

# Z Similarity Measure Among Fuzzy Sets

**Zoran Mitrović**

Associate Professor  
Faculty of Mechanical Engineering  
University of Belgrade

**Srdan Rusov**

Associate Professor  
Faculty of Traffic and Transport Engineering  
University of Belgrade

*The existing similarity measures between fuzzy sets have been analyzed at the beginning of the paper. The existing measures have been defined as the similarity measures of the two fuzzy sets. The constraints these measures include are that they are applied for the fuzzy sets whose membership functions are discontinuous. Likewise, the properties of the existing measures have been analyzed. It has been noted that in most frequent cases not even the basic properties these measures should fulfill have been fulfilled. On the basis of the existing measures, the similarity measure  $Z_{A,B}$  between two fuzzy sets has been analyzed. This measure fulfills the basic properties. In certain cases this measure gives the indefinite result  $\infty/\infty$ . The solution to this problem has been presented in this paper. With this solution, the measure  $Z_{A,B}$  has been determined to the full in each case, and the constraints do not affect the quality of the solution. The new measure with which we can determine the similarity measure of more fuzzy sets has been defined using the measure  $Z_{A,B}$ . With the introduction of the constraints for the new similarity measure among more fuzzy sets the new measure provides good results in all cases.*

**Keywords:** similarity measures, fuzzy sets,  $Z_{A,B}$  measure, Z measure.

## 1. INTRODUCTION

In the cases when we can not precisely enough describe a phenomenon or a process, we introduce certain assumptions, i.e. simplifications. This is a classical way of obtaining a model of an object or a process. The question then arises whether by introducing this assumption we have formed the model correctly enough, i.e. whether some of the basic properties have been neglected.

However, the problem may arise in the case we have modeled an object well enough. We do this most frequently in general numbers. When we have to make specific calculations, we are not sure which values certain parameters should have. In this case fuzzy numbers may help.

If we have more indefinite values in one model, then the need to introduce more fuzzy numbers arises. The question arises whether all these fuzzy numbers (sets) are mutually independent or connected. In order to determine the degree of their correlation it is necessary to introduce certain values which can "measure" this correlation.

For this purpose similarity measures between or among fuzzy sets have been introduced. The first papers in this field appeared in 1980s [1]. Afterwards, different authors proposed various similarity measures of fuzzy sets. In 1993 Pappis and Karacapilidis [2] were among

the first ones to make a classification of the measures known up to then. They made a comparison of these measures analyzing their advantages and disadvantages. Later, new, improved measures were defined in [3]. Likewise, three more measures were introduced in [4]. The detailed analysis was conducted for these measures.

The problem of similarity between fuzzy sets was also analyzed from the mathematical point of view, using the correlation theory [5]. In this case, the correlation quotient was used as a similarity measure of fuzzy sets.

In mid 1990s another comparison of known measures was made [6]. Certain cases of solving the disadvantages of existing measures were presented in [7] and [8] as well.

## 2. BASIC DEFINITIONS

Let us assume that  $A_i$  ( $i=1,2,\dots,m$ ) are fuzzy sets. With  $X_i$  we can denote the universe of discourse for each of the fuzzy sets  $A_i$ , i.e.  $X_i = \{x_1, x_2, \dots, x_n\}$  when fuzzy sets are represented by their membership functions, then

$$A_i = \{ \mu_{A_i}(x), a_{1i} \leq x \leq a_{2i} \}, \quad (1)$$

where  $\mu_{A_i}(x)$  is a membership function of fuzzy set  $A_i$ , for which  $\mu_{A_i}(x) : x \rightarrow [0,1]$ ,  $\forall i=1,2,\dots,m$ . Measures  $a_{1i}$  and  $a_{2i}$  are such that:  $a_{1i} \leq a_{2i}$ ;  $\mu_{A_i}(x) = 0$   $\forall x < a_{1i} \wedge \forall x > a_{2i}$ .

Originally, the distance function has been introduced into [1] as

Received: May 2006, Accepted: August 2006

Correspondence to: Zoran Mitrović

Faculty of Mechanical Engineering,  
Kraljice Marije 16, 11120 Belgrade 35, Serbia and Montenegro  
E-mail: zmitrovic@mas.bg.ac.yu

$$d_r(a,b) = \left[ \sum_{i=1}^n |a_i - b_i|^r \right]^{1/r}, \quad (2)$$

and it represents the distance of  $r^{\text{th}}$  order between points  $a$  and  $b$ , in  $n$ -dimensional space. In special cases, i.e. for  $r=1$  and  $r=\infty$ , and

$$d_1(a,b) = \sum_{i=1}^n |a_i - b_i|, \quad (3)$$

$$d_\infty(a,b) = \max_i |a_i - b_i|. \quad (4)$$

These relations are necessary since they help to define similarity measures between fuzzy sets which, in relations (2), (3) and (4), are represented as points.

There are several approaches to defining the similarity measures between fuzzy sets [6].

The first approach is such that it involves geometric distances in  $n$ -dimensional space. These are measures based on the model of geometric distance. These measures are applied only to define the similarity measures of two fuzzy sets. Let us assume that the fuzzy sets are  $A$  and  $B$ .

These measures are

$$L_{A,B} = 1 - \max_i |a_i - b_i|, \quad i=1,2,\dots,n \quad (5)$$

$$L_{A,B} = 1 - d_\infty(a,b), \quad (6)$$

$$L_{A,B} = 1 - \frac{\sum_{i=1}^n |a_i - b_i|}{\sum_{i=1}^n |a_i + b_i|}. \quad (7)$$

Analyzing the disadvantages of these measures, the similarity between fuzzy sets  $A$  and  $B$  has been given in [6] in the following way

$$W_{A,B} = 1 - \frac{1}{n} \sum_{i=1}^n |a_i - b_i|. \quad (8)$$

The second approach to defining similarity measures between fuzzy sets is when they are based on the so called set-theoretic approach. It is then assumed that fuzzy sets are defined through their continuous membership functions ( $\mu_{A_i}(x)$ ).

If the scalar cardinality (power) of fuzzy subset  $A$  is defined as

$$|A| = \int_{-\infty}^{\infty} \mu_{A_i}(x) dx, \quad (9)$$

then, e.g. the similarity measure of fuzzy sets  $A$  and  $B$  is as follows

$$S_{A,B} = 1 - \frac{|A \cap B|}{|A \cup B|}, \quad (10)$$

or

$$S_{A,B} = 1 - \sup_{x \in X} \mu_{A \cap B}(x). \quad (11)$$

By analyzing the previous measures, Pappis suggested that relations (10) and (11) should be

modified in the following way, so that the new measure should be

$$M_{A,B} = 1 - S_{A,B} = \frac{|A \cap B|}{|A \cup B|}, \quad (12)$$

while in the continuous membership functions this measure is

$$T_{A,B} = 1 - S_{A,B} = \sup_{x \in X} \mu_{A \cap B}(x). \quad (13)$$

The third approach to forming similarity measure between fuzzy sets is based on the so called matching function  $S$  [6]. In this case, vectors  $a$  and  $b$ , which are representatives of fuzzy sets  $A$  and  $B$ , are observed.

Then the similarity measure between fuzzy sets  $A$  and  $B$  can be defined as follows

$$S(a,b) = \frac{a \cdot b}{\max(a \cdot a, b \cdot b)}. \quad (14)$$

This measure has retained its original form, but is frequently denoted as  $P_{A,B} = S(a,b)$ .

On the basis of the stated measures, the forming of measure  $Z$  will be shown later. It is therefore necessary to explain the properties of these measures, i.e. their advantages and disadvantages.

Bearing in mind the three different approaches to solving this problem, the properties to be considered refer to measures  $W_{A,B}$ ,  $T_{A,B}$  and  $P_{A,B}$ .

### 3. PROPERTIES OF SIMILARITY MEASURES BETWEEN FUZZY SETS A AND B

The similarity measure of the sets  $A$  and  $B$ , based on geometric distances ( $W_{A,B}$ , [6]) has the following properties:

$$(W1) \quad W_{A,B} = W_{B,A},$$

$$(W2) \quad A = B \Leftrightarrow W_{A,B} = 1,$$

$$(W3) \quad A \cap B = 0 \Leftrightarrow W_{A,B} = 0,$$

$$(W4) \quad W_{A,A} = 1,$$

$$(W5) \quad W_{A,A} = 0 \Leftrightarrow A = I \vee A = 0,$$

$$(W6) \quad A \sim B \Rightarrow W_{A \cup C, B \cup C} \approx W_{A,B}.$$

These are just the basic properties of this measure. However, it can be concluded at first sight that some of these properties have not been fulfilled. We should bear in mind that each of the measures should fulfill these basic properties.

In this case (W3) and (W6) are the properties which are not always fulfilled. E.g. there are cases where  $A \cap B = 0$  but then  $W_{A,B} \neq 0$ . Generally, (W6) should also be valid, but it can be shown that it is not.

Regarding the basicity of these properties, which are not always fulfilled, it may be concluded that this similarity measure between fuzzy sets can not be generally accepted.

In the cases when we observe the similarity measure between fuzzy sets  $A$  and  $B$  based on the set-theoretic

approach, this measure is  $T_{A,B}$ . It should fulfill the following properties

- (T1)  $T_{A,B} = T_{B,A}$ ,
- (T2)  $A = B \Leftrightarrow T_{A,B} = 1$ ,
- (T3)  $A \cap B = 0 \Leftrightarrow T_{A,B} = 0$ ,
- (T4)  $A \sim B \Rightarrow T_{A \cup C, B \cup C} \approx T_{A,B}$ .

On the basis of these fundamental properties, it may be concluded that this measure is not suitable for use in most cases. It may be very easily concluded that the characteristic (T2) has not been fulfilled generally. This disadvantage may be overcome if the membership functions are standardized. As this is not always the case, this measure is not acceptable generally since the characteristic (T2) must be fulfilled.

The property (T4) is not always fulfilled either, bearing in mind the necessity for the measure (T4) to be fulfilled, which is not always the case, the measure  $T_{A,B}$  cannot be accepted as a measure, which generally provides a fairly accurate picture of the similarity of two fuzzy sets  $A$  and  $B$ .

Bearing in mind the similarity measures based on matching function  $S$ , then the measure  $P_{A,B}$  is defined. The properties of this measure are similar to the previous properties, i.e.

- (P1)  $P_{A,B} = P_{B,A}$ ,
- (P2)  $A = B \Leftrightarrow P_{A,B} = 1$ ,
- (P3)  $A \cap B = 0 \Leftrightarrow P_{A,B} = 0$ ,
- (P4)  $A \sim B \Rightarrow P_{A \cup C, B \cup C} \approx P_{A,B}$ .

In the case of this measure we may notice that the majority of basic properties have been fulfilled. It means that the basic disadvantages of the previous measures have been overcome with this measure. However, in this case the property (P4) need not always be fulfilled either.

Bearing in mind all the properties which occur with the presented measures, it is essential to define a measure which will at least fulfill the basic properties. Likewise, it is necessary to generalize this measure for the case of similarity of more sets as well. The attempt to introduce this measure can be found in [8].

#### 4. Z SIMILARITY MEASURES BETWEEN FUZZY SETS AND THEIR PROPERTIES

##### 4.1. Z similarity measure between two fuzzy sets

We are analyzing sets  $A$  and  $B$  defined by their continuous membership functions  $\mu_A$  and  $\mu_B$ . Their universe of discourse are  $X_A$  and  $X_B$  respectively, and they need not be equal, i.e.  $X_A \neq X_B$ .

The set of points of cross section of the membership functions  $\mu_A$  and  $\mu_B$  is defined by

$$K = \{k_i : \mu_A(k_i) = \mu_B(k_i), k_i < k_{i+1}, i = 1, 2, \dots, r\} \quad (15)$$

which means that there are  $r$  points of cross section of these two membership functions

Let us assume that  $p$  and  $q$  are such numbers so that

$$\mu_A(x) > 0 \wedge \mu_B(x) > 0, \quad \forall x \in (p, q). \quad (16)$$

Since the universe of discourse of fuzzy sets  $A$  and  $B$  is defined by  $X_A$  and  $X_B$ , i.e.

$$X_A = \{x, a_1 \leq x \leq a_2\}, \quad (17)$$

$$X_B = \{x, b_1 \leq x \leq b_2\}, \quad (18)$$

it may be concluded that  $p = a_1$  and  $q = b_2$ .

On the basis of [8], the similarity measure between two fuzzy sets ( $A$  and  $B$ ) can be expressed as follows

$$Z_{A,B} = \frac{2 \int_p^q \mu_C(x) dx}{\int_{a_1}^{a_2} \mu_A(x) dx + \int_{b_1}^{b_2} \mu_B(x) dx}. \quad (19)$$

The membership function  $\mu_C(x)$  can be defined as

$$\mu_C(x) = \begin{cases} \mu_A(x), & \mu_A(x) \leq \mu_B(x) \\ \mu_B(x), & \mu_A(x) > \mu_B(x) \end{cases}, \quad \forall x \in X_A, X_B. \quad (20)$$

When defining the value  $\int_p^q \mu_C(x) dx$ , we should

bear in mind that the calculation interval is divided into several parts and that it depends on  $r$  - i.e. on the number of cross section points of the membership function. It means that

$$\int_p^q \mu_C(x) dx = \int_p^{k_1} \mu_C(x) dx + \int_{k_1}^{k_2} \mu_C(x) dx + \dots + \int_{k_r}^q \mu_C(x) dx \quad (21)$$

The properties the stated measure  $Z_{A,B}$  fulfills are

- (Z1)  $Z_{A,B} = Z_{A,B}$ ,
- (Z2)  $A = B \quad (A \neq 0 \wedge B \neq 0) \Leftrightarrow Z_{A,B} = 1$ ,
- (Z3)  $A \cap B = 0 \Leftrightarrow Z_{A,B} = 0$ ,
- (Z4)  $A \sim B \Rightarrow Z_{A \cup C, B \cup C} \approx Z_{A,B}$ .

Note: If some universe of discourse is such that  $p = -\infty$  or  $q = \infty$ , then it most frequently occurs that e.g. on the interval  $(p, k_1)$  or on the interval  $(k_r, q)$  the following is fulfilled

$$\mu_A(x) = 1 \vee \mu_B(x) = 1. \quad (22)$$

In this case, when calculating the measure  $Z_{A,B}$ , using the expression (14) we obtain the indefinite expression  $\infty/\infty$ .

In order to overcome this, constraints must be imposed for the application of the expression (14). These constraints may be imposed on the intervals

$$x \in (p, k_1) \wedge x \in (k_r, q), \quad (23)$$

on which

$$\mu_A(x) = \mu_B(x) = 1. \quad (24)$$

In this case the set also has an unlimited number of elements, i.e.  $r = \infty$ . Then, on the intervals where (23) and (24) are valid it must be as follows

$$\mu_C(x) = 0. \quad (25)$$

so that  $Z_{A,B}$  is a final number.

#### 4.2. Z Similarity measure among fuzzy sets

Let us assume that  $A_i (i=1,2,\dots,n)$  are fuzzy sets, defined with their continuous membership functions  $\mu_{A_i}(x)$ . If we assume that the universe of discourse is  $X_i$ , for each set  $A_i$ , respectively. It means that

$$X_i = \{x, a_{1i} \leq x \leq a_{2i}\}. \quad (26)$$

On the basis of the previous discussion, values  $p$  and  $q$  can be defined as

$$p = \min_i (a_{1i}), \quad (27)$$

and

$$q = \max_i (a_{2i}). \quad (28)$$

The set of cross section points of the membership functions are

$$K = \left\{ k_l : \mu_{A_i}(k_l) = \mu_{A_j}(k_l), \right. \\ \left. \forall i, j = 1, 2, \dots, n, k_l < k_{l+1}, l = 1, 2, \dots, r \right\}. \quad (29)$$

Now the Z similarity measure among fuzzy sets is defined with

$$Z_A = \frac{n \int_p^q \mu_C(x) dx}{\sum_{i=1}^n \int_{a_{1i}}^{a_{2i}} \mu_{A_i}(x) dx}. \quad (30)$$

The membership function  $\mu_C(x)$ , in this case is defined as

$$\mu_C(x) = \min_i (\mu_{A_i}(x)), \quad \forall i = 1, 2, \dots, n, \quad (31)$$

where we should take into consideration that this function is determined on each interval  $(k_l, k_{l+1})$ ,  $\forall l = 0, 1, 2, \dots, r$ .

The value in the numerator of the expression (30) is determined in completely the same way as for the two sets, using (20).

In order to obtain the final value of the similarity measure among fuzzy sets  $Z_A$ , i.e., so that it would not

be  $\infty/\infty$ , similar constraints should be imposed as in the case of the measure  $Z_{A,B}$ . It means that, on the intervals  $(k_l, k_{l+1})$  on which

$$\mu_{A_i}(x) = 1 \quad \forall i = 1, 2, \dots, n \quad (32)$$

we should take into account that

$$\mu_C(x) = 0. \quad (33)$$

In this case the measure  $Z_A$  is the final number.

#### 5. CONCLUSION

1. In the introductory part we discussed the need to introduce similarity measures between fuzzy sets. The chronological analysis of the scientific views in this field has also been presented. Having in mind the frequent appearance of generalization of the results in this field, the need for further generalization of the existing results arises, as well as the improvement of the existing solutions. It is particularly important to provide new solutions when the existing ones are not good enough.
2. The existing similarity measures are frequently connected with the fuzzy numbers which are defined with discontinuous membership functions. Having in mind the limited technical application of these fuzzy numbers it was necessary to define the similarity measures between fuzzy sets determined with their continuous membership function.
3. The properties of the existing similarity measures have been analyzed. It may be concluded that the basic properties, which should be fulfilled by these measures, are most frequently not fulfilled.
4. The  $Z_{A,B}$  similarity measure between two fuzzy sets has been separately analyzed, when these sets are determined by its continuous membership function.
5. Although the measure  $Z_{A,B}$  does not fulfill all basic properties which should be fulfilled by a certain measure, there are cases where this measure has not been determined to the full, i.e. it has the value  $\infty/\infty$ . The solution to this problem has been given.
6. The similarity measure among more fuzzy sets, i.e.  $Z_A$ ,  $i = 1, 2, \dots, n$  has been defined using the introduced measure  $Z_{A,B}$ . Likewise, the solution when this measure has an indefinite value, has been given.

#### REFERENCES

- [1] Zwick, R., Carlstein, E., Budescu, D.: Measures of similarity among fuzzy concepts: a comparative analysis, *Internat. J. Approximate Reasoning* 1, pp. 221-242, 1987.
- [2] Pappis, C., Karacapilidis, N.: A comparative assessment of measures of similarity of fuzzy values, *Fuzzy Sets and Systems* 56, pp. 171-174, 1993.

- [3] Hyung, L., Song, Y., Lee, K.: Similarity measure between fuzzy sets and between elements, Fuzzy Sets and Systems 62, pp. 291-293, 1994.
- [4] Wang, W.: New similarity measures on fuzzy sets and on elements, Fuzzy Sets and Systems 85, Vol. 3, pp. 305-309, 1997.
- [5] Gerstenkorn, T., Man'ko, J.: Correlation of intuitionistic fuzzy sets, Fuzzy Sets and Systems 44, pp. 39-43, 1991.
- [6] Chen, S., Yeh, M., Hsiao, P.: A comparison of similarity measures of fuzzy values, Fuzzy Sets and Systems 72, Vol. 1, pp. 79-89, 1995.
- [7] Mitrović, Z.: Measures of connection of fuzzy sets and their elements, Transactions, Vol. XXVII, issue 2, pp. 13-15, 1998.
- [8] Mitrović, Z., Rusov, S., Mladenović, N.: About similarity measures between fuzzy sets, Bulletins for Applied & Computer Mathematics, No. CVII/2005, pp. 179-184, 2005.

---

## Z МЕРЕ ПОВЕЗАНОСТИ ИЗМЕЂУ FUZZY СКУПОВА

**Зоран Митровић, Срђан Русов**

У раду се најпре анализирају постојеће мере повезаности између fuzzy скупова. Постојеће мере дефинисане су као мере повезаности два fuzzy скупа. Ограничења која ове мере имају су да се оне примењују за fuzzy скупове чије функције припадности су прекидне. Такође, анализирани су особине постојећих мера. Уочено је да најчешће ни основне особине, које би ове мере требало да задовољавају, нису испуњене. На основу постојећих мера, анализирана је и  $Z_{A,B}$  - мера повезаности између два fuzzy скупа. Ова мера задовољава основне особине. У одређеним случајевима ова мера даје неодређени резултат облика  $\infty/\infty$ . У раду је дато решење овог проблема. Овим решењем мера  $Z_{A,B}$ , у потпуности је одређена у сваком случају, а ограничења која су наметнута у поступку њеног израчунавања, не утичу на квалитет решења. Користећи меру  $Z_{A,B}$  дефинисана је нова мера којом се може утврдити повезаност између више fuzzy скупова. Са уведеним ограничењима за меру  $Z_{A,B}$ , која је прилагођена за нову меру повезаности између више fuzzy скупова, нова мера даје квалитативно добре резултате у свим случајевима.