Weighted Fuzzy Similarity Measure Based on Tangent Function and its Application to Medical Diagnosis

ARTICLE - FEBRUARY 2015
DOI: 10.15680/IJRSET.2015.0402023

2 AUTHORS:

Surapati Pramanik
Nandalal Ghosh B.T. College, Panpur, P.O.-...
74 PUBLICATIONS 148 CITATIONS

Kalyan Mondal
West Bengal State University
17 PUBLICATIONS 33 CITATIONS
Weighted Fuzzy Similarity Measure Based on Tangent Function and its Application to Medical Diagnosis

Surapati Pramanik¹, Kalyan Mondal²
Assistant Professor, Department of Mathematics, Nandalal Ghosh B.T. College, Panpur, PO-Narayanpur, and District: North 24 Parganas, West Bengal, India¹
Assistant Teacher, Department of Mathematics, Birnagar High School (HS), Birnagar, Ranaghat, and District: Nadia, West Bengal, India².

ABSTRACT: In this paper, the weighted tangent similarity measure of fuzzy sets is proposed and its properties are studied. The concept of the weighted tangent similarity measure of fuzzy sets is a decision making tool which is characterized by the degree of membership function, degree of non-membership function (sum of this two components is equal to one). Finally, using this proposed method, an application on medical diagnosis is given for the applicability of the proposed approach.

KEYWORDS: Tangent similarity measure, weighted tangent similarity measure, fuzzy sets, decision making, and medical diagnosis.

I. INTRODUCTION

It is recognized that uncertainty plays an important role in modelling real world problems. Researchers recognize the need to bridge the gap between mathematical models and uncertainty and their empirical interpretations. The reflection of this gap can be found in problems of operations research, mathematics, biological, cognitive, and social sciences as well as modern technology, medicine and other applied sciences. The need to bridge the gap between a mathematical model and experience is well addressed by Max Black [1] in 1937. Zadeh [2] expressed the same need in 1962. In 1965, Zadeh [3] proposed the new paradigm of mathematics based on the very concept of fuzzy sets. When the new paradigm was proposed [3], the usual process of a paradigm shift [4] begins. For details of scientific paradigm you may consult the highly influential book namely, “The Structure Of Scientific Revolutions” “authored by Kuhn [4]. The paradigm shift is still continuing. It is reflected in the challenges thrown by the theory of fuzzy sets to the very foundation of science i.e. the Aristotelian two-valued logic. The new paradigm has a greater capability to deal with human decision making, machine intelligence, etc. The concept of Fuzzy set [3] generalizes the Cantor set discovered by Smith [5] in 1874 and introduced by German mathematician Cantor [6] in 1883. In fuzzy set theory, membership and non-membership degrees are complementary, i.e., the sum of membership and non-membership degrees of an element belonging in a fuzzy set is equal to one. The fuzzy set theory facilitates to solve various real world problems involving partially unknown information.

Literature review reflects that the studies on fuzzy similarity measures are mostly theoretic [7], [8], [9], [10], [11]. Fuzzy similarity measures are applied to image processing, [12], [13], fuzzy reasoning, [14], medical diagnosis [15], etc. Kakati has established a new similarity measure [16] for fuzzy sets using the extended definition of complementation [17] based on reference function. Kakati has proved the validity [18] of the new similarity measure [16] with the help of traditional Hamming Distance and Euclidean Distance measures and applied it to medical reasoning.
In this paper the authors have proposed a new similarity measure of fuzzy sets namely, weighted fuzzy similarity measure based on tangent function and studied its basic properties and applied it to medical diagnosis. The rest of the paper is organized as follow:

Section II presents the concepts of FSs, tangent similarity measures and weighted tangent similarity measures for fuzzy sets. Section III presents decision making methodology based on weighted tangent similarity measures. Section IV is devoted to present an example on medical diagnosis using the proposed approach. Finally, the conclusion of the paper and scope of future work are presented in Section V.

II. MATHEMATICAL PRELIMINARIES


Definition 2.1: Fuzzy set: A fuzzy set A in a universe of discourse X is defined as the following set of pairs 

\[ A = \{ x, \mu_A(x) : x \in X \}. \]

Here, \( \mu_A(x) : x \rightarrow [0,1] \) is a mapping called the membership value of \( x \in X \) in a fuzzy set A.

Definition 2.2: The value \( \mu_A(x) \) is called the degree of non-membership of the element \( x \in X \) to the fuzzy set.

Definition 2.3: Fuzzy Number: A fuzzy number is an extension of a regular number such that it does not refer to one single value but related to a connected set of possible values, where each possible value has its own weight between 0 and 1. This weight is called the membership function. Thus a fuzzy number is a normal fuzzy set and convex set.

Definition 2.4: Hamming distance between two fuzzy sets A and B is defined as

\[ H(A, B) = \frac{1}{n} \sum_{i=1}^{n} | \mu_A(x_i) - \mu_B(x_i) |. \]

Definition 2.5: Tangent similarity measure for fuzzy sets

The authors propose the tangent similarity measure for fuzzy sets in the following way. Let \( P = (\rho(x_i), \varphi(x_i)) \) and \( Q = (\sigma(x_i), \psi(x_i)) \) be two fuzzy numbers. Now tangent similarity function \( T_{FS}(P, Q) \) which measures the similarity between two vectors \( P \) and \( Q \) based only on the direction, ignoring the impact of the distance between them can be presented as:

\[ T_{FS}(P, Q) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\tan \left( \frac{\rho(x_i)}{\sigma(x_i)} \right)}{8} \right) \]

Proposition 1. The defined tangent similarity measure \( T_{FS}(P, Q) \) between two fuzzy numbers \( P \) and \( Q \) satisfies the following properties:

1. \( T_{FS}(P, Q) = 1 \)
2. \( T_{FS}(P, Q) = 1 \) if and only if \( P = Q \)
3. \( T_{FS}(P, Q) = T_{FS}(Q, P) \)
4. If \( R \) is a FS in \( X \) and \( P = Q \) then \( T_{FS}(P, R) = T_{FS}(P, Q) \) and \( T_{FS}(P, R) = T_{FS}(Q, R) \)

Proofs:

(1) As the membership, non-membership function of the fuzzy set are in \([0,1]\) and the value of the tangent function are within \([0,1]\), the similarity measure based on tangent function also is within \([0,1]\). Hence \( T_{FS}(P, Q) = 1 \)

(2) For any two fuzzy sets \( P \) and \( Q \) if \( P = Q \) this implies \( \rho(x_i) = \sigma(x_i) \) and \( \varphi(x_i) = \psi(x_i) \). Hence \( \frac{\rho(x_i)}{\sigma(x_i)} = \frac{\varphi(x_i)}{\psi(x_i)} = 0 \),

Thus \( T_{FS}(P, Q) = 1 \)

Conversely, If \( T_{FS}(P, Q) = 1 \) then \( \frac{\rho(x_i)}{\sigma(x_i)} = \frac{\varphi(x_i)}{\psi(x_i)} = 0 \), since \( \tan(0) = 0 \).

So we can write \( P = x_i \cdot q(x_i) \), \( P = x_i \cdot q(x_i) \). Hence \( P = Q \).
This proof is obvious.

(4) If $P \subseteq Q \subseteq R$ then $T_P(x_i) \leq T_Q(x_i) \leq T_R(x_i)$, and $F_P(x_i) \leq F_Q(x_i) \leq F_R(x_i)$ for $x_i \in X$.

Now we have the following inequalities:

$$\left| T_P(x_i) - T_Q(x_i) \right|, \left| T_P(x_i) - T_R(x_i) \right|, \left| T_Q(x_i) - T_R(x_i) \right|;$$

$$\left| F_P(x_i) - F_Q(x_i) \right|, \left| F_P(x_i) - F_R(x_i) \right|, \left| F_Q(x_i) - F_R(x_i) \right|.$$

Thus $T_{FS}(P, R) \leq T_{FS}(P, Q)$ and $T_{FS}(P, R) \leq T_{FS}(Q, R)$. Since tangent function is increasing in the interval $0, \frac{\pi}{4}$.

**Definition 2.6:** Weighted tangent similarity measure for fuzzy sets

The authors propose weighted tangent similarity measure for fuzzy sets in the following way.

Let $P = (\rho(x_i), \rho(x_i))$ and $Q = (\omega(x_i), \omega(x_i))$ be two fuzzy numbers. Now we present weighted tangent similarity function ($T_{WFS}(P, Q)$) which measures the similarity between two vectors $P$ and $Q$ based only on the direction, ignoring the impact of the distance between them can be presented as:

$$T_{WFS}(P, Q) = \frac{1}{n} \sum_{i=1}^{n} w_i \left( \tan \left( \frac{\rho(x_i) - \omega(x_i)}{8} \right) \right)$$

Where, $0 \leq w_i \leq 1$, $i = 1, 2, \ldots, n$; and $\sum_{i=1}^{n} w_i = 1$. Definition 2.6 converts to definition 2.5 if $w_i = \frac{1}{n}$.

**Proposition 2.** The defined tangent similarity measure $T_{WFS}(P, Q)$ between two fuzzy numbers $P$ and $Q$ satisfies the following properties:

5. $T_{WFS}(P, Q) = 1$
6. $T_{WFS}(P, Q) = 1$ if and only if $P = Q$
7. $T_{WFS}(P, Q) = T_{WFS}(Q, P)$
8. If $R$ is a FS in $X$ and $P \subseteq Q \subseteq R$ then $T_{WFS}(P, R) \leq T_{WFS}(P, Q) \leq T_{WFS}(P, R)$

**Proofs:**

(5) As the membership, non-membership function of the fuzzy set are in $[0, 1]$ and the value of the tangent function are within $[0, 1]$, the similarity measure based on tangent function also is within $[0, 1]$. Hence $0 \leq T_{WFS}(P, Q) \leq 1$.

(6) For any two fuzzy sets $P$ and $Q$ if $P = Q$ this implies $\rho(x_i) = \omega(x_i), \quad \rho(x_i) = \omega(x_i)$. Hence $\left| \rho(x_i) - \omega(x_i) \right| = 0$, $\left| \rho(x_i) - \omega(x_i) \right| = 0$, Thus $T_{WFS}(P, Q) = 1$.

Conversely, if $T_{WFS}(P, Q) = 1$ then $\left| \rho(x_i) - \omega(x_i) \right| = 0, \left| \rho(x_i) - \omega(x_i) \right| = 0$, since $\tan(0) = 0$.

So we can write $\rho(x_i) = \omega(x_i), \quad \rho(x_i) = \omega(x_i)$. Hence $P = Q$.

(7) This proof is obvious.

(8) If $P \subseteq Q \subseteq R$ then $T_P(x_i) \leq T_Q(x_i) \leq T_R(x_i)$, and $F_P(x_i) \leq F_Q(x_i) \leq F_R(x_i)$ for $x_i \in X$.

Now we have the following inequalities:

$$\left| T_P(x_i) - T_Q(x_i) \right|, \left| T_P(x_i) - T_R(x_i) \right|, \left| T_Q(x_i) - T_R(x_i) \right|;$$

$$\left| F_P(x_i) - F_Q(x_i) \right|, \left| F_P(x_i) - F_R(x_i) \right|, \left| F_Q(x_i) - F_R(x_i) \right|.$$
Thus $T_{WF}(P, R)$, $T_{WF}(P, Q)$, and $T_{WF}(Q, R)$. Since tangent function is increasing in the interval $0, \frac{\pi}{4}$.

III. FUZZY DECISION MAKING BASED ON TANGENT FUNCTION

Let $A_1, A_2, ..., A_m$ be a discrete set of candidates, $C_1, C_2, ..., C_n$ be the set of criteria of each candidate, and $D_1, D_2, ..., D_k$ are the alternatives of each candidate. The decision-maker provides the ranking of alternatives with respect to each candidate. The ranking presents the performances of candidates $A_i (i = 1, 2, ..., m)$ against the criteria $C_j (j = 1, 2, ..., n)$. The values associated with the alternatives for MADM problem can be presented in the following two decision matrices (see Table 1 and Table 2).

Table 1: The relation between candidates and attributes

<table>
<thead>
<tr>
<th>$A_i$</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>...</th>
<th>$C_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>$[AC]_{i1}$</td>
<td>$[AC]_{i2}$</td>
<td>...</td>
<td>$[AC]_{in}$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>$[AC]_{21}$</td>
<td>$[AC]_{22}$</td>
<td>...</td>
<td>$[AC]_{2n}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$A_m$</td>
<td>$[AC]_{m1}$</td>
<td>$[AC]_{m2}$</td>
<td>...</td>
<td>$[AC]_{mn}$</td>
</tr>
</tbody>
</table>

Table 2: The relation between attributes and alternatives

<table>
<thead>
<tr>
<th>$C_j$</th>
<th>$D_1$</th>
<th>$D_2$</th>
<th>...</th>
<th>$D_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>$[CD]_{11}$</td>
<td>$[CD]_{12}$</td>
<td>...</td>
<td>$[CD]_{1k}$</td>
</tr>
<tr>
<td>$C_2$</td>
<td>$[CD]_{21}$</td>
<td>$[CD]_{22}$</td>
<td>...</td>
<td>$[CD]_{2k}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$C_n$</td>
<td>$[CD]_{n1}$</td>
<td>$[CD]_{n2}$</td>
<td>...</td>
<td>$[CD]_{nk}$</td>
</tr>
</tbody>
</table>

Here $[AC]_{ij}$ and $[CD]_{ik}$ are all fuzzy numbers.

The steps of decision making corresponding to fuzzy number based on tangent function are presented as following steps.

Step 1: Determination of the relation between candidates and attributes:
Each candidate $A_i (i = 1, 2, ..., m)$ having the attribute $C_j (j = 1, 2, ..., n)$. The corresponding relational values between candidates and their attributes are presented in terms of fuzzy numbers as follows (see Table 3):

Table 3: Relation between candidates and attributes in terms of fuzzy numbers

<table>
<thead>
<tr>
<th>$A_i$</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>...</th>
<th>$C_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>${T_{11}, F_{11}}$</td>
<td>${T_{12}, F_{12}}$</td>
<td>...</td>
<td>${T_{1n}, F_{1n}}$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>${T_{21}, F_{21}}$</td>
<td>${T_{22}, F_{22}}$</td>
<td>...</td>
<td>${T_{2n}, F_{2n}}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$A_m$</td>
<td>${T_{m1}, F_{m1}}$</td>
<td>${T_{m2}, F_{m2}}$</td>
<td>...</td>
<td>${T_{mn}, F_{mn}}$</td>
</tr>
</tbody>
</table>

Step 2: Determination of the relation between attributes and alternatives:
The relation between attributes $C_i (i = 1, 2, ..., n)$ and alternatives $D_t (t = 1, 2, ..., k)$ is presented as follows (see Table 4):

Table 4: The relation between attributes and alternatives in terms of fuzzy numbers

<table>
<thead>
<tr>
<th>$C_i$</th>
<th>$D_1$</th>
<th>$D_2$</th>
<th>...</th>
<th>$D_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>${T_{11}, F_{11}}$</td>
<td>${T_{12}, F_{12}}$</td>
<td>...</td>
<td>${T_{1k}, F_{1k}}$</td>
</tr>
<tr>
<td>$C_2$</td>
<td>${T_{21}, F_{21}}$</td>
<td>${T_{22}, F_{22}}$</td>
<td>...</td>
<td>${T_{2k}, F_{2k}}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$C_n$</td>
<td>${T_{n1}, F_{n1}}$</td>
<td>${T_{n2}, F_{n2}}$</td>
<td>...</td>
<td>${T_{nk}, F_{nk}}$</td>
</tr>
</tbody>
</table>

Step 3: Determination of the criteria weight structure of similarity measure:
In the diagnosis process, decision maker may often encounter with unknown criteria weights. It may happen that the importance of the criteria is different. Therefore it is necessary to determine reasonable criteria weight for similarity measures.

Step 4: Determination of the co-relation measure between two relations:
Determine the correlation measure between the Table 3 and the Table 4 using $T_{WF}(P, Q)$ (from the equation 2).

Step 5: Ranking the alternatives:
Ranking the alternatives corresponding to each candidate is prepared as the descending order of correlation measures. Highest value indicates the best alternative for corresponding candidate.

Step 6: End

IV. NUMERICAL EXAMPLE ON MEDICAL DIAGNOSIS BASED ON TANGENT FUNCTION

Let us consider a numerical example on medical diagnosis. Medical diagnosis consists of a large amount of uncertainties and increased volume of information available to physicians from new medical technologies. The process of classifying different set of symptoms is under a single name of a disease. The proposed similarity measure among the patients Vs symptoms and symptoms Vs diseases gives proper medical diagnosis. The main feature of this proposed method is that it considers membership, non-membership degree by taking one time inspection for diagnosis.

Now, an example of a medical diagnosis will be presented. Example: Let $P = \{P_1, P_2, P_3, P_4\}$ be a set of patients, $D = \{\text{Viral Fever, Malaria, Typhoid, Stomach problem, Chest problem}\}$ be a set of diseases and $S= \{\text{Temperature, Headache, Stomach pain, cough, Throat pain}\}$ be a set of symptoms. Our solution is to examine the patients which in turn gives arise to membership and non-membership function for each patient.

Step 1: Determination of the relation between candidates and attributes:
Four patients ($P_1, P_2, P_3, \text{and} P_4$) have the symptoms temperature, Headache, Stomach pain, cough, and Throat pain. They feel illness. With the help of expert assessments, we tabulate the relational values between patients and their symptoms as follows (see Table 5).

<table>
<thead>
<tr>
<th>Relation-1</th>
<th>Temperature</th>
<th>Headache</th>
<th>Stomach pain</th>
<th>cough</th>
<th>Throat pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1$</td>
<td>(0.8, 0.2)</td>
<td>(0.6, 0.4)</td>
<td>(0.5, 0.5)</td>
<td>(0.6, 0.4)</td>
<td>(0.4, 0.6)</td>
</tr>
<tr>
<td>$P_2$</td>
<td>(0.8, 0.2)</td>
<td>(0.4, 0.6)</td>
<td>(0.6, 0.4)</td>
<td>(0.7, 0.3)</td>
<td>(0.7, 0.3)</td>
</tr>
<tr>
<td>$P_3$</td>
<td>(0.7, 0.3)</td>
<td>(0.7, 0.3)</td>
<td>(0.6, 0.4)</td>
<td>(0.5, 0.5)</td>
<td>(0.5, 0.5)</td>
</tr>
<tr>
<td>$P_4$</td>
<td>(0.6, 0.4)</td>
<td>(0.5, 0.5)</td>
<td>(0.7, 0.3)</td>
<td>(0.8, 0.2)</td>
<td>(0.7, 0.3)</td>
</tr>
</tbody>
</table>

Step 2: Determination of the relation between attributes and alternatives:
Every disease has some symptoms. There are some diseases whose symptoms are more or less same. So, in medical diagnosis, there is a confusion to detect exact disease of a patient. That’s why, it is important to set up the relations between symptoms and diseases in uncertain environment. Here, the relations are presented in fuzzy numbers as follows (see Table 6).

<table>
<thead>
<tr>
<th>Relation-2</th>
<th>Viral Fever</th>
<th>Malaria</th>
<th>Typhoid</th>
<th>Stomach problem</th>
<th>Chest problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>(0.7, 0.3)</td>
<td>(0.6, 0.4)</td>
<td>(0.4, 0.6)</td>
<td>(0.4, 0.6)</td>
<td>(0.6, 0.4)</td>
</tr>
<tr>
<td>Headache</td>
<td>(0.5, 0.5)</td>
<td>(0.5, 0.5)</td>
<td>(0.3, 0.7)</td>
<td>(0.2, 0.8)</td>
<td>(0.3, 0.7)</td>
</tr>
<tr>
<td>Stomach pain</td>
<td>(0.3, 0.7)</td>
<td>(0.1, 0.9)</td>
<td>(0.1, 0.9)</td>
<td>(0.9, 0.1)</td>
<td>(0.2, 0.8)</td>
</tr>
<tr>
<td>Cough</td>
<td>(0.8, 0.2)</td>
<td>(0.7, 0.3)</td>
<td>(0.8, 0.2)</td>
<td>(0.4, 0.6)</td>
<td>(0.8, 0.2)</td>
</tr>
<tr>
<td>Throat pain</td>
<td>(0.6, 0.4)</td>
<td>(0.3, 0.7)</td>
<td>(0.3, 0.7)</td>
<td>(0.2, 0.8)</td>
<td>(0.6, 0.4)</td>
</tr>
</tbody>
</table>

Step 3: Determination of the weight structure of each similarity measure:
Weight structure of each criterion for proposed similarity measure is determined by expert (doctor/ medical practitioner) as follows:

\[ w_1 = 0.225, w_2 = 0.195, w_3 = 0.200, w_4 = 0.190, w_5 = 0.190 \]

Step 4: Determination of the co-relation measure between two relations:
Using equation 2 (tangent function) we calculate correlation measures between Relation-1 and Relation-2 as follows (see Table 7).
Step 5: Ranking the alternatives:
The highest correlation measure from the Table 7 gives proper medical diagnosis. Therefore, patient P₁ suffers from Malaria, P₂ suffers from Viral fever, P₃ suffers from Viral fever and P₄ suffers from viral fever.

V. CONCLUSION

In this paper, we have proposed a tangent similarity measure approach of fuzzy sets and proved some of their basic properties. We have presented an application of weighted tangent similarity measure of fuzzy sets in medical diagnosis problem. In the future work, we will extend this tangent similarity measure to fuzzy multi sets.

REFERENCES


BIOGRAPHY

Dr. Surapati Pramanik (Ph. D., M.SC., M. Ed.) did his B. Sc. and M. Sc. in Mathematics from University of Kalyani. He received Ph. D. in Mathematics in 2010 from Bengal Engineering and Science University (BESU) Shibpur, India. He is currently an Assistant Professor of Mathematics at the Nandalal Ghosh B. T. College, Panpur, P.O.-Narayanpur, West Bengal, India. He has authored/co-authored more than 60 research papers in international peer reviewed journals; He authored mathematics method book for B. Ed. Courses from Aheli Publisher, Kolkata, India. His research interests...
include optimization, soft computing, grey system theory, neutrosophic decision making, rough sets, mathematics education, comparative education, international relation.

Kalyan Mondal (M.Sc., B. Ed.) passed B. Sc. Honours and M. Sc in Mathematics in 2001 and 2003 respectively from the University of Calcutta and University of Kalyani. Currently, he is an assistant Teacher of mathematics at Birnagar High School (HS), Birnagar, Ranaghat, Nadia, Pin Code: 741127, West Bengal, India. He has co-authored more than 10 research papers in international journals. His field of research interests includes fuzzy goal programming, grey system theory, Fuzzy decision making, intuitionistic fuzzy sets, neutrosophic sets, rough neutrosophic sets, and neutrosophic decision making.