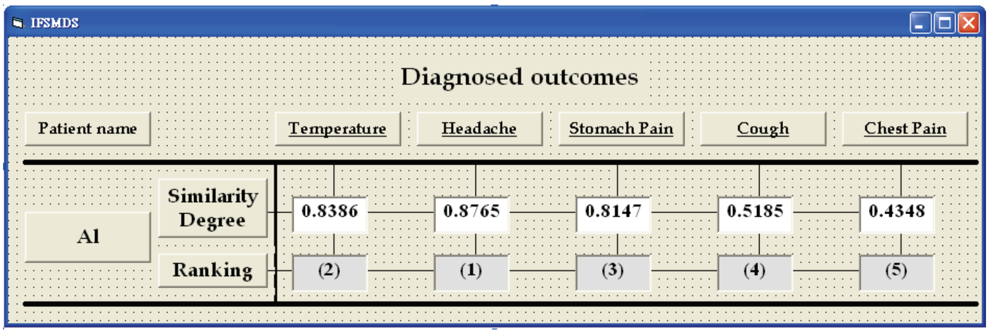

Fig. 7. The outcomes and ranking of the medical diagnosis for patient Al.

7. Conclusions

This paper has developed an integrated similarity measures model for medical pattern recognition applications. The integrated operation involves an improvement of Ye[22], an extension of Hung et al.[33], and the sum of pro and con concept. As expected, the proposed method can be used to cluster the decision samples between some patterns and the sample according to the degree of similarity. In this manner the usefulness of the different similarity measures for IFSs may be compared and analyzed in various medical cases. In addition, in order to make computing and ranking the results much easier and to increase the recruiting productivity, we have developed a computer-based decision support system to effectively aid decision maker with handling IFSs similarity measure problems. In the future, we also hope this research may extend the proposed approach to evaluate and study more other practical cases of medical engineering or management science in an uncertain environment.

Appendix

Proof for (P4): For new cosine similarity measure

For three IFSs with the following expressions

$$A=\left(b_1-x,b_2+y,b_3+x-y\right),\quad B=\left(b_1,b_2,b_3\right),$$

$$C=\left(b_1+z,b_2-w,b_3-z+w\right) \qquad \text{(A.1)}$$

with $A\le B\le C$, we try to prove that

$$\frac{A\bullet B}{\sqrt{A\bullet A}\sqrt{B\bullet B}}\ge\frac{A\bullet C}{\sqrt{A\bullet A}\sqrt{C\bullet C}} \qquad \text{(A.2)}$$

where $b_j\ge 0$ for $j=1,2,3$ and $b_1+b_2+b_3=1$, $x\ge 0$, $z\ge 0$, $y\ge 0$, and $w\ge 0$ such that all components of A, B and C are non-negative. We assume an auxiliary function that

$$
\begin{aligned}
& f(x,y,z,w) \\
& = \left[(b_1 - x) b_1 + (b_2 + y) b_2 + (b_3 + x - y) b_3 \right] \\
& \quad \cdot \sqrt{(b_1 + z)^2 + (b_2 - w)^2 + (b_3 - z + w)^2} \\
& \quad - \left[(b_1 - x)(b_1 + z) + (b_2 + y)(b_2 - w) \right. \\
& \quad \left. + (b_3 + x - y)(b_3 - z + w) \right] \cdot \sqrt{b_1^2 + b_2^2 + b_3^2} \,. \quad \text{(A.3)}
\end{aligned}
$$

Our goal is to prove that

$$f(x,y,z,w) \geq 0 \,. \quad \text{(A.4)}$$

First, we will consider

$$
\begin{aligned}
& f(0,0,0,w) \\
& = \sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2} \cdot \left(b_1^2 + b_2^2 + b_3^2 \right) \\
& - \left[b_1^2 + b_2 (b_2 - w) + b_3 (b_3 + w) \right] \cdot \sqrt{b_1^2 + b_2^2 + b_3^2} \,, \quad \text{(A.5)}
\end{aligned}
$$

to show that $f(0,0,0,w) \geq 0$. We know that $f(0,0,0,0) = 0$ and

$$
\frac{\partial}{\partial w} f(0,0,0,w) = \left(b_1^2 + b_2^2 + b_3^2 \right)
\frac{2w - b_2 + b_3}{\sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2}} + (b_2 - b_3) \cdot \sqrt{b_1^2 + b_2^2 + b_3^2} \,. \quad \text{(A.6)}
$$

We try to show $\frac{\partial}{\partial w} f(0,0,0,w) \geq 0$ that is

$$
(b_2 - b_3) \sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2} + (2w + b_3 - b_2) \sqrt{b_1^2 + b_2^2 + b_3^2} \geq 0 \,. \quad \text{(A.7)}
$$

We divide the problem into two cases: (a) $b_3 \geq b_2$, and (b) $b_3 < b_2$.

If $b_3 \geq b_2$, Eq. (A.7) is equivalent to

$$
(2w + b_3 - b_2)^2 \left(b_1^2 + b_2^2 + b_3^2 \right) - (b_3 - b_2)^2 \left[b_1^2 + (b_2 - w)^2 + (b_3 + w)^2 \right] \geq 0 \,. \quad \text{(A.8)}
$$

We can simplify the left hand side of Eq. (A.8) as $2\left[2b_1^2 + (b_2 + b_3)^2 \right] \left[w^2 + w(b_3 - b_2) \right]$ so case (a) is proved.

If $b_3 < b_2$, Eq. (A.7) is equivalent to

$$
b_2 - b_3 \geq (b_2 - b_3 - 2w) \frac{\sqrt{b_1^2 + b_2^2 + b_3^2}}{\sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2}} \,. \quad \text{(A.9)}
$$

If $\sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2} \geq \sqrt{b_1^2 + b_2^2 + b_3^2}$, then Eq. (A.9) is valid, since $w \geq 0$.

We know that

$$
\sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2} \geq \sqrt{b_1^2 + b_2^2 + b_3^2} \quad \text{(A.10)}
$$

is equivalent to $b_3 + w \geq b_2$. Hence, for $b_3 < b_2$, we further divide them into two sub-cases: $b_3 + w \geq b_2$ and $b_2 > b_3 + w$.

When $b_2 > b_3 + w$, we define an auxiliary function, say $T(w)$, for $0 \leq w \leq b_2 - b_3$;

$$
\begin{aligned}
T(w) &= (b_2 - b_3) \sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2} \\
&\quad + 2w \sqrt{b_1^2 + b_2^2 + b_3^2} - (b_2 - b_3) \sqrt{b_1^2 + b_2^2 + b_3^2} \,. \quad \text{(A.11)}
\end{aligned}
$$

We try to verify that

$$T(w) \geq 0 \,. \quad \text{(A.12)}$$

We have $T(0) = 0$ and

$$
T'(w) = \frac{(2w + b_3 - b_2)(b_2 - b_3)}{\sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2}} + 2\sqrt{b_1^2 + b_2^2 + b_3^2} \,. \quad \text{(A.13)}
$$

We want to show that

$$T'(w) \geq 0 \,. \quad \text{(A.14)}$$

If $2w + b_3 - b_2 \geq 0$, we know that Eq. (A.14) is valid. Then $T(w) \geq 0$ that is Eq. (A.12) is proved.

On the other hand, if $2w + b_3 - b_2 < 0$, to verify Eq. (A.14) is equivalent to

$$
2\sqrt{b_1^2 + b_2^2 + b_3^2} \sqrt{b_1^2 + (b_2 - w)^2 + (b_3 + w)^2} \geq (b_2 - b_3 - 2w)(b_2 - b_3) \,. \quad \text{(A.15)}
$$

We know that Eq. (A.15) is equivalent to

$$
4\left(b_1^2 + b_2^2 + b_3^2 \right) \left(b_1^2 + (b_2 - w)^2 + (b_3 + w)^2 \right) - \left((b_2 - w) - (b_3 + w) \right)^2 (b_2 - b_3)^2 \geq 0 \,. \quad \text{(A.16)}
$$

We can simplify the left hand side of Eq. (A.16) as

$$4b_1^2\left(b_1^2+b_2^2+b_3^2\right)+3\left(b_1^2+b_2^2+b_3^2\right)\left(\left(b_2-w\right)^2+\left(b_3+w\right)^2\right)$$
$$+\left(b_1^2+2b_2b_3\right)\left(\left(b_2-w\right)^2+\left(b_3+w\right)^2\right)+2\left(b_2-w\right)$$
$$\cdot\left(b_3+w\right)\left(b_2-b_3\right)^2>0 \quad \text{(A.17)}$$

From Eq. (A.17), we obtain that Eq. (A.16) is valid so Eqs. (A.15, A.14, A.12) are verified for the case $b_2>b_3+2w$; we thus finish the proof of Eq. (A.7). Up to now, we have show that $f\left(0,0,0,w\right)\geq 0$ for $w\geq 0$ and w satisfies the condition that all components of A , B and C are non-negative.

Second, we will consider

$$f\left(0,y,0,w\right)$$
$$=\left[b_1^2+\left(b_2+y\right)b_2+\left(b_3-y\right)b_3\right]\sqrt{b_1^2+\left(b_2-w\right)^2+\left(b_3+w\right)^2}$$
$$-\left[b_1^2+\left(b_2+y\right)\left(b_2-w\right)+\left(b_3-y\right)\left(b_3+w\right)\right]\sqrt{b_1^2+b_2^2+b_3^2} \quad \text{(A.18)}$$

and then we will prove

$$f\left(0,y,0,w\right)\geq 0\,. \quad \text{(A.19)}$$

We derive that

$$\frac{\partial}{\partial y}f\left(0,y,0,w\right)$$
$$=\left(b_2-b_3\right)\sqrt{b_1^2+\left(b_2-w\right)^2+\left(b_3+w\right)^2}$$
$$+\left(2w+b_3-b_2\right)\sqrt{b_1^2+b_2^2+b_3^2} \quad \text{(A.20)}$$

We know that Eq. (A.20) is exact the same as Eq. (A.7) such that we obtain that $\frac{\partial}{\partial y}f\left(0,y,0,w\right)\geq 0$ together with $f\left(0,0,0,w\right)\geq 0$ so it yields that Eq. (A.19) is verified.

Third, we try to prove that

$$\frac{b_1\left(b_1-x\right)+b_2^2+b_3\left(b_3+x\right)}{\sqrt{b_1^2+b_2^2+b_3^2}\sqrt{\left(b_1-x\right)^2+b_2^2+\left(b_3+x\right)^2}}$$
$$\geq\frac{\left(b_1-x\right)\left(b_1+z\right)+b_2^2+\left(b_3+x\right)\left(b_3-z\right)}{\sqrt{\left(b_1-x\right)^2+b_2^2+\left(b_3+x\right)^2}\sqrt{\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2}}, \quad \text{(A.21)}$$

such that we will try to show that

$$\left[b_1\left(b_1-x\right)+b_2^2+b_3\left(b_3+x\right)\right]\sqrt{\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2}$$
$$\geq\sqrt{b_1^2+b_2^2+b_3^2}\left[\left(b_1-x\right)\left(b_1+z\right)+b_2^2+\left(b_3+x\right)\left(b_3-z\right)\right]$$
$$. \quad \text{(A.22)}$$

We will prove that

$$F\left(z\right)\geq 0 \quad \text{(A.23)}$$

with

$$F\left(z\right)=\sqrt{b_1^2+b_2^2+b_3^2}\sqrt{\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2}$$
$$-\left[b_1\left(b_1+z\right)+b_2^2+b_3\left(b_3-z\right)\right]. \quad \text{(A.24)}$$

We know that

$$\frac{\partial}{\partial z}F\left(z\right)$$
$$=\left(2z+b_1-b_3\right)\frac{\sqrt{b_1^2+b_2^2+b_3^2}}{\sqrt{\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2}}-\left(b_1-b_3\right). \quad \text{(A.25)}$$

Our goal is to show that

$$\frac{\partial F}{\partial z}>0\,. \quad \text{(A.26)}$$

If $b_1-b_3\geq 0$, then we know that

$$\left(2z+b_1-b_3\right)^2\left(b_1^2+b_2^2+b_3^2\right)-\left(b_1-b_3\right)^2$$
$$\cdot\left[\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2\right]$$
$$=2z\left(b_1-b_3\right)\left[\left(b_1+b_3\right)^2+2b_2^2\right]$$
$$+2z^2\left(b_1^2+2b_2^2+b_3^2+4b_1b_3\right)>0\,, \quad \text{(A.27)}$$

such that $\frac{\partial F}{\partial z}>0$ is proved.

We compare $\sqrt{b_1^2+b_2^2+b_3^2}$ and $\sqrt{\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2}$ to derive that

$$\sqrt{b_1^2+b_2^2+b_3^2}\geq\sqrt{\left(b_1+z\right)^2+b_2^2+\left(b_3-z\right)^2} \quad \text{(A.28)}$$

if and only if $b_3 - b_1 \geq z$.
If $b_1 < b_3$ and $b_3 - b_1 < z$, then it follows that

$$\frac{\sqrt{b_1^2 + b_2^2 + b_3^2}}{\sqrt{(b_1 + z)^2 + b_2^2 + (b_3 - z)^2}} \leq 1 \qquad \text{(A.29)}$$

such that based on Eq. (A.29), we find

$$(2z + b_1 - b_3)\frac{\sqrt{b_1^2 + b_2^2 + b_3^2}}{\sqrt{(b_1 + z)^2 + b_2^2 + (b_3 - z)^2}} - (b_1 - b_3)$$

$$\geq (2z + b_1 - b_3) - (b_1 - b_3) = 2z \geq 0. \qquad (30)$$

From the above discussion, we will divide the problem into three cases: (C1) $b_1 - b_3 \geq 0$, (C2) $0 < b_3 - b_1 < z$, and (C3) $b_3 - b_1 \geq z \geq 0$. We already used algebraic method for cases (C1) and (C2), the assertion of Eq. (A.26) is verified. Next, we will use analytical method for case (C3). We define an auxiliary function, say $H(z)$ with

$$H(z) = (2z + b_1 - b_3)\sqrt{b_1^2 + b_2^2 + b_3^2}$$

$$+(b_3 - b_1)\sqrt{(b_1 + z)^2 + b_2^2 + (b_3 - z)^2} \qquad \text{(A.31)}$$

for $0 \leq z \leq b_3 - b_1$. We obtain that

$$H'(z) = 2\sqrt{b_1^2 + b_2^2 + b_3^2} + (b_3 - b_1)$$

$$\cdot\frac{2z + b_1 - b_3}{\sqrt{(b_1 + z)^2 + b_2^2 + (b_3 - z)^2}}. \qquad \text{(A.32)}$$

Owing to

$$4(b_1^2 + b_2^2 + b_3^2)\left[(b_1 + z)^2 + b_2^2 + (b_3 - z)^2)\right]$$

$$-(b_3 - b_1)^2\left[(b_1 + z) - (b_3 - z)\right]^2$$

$$= (b_1 + z)^2\left[4(b_1^2 + b_2^2 + b_3^2) - (b_3 - b_1)^2\right]$$

$$+4(b_1^2 + b_2^2 + b_3^2)b_2^2 + (b_3 - z)^2$$

$$\cdot\left[4(b_1^2 + b_2^2 + b_3^2) - (b_3 - b_1)^2\right]$$

$$+2(b_1 + z)(b_3 - z)(b_3 - b_1)^2 > 0, \qquad \text{(A.33)}$$

it yields that $H'(z) > 0$ with $H(0) = 0$. Hence, for $0 \leq z \leq b_3 - b_1$, we derive

$$H(z) > 0 \qquad \text{(A.34)}$$

Now, we combine the results of Eqs. (A.27, A.30, A.34). We derive that Eq. (A.26) of $\frac{\partial F}{\partial z} > 0$ is proved. From $F(0) = 0$, it follows that our goal of Eq. (A.23) with $F(z) \geq 0$ is verified.

Fourth, we will show that

$$G(x,z) \geq 0 \qquad \text{(A.35)}$$

where

$$G(x,z) =$$

$$\left[b_1(b_1 - x) + b_2^2 + b_3(b_3 + x)\right]\sqrt{(b_1 + z)^2 + b_2^2 + (b_3 - z)^2}$$

$$-\sqrt{b_1^2 + b_2^2 + b_3^2}\left[(b_1 - x)(b_1 + z) + b_2^2 + (b_3 + x)(b_3 - z)\right].$$

$$\text{(A.36)}$$

We find that

$$\frac{\partial}{\partial x}G(x,z) = (b_3 - b_1)\sqrt{(b_1 + z)^2 + b_2^2 + (b_3 - z)^2}$$

$$+(2z + b_1 - b_3)\sqrt{b_1^2 + b_2^2 + b_3^2}. \qquad \text{(A.37)}$$

If we compare Eqs. (A.25) and (A.37), then we know that $\frac{\partial F}{\partial z}$ and $\frac{\partial}{\partial x}G(x,z)$ have the same sign. We have already proved that Eq. (A.26) of $\frac{\partial F}{\partial z} > 0$ is verified. Therefore, it follows that $\frac{\partial}{\partial x}G(x,z) > 0$. We recall that $G(0,z) = F(z) \geq 0$ and then we derive $G(x,z) > 0$.

Based on the above discussion, we show that Eq. (A.2) is proved.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC102-2410-H-241-010-MY2.

References

1. L. A. Zadeh, Fuzzy sets, *Inform. Control* **8**(3) (1965) 338–356.

2. K. T. Atanassov, Intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **20**(1) (1986) 87–96.
3. S. M. Chen, Measures of similarity between vague sets, *Fuzzy Set. Syst.* **74**(2) (1995) 217–223.
4. S. M. Chen, Similarity measures between vague sets and between elements, *IEEE T. Syst. Man. Cyb.* **27**(1) (1997) 153–158.
5. D. H. Hong and C. Kim, A note on similarity measures between vague sets and between elements, *Inform. Science* **115** (1999) 83–96.
6. E. Szmidt and J. Kacprzyk, Distances between intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **114**(3) (2000) 505–518.
7. E. Szmidt and J. Kacprzyk, An application of intuitionistic fuzzy set similarity measures to a multi-criteria decision making problem, *Lect. Notes Comput. Sc.* **4029** (2006) 314–323.
8. D. F. Li and C. T. Cheng, New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions, *Pattern Recogn. Lett.* **23** (2002) 221−225.
9. Z. Liang and P. Shi, Similarity measures on intuitionistic fuzzy sets, *Pattern Recogn. Lett.* **24** (2003) 2687–2693.
10. H. B. Mitchell, On the Dengfeng–Chuntian similarity measure and its application to pattern recognition, *Pattern Recogn. Lett.* **24** (2003) 3101–3104.
11. W. L. Hung and M. S. Yang, Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance, *Pattern Recogn. Lett.* **25** (2004) 1603–1611.
12. P. Grzegorzewski, Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric, *Fuzzy Set. Syst.* **148** (2004) 319–328.
13. W. Liu, New similarity measures between intuitionistic fuzzy sets and between elements, *Math. Comput. Model.* **42** (2005) 61–70.
14. W. Wang and X. Xin, Distance measure between intuitionistic fuzzy sets, *Pattern Recogn. Lett.* **26** (2005) 2063–2069.
15. C. G. Zhang and H. Y. Fu, Similarity measures on three kinds of fuzzy sets, *Pattern Recogn. Lett.* **27** (2006) 1307–1317.
16. Z. S. Xu, Some similarity measures of intuitionistic fuzzy sets and their applications to multiple attribute decision making, *Fuzzy Optim. Decis. Ma.* **6** (2007) 109–121.
17. J. H. Park, J. S. Park, Y. C. Kwun and K. M. Lim, New Similarity Measures on Intuitionistic Fuzzy Sets, *Fuzzy Information and Engineering (ICFIE)*, **ASC 40** (2007) 22–30.
18. W. L. Hung and M. S. Yang, Similarity measures of intuitionistic fuzzy sets based on L_p metric, *Int. J. Approx. Reason.* **46** (2007) 120–136.
19. Z. S. Xu and M. Xia, Distance and similarity measures for hesitant fuzzy sets, *Inform. Sciences* **181** (2011) 2128−2138.
20. B. Yusoff, I. Taib, L. Abdullah and A. F. Wahab, A new similarity measure on intuitionistic fuzzy sets, *International Journal of Computational and Mathematical Sciences* **5**(2) (2011) 70–77.
21. S. Mukherjee and K. Basu, Solution of a class of intuitionistic fuzzy assignment problem by using similarity measures, *Knowl.-Based Syst.* **27** (2012) 170–179.
22. J. Ye, Cosine similarity measures for intuitionistic fuzzy sets and their applications, *Math. Comput. Model.* **53**(1–2) (2011) 91–97.
23. S. K. De, R. Biswas and A. R. Roy, Some operations on intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **114** (2000) 477–484.
24. M. S. Yang and K. L. Wu, A similarity-based robust clustering method, *IEEE T. Pattern Anal.* **26** (2004) 434–448.
25. Y. H. Li, D. L. Olson and Q. Zheng, Similarity measures between intuitionistic fuzzy (vague) sets: a comparative analysis, *Pattern Recogn. Lett.* **28** (2007) 278–285.
26. L. Gau and D. J. Buehrer, Vague sets, *IEEE T. Syst. Man. Cyb.* **23** (1993) 610–614.
27. H. Bustince and P. Burillo, Vague sets are intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **79** (1996) 403-405.
28. K. T. Atanassov, More on intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **33** (1989) 37–46.
29. K. T. Atanassov, *Intuitionistic Fuzzy Sets: Theory and Applications*, Physica-Verlag, Heidelberg, 1999.
30. A. Bhattacharya, On a measure of divergence of two multinomial populations, *Sankhya* **7** (1946) 401–406.
31. G. Salton and M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, New York, 1983.
32. T. Gerstenkorn and J. Manko, Correlation of intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **44** (1991) 39–43.
33. K.-C. Hung, Gino K.Yang, Peter Chu and T.-H. Jin, An enhanced method and its application for fuzzy multi-criteria decision-making based on vague sets, *Comput. Aided Design* **40**(4) (2008) 447–454.
34. E. Szmidt and J. Kacprzyk, Entropy of intuitionistic fuzzy sets, *Fuzzy Set. Syst.* **118** (2001a) 467–477.
35. E. Szmidt and J. Kacprzyk, Intuitionistic fuzzy sets in intelligent data analysis for medical diagnosis, *Lect. Notes Comput. Sc.* **2074** (2001b) 263–271.
36. L. S. Ong, B. Shepherd, L. C. Tong, F. S. Choen, Y. H. Ho, C. L. Tang, Y. S. Ho and K. Tan, The colorectal cancer recurrence support (CARES) system, *Artificial Intelligence Medicine* **11** (1997) 175–188.