A New Similarity Measure on Intuitionistic Fuzzy Sets

Binyamin Yusoff, Imran Taib, Lazim Abdullah and Abd Fatah Wahab

Abstract—Intuitionistic fuzzy sets as proposed by Atanassov, have gained much attention from past and latter researchers for applications in various fields. Similarity measures between intuitionistic fuzzy sets were developed afterwards. However, it does not cater the conflicting behavior of each element evaluated. We therefore made some modification to the similarity measure of IFS by considering conflicting concept to the model. In this paper, we concentrate on Zhang and Fu’s similarity measures for IFSs and some examples are given to validate these similarity measures. A simple modification to Zhang and Fu’s similarity measures of IFSs was proposed to find the best result according to the use of degree of indeterminacy. Finally, we mark up with the application to real decision making problems.

Keywords—Intuitionistic fuzzy sets, similarity measures, multi-criteria decision making.

I. INTRODUCTION

THE theory of fuzzy sets proposed by Zadeh [26] has successfully applied in numerous fields such as engineering, finance, biology, and etc. In fuzzy set theory, the degree of belonging of element to the set is represented by a membership value in the real interval [0, 1] and there exists degree of non-membership which is complementary in nature. From latter point of view, it is true and acceptable that grade of membership and non-membership are complementary. Conversely in Atanassov [1] critical sense, some hesitation degree needs to be introduced (Atanassov [30]) in the concept of Intuitionistic Fuzzy Sets (IFSs). Whist Bustince and Burillo [8] found that this notion coincides with the notion of vague sets proposed by Gau and Buehrer [14], the IFSs make descriptions of the objective world become more realistic, practical, and accurate, making it very promising. Instead of using fuzzy approach, past researchers have studied IFSs to be applied in variety area such as decision making problems [22], medical diagnostics [11] and pattern recognition [12] and seem to be more popular than fuzzy sets in recent years. A similarity measure is used for estimating the degree of similarity between two sets. Based on similarity measures that benefiting to some areas, such as pattern recognition, machine learning, decision making and market prediction, huge methods to measure similarity between fuzzy sets have been proposed and studied in recent years (see [10]; [18]; [21]). For that purposes, other similarity measures for IFSs/vague sets have been proposed recently as a generalization of fuzzy set ([9]; [10]; [13]; [15]; [18]; [19]; [20]; [23]; [27]; [29]). Although enormous studies have been done in measuring similarity, most of them just reflect the difference between degree of membership and degree of non-membership and affect to similarity measure without considering the degree of indeterminacy [29]. Li et al. [19] made a comparative analysis for similarity measures between IFSs/vague sets has found inadequate conditions for similarity measures and was said to be inefficient though it covers the degree of membership and degree of non-membership. Therefore, the degree of indeterminacy was introduced as an effective method to cover such problem of existing similarity measures between IFSs/vague sets ([24]; [25]; [28]; [29]).

Zhang and Fu [28] introduced some new similarity measures for IFSs, fuzzy rough sets and rough fuzzy sets with a new similarity measure for IFSs by considering degree of indeterminacy to improve distinguish precision. Numerical examples given by [28] show the effectiveness of proposed method from previous methods and claimed that all parameters of IFS are fully utilized for measures similarity. However in certain cases, this method is inefficient and need some more modification for better results. In this paper, Zhang and Fu’s similarity measures are reviewed and some examples are given to show the inefficentness for such certain cases. Some modifications of Zhang and Fu’s method have been made and the results were illustrated in decision making problems.

The meaning of conflict can be defined as a state of clash or discord caused by the actual resistance of needs, values and interests. A conflict can be internal (individual conflict) or external (two or more individuals). Conflict as another perspective may be able to explain disagreement amongst individuals, groups, or organizations. From conflicting perspective, each element has two sides of attribute such as positive and negative, bad and good, strong and weak and etc. This concept has been well accepted and authorized by Ying Yang’s theories and become rotundity when the both side turn
into complementary. Ying Yang bipolar logic has been expanding through basic Ying Yang concept. Zhang et al. [27] said that any product can have both good and/or bad aspects. Every relation between two agents or agencies is the equilibrium of conflict and common interests even for a married couple or for two allied countries. These are among a few examples to lead a belief that there exists conflicting in bipolar. The same analogy recommends that conflict also exist in intuition, which involve positive and negative elements.

The rest of this paper is organized as follows. In Section 2, basic notions and definitions of IFSs are reviewed. In Section 3, we point out Zhang and Fu’s method for measures similarity and give some cases to show ineffectiveness of the method. In Section 4, some modification of the method is proposed. Later in section 5, the proposed similarity measures are applied in a problem related decision making. The final section is conclusion.

II. BASIC CONCEPTS OF INTUITIONISTIC FUZZY SETS

In the following, we recall basic notions and definitions of IFSs which can be found in [1] – [4].

Let $X$ be the universe of discourse. An intuitionistic fuzzy set $A$ in $X$ is an object having the form

$$A = \{(x, \mu_A(x), \nu_A(x)) \mid x \in X\}$$

where $\mu_A(x), \nu_A(x) : X \rightarrow [0,1]$ denote membership function and non-membership function, respectively, of $A$ and satisfy $0 \leq \mu_A(x) + \nu_A(x) \leq 1$ for every $x \in X$.

$\mu_A(x)$ is the lowest bound of membership degree derived from proofs of supporting $x$; $\nu_A(x)$ is the lowest bound of non-membership degree derived from proofs of rejecting $x$. It is clear that the membership degree of IF set $A$ has been restricted in $[\mu_A(x), 1-\nu_A(x)]$ which is a subinterval of $[0,1]$.

Obviously, each fuzzy set $A$ in $X$ could be represented as the following IFS:

$$A = \{(x, \mu_A(x), \nu_A(x)) \mid x \in X\}$$

For each IFS $A$ in $X$ we call

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$$

as the intuitionistic index of $x$ in $A$. It is hesitation degree (or degree of indeterminacy) of $x$ to $A$. It is obvious that $0 \leq \pi_A(x) \leq 1$ for each $x \in X$. For example, let $A$ be an IFS with membership function $\mu_A(x)$ and non-membership function $\nu_A(x)$, respectively. If $\mu_A(x) = 0.5$ and $\nu_A(x) = 0.3$, then we have $\mu_A(x) = 1 - 0.5 - 0.3 = 0.2$. It could be interpreted as the degree that the object $x$ belongs to the IFS $A$ is 0.5, the degree that the object $x$ does not belong to the IFS $A$ is 0.3 and the degree of hesitation is 0.2. Thus, IFS $A$ in $X$ can be expressed as

$$A = \{(x, \mu_A(x), \nu_A(x), \pi_A(x)) \mid x \in X\}$$

If $A$ is an ordinary fuzzy set, then $\pi_A(x) = 1 - \mu_A(x) - (1 - \nu_A(x)) = 0$ for each $x \in X$. It means that the third parameter $\pi_A(x)$ can not be casually omitted if $A$ is a general IFS, not an ordinary fuzzy set. Therefore, the representation of IFS should consider all three parameter in calculating the degree of similarity between IFSs.

For $A, B \in IFS(X)$, Atanassov [1] defined the notion of containment as follows:

$$A \subseteq B \iff \mu_A(x) \leq \mu_B(x) \text{ and } \nu_A(x) \geq \nu_B(x), x \in X.$$ 

Similarity measure is a term that measures the degree of similarity between IFSs. As an important content in fuzzy mathematics, similarity measures between IFSs have gained much attention for their wide applications in real world, such as pattern recognition, machine learning, decision making and market prediction.

Dengfeng and Chuntian [12] introduced the following definition of similarity measure between IFSs as follows:

**Definition 1.** A mapping $S : IFSs(X) \times FSs(X) \rightarrow [0,1]$, $IFSs(X)$ denotes the set of all IFSs($X$) in $X = \{x_1, x_2, ..., x_n\}$. $S(A, B)$ is said to be the degree of similarity between $A \in IFSs(X)$ and $B \in IFSs(X)$, if $S(A, B)$ satisfies the properties condition – (P1-P5)

- **P1:** $S(A, B) \in [0,1]$.
- **P2:** $S(A, B) = 1 \iff A = B$,
- **P3:** $S(A, B) = S(B, A)$,
- **P4:** $S(A, C) \leq S(A, B)$ and $S(A, C) \leq S(B, C)$
  
  if $A \subseteq B \subseteq C, C \in IFSs(X)$,
- **P5:** $S(A, B) = 0 \iff A = \Phi$ and $B = \bar{A}$, or $A = \bar{B}$ and $B = \Phi$.

But this definition has some limitations. So, Mitchell [20] gave a simple modification of it by replacing (P2) with a strong version (P2’) as follows.

**P2’:** $S(A, B) = 1 \iff A = B$.

This definition proved to be more reasonable than Dengfeng and Chuntian’s (Mitchell [20]).
III. ZHANG-FU’S SIMILARITY MEASURES

Zhang and Fu [28] developed some similarity measures between three fuzzy sets that are IFSs, FRSSs and RFSs. They stated that IFSs, FRSSs and RFSs were L-fuzzy sets with L being a special fuzzy lattice, then, defined the similarity measure between L-fuzzy sets and its elements. Here, we refer all similarity measures proposed by Zhang and Fu [28] as similarity measures between IFSs and notation used in (Zhang and Fu [28]) is changed so that it consistent with those in this paper. Zhang and Fu [28] defined similarity measures of IFSs as follow:

\begin{definition}
Let $A, B \in IFS(X)$, $X = \{x_1, x_2, \ldots, x_n\}$. If $V_A(x) = [\mu_A(x), 1 - \nu_A(x)]$, $V_B(x) = [\mu_B(x), 1 - \nu_B(x)]$, are the IFS values of $x$ in $A$ and $B$ respectively. Then the similarity degree of $A$ and $B$ can be evaluated by the function $S$.

$$S_{ZF}(A, B) = \frac{1}{n} \sum_{i=1}^{n} MV(V_A(x), V_B(x))$$

where $MV(x, x') = \frac{1}{2} \left( \frac{|\mu_A(x) - \mu_B(x)| + |1 - \nu_A(x) - (1 - \nu_B(x))|}{2} \right)$

We can see that the above similarity measure only considering $\mu(x)$ and $\nu(x)$. In fact, $S_{ZF}(A, B)$ similar to Zhizhen and Pengfei’s similarity measure, $S^*_{ZF}(A, B)$ (refer to Zhizhen and Pengfei [29]). Although in most cases the similarity measure $S_{ZF}(A, B)$ or $S^*_{ZF}(A, B)$ gives intuitively satisfying results, there are some cases in which this is not true. The following example shows one such case.

Example 1: Assume that there are two alternatives denoted with IFSs in $X = \{x_1, x_2, x_3\}$. Two alternatives $A_1$ and $A_2$ are denoted as follows:

$A_1 = \{(x_1, 0.2, 0.6), (x_2, 0.2, 0.6), (x_3, 0.2, 0.5)\}$

$A_2 = \{(x_1, 0.4, 0.6), (x_2, 0.2, 0.6), (x_3, 0.3, 0.3)\}$

Assume that a reference $B = \{(x_1, 0.3, 0.7), (x_2, 0.3, 0.5), (x_3, 0.1, 0.4)\}$ is given. Then, it is obvious that $S_{ZF}(A_1, B) = S_{ZF}(A_2, B)$. So, the alternatives $A_1$ and $A_2$ cannot be differentiated using the above method. Hence, we cannot obtain correct results.

Then, by considering degree of indeterminacy (hesitation degree) Zhang and Fu [28] defined a new similarity measure between IFSs as follow:

\begin{definition}
Let $A, B \in IFS(X)$, $X = \{x_1, x_2, \ldots, x_n\}$. If $V_A(x) = [\mu_A(x), 1 - \nu_A(x)]$, is the fuzzy values of $x$ in the IFS $A$ and $V_B(x) = [\mu_B(x), 1 - \nu_B(x)]$, is the fuzzy values of $x$ in the IFS $B$. Then the degree of similarity between the IFS $A$ and $B$ can be evaluated by the function $S$.

$$S_{ZF}(A, B) = \frac{1}{n} \sum_{i=1}^{n} MV(V_A(x_i), V_B(x_i))$$

where

$$MV(x, x') = \frac{1}{2} \left( \frac{|\delta_A(x) - \delta_B(x)| + |\mu_A(x) - \mu_B(x)|}{2} \right)$$

where $\delta_A(x) = \mu_A(x) + (1 - \mu_A(x))\nu_A(x)$

According to [28], the larger the value of $\pi_A(x)$, the more the degree of unknown for $x$. Especially, if $\pi_A(x) = 1$, we know nothing for $x$; if $1 - \nu(x) = \mu(x)$, then IFS $A$ is a fuzzy set; if $1 - \nu(x) = \mu(x) = 1$ (or 0), then $A$ is a common set and $x \in A$ (or $x \notin A$). The priori knowledge is considered when defining $\delta(x)$ and $\pi(x)$, it can be interpreted by the voting model. A fuzzy value $[0.4, 0.8]$ can be interpreted as “the vote for resolution as above, besides 4 against, it is possible that there is 0.4 $\times$ 4 favor in the 4 abstention”. Similarly, $\pi(x) = 0.2 + 0.2(1 - 0.4 - 0.2) = 0.28$ can be interpreted as “considering the vote for resolution as above, besides 4 in favor, it is possible that there is 0.4 $\times$ 4 favor in the 4 abstention”. In most cases the similarity measure $S_{ZF}(A, B)$ gives intuitively satisfying results, there are situation in which this is not true. The following example illustrates one such case.

Example 2: Assume that there are two alternatives $A_1$ and $A_2$ denoted with IFSs in $X = \{x_1, x_2, x_3\}$. Two alternatives are denoted as follows:

$A_1 = \{(x_1, 0.2, 0.4), (x_2, 0.2, 0.4), (x_3, 0.2, 0.5)\}$

$A_2 = \{(x_1, 0.3, 0.3), (x_2, 0.2, 0.4), (x_3, 0.3, 0.3)\}$

Assume that a sample $B = \{(x_1, 0.3, 0.3), (x_2, 0.3, 0.5), (x_3, 0.1, 0.6)\}$ is given. Then, it is obvious that $S_{ZF}(A_1, B) = S_{ZF}(A_2, B) = 0.92$. So, the patterns cannot be differentiated using the above Zhang and Fu’s method. Hence, we cannot obtain correct recognition results.

IV. NEW SIMILARITY MEASURES FOR INTUITIONISTIC FUZZY SETS

In order to deal with this problem, we modified Zhang and Fu’s method by propose a new similarity measure between IFSs by considering $1 - \delta(x) - \pi(x)$ (we state that as the remainder of hesitation degrees of IFSs) in Zhang and Fu’s definition as follows:
Definition 4. Let $A, B \in \text{IFS}(X), X = \{x_1, x_2, ..., x_n\}$. If $V_A(x) = \{\mu_A(x), 1 - \nu_A(x)\}$ is the fuzzy values of $x$ in the IFS $A$ and $V_B(x) = \{\mu_B(x), 1 - \nu_B(x)\}$ is the fuzzy values of $x$ in the IFS $B$. Then the degree of similarity between the IFS $A$ and $B$ can be evaluated by the function $S_{ZF(mod)}$.

$$S_{ZF(mod)}(A, B) = \frac{1}{n} \sum_{x \in X} |\delta_A(x) - \delta_B(x)|$$

where $\delta_A(x) = \mu_A(x) + (1 - \mu_A(x) - \nu_A(x))\mu_A(x)$ and $\beta_A(x) = 1 - \delta_A(x) - \alpha_A(x)$.

We introducing another parameter $\beta_A(x)$ to represent remainder of hesitancy degree where $\delta_A(x) + \alpha_A(x) + \beta_A(x) = 1$ describing the whole information in IFSs. This parameter should take into consideration as an alternative to differentiate the result precisely.

From Definition 4, we obtained the following Theorem 1 and 2.

Theorem 1. $S_{ZF(mod)}(A, B)$ is a degree of similarity in the Mitchell’s sense between two IFSs $A$ and $B$ in $X = \{x_1, x_2, ..., x_n\}$.

Proof. Obviously, $S_{ZF(mod)}(A, B)$ satisfies (P1) and (P3). As to (P2‘) and (P4), we give the following proof.

(P2‘): It is obvious that $\delta_A(x) = \delta_B(x), \alpha_A(x) = \alpha_B(x)$ and $\beta_A(x) = \beta_B(x)$ from Eq. (3). Therefore $A = B$.

(P4): Since $A \subseteq B \subseteq C$, we have $\delta_A(x) \leq \delta_B(x) \leq \delta_C(x), \alpha_A(x) \geq \alpha_B(x) \geq \alpha_C(x)$ and $\beta_A(x) \geq \beta_B(x) \geq \beta_C(x)$ for any $x \in X$. Then we have

$$|\delta_A - \delta_B|/2 + |\alpha_A - \alpha_B|/2 + |\beta_A - \beta_B|/2$$

$$\leq |\delta_A - \delta_C|/2 + |\alpha_A - \alpha_C|/2 + |\beta_A - \beta_C|/2$$

$$= 1/2n(|\delta_A - \delta_C| + |\alpha_A - \alpha_C| + |\beta_A - \beta_C|)$$

So we have

Therefore, $S_{ZF(mod)}(A, B) \geq S_{ZF(mod)}(A, C)$. □

In the similar way, it is easy to prove $S_{ZF(mod)}(B, C) \geq S_{ZF(mod)}(A, C)$.

Theorem 2. Assume that $S_{ZF(b)}(A, B)$ and $S_{ZF(mod)}(A, B)$ are given in Eq. (2) and Eq. (3), respectively. Then we have $S_{ZF(b)}(A, B) \geq S_{ZF(mod)}(A, B)$.

Proof. Let $x_i = X$. Since

$$|\delta_A - \delta_B|/2 + |\alpha_A - \alpha_B|/2$$

$$\leq |\delta_A - \delta_C|/2 + |\alpha_A - \alpha_C|/2 + |\beta_A - \beta_B|/2,$$

and $1/2n(|\delta_A - \delta_B| + |\alpha_A - \alpha_B|)$

$$\leq 1/2n(|\delta_A - \delta_C| + |\alpha_A - \alpha_C| + |\beta_A - \beta_B|)$$

We have

$$\geq 1 - 1/2n(|\delta_A - \delta_C| + |\alpha_A - \alpha_C| + |\beta_A - \beta_B|)$

Therefore, $S_{ZF(b)}(A, B) \geq S_{ZF(mod)}(A, B)$. □

Example 3: We consider the two alternatives $A_i$ and $A_j$ and the reference $B$ as discussed in Example 2. Then, using Eq. (2), we find $S_{ZF(b)}(A_i, B) = 0.8762$ and $S_{ZF(mod)}(A_j, B) = 0.8802$ and thus obtain the satisfying result $S_{ZF(b)}(A_i, B) \leq S_{ZF(mod)}(A_j, B)$. Hence, we obtain correct results.

V. NUMERICAL EXAMPLE

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 used by [28] 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 the colorectal cancer. The patient, who is in the 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 decided, then the state information can be utilized to choose an appropriate treatment. A physician can subjectively judge the belongingness of each patient in the output classes.

Let $A$ be an attribute set of a patient and the main 5 attributes (the change of habit and character of stool, bellyache, ictus sileus, chronic sileus, anemia) used to
quantify the attribute, respectively denoted a, b, c, d, e. As these characters usually are language variables, for every character, IFSs function established by fuzzy method or probability method, and obtains their character values. Let a colorectal cancer patient whose 5 attributes quantify as $A = \{[0.3, 0.5], [0.4, 0.6], [0.6, 0.8], [0.5, 0.9], [0.9, 1]\}$, $A_1, A_2, A_3$ and $A_4$ are the attribute sets of the samples denoted metastasis, recurrence, bad and well, shown in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Attribute Sets of the Sample</th>
<th>A_1</th>
<th>A_2</th>
<th>A_3</th>
<th>A_4</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
</tr>
<tr>
<td>b</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
</tr>
<tr>
<td>c</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
</tr>
<tr>
<td>d</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
</tr>
<tr>
<td>e</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
<td>[0,1]</td>
</tr>
</tbody>
</table>

Using the method in Definition 1, Definition 2 and proposed modified method (Definition 4), we compute the similarity measure between the patient $A$ and the sample $A_1, A_2, A_3$ and $A_4$: $A = \{[0.3, 0.5], [0.4, 0.6], [0.6, 0.8], [0.5, 0.9], [0.9, 1]\}$. For comparison, we show the final results of the above similarity measures as in Table 2.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{Zhang}(A, A_1)$</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>$S_{Zhang}(A, A_2)$</td>
<td>0.90</td>
<td>1.00</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>$S_{Zhang}(A, A_3)$</td>
<td>0.50</td>
<td>0.50</td>
<td>1.00</td>
<td>0.50</td>
</tr>
<tr>
<td>$S_{Zhang}(A, A_4)$</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
</tbody>
</table>

From the above result, since $S(A, A_1) = S(A, A_2)$, we could not classify $A$ as metastasis patient or recursive patient in Definition 1. It shows the similarity measure $S_{Zhang}(A, B)$ could not distinguish $A$ is similar to $A_1$ or $A_2$. While using the method in Definition 3 and Definition 4, A could be distinguished correctly. But in some cases, Definition 3 cannot work well and differentiate $A$ accurately. From the above result, we can say that mode $A$ is supposed to be a $A_1$ (metastasis patient). It shows that similarity measures Definition 4 could improved method in Definition 3 in case such example mentioned above and also could improve distinguish precision. Therefore, enhanced the capability of classification and make a decision correctly.

### VI. CONCLUSION

The similarity measures between IFSs have been developed by some researchers. Although some existing similarity measures provide an effective way to deal with IFSs in some cases, it obtains unreasonable results using those measures in some cases. Therefore, a new similarity measure was proposed to differentiate different IFSs and a modification of Zhang-Fu similarity measure was made to overcome the drawbacks of some existing methods that dealing with problems in an effective and reasonable way.

### REFERENCES