Expert Systems With Applications 60 (2016) 62–80

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

# An activation detection based similarity measure for intuitionistic fuzzy sets

Shing-Chung Ngan*

*Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong*

ARTICLE INFO

*Article history:*
Received 12 January 2016
Revised 29 April 2016
Accepted 30 April 2016
Available online 30 April 2016

*Keywords:*
Fuzzy sets
Intuitionistic fuzzy sets
Similarity measure
Pattern recognition

ABSTRACT

Intuitionistic fuzzy sets (IF-sets), with mechanisms to represent both the degree of membership and hesitancy of a given entity with respect to a concept under consideration, have been proven to be a useful extension to Zadeh's fuzzy set theory. Noteworthy efforts by various researchers have been devoted to defining a robust similarity measure for a given pair of IF-sets, as we often need to quantify the similarity between given entities in application domains ranging from medical diagnosis to multiple criteria decision making. These efforts have shown that it is highly non-trivial to construct a truly robust IF-set similarity measure with easy-to-understand interpretations. In this article, grounded on native concepts from activation detection in medical image analysis, a model for determining the degree of similarity between IF-sets is proposed. An IF-set similarity measure (termed the activation detection based similarity measure) is then systematically built from this model. We show that the proposed measure produces results that are intuitively appealing, easy to understand, and can be robustly interpreted. Moreover, we demonstrate that the proposed measure obeys standard conventions regarding set definition in the classical setting, and is equivalent to the Jaccard's similarity measure as we transition from the intuitionistic fuzzy setting to the classical setting. The source code of the numerical implementation of the proposed measure is available from the author upon request.



## 1. Introduction

As a useful extension to Zadeh's fuzzy sets (Zadeh, 1965), intuitionistic fuzzy sets (IF-sets, or equivalently IFS) were first introduced and further developed in a series of papers by Atanassov (1986, 1989, 1994a, 1994b). The concept of vague sets was separately proposed by Gau and Buehrer (1993) and was later determined to be equivalent to IFS (Bustince & Burillo, 1996). Elements of ideas concerning IFS could also be found in Narin'yani (1980). A prototypical intuitionistic fuzzy set $A$ is characterized by two functions $\mu_A$ and $\nu_A$, assigning to any given element $s$ from the universe of discourse the degree of membership $\mu_A(s)$ and the degree of non-membership relative $\nu_A(s)$ to the IFS respectively. As it offers a practical mechanism to represent both the degree of membership (via $\mu_A(s)$) and hesitancy (via $\pi_A(s) \equiv 1 - \mu_A(s) - \nu_A(s)$), IFS proves to be a versatile tool for modeling real-life situations and has been applied to a wide range of domains, such as multiple criteria decision making (Chen, 2013; Hong & Choi, 2000; Liu & Wang, 2007; Szmidt & Kacprzyk, 2002), clustering (Pelekis, Lakovidis, Kotsifakos, & Kopanakis, 2008; Xu et al., 2008) and pattern recognition (Hung & Yang, 2008).

In addition to its contributions in terms of applications to various scientific domains, IFS itself has also been undergoing intense theoretical developments in the past two to three decades. A monograph by Atanassov (2012) has provided a concise and contemporary survey of the fundamental results in IFS. These fundamental results include, for example, geometric interpretations of IF-sets (e.g., Antonov (1995); Atanassova (2010); Danchev (1995); Szmidt and Kacprzyk (2004)), operators on and relations in IF-sets (e.g., Atanassov (2001, 2010b, 2010c, 2010d); Atanassov and Ban (2000); Burillo and Bustince (1995a, 1995b); Cornelis, Deschrijver, and Kerre (2004); De, Biswas, and Roy (2000); Parvathi, Vassilev, and Atanassov (2009)), and norms over IF-sets (e.g., Atanassov, (2010a); Tanev, (1995)). Meanwhile, with theoretical results in fuzzy set-theoretic measures as a basis (Butnariu & Klement, 1993; Wang & Klir, 1995), substantive works have been performed in developing intuitionistic fuzzy set-theoretic measures (e.g., see Ban (2006); Ban and Gal (2001)), leading to extensions such as an information measure on intuitionistic fuzzy sets, and notions such as a measurable entropy on intuitionistic fuzzy dynamic systems (Ban, 2006). Active research also takes place in generalizing IFS, such as interval-valued intuitionistic fuzzy sets (e.g.,

* Corresponding author. AC-1 P6603, 83 Tat Chee Avenue, City University of Hong Kong, Kowloon Tong, Kowloon, Hong Kong. Fax: (852) 3442 0172.
*E-mail address:* scngan@cityu.edu.hk

http://dx.doi.org/10.1016/j.eswa.2016.04.037
0957-4174/

see Atanassov (1994b); Atanassov and Gargov (1989); Chen (2015); Düğenci (2016) and the references therein).

Among the various types of measures for IFS, similarity measure is a category of particular interest, since many machine learning techniques require the quantification of similarity of given objects as a key ingredient in achieving their computation objectives – e.g., in hierarchical cluster analysis (Chen, Xu, & Xia, 2013) and in classification (Vlachos & Sergiadis, 2007). Thus, there have been persistent research activities in developing similarity measures for IFS, with the work of Szmidt and Kacprzyk (2000) among the early efforts. Specifically, Szmidt & Kacprzyk extended several frequently used distance measures such as the Euclidean distance and the Hamming distance to the IFS setting by manipulating with the values $(\mu_A(s) - \mu_B(s))$, $(\nu_A(s) - \nu_B(s))$, and $(\pi_A(s) - \pi_B(s))$, where $A$ and $B$ are the IFSs under consideration. This resulted in, for example, a normalized Hamming distance between $A$ and $B$:

$$d(A, B) \equiv \frac{1}{2n}\sum_{i=1}^{n}(|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|) \quad (1)$$

with the universe of discourse $U \equiv \{x_1, x_2, ..., x_n\}$. Meanwhile, Li and Cheng (2002) introduced alternative similarity measures with formulations based on $(\varphi_A(s)\text{-}\varphi_B(s))$, where $\varphi(s)$ is the median value of the interval $[\mu(s), \psi(s)]$ (with $\psi(s) \equiv 1-\nu(s)$), leading to the formula

$$S_d^p(A, B) \equiv 1 - \frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}(\phi_A(x_i) - \phi_B(x_i))^p} \quad (2)$$

with $1 \le p < \infty$ being some fixed constant. Liang and Shi (2003) and Mitchell (2003) pointed out that Li & Cheng's measures could lead to counter-intuitive outcomes in some cases, while Wang and Xin (2005) also identified problems with the approach of Szmidt & Kacprzyk. Thus, new similarity measures were separately proposed in these three papers in their efforts to mitigate the observed deficiencies. For example, Liang & Shi introduced additional ingredients such as the lengths and the median values of certain sub-intervals to their similarity measure formula, and Mitchell postulated the similarity $S_{mod}(A,B)$ between $A$ and $B$ as the average of $S(\mu_A,\mu_B)$ and $S(\psi_A,\psi_B)$, where

$$S(\mu_A, \mu_B) \equiv 1 - \frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)|^p} \quad (3)$$

and $S(\psi_A,\psi_B)$ was similarly defined. Both Hung and Yang (2004) and Grzegorzewski (2004) proposed IFS similarity measure based on the Hausdorff distance $H(I_A(s),I_B(s))$ between intervals $I_A(s) \equiv [\mu_A(s), \psi_A(s)]$ and $I_B(s) \equiv [\mu_B(s), \psi_B(s)]$, leading e.g., to formulas like

$$S_l(A, B) \equiv 1 - d_H(A, B) \quad (4)$$

with

$$d_H(A, B) \equiv \frac{1}{n}\sum_{i=1}^{n}H(I_A(x_i), I_B(x_i))$$

and

$$H(I_A(x_i), I_B(x_i)) \equiv \max\{|\mu_A(x_i) - \mu_B(x_i)|, |\psi_A(x_i) - \psi_B(x_i)|\}.$$

Subsequently, Chen (2007) pointed out some problems in the Hausdorff measure approach by demonstrating some counter-examples. To extend their previous work, Hung and Yang (2007a) then postulated several similarity measures induced by the $L_p$ metric, with the Hausdorff measure approach becoming a special case. Li, Olson, and Qin (2007) and Xu and Chen (2008) respectively gave comprehensive overview and comparative analysis of similarity measures for IFS and highlighted the counter-intuitive cases for the various measures they analyzed.

More recently, Ye (2011) defined a cosine similarity measure for IFS by extending an analogous formula for the fuzzy sets:

$$C_{IFS}(A, B) \equiv \frac{1}{n}\sum_{i=1}^{n}\frac{\mu_A(x_i)\cdot\mu_B(x_i) + \nu_A(x_i)\cdot\nu_B(x_i)}{\sqrt{\mu_A^2(x_i) + \nu_A^2(x_i)}\cdot\sqrt{\mu_B^2(x_i) + \nu_B^2(x_i)}} \quad (5)$$

Hwang, Yang, Hung, and Lee (2012), meanwhile, proposed a new IFS similarity measure involving fuzzy measure-theoretic concepts via the Sugeno integral. Last but not least, Boran and Akay (2014) recently introduced a two-parameter IFS similarity measure with the parameters $p$ and $t$ representing the $L_p$ norm and the uncertainty level respectively:

$$S_E^p(A, B) \equiv 1 - D(A, B) \quad (6)$$

where

$$D(A, B) \equiv \sqrt[p]{\frac{1}{2n(t+1)^p}} \times \sqrt[p]{\sum_{i=1}^{n}\{|t\cdot(\mu_A(x_i) - \mu_B(x_i)) - (\nu_A(x_i) - \nu_B(x_i))|^p + |t\cdot(\nu_A(x_i) - \nu_B(x_i)) - (\mu_A(x_i) - \mu_B(x_i))|^p\}}$$

Furthermore, Boran & Akay performed a comparative analysis of their two-parameter measure against the measures previously proposed by Li & Cheng, Mitchell, Liang & Shi, Hung & Yang and Ye respectively.

In addition to the above-mentioned research articles, a monograph by Szmidt (2014) has provided a highly informative survey and in-depth analysis of the major IFS similarity measures in the literature. The ideas of two-terms (i.e., membership and non-membership values) and three-term (i.e., membership, non-membership and hesitancy values) representations of IF-sets have been extensively discussed and employed in Szmidt (Szmidt, 2014) to characterize existing IFS distance and similarity measures. Further, a detailed analysis of existing works (e.g., Bustince and Burillo (1995); Gerstenkorn and Manko (1991); Hong and Hwang (1995); Hung and Wu (2002, 2007b); Szmidt and Kacprzyk (2010); Zeng and Li (2007)) concerning the correlation of IF-sets, a measure closely related to similarity and distance, has also been provided in Szmidt (2014), again in terms of the two-terms and three-terms representations.

Based on the above bird's eye view of the research landscape of IFS similarity measures and of the noteworthy efforts by the above-mentioned and other researchers, one could reasonably conclude that (i) it is highly non-trivial to construct a truly robust IFS similarity measure, and that (ii) there is a need of more research in identifying an easier-to-understand similarity measure for IFS. (i) is indicative by the identification of counter-intuitive examples by researchers for the various existing measures. Meanwhile, (ii) can be supported by the observation that most of the existing measures, when proposed, were first and foremost postulated at the "formula" level (e.g., Eqs. (1–6)). While their respective authors interpreted these measures subsequently, it could be said that the interpretations were done in a post-hoc manner, and because of that, the underlying native effects of these measures on IF-sets were arguably not easy to be fully grasped and understood. It is likened to a scenario in mechanical science, in which the overall