

Article

# Analysis of a Similarity Measure for Non-Overlapped Data

Sanghyuk Lee <sup>1,2,3,\*</sup>, Jaehoon Cha <sup>1</sup>, Nipon Theera-Umpon <sup>3,4</sup> and Kyeong Soo Kim <sup>1,2</sup>

<sup>1</sup> Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, Xi'an 215123, China; jaehoon.cha@xjtlu.edu.cn (J.C.); kyeongsoo.kim@xjtlu.edu.cn (K.S.K.)

<sup>2</sup> Centre for Smart Grid and Information Convergence, Xi'an Jiaotong-Liverpool University, Xi'an 215123, China

<sup>3</sup> Biomedical Engineering Centre, Chiang Mai University, Chiang Mai 50200, Thailand; nipon.t@cmu.ac.th

<sup>4</sup> Department of Electrical Engineering, Faculty of Engineering Chiang Mai University, Chiang Mai 50200, Thailand

\* Correspondence: sanghyuk.lee@xjtlu.edu.cn; Tel.: +86-512-8816-1415

Academic Editor: Shu-Ching Chen

Received: 19 September 2016; Accepted: 1 May 2017; Published: 9 May 2017

**Abstract:** A similarity measure is a measure evaluating the degree of similarity between two fuzzy data sets and has become an essential tool in many applications including data mining, pattern recognition, and clustering. In this paper, we propose a similarity measure capable of handling non-overlapped data as well as overlapped data and analyze its characteristics on data distributions. We first design the similarity measure based on a distance measure and apply it to overlapped data distributions. From the calculations for example data distributions, we find that, though the similarity calculation is effective, the designed similarity measure cannot distinguish two non-overlapped data distributions, thus resulting in the same value for both data sets. To obtain discriminative similarity values for non-overlapped data, we consider two approaches. The first one is to use a conventional similarity measure after preprocessing non-overlapped data. The second one is to take into account neighbor data information in designing the similarity measure, where we consider the relation to specific data and residual data information. Two artificial patterns of non-overlapped data are analyzed in an illustrative example. The calculation results demonstrate that the proposed similarity measures can discriminate non-overlapped data.

**Keywords:** similarity measure; overlapped data; non-overlapped data; neighbor data

---

## 1. Introduction

### 1.1. Background and Motivation

A fuzzy set was defined and summarized with respect to an ordinary set [1,2]; the fuzzy set theory provides a fundamental background for controller design, signal processing, pattern recognition, and other related areas [3–5]. Among the research on the fuzzy set theory, the analysis of data uncertainties has been carried out by numerous researchers through fuzzy data sets [6–8]. Of those researches based on fuzzy data sets, the similarity measure design problem—i.e., the design of a measure evaluating the degree of similarity between two data sets [8–15]—has been attracting an increasing amount of attention from the research community due to the increasing number of applications, including data mining, pattern recognition, and clustering [16,17]. The design of the similarity measure based on fuzzy numbers is convenient, but it can only use triangular or trapezoidal fuzzy membership functions [11,12]. If we design a similarity measure based on a distance measure, we can use a general fuzzy membership function without any limit on its shape [13–15].

Note that most conventional similarity measures—whether based on fuzzy number or distance measure—can be applied to overlapped data sets only [8,11,13,14,18–21]. By “overlapped data sets” we mean data sets with the same support. Similarity measure design and investigations of its relation to dissimilarity have been carried out in the literature. Some similarity measure results are designed based on fuzzy numbers [11,12], and axiomatic similarity measures have been defined [18,19]. In addition, some similarity measures have been designed to work with non-overlapped data sets as well [22–24]. However, research on non-overlapped data sets is increasingly becoming important in big data analysis, especially in atypical data analysis [23,24].

In our previous research, analysis was carried out on overlapped and non-overlapped data, and intuitionistic fuzzy sets were introduced with artificial data [20]. Similarity measure design based on fuzzy numbers and similarity measure applications to intuitionistic fuzzy sets have also been studied [21]. The obtained results were summarized in a follow-up paper, in which similarity measures on intuitionistic fuzzy sets were proposed and proved [22].

In this paper, we present a methodology for the analysis of the similarity of non-overlapped data distributions. If conventional similarity measures are applied, there is no discrimination on different non-overlapped data distribution pairs, which results in the same value. We preprocessed non-overlapped data distributions to obtain information relations as a new measure between neighbor information. This represents information on how distributions are related to each other, and it allows conventional similarity to be applied. This procedure is also effective for similarity measures on overlapped data distributions. Another approach is to consider neighbor information, where a similarity measure is derived and verified with an illustrative example.

The results of the similarity of non-overlapped data suggest that the measure can be applicable to atypical data analysis. The approach is extensible to big data analysis based on the rationale that the neighbor information of each data is closely related to the similarity of each data.

## 1.2. Data Description

We considered a sequence of data, each element of which can take either a continuous or a discrete value. A data sequence provides information on facts. For each data sequence, its data distribution is represented by various structures. We designed similarity measures for data that have been shown as overlapped data in previous works [13–15]. In the following, we explain several types of data with examples:

- Two data,  $f(x_i)$  and  $g(y_j)$  for  $x_i, y_j \in X$ ,  $X$  denotes a universe of discourse.  $f(x_i)$  and  $g(y_j)$  have values at the same support  $x_i = y_j$  whether it is same or not. It means direct operation such as summation or subtract is possible between two values.
- On the other hand, they are classified as non-overlapped data. It is rather difficult to attain operation results between two data in different supports. In this paper, we propose a similarity design for such non-overlapped data with the help of preprocessing.
- Atypicality is one of the characteristics in big data [23,24]. Huge amount of data constitute different types of structure. Hence, it is challenging to analyze the similarity between different structures of data, even though they have the same meaning and fact.
- In general, data—especially big data—provide a large amount of information, and groups of data are located close to or far from each other geometrically. The information analysis on neighbor data is used to design the non-overlapped data in this paper.

In this paper, we illustrate the usefulness of the proposed similarity measure on the non-overlapped data with the example data. Data distribution can be continuous or discrete; in this paper, however, only discrete data are considered for the ease of explanation of the computational procedure.

We begin with the review of the preliminary results on the similarity measure and distance measure using two discrete data sets in Section 2. In Section 3, the similarity measure for non-overlapped data is designed with neighbor information, and an explicit similarity measure is proposed and proved.

In order to illustrate the usefulness of the proposed similarity measure, example data distributions are introduced and analyzed. Section 4 concludes our work in this paper.

## 2. Preliminaries on Similarity Measure

The axiomatic definition of a similarity measure is given by Liu [8]; based on this definition, a similarity measure can be designed explicitly using a distance measure such as the Manhattan distance. From the definition, numerous properties for a similarity measure between data sets can be derived.

**Definition 1.** [8] A real function  $s: F^2 \rightarrow R^+$  is called a similarity measure if  $s$  satisfies the following properties:

- (S1)  $s(A, B) = s(B, A)$ , for  $\forall A, B \in F(X)$
- (S2)  $s(D, D^C) = 0$ , if and only if  $D \in P(X)$
- (S3)  $s(C, C) = \max_{A, B \in F(X)} s(A, B)$ , for  $\forall C \in F(X)$
- (S4)  $A, B, C \in F(X)$ , if  $A \subset B \subset C$ , then  $s(A, B) \geq s(A, C)$  and  $s(B, C) \geq s(A, C)$

where  $R^+ = [0, \infty)$ ,  $X$  is the universal set,  $F(X)$  denotes the class of all fuzzy sets of  $X$ ,  $P(X)$  is the class of all crisp sets of  $X$ , and  $D^C$  is the complement of  $D$ . From Definition 2.1, numerous similarity measures can be derived.

Distance is needed to represent the similarity measure explicitly; Liu also introduced the distance measure with axiomatic definition in [8].

**Definition 2.** [8] A real function  $d: F^2 \rightarrow R^+$  is called a distance measure on  $F$  if  $d$  satisfies the following properties:

- (D1)  $d(A, B) = d(B, A)$ , for  $\forall A, B \in F(X)$
- (D2)  $d(A, A) = 0$ ,  $A \in F(X)$
- (D3)  $d(D, D^C) = \max_{A, B \in F(X)} d(A, B)$ ,  $D \in F(X)$
- (D4)  $A, B, C \in F(X)$ , if  $A \subset B \subset C$ , then  $d(A, B) \leq d(A, C)$  and  $d(B, C) \leq d(A, C)$ .

Generally, Manhattan distance is commonly used as a distance measure between fuzzy sets  $A$  and  $B$  [22]:

$$d(A, B) = \frac{1}{n} \sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)|$$

where  $X = \{x_1, x_2, \dots, x_n\}$ ,  $|k|$  is the absolute value of  $k$ , and  $\mu_A(x)$  denotes the membership function of  $A \in F(X)$ . Conventional similarity measures are illustrated in the following theorems. Theorem 1 is a similarity measure equation with Manhattan distance [22].

**Theorem 1.**  $\forall A, B \in F(X)$ , if  $d$  is the Manhattan distance measure, then:

$$s(A, B) = 1 - d(A, A \cap B) - d(B, A \cap B) \quad (1)$$

is a normalized similarity measure between sets  $A$  and  $B$ .

**Proof.** The proof of Theorem 1 is given in Appendix A.  $\square$

Besides the similarity measure given by Theorem 1, other similarity measures can be defined as well, which is illustrated in Theorem 2.

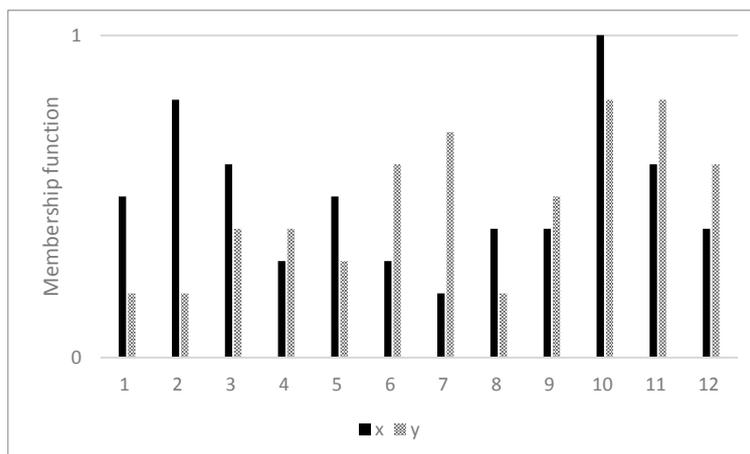
**Theorem 2.**  $\forall A, B \in F(X)$  if  $d$  satisfies the Manhattan distance measure, then:

$$s(A, B) = 2 - d((A \cap B), [1]_X) - d((A \cap B), [0]_X) \quad (2)$$

is another normalized similarity measure between sets  $A$  and  $B$ ; note that given sets  $A$  and  $B$ , Equation (2) results in the same measure value as Equation (1).

**Proof.** The proof of Theorem 2 is given in Appendix B. □

Now we consider the overlapped discrete data distributions shown in Figure 1, where two data sets  $X$  and  $Y$  are distributed over the same support in the universe of discourse with 12 data, each with different magnitudes:  $X = \{x(i): 0.5, 0.8, 0.6, 0.3, 0.5, 0.3, 0.2, 0.4, 0.4, 1.0, 0.6, 0.4\}$  and  $Y = \{y(i): 0.2, 0.2, 0.4, 0.4, 0.3, 0.6, 0.7, 0.2, 0.5, 0.8, 0.8, 0.6\}$ , for  $i = 1, 2, \dots, 12$ . The similarity calculation between diamond and circle data is carried out with Equations (1) and (2). We explain the computation results in the following. For more details of the computation, readers are referred to [25].



**Figure 1.** Overlapped discrete data distribution.

Applying the similarity measure Equation (1) to the data distributions in Figure 1, we obtain:

$$\begin{aligned}
 s(X, Y) &= 1 - d(X, X \cap Y) - d(Y, X \cap Y) \\
 &= 1 - \frac{1}{12}(2.9) = 0.758
 \end{aligned}
 \tag{3}$$

With the similarity measure in Equation (2), we obtain the same result as follows:

$$\begin{aligned}
 s(X, Y) &= 2 - d((X \cap Y), [1]_X) - d((X \cup Y), [0]_X) \\
 &= 2 - \frac{1}{12}(7.5 + 7.4) = 0.758
 \end{aligned}
 \tag{4}$$

The computational procedures are clear.

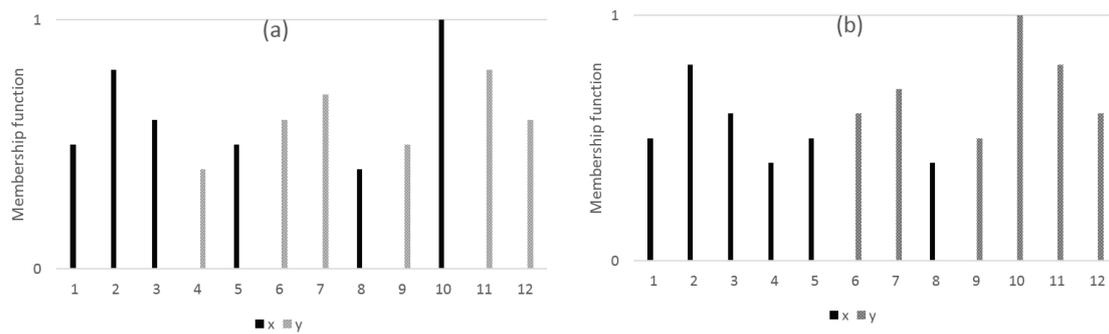
Hence, similarity measures in Equations (1) and (2) show their effectiveness.

For comparison, example non-overlapped data distributions are shown in Figure 2. Each of twelve-member sets  $X$  and  $Y$  has six non-zero membership values, which are distributed differently and without any overlap, but they have the same magnitude distributions as those in Figure 1.

Now, the similarity measure is applied to calculate similarity for non-overlapped data. We demonstrate the usefulness of the similarity measure with examples. Consider the discrete data shown in Figure 2, where two data sets  $X$  and  $Y$  are illustrated with different combinations as follows:

- Figure 2a:  $X = \{x(i): 0.5, 0.8, 0.6, 0.0, 0.5, 0.0, 0.0, 0.4, 0.0, 1.0, 0.0, 0.0\}$  and  $Y = \{y(i): 0.0, 0.0, 0.0, 0.4, 0.0, 0.6, 0.7, 0.0, 0.5, 0.0, 0.8, 0.6\}$
- Figure 2b:  $X = \{x(i): 0.5, 0.8, 0.6, 0.4, 0.5, 0.0, 0.0, 0.4, 0.0, 0.0, 0.0, 0.0\}$  and  $Y = \{y(i): 0.0, 0.0, 0.0, 0.0, 0.0, 0.6, 0.7, 0.0, 0.5, 1.0, 0.8, 0.6\}$

where there are no overlaps in positions between non-zero elements of  $X$  and  $Y$ . Next, the conventional similarity measures in Equations (1) and (2) are applied to calculate the similarity between  $X$  and  $Y$  in Figure 2.



**Figure 2.** Two data distribution between X and Y : (a) rather mixed; (b) slightly mixed.

For the data shown in Figure 2a, the similarity calculation between X and Y is derived with Equation (1) as follows:

$$\begin{aligned}
 s(X, Y) &= 1 - d(X, X \cap Y) - d(Y, X \cap Y) \\
 &= 1 - d(X, [0]_X) - d(Y, [0]_X) \\
 &= 1 - \frac{1}{12} \sum_i \text{height}(x(i)) - \frac{1}{12} \sum_i \text{height}(y(i))
 \end{aligned} \tag{5}$$

Note that the similarity calculation for the data shown in Figure 2b results in the same value as that obtained for Figure 2a as long as the data magnitudes are the same.

Because there is no intersection between the two distributions,  $X \cap Y = [0]_X$ , the similarity measure calculations using Equation (2) for the data shown in Figure 2 are the same as shown below:

$$\begin{aligned}
 s(X, Y) &= 2 - d((X \cap Y), [1]_X) + d((X \cap Y), [0]_X) \\
 &= 2 - d([0]_X, [1]_X) + d((X \cap Y), [0]_X) = 2 - 1 - d((X \cap Y), [0]_X) \\
 &= 1 - \frac{1}{12} \sum_i \text{height}(x(i) + y(i))
 \end{aligned} \tag{6}$$

Results in Equations (5) and (6) suggest that the similarity between the two non-overlapped data sets is related to the summation of their magnitudes and the independence of their distributions.

Note that the similarity measures in Equations (1) and (2) do not discriminate between the two different non-overlapping data distributions shown in Figure 2. This is because the operations in the definitions of the similarity measures are based on minimum/maximum value comparison in the same support such as  $A \cap B$  or  $A \cup B$ . Therefore, the design of the similarity measures on non-overlapped data distributions requires a different approach.

### 3. Similarity Measure on Non-Overlapped Data

Two approaches for the similarity measure on non-overlapped data distributions are now proposed. First, a conventional similarity measure approach is applied in a novel way to non-overlapping data. In order to apply similarity measures given in Equations (1) or (2) to non-overlapped data distribution (Figure 2), the original data distribution is expressed as an information-related diagram with the proposed measure. Having obtained the information-related graph, we can find the similarity measure between this and the original data distribution.

Second, a new design of a similarity measure on non-overlapping data is carried out. It is based on neighbor information.

### 3.1. Data Transformation and Application to Similarity Measure

In order to apply a conventional similarity measure to non-overlapped data, we consider both the distance between two different patterns and the difference between two membership values, i.e., for each  $x(i) > 0, i = 1, 2, \dots, 12$ :

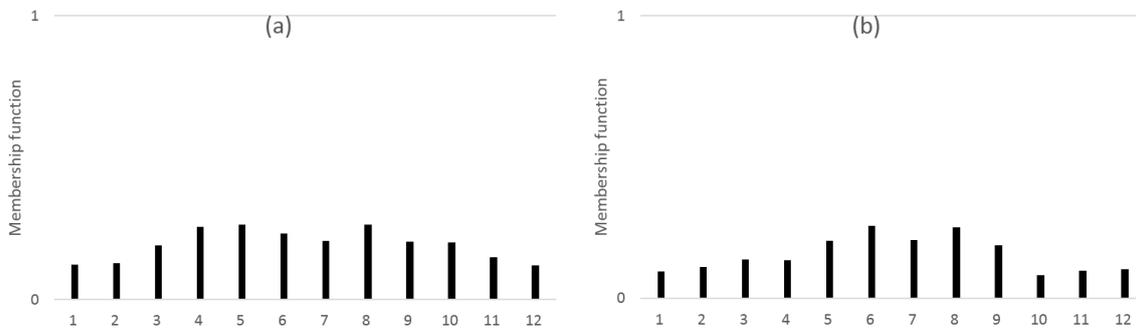
$$\sum_{\forall j \in [1,2,\dots,12] \text{ s.t. } y(j) > 0} \frac{1}{|i-j|+1} (1 - |x(i) - y(i)|) \quad (7)$$

and for each  $y(i) > 0, i = 1, 2, \dots, 12$ :

$$\sum_{\forall j \in [1,2,\dots,12] \text{ s.t. } x(j) > 0} \frac{1}{|i-j|+1} (1 - |x(j) - y(i)|) \quad (8)$$

where each difference should be satisfied between different patterns. From calculations of the proposed preprocessing, six data are obtained from Equations (7) and (8), respectively.

Following the calculation process described and applying it to each pattern in turn, we obtain the information relationships between different patterns shown in Figure 3.



**Figure 3.** Data preprocessing with Equations (7) and (8) for: (a) Figure 2a; and (b) Figure 2b.

Next, we apply the similarity measure between the previous pattern of Figure 2 and preprocessing results of Figure 3. Here, the previous pattern is considered in terms of the whole distribution.

Now, conventional similarity measures given in Equations (1) and (2) can be applied to the newly derived data distributions shown in Figure 4, i.e., the union of  $X$  and  $Y$  (i.e., the black bars in the figure and denoted  $d$  below) and the preprocessing results from Figure 3 (i.e., the gray bars in the figure and denoted  $p$  below). Equations (1) and (2) become Equations (9) and (10) as follows:

$$s(d, p) = 1 - d(d, d \cap p) - d(p, d \cap p) \quad (9)$$

$$s(d, p) = 2 - d((d \cap p), [1]_X) + d((d \cap p), [0]_X) \quad (10)$$

The calculation results of Equations (9) and (10) for Figure 4a are given by:

$$\begin{aligned} s(d, p) &= 1 - d(d, d \cap p) + d(p, d \cap p) \\ &= 1 - \frac{1}{12} (|0.5 - 0.1232| + |0.8 - 0.1284| + |0.6 - 0.1912| + |0.4 - 0.2554| \\ &\quad + |0.5 - 0.2632| + |0.6 - 0.2328| + |0.7 - 0.206| + |0.4 - 0.2628| + |0.5 - 0.2045| \\ &\quad + |1.0 - 0.2004| + |0.8 - 0.1504| + |0.6 - 0.12|) \\ &= 0.5782 \end{aligned}$$

$$\begin{aligned}
 s(d, p) &= 2 - d((d \cap p), [1]_X) + d((d \cap p), [0]_X) \\
 &= 2 - \frac{1}{12}(0.8768 + 0.8716 + 0.8088 + 0.7446 + 0.7368 + 0.7672 + 0.794 + 0.7372 \\
 &\quad + 0.7955 + 0.7996 + 0.8496 + 0.88) \\
 &\quad - \frac{1}{12}(0.5 + 0.8 + 0.6 + 0.4 + 0.5 + 0.6 + 0.7 + 0.4 + 0.5 + 1.0 + 0.8 + 0.6) = 0.5782
 \end{aligned}$$

The same result is obtained for both cases.

Applying Equations (9) and (10), we obtain the similarity measure for Figure 4b as follows:

$$\begin{aligned}
 s(d, p) &= 1 - d(d, d \cap p) - d(p, d \cap p) \\
 &= 1 - \frac{1}{12}(|0.5 - 0.094| + |0.8 - 0.1099| + |0.6 - 0.1371| + |0.4 - 0.1354| \\
 &\quad + |0.5 - 0.2021| + |0.6 - 0.2572| + |0.7 - 0.206| + |0.4 - 0.2517| \\
 &\quad + |0.5 - 0.1879| + |1.0 - 0.0813| + |0.8 - 0.0963| + |0.6 - 0.1015|) = 0.5384
 \end{aligned}$$

$$\begin{aligned}
 s(d, p) &= 2 - d((d \cap p), [1]_X) - d((d \cup p), [0]_X) \\
 &= 2 - \frac{1}{12}(0.906 + 0.8901 + 0.8629 + 0.8646 + 0.7979 + 0.7428 + 0.794 + 0.7483 \\
 &\quad + 0.8121 + 0.9187 + 0.9037 + 0.8985) - \frac{1}{12}(0.5 + 0.8 + 0.6 + 0.4 \\
 &\quad + 0.5 + 0.6 + 0.7 + 0.4 + 0.5 + 1.0 + 0.8 + 0.6) = 0.5384
 \end{aligned}$$

From a heuristic point of view, it is clear that the similarity between X and Y in Figure 2a is higher than that in Figure 2b. The calculation results clearly show this similarity.

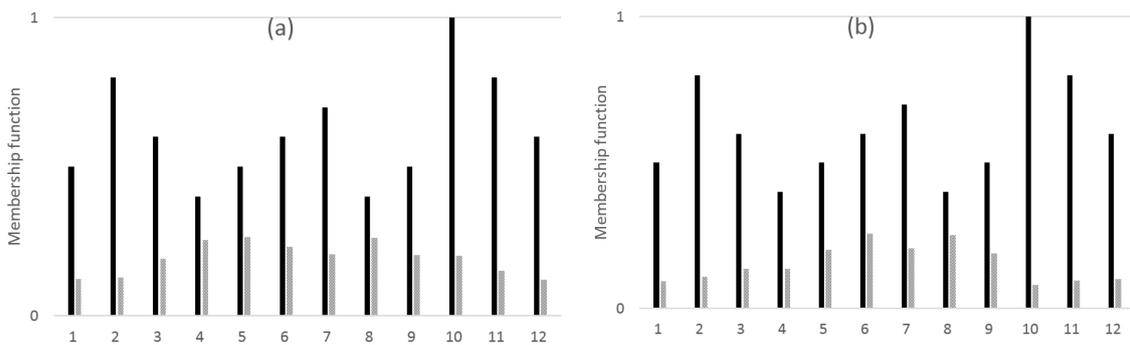


Figure 4. (a) Data distributions with Figures 2a and 3a; (b) data distributions with Figures 2b and 3b.

### 3.2. Similarity Measure Design Using Neighbor Information

Here, we give a brief introduction of the similarity measure on non-overlapped data proposed in [21,22] and compare its results with those obtained in Section 3.1. We assume that the similarity measure is affected from neighbor data information as in Section 3.1. Theorem 3 provides a similarity measure on non-overlapped data, which has been proved in [21,22].

**Theorem 3.** Given fuzzy sets A and B, let  $\tilde{A}$  and  $\tilde{B}$  be their supports, respectively. If d is the Manhattan distance measure, then:

$$s(A, B) = 1 - |s_{\tilde{A}} - s_{\tilde{B}}| \tag{11}$$

is a similarity measure between set A and B, where:

$$S_{\tilde{A}} = d\left(\left(\tilde{A} \cap (A \cup B)\right), [1]_X\right)$$

and:

$$S_{\tilde{B}} = d\left(\left(\tilde{B} \cap (A \cup B)\right), [1]_X\right)$$

The similarity measure given in Equation (10) is designed using a distance measure. Equation (10) is applied to the non-overlapped data in Figure 2, and calculation results are given below.

For the data distribution shown in Figure 2a:

$$s(A, B) = 1 - \left| d\left(\left(\tilde{A} \cap (A \cup B)\right), [1]_X\right) - d\left(\left(\tilde{B} \cap (A \cup B)\right), [1]_X\right) \right| = 0.967 \quad (12)$$

For the data distribution shown in Figure 2b:

$$s(A, B) = 1 - \left| d\left(\left(\tilde{A} \cap (A \cup B)\right), [1]_X\right) - d\left(\left(\tilde{B} \cap (A \cup B)\right), [1]_X\right) \right| = 0.833 \quad (13)$$

The calculation results show that the similarity measure designed for non-overlapped data gives similar results; the value of the similarity measure on the data shown in Figure 4a is higher than that on the data shown in Figure 4b, i.e., 0.967 versus 0.833. Calculation results are shown in Appendix C. Therefore, the similarity measure given in Equation (10) enables the comparison of similarity on non-overlapped data. Note that the first distribution pair shows a higher similarity even when we decide it heuristically.

#### 4. Conclusions

In this paper, we have designed similarity measures on non-overlapped data and demonstrated their effectiveness with illustrative examples.

First, we presented conventional similarity measures on overlapped data and shown their usefulness with an overlapped data distribution. Calculation results have shown that they are effective for overlapped data. However, similarity calculation results on non-overlapped data are the same even for different distribution pairs, which means that they cannot provide discrimination on non-overlapped data. From the example, therefore, we have concluded that we cannot directly apply those similarity measures designed for overlapped data to non-overlapped data.

Second, in order to utilize conventional similarity measures, we proposed a new measure with nearest information and derived information relation. The result is illustrated in Figure 3. This makes it possible to apply conventional similarity measures to non-overlapped data. Calculation results show the difference for data distributions shown in Figure 2. We have also proposed another similarity measure based on neighbor information. Its effectiveness is also demonstrated through an illustrative example. Note that, even though the calculation results are not identical to those of the first approach, they show a similar pattern in discriminating between two different data distributions.

The results from this paper show that it is possible to design a similarity measure applicable to both overlapped and non-overlapped data, which can discriminate between two different data distributions even in the case of non-overlapped data.

Note that our work reported in this paper could be extended for applications in the analysis of big data in the future.

**Acknowledgments:** This research was financially supported by the Centre for Smart Grid and Information Convergence (CeSGIC) at Xian Jiaotong-Liverpool University.

**Author Contributions:** Sanghyuk Lee and Nipon Theera-Umpon developed the concept and drafted the manuscript. Jaehoon Cha derived preprocessing Equations (7) and (8). Kyeong Soo Kim checked and clarified the results and made major revisions to the manuscript. All authors read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### Appendix A. Proof of Theorem 1

Because the detailed proof of Theorem 1 is given in [13], here we provide a brief summary of the proof.

(S1): It is clear from Equation (1) itself, hence  $s(A, B) = s(B, A)$  is satisfied.

(S2):

$$\begin{aligned} s(D, D^C) &= 1 - d(D, D \cap D^C) - d(D^C, D \cap D^C) \\ &= 1 - d(D, [0]_X) - d(D^C, [0]_X) = 0 \end{aligned}$$

is clear. Hence, (S2) is satisfied.

(S3): It is also clear because:

$$\begin{aligned} s(C, C) &= 1 - d(C, C \cap C) - d(C, C \cap C) \\ &= 1 - d(C, C) - d(C, C) = 1. \end{aligned}$$

(S4): From Equation (1), because:

$$d(A, A \cap C) = d(A, A \cap B) \text{ and } d(C, A \cap C) \geq d(B, A \cap B)$$

it is guaranteed that  $s(A, C) \leq s(A, B)$ .

Similarly, because:

$$d(A, A \cap C) = d(B, B \cap C) \text{ and } d(C, A \cap C) \geq d(C, B \cap C)$$

$s(A, C) \leq s(B, C)$  is also satisfied.

## Appendix B. Proof of Theorem 2

The proof of Theorem 2 is similar to that of Theorem 1.

(S1) It is clear from Equation (2) itself, hence  $s(A, B) = s(B, A)$ .

(S2) Because:

$$\begin{aligned} s(D, D^C) &= 2 - d((D \cap D^C), [1]_X) - d((D \cup D^C), [0]_X) \\ &= 2 - d([0]_X, [1]_X) - d([1]_X, [0]_X) = 0 \end{aligned}$$

(S2) is satisfied.

(S3) This property is satisfied because:

$$\begin{aligned} s(C, C) &= 2 - d((C \cap C), [1]_X) - d((C \cup C), [0]_X) \\ s(C, C) &= 2 - d(C, [1]_X) - d(C, [0]_X) \\ &= 2 - 1 = 1. \end{aligned}$$

(S4) From Equation (2), because:

$$d((A \cap B), [1]_X) = d((A \cap C), [1]_X) \text{ and } d((A \cup B), [0]_X) \leq d((A \cup C), [0]_X)$$

it is guaranteed that  $s(A, C) \leq s(A, B)$ . Similarly, because:

$$d((A \cap C), [1]_X) \geq d((B \cap C), [1]_X) \text{ and } d((A \cup C), [0]_X) = d((B \cup C), [0]_X)$$

$s(A, C) \leq s(B, C)$  is also satisfied.

### Appendix C. Derivation of Equations (12) and (13)

Note that the calculation procedures are provided in [22]. The computation result is as follows:

$$\begin{aligned} s(A, B) &= 1 - \left| d\left(\left(\tilde{A} \cap (A \cup B)\right), [1]_X\right) - d\left(\left(\tilde{B} \cap (A \cup B)\right), [1]_X\right) \right| \\ &= 1 - 1/6 \left| ((1 - 0.5) + (1 - 0.8) + (1 - 0.6) + (1 - 0.5) + (1 - 0.4) \right. \\ &\quad \left. + (1 - 1)) - ((1 - 0.4) + (1 - 0.6) + (1 - 0.7) + (1 - 0.5) + (1 - 0.8) \right. \\ &\quad \left. + (1 - 0.6)) \right| = 1 - 1/6 |2.2 - 2.4| = 0.967 \end{aligned}$$

$$\begin{aligned} s(A, B) &= 1 - \left| d\left(\left(\tilde{A} \cap (A \cup B)\right), [1]_X\right) - d\left(\left(\tilde{B} \cap (A \cup B)\right), [1]_X\right) \right| \\ &= 1 - 1/6 \left| ((1 - 0.5) + (1 - 0.8) + (1 - 0.6) + (1 - 0.4) + (1 - 0.5) \right. \\ &\quad \left. + (1 - 0.4)) - ((1 - 0.6) + (1 - 0.7) + (1 - 0.5) + (1 - 1) + (1 - 0.8) \right. \\ &\quad \left. + (1 - 0.6)) \right| = 1 - 1/6 |2.8 - 1.8| = 0.833. \end{aligned}$$

### References

- Zadeh, L.A. Fuzzy sets and systems. In Proceedings of the Symposium on System Theory; Polytechnic Institute of Brooklyn: New York, NY, USA, 1965; pp. 29–37.
- Dubois, D.; Prade, H. *Fuzzy Sets and Systems*; Academic Press: New York, NY, USA, 1988.
- Kovacic, Z.; Bogdan, S. *Fuzzy Controller Design: Theory and Applications*; CRC Press: Boca Raton, FL, USA, 2005.
- Plataniotis, K.N.; Androutsos, D.; Venetsanopoulos, A.N. Adaptive Fuzzy systems for Multichannel Signal Processing. *Proc. IEEE* **1999**, *87*, 1601–1622. [[CrossRef](#)]
- Fakhar, K.; El Aroussi, M.; Saidi, M.N.; Aboutajdine, D. Fuzzy pattern recognition-based approach to biometric score fusion problem. *Fuzzy Sets Syst.* **2016**, *305*, 149–159. [[CrossRef](#)]
- Pal, N.R.; Pal, S.K. Object-background segmentation using new definitions of entropy. *IEEE Proc.* **1989**, *36*, 284–295. [[CrossRef](#)]
- Kosko, B. *Neural Networks and Fuzzy Systems*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1992.
- Liu, X. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets Syst.* **1992**, *52*, 305–318.
- Bhandari, D.; Pal, N.R. Some new information measure of fuzzy sets. *Inf. Sci.* **1993**, *67*, 209–228. [[CrossRef](#)]
- De Luca, A.; Termini, S. A Definition of nonprobabilistic entropy in the setting of fuzzy entropy. *J. Gen. Syst.* **1972**, *5*, 301–312.
- Hsieh, C.H.; Chen, S.H. Similarity of generalized fuzzy numbers with graded mean integration representation. In Proceedings of the 8th International Fuzzy Systems Association World Congress, Taipei, Taiwan, 17–20 August 1999; Volume 2, pp. 551–555.
- Chen, S.J.; Chen, S.M. Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. *IEEE Trans. Fuzzy Syst.* **2003**, *11*, 45–56. [[CrossRef](#)]
- Lee, S.H.; Pedrycz, W.; Sohn, G. Design of Similarity and Dissimilarity Measures for Fuzzy Sets on the Basis of Distance Measure. *Int. J. Fuzzy Syst.* **2009**, *11*, 67–72.
- Lee, S.H.; Ryu, K.H.; Sohn, G.Y. Study on Entropy and Similarity Measure for Fuzzy Set. *IEICE Trans. Inf. Syst.* **2009**, *E92-D*, 1783–1786. [[CrossRef](#)]
- Lee, S.H.; Kim, S.J.; Jang, N.Y. Design of Fuzzy Entropy for Non Convex Membership Function. In *Communications in Computer and Information Science*; Springer: Berlin, Germany, 2008; Volume 15, pp. 55–60.
- Dengfeng, L.; Chuntian, C. New similarity measure of intuitionistic fuzzy sets and application to pattern recognitions. *Pattern Recognit. Lett.* **2002**, *23*, 221–225. [[CrossRef](#)]
- Li, Y.; Olson, D.L.; Qin, Z. Similarity measures between intuitionistic fuzzy (vague) set: A comparative analysis. *Pattern Recognit. Lett.* **2007**, *28*, 278–285. [[CrossRef](#)]
- Couso, I.; Garrido, L.; Sanchez, L. Similarity and dissimilarity measures between fuzzy sets: A formal relational study. *Inf. Sci.* **2013**, *229*, 122–141. [[CrossRef](#)]
- Li, Y.; Qin, K.; He, X. Some new approaches to constructing similarity measures. *Fuzzy Sets Syst.* **2014**, *234*, 46–60. [[CrossRef](#)]

20. Lee, S.; Sun, Y.; Wei, H. Analysis on overlapped and non-overlapped data. In Proceedings of the Information Technology and Quantitative Management (ITQM2013), Suzhou, China, 16–18 May 2013; Volume 17, pp. 595–602.
21. Lee, S.; Wei, H.; Ting, T.O. Study on Similarity Measure for Overlapped and Non-overlapped Data. In Proceedings of the Third International Conference on Information Science and Technology, Yangzhou, China, 23–25 March 2013.
22. Lee, S.; Shin, S. Similarity measure design on overlapped and non-overlapped data. *J. Cent. South Univ.* **2014**, *20*, 2440–2446. [[CrossRef](#)]
23. Host-Madison, A.; Sabeti, E. Atypical Information Theory for real-valued data. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 666–670.
24. Host-Madison, A.; Sabeti, E.; Walton, C. Information Theory for Atypical Sequence. In Proceedings of the 2013 IEEE Information Theory Workshop (ITW), Sevilla, Spain, 9–13 September 2013; pp. 1–5.
25. Pemmaraju, S.; Skiena, S. *Computational Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*; Cambridge University: Cambridge, UK, 2003.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).