Knowledge-Based Systems xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Similarity measure for vague sets based on implication functions

Q1 Wenyi Zeng [a,*], Yibin Zhao [a,b], Yundong Gu [c]

Q2 [a] College of Information Science and Technology, Beijing Normal University, Beijing 100875, PR China
[b] Department of Basic Courses, Institute of Disaster Prevention Science and Technology, Yanjiao, Beijing 065201, PR China
[c] School of Mathematic and Physics, North China Electric Power University, Beijing 102206, PR China

ARTICLE INFO

Article history:
Received 10 April 2015
Revised 10 October 2015
Accepted 21 November 2015
Available online xxx

Keywords:
Vague set
Fuzzy set
Similarity measure
Implication function
Pattern recognition

ABSTRACT

The similarity measure of vague sets is an important index in intelligent systems. Based on the implication function, this paper investigates the similarity measure of vague sets (or elements) and proposes new formulas to calculate the similarity measure of vague sets (or elements). A comparison with the existing similarity measures for vague sets and two applications show that our proposed method is reasonable and valid.

1. Introduction

Since Zadeh [29] first introduced fuzzy set theory, many approaches and theories addressing imprecision and uncertainty have been proposed. Some of these theories, such as Atanassov's intuitionistic fuzzy sets theory [1] and Zadeh's generalized theory of uncertainty (GTU) [30], are extensions of classic fuzzy set theory and treat uncertainty and information from a much broader perspective. Another well-known generalization of an ordinary fuzzy set is based on vague sets and was introduced by Gau and Buehrer [15] in 1993. Since then, the theory of vague sets has attracted the attention of many researchers seeking to address imperfectly defined facts and data, including situations with imprecise knowledge. Some authors have investigated the topic and obtained meaningful conclusions. For example, Chen [8] used vague sets to analyze fuzzy system reliability, Dug and Choi [13] and Kuo et al. [19] investigated multi-criteria fuzzy decision-making based on vague sets, Ye [28] proposed an improved method of multi-criteria fuzzy decision-making based on vague sets, Demirci and Eken [11] introduced the theory of vague complement operation, and Gottwald [16] investigated mathematical fuzzy logic as a tool to treat vague information, Feng et al. [14] proposed a vague-rough set approach for extracting knowledge under uncertain environment. As a powerful tool for describing imprecise data, vague sets theory has been extensively applied in many fields such as pattern recognition, machine learning, fuzzy decision-making and so on. Moreover, some researchers have investigated the connection between intuitionistic fuzzy sets and vague sets. For more details, please see [4].

The similarity measure of fuzzy sets is an important component of fuzzy set theory. The similarity measure indicates the degree of similarity between two fuzzy sets. Wang [25] first proposed the concept of the similarity measure in fuzzy set theory and computational formula for that purpose. Since then, the similarity measure of fuzzy sets has been further investigated and extensively applied in many fields such as fuzzy clustering, image processing, fuzzy reasoning, and fuzzy neural networks [9,25]. For example, Balopoulos et al. [2] investigated similarity measures for fuzzy operators, Hung and Yang [17] investigated the J-divergence of intuitionistic fuzzy sets and its application to pattern recognition, and Zeng and Guo [31] investigated the similarity measure of interval-valued fuzzy sets, Wu and Mendal [26] investigated the uncertainty measures for interval type-2 fuzzy sets. Using these applications of the similarity measure, some researchers extended it to vague sets theory. The similarity measure of vague sets indicates the vague sets' degree of similarity. For example, Chen [6,7] investigated the similarity measure and the weighted similarity measure between vague sets and between elements, respectively, Dug and Chul [12] improved Chen's [7] similarity measure expressions and proposed a new method to calculate the similarity measure between vague sets in 1999, Li et al. [21] investigated the distance between vague sets and proposed an improved similarity measure for vague sets, Xu and Wei [27] applied the similarity measure between vague sets in fuzzy classification, Zhang and Jiang [34] investigated the non-probabilistic entropy of a vague set, Zhang et al. [33]

* Corresponding author. Tel.: +86 10 58805665; fax: +86 10 58808313.
E-mail address: zengwy@bnu.edu.cn (W. Zeng).

http://dx.doi.org/10.1016/j.knosys.2015.11.015
0950-7051/© 2015 Published by Elsevier B.V.

investigated the framework for comparing two interval sets by inclusion measures.

It is well known that fuzzy implication functions play a fundamental role in fuzzy logic, approximate reasoning and applications such as fuzzy control, fuzzy relational equation, fuzzy DI-subsethood (inclusion) measure and image processing. Thus the implication functions have been extensively studied by many researchers in theoretical and practical applications. For example, Bustince [5] investigated the indicator of inclusion grade for interval-valued fuzzy sets based on implication; De Baets and Kerre [10] investigated fuzzy inclusion and its inverse problem; Mas et al. [22] studied the law of importation for two types of implications; Pei [23] investigated the unified full implication inference algorithms of fuzzy reasoning, Jin et al. [18] investigated certainty rule base and its inference method; and Zhou et al. [35] characterized intuitionistic fuzzy rough sets based on intuitionistic fuzzy implications, Beliakov et al. [3] investigated the properties relating to consensus measures and proposed two general models built component-wise from aggregation functions and fuzzy implications, Zhai et al. [32] investigated the semantical and syntactical characteristics of fuzzy decision implications. Until now, the existing similarity measures for vague sets have been defined based on the distance between the vague sets. Considering the importance of implication functions, it is necessary to investigate the similarity measure of vague sets from different perspectives. Based on the use of implication function, this paper investigates similarity measures of vague sets, proposes new formulas to calculate the similarity measures of vague sets or elements, performs sensitive analysis of these similarity measures based on implication functions and compares the existing similarity measures of vague sets with our proposed similarity measures. Finally, we investigate the classification and decision making based on the similarity measure of vague sets.

The rest of this work is organized as follows. In Section 2, we review some basic ideas about the theory of vague sets. In Section 3, we propose similarity measures between vague sets and between elements based on the implication function, and we make sensitive analysis and comparisons between the existing similarity measures of vague sets and our proposed similarity measures of vague sets. In Section 4, we investigate classification and decision making based on the similarity measure of vague sets. The final section is the conclusion.

2. Basic ideas

Throughout this paper, we use U, $U = \{u_1, u_2, u_3, \dots, u_n\}$, to denote the discourse set, $V(U)$ stands for the set of all vague subsets in U, and A expresses a vague set.

Let $L = [0, 1]$ and $[L]$ be the set of all closed subintervals of the interval [0, 1]. Especially for an arbitrary element $a \in [0, 1]$, we assume that a is the same as $[a, a]$, namely, $a = [a, a]$. Then, according to Zadeh's extension principle [29], for any $\overline{a} = [a^-, a^+], \overline{b} = [b^-, b^+] \in [L]$, we can popularize some operators such as $\bigvee, \bigwedge$, and c to $[L]$ and have $\overline{a} \bigvee \overline{b} = [a^- \bigvee b^-, a^+ \bigvee b^+]$, $\overline{a} \bigwedge \overline{b} = [a^- \bigwedge b^-, a^+ \bigwedge b^+]$, $\overline{a}^c = [1 - a^+, 1 - a^-]$, $\bigvee_{t \in W} \overline{a_t} = [\bigvee_{t \in W} a_t^-, \bigvee_{t \in W} a_t^+]$ and $\bigwedge_{t \in W} \overline{a_t} = [\bigwedge_{t \in W} a_t^-, \bigwedge_{t \in W} a_t^+]$, where W denotes an arbitrary index set. Furthermore, we have $\overline{a} = \overline{b} \Longleftrightarrow a^- = b^-, a^+ = b^+$, $\overline{a} \le \overline{b} \Longleftrightarrow a^- \le b^-, a^+ \le b^+$, and $\overline{a} < \overline{b} \Longleftrightarrow \overline{a} \le \overline{b}$ and $\overline{a} \ne \overline{b}$; there is then a minimal element $\overline{0} = [0, 0]$ and a maximal element $\overline{1} = [1, 1]$ in $[L]$.

A vague set A in U is characterized by a truth-membership function t_A and a false-membership function f_A, t_A: $U \to [0, 1]$, f_A: $U \to [0, 1]$, where $t_A(u_i)$ is a lower bound on the grade of membership of u_i derived from the evidence for u_i, $f_A(u_i)$ is a lower bound on the negation of u_i derived from the evidence against u_i, and $t_A(u_i) + f_A(u_i) \le 1$. The grade of membership of u_i in the vague set A is bounded by a subinterval $[t_A(u_i), 1 - f_A(u_i)]$ of [0, 1]. Simply expressed, $A(u_i) = [t_A(u_i), 1 - f_A(u_i)]$.

For every $u_i \in U$, if $A(u_i) = [t_A(u_i), 1 - f_A(u_i)]$, then $\pi_A(u_i) = 1 - t_A(u_i) - f_A(u_i)$ indicates the vague degree of u_i for the vague set A. It expresses the measure of u_i for the unknown information of the vague set A. We call $D(u_i) = t_A(u_i) - f_A(u_i)$ the superior function for the vague set A for the element u_i, then $D(u_i)$ indicates the absolute supporting degree of u_i for the vague set A.

When U is discrete, $U = \{u_1, u_2, u_3, \dots, u_n\}$, a vague set A can be written as $A = \sum_{i=1}^{n} [t_A(u_i), 1 - f_A(u_i)]/u_i$, $u_i \in U, i = 1, 2, \dots, n$; when the universe of discourse U is continuous, a vague set A can be written as $A = \int_U [t_A(u), 1 - f_A(u)]/u$, $u \in U$. Specifically, $U = \sum_{i=1}^{n} [1, 1]/u_i$, $\emptyset = \sum_{i=1}^{n} [0, 0]/u_i$, $u_i \in U, i = 1, 2, \dots, n$.

For example, let U be the universe of discourse, $U = \{6, 7, 8, 9, 10\}$. A vague set "LARGE" of U may be defined by

$$\begin{aligned} \text{LARGE} = {} & [0.1, 0.2]/6 + [0.3, 0.5]/7 + [0.6, 0.8]/8 \\ & + [0.9, 1]/9 + [1, 1]/10 \end{aligned}$$

If $A, B \in V(U)$, $u \in A$, and if we let x and y be two vague values such that $x = [t_x, 1 - f_x], y = [t_y, 1 - f_y]$, where $0 \le t_x \le 1 - f_x \le 1$, and $0 \le t_y \le 1 - f_y \le 1$, then the following operations can be founded on Gau [15].

$A = B$ iff $t_A(u) = t_B(u)$ and $f_A(u) = f_B(u), \forall u \in U$;
$A \subseteq B$ iff $t_A(u) \le t_B(u)$ and $f_A(u) \ge f_B(u), \forall u \in U$;
$x = y$ iff $t_x = t_y$ and $f_x = f_y$;
$x \subseteq y$ iff $t_x \le t_y$ and $f_x \ge f_y$.

Definition 1. Let x, y and z be three vague values such that $x = [t_x, 1 - f_x]$, $y = [t_y, 1 - f_y]$ and $z = [t_z, 1 - f_z]$. A real function

$$\begin{aligned} i : [L] \times [L] & \to [0, 1] \\ (x, y) & \mapsto i(x, y) \end{aligned}$$

is called the inclusion measure of the vague values x to y, if i satisfies the following properties:

(i1) $i([1, 1], [0, 0]) = 0$;
(i2) $t_x \le t_y, f_x \ge f_y \Leftrightarrow i(x, y) = 1$;
(i3) If $t_x \le t_y \le t_z$, and $f_x \ge f_y \ge f_z$, then $i(z, x) \le min(i(y, x), i(z, y))$.

where $L = [0, 1]$ and $[L]$ is the set of all closed subintervals of the interval [0, 1].

Definition 2. For every $A, B \in V(U)$, a real function

$$\begin{aligned} I : V(U) \times V(U) & \to [0, 1] \\ (A, B) & \mapsto I(A, B) \end{aligned}$$

is called the inclusion measure of the vague set A to B if I satisfies the following properties:

(I1) $I(X, \emptyset) = 0$;
(I2) $A \subseteq B \Leftrightarrow I(A, B) = 1$;
(I3) $\forall A, B, C \in V(U)$, if $A \subseteq B \subseteq C$, then $I(C, A) \le min(I(B, A), I(C, B))$.

Definition 3 ([24]). A real function R: $[0, 1] \times [0, 1] \to [0, 1]$ is called a fuzzy implication (or implication function) if it is monotone non-increasing in the first argument, i.e. $R(x, z) \ge R(y, z)$ for $x \le y$ and monotone non-decreasing in the second argument, i.e. $R(x, u) \le R(x, v)$ for $u \le v$, and satisfies the boundary conditions, $R(0, 0) = 1, R(1, 1) = 1$ and $R(1, 0) = 0$.

Furthermore, we list some properties related to fuzzy implication.

(1) Neutrality of truth when $R(1, t) = t, \forall t \in [0, 1]$;
(2) Identity principle when $R(t, t) = 1, \forall t \in [0, 1]$;
(3) Ordering property when $R(a, b) = 1$, if and only if $a \le b$.

The Lukasiwicz implication function $R_{Lu}(a, b) = (1 - a + b) \wedge 1$ satisfies each of these properties in the above.

Definition 4. Let x, y and z be three vague values such that $x = [t_x, 1 - f_x]$, $y = [t_y, 1 - f_y]$ and $z = [t_z, 1 - f_z]$. A real function

Please cite this article as: W. Zeng et al., Similarity measure for vague sets based on implication functions, Knowledge-Based Systems (2015), http://dx.doi.org/10.1016/j.knosys.2015.11.015