

Similarity Measure Based on the Belief Function Theory: Application in a Decision-Making Process

M. Khalaj^{1,*}, F. Khalaj^{2,*}

¹Department of Industrial Engineering, Robat Karim Branch, Islamic Azad University, Tehran, Iran

²Department of Mathematics and statistics, Robat Karim Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Article History

Received 19 May 2020

Accepted 04 Jan 2021

Keywords

Belief function theory

Belief set

Similarity measure

Decision-making

ABSTRACT

This study considers a new aspect of the belief function theory to define a belief set, which is characterized by truth, uncertainty and falsity belief degrees as a 3D vector representation. Then, based on the implication of a belief set, one of the similarity measures (i.e., Cosine, Jaccard and Dice) between two belief sets is defined. Furthermore, the weighted similarity measure of these different species between each alternative and ideal alternative is presented in order to rank alternatives and determine the best one. Finally, a comparison between similarity measures and an application of a new method based on similarity measures between two belief sets in the decision-making process is calculated to show the capability and validity of the proposed method.

© 2021 The Authors. Published by Atlantis Press B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

1. INTRODUCTION

The similarity measure is an important operation to evaluate similarity between two sets or objects. Many researchers are interested to investigate the different distance measure in order to take the similarity in the uncertain situations with imprecise knowledge. It can help to extract information under uncertain environment as a powerful tool to describe imprecise and insufficient data. It has been extensively applied in many files. The correlation coefficient, rank correlation coefficient, cosine, Dice and Jaccard similarity are some of the important distance measures for evaluating the similarity between objects. Fuzzy set theory and belief function theory (Dempster–Shafer theory [DST]) are powerful theories to address the imprecise and uncertain condition. Since these theories have been introduced, many studies are proposed related to imprecise and uncertain modeling. Due to these capabilities in uncertainty reasoning, several generalizations have been proposed to measure correlation, distance or similarity of evidence in a real system.

In a belief function theory, a similarity/dissimilarity measure is an important subject, which is used for two bodies of evidence as a Dempster's conflict factor (K value), which is used in a combination rule. It is an important and effective index commonly used to measure similarity/dissimilarity between two bodies of evidence. Several methods are proposed to modify the conflict degree and improve combination rule by decreasing its restriction or to correct the evidential value before [1–10]. Similarity/dissimilarity in belief functions is usually used to measure the relevance of evidence in order to manage existing conflict and have been applied in different applications, such as [11–13,14], classification [15,16], risk assessment, [17–20], others [21–23]. Furthermore, in the field of a fuzzy theory, for example, Wang [24] introduced the concept of similarity measure in a fuzzy set theory. Yingfang *et al.* [25] presented three computations method to calculate the similarity degree between an interval-valued fuzzy set, which is a generalization of a fuzzy set. Zeng *et al.* [26] proposed new formulas to evaluate the similarity measure of a vague set, which is a generalization of the fuzzy set. Rıdvan and Ahmet [27] defined different distance and similarity measures between neutrosophic soft sets, which is another generalization of a fuzzy theory and showed the application of the proposed method to rank all possible alternatives in a multi-criteria decision-making (MCDM) problem. Ismata and Tabasamb [28] defined a new measure of the similarity between two intuitionistic fuzzy sets and its application in MCDM. Negan [29] presented a model to determine the degree of the similarity based on an intuitionistic fuzzy value and its application in a medical image analysis. Qinrong and Xiao [30] developed a flexible and effective method to solve the problems of decision-making in an interval-valued intuitionistic fuzzy soft set based on distance measures. Fei *et al.* [31] purposed an ELECTRE model based MCDM method to evaluate information which is expressed by DST. Fei *et al.* [32] expended the VIKOR method by DST to solve a problem of supplier selection.

Because of the relationship between fuzzy set theory and belief function theory, several methodologies have been concentrated on generalization the belief function theory to fuzzy sets or vice versa. In the fields of distance and similarity measure, Jousselme and Maupin [33]

*Corresponding author. Email: khalaj82@gmail.com

proposed a distance and similarity measure in classical belief function theory and shown a generalization of them to a fuzzy belief function. Bouchard *et al.* [34] defined a positive definite similarity measure then presented a distance measure between the intuitionistic fuzzy set and belief function theory. Song *et al.* [35] proposed the distance measure between intuitionistic fuzzy belief functions based on the Euclidean distance. Feng *et al.* [36] defined a pair of belief and plausibility functions based on intuitionistic fuzzy lower and upper approximation operators. Yang and Han [9] presented the total uncertainty degree of a body of evidence by the average distance between the belief interval. Song *et al.* [23] proposed two similarity measures using the Jousselme's distance measure and cosine similarity for intuitionistic fuzzy sets in the framework of the belief function theory and shown its application in MCDM problems and pattern recognition.

In essential, most of the abovementioned studies on the belief function theory are the generalization of another technique and theories to propose the distance measure, especially fuzzy set theory. But not the definition in the framework of the belief function theory, straightly. On the other hand, so far, there is no investigation on the mathematical framework to calculate the distance measure between different belief degrees, which is shown similarity/dissimilarity between bodies of evidence directly. However, the relation between the belief function theory using different theory or mathematical framework can be rarely affected to improve the application of a distance measure in various fields; however, these do not involve a similarity measure among expert's belief degrees about an objective or hypothesis as a crucial matter in belief functions. Therefore, the distance measure based on belief functions may play an important role to analyze a similarity/dissimilarity measure between the objectives. It can be considered based on belief degrees, which are included, belief degrees of truth, falsity or unknown about a proposition, by using three belief functions (i.e., belief, uncertainty and disbelief functions), which is defined in the theory of belief function.

Therefore, in this work, a new definition of belief measure for an objective, based on the belief function is directly proposed in the framework of the DS T. It is called belief set and characterized by the truth of a belief value, the uncertainty of a belief value (i.e., ambiguity between to be truth or falsity) and the falsity of a belief value about each focal element. This definition is based on the concept of belief functions and proposed without the switching by another theory (e.g., fuzzy theory). It is investigated as a triple set of belief degrees about a matter and each body of evidence is considered as a triple evidence vector and is contained three elements (i.e., belief, uncertainty and disbelief functions). Then, from the point of view of an evidence vector, similarity or dissimilarity between two bodies of evidence is proposed to describe a degree of conflict. For this purpose, the relationship between vector measure is considered to evaluate similarities and differences of the body of evidence, using the different type of the similarity and distance measure. The rest of the paper is organized as follows: Section 2 provides the preliminaries belief function theory and proposed a new definition of a belief set. Section 3 presents a brief recall of a variation coefficient similarity measure. Section 4 proposes new similarity measures upon a body of evidence, and then presents a numerical example and application of the proposed method in the decision-making problem. Section 5 is the conclusion of this paper.

DS-VIKOR: A New Multi-criteria Decision-Making Method for Supplier Selection, *International Journal of Fuzzy Systems*, Liguo Feim Yong Deng, Yong Hu

2. THEORY OF BELIEF FUNCTIONS

2.1. Preliminaries of Belief Function Theory

The theory of belief functions is a generalization of Bayesian theory of subjective probability. It was first proposed by Dempster [37] and then developed by Shafer [38]. The belief function theory is based on degrees of belief from multiple sources of evidence rather than estimating probabilities that a hypothesis is true. The DST estimates how close the evidence proves the truth of that hypothesis [39]. From [39,38] some basic concept are reviewed.

Definition 2.1.1. Suppose that X is a finite discrete set of all probable elements, which consists of mutually exclusive and exhaustive factors with n elements, $X = \{x_1, x_2, \dots, x_n\}$. Frame of discernment or power set is denoted by $P(X)$ or 2^X . It is a set of all subsets of X including the empty set and X itself, where $P(X) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_N\}, \dots, \{x_1 \cup x_2\}, \dots, X\}$

Definition 2.1.2. Let $H_1 = \{x_1\}, H_2 = \{x_2\}, \dots, H_n = \{x_n\}, \dots, H_{2^n} = \{x_1, x_2, \dots, x_n\}$ are subsets of the power set or focal elements, which is called body of evidence. The basic probability assignments (or belief mass functions) for each body of evidence are denoted by $m(H_i)$, ($i = 1, 2, \dots, 2^n$) to the subset H_i on the frame of discernment. It represents the level of evidence in support of hypotheses or alternative H_i and measures the degree of belief for one question from subjective probabilities for a related question where the evidence supports H_i , as defined by

$$m(H_i) : 2^X \rightarrow [0, 1] \quad (1)$$

The conditions of this function are as follows:

$$\sum_{H_i \in 2^X} m(H_i) = 1 \quad (2)$$

$$m(\emptyset) = 0 \quad (3)$$

Definition 2.1.3. Let $H_i, H_j \in P(X)$, where $m(H_i) \geq 0$ and $m(H_j) \geq 0$. The belief function $bel(H_i)$ and plausibility function $pl(H_i)$ are defined by

$$bel(H_i) = \sum_{H_j \subseteq H_i} m(H_j) \quad (4)$$

$$pl(H_i) = \sum_{H_j \cap H_i \neq \emptyset} m(H_j) \quad (5)$$

The belief function $bel(H_i)$ expresses the justified specific support for the proposition or focal element H_j and the plausibility function $pl(H_i)$, shows the total belief or the potential specific support for H_i . The degree of imprecision for H_i is represented as, *uncertainty* = $pl(H_i) - bel(H_i)$ or difference between the belief and plausibility.

Definition 2.1.4. Two independent basic belief assignment $m_1(H_i)$ and $m_2(H_i)$ can be combined using the Dempster's rule and denoted as $m_1 \oplus m_2$:

$$m_1 \oplus m_2(H_i) = \frac{1}{1-k} \sum_{H_j, H_k \subseteq P(X), H_j \cap H_k = H_i} m_1(H_j) m_2(H_k) \quad (6)$$

where

$$k = \sum_{H_j, H_k \subseteq P(X), H_j \cap H_k = \emptyset} m_1(H_j) m_2(H_k) \quad (7)$$

Management of conflict is a central problem in the belief function theory to measure the difference between the two bodies of evidence. In the classical belief function or the D-ST and combination rule by Eq. (6), k is an effective item to measure of conflict in evidence fusion, which is commonly used. But, as mentioned in the introduction, because of some shortcoming of classical conflict index to compare the information contained in different bodies of evidence, different method have been widely extended and proposed to approximation and measurement of conflicting evidence in the various area, in terms of research.

Definition 2.1.5. Four basic evidential belief functions used to represent the belief functions can be shown by belief (bel) as a degree of belief; disbelief (dis) as a degree of disbelief; uncertainty (u) as a degree of uncertainty and plausibility (pl) as a degree of plausibility, where (bel) and (pl) functions are defined as Eqs. (4) and (5) and $dis(H_i) = 1 - pl(H_i)$, or $dis(H_i) = 1 - u(H_i) - bel(H_i)$, where $bel(H_i) + u(H_i) + dis(H_i) = 1$.

Definition 2.1.6. The Dempster's rule of a combination is used to obtain with the corresponding evidential belief function as follows:

$$bel_1(H_i) \oplus bel_2(H_i) = \frac{bel_1(H_i) bel_2(H_i) + bel_1(H_i) u_2(H_i) + bel_2(H_i) u_1(H_i)}{1 - bel_1(H_i) dis_2(H_i) - dis_1(H_i) bel_2(H_i)} \quad (8)$$

$$dis_1(H_i) \oplus dis_2(H_i) = \frac{dis_1(H_i) dis_2(H_i) + dis_1(H_i) u_2(H_i) + dis_2(H_i) u_1(H_i)}{1 - bel_1(H_i) dis_2(H_i) - dis_1(H_i) bel_2(H_i)} \quad (9)$$

$$u_1(H_i) \oplus u_2(H_i) = \frac{u_1(H_i) u_2(H_i)}{1 - bel_1(H_i) dis_2(H_i) - dis_1(H_i) bel_2(H_i)} \quad (10)$$

2.2. Basic Idea of Defining a Belief Set

Based on the Definitions 2.1.5 and 2.1.6, the information content in each body of evidence can be defined upon three belief functions. Thus, the idea of this paper is to obtain degrees of belief for one object from subjective probability, where pieces of evidence appear as the degree of true, false and unknown situation that are measured through the different information.

Definition 2.2.1. Let H_i is a subset of the X as a frame of discernment or power set. A belief set of each body of evidence $A \in 2^X$ is characterized by belief function bel_A , uncertainty function u_A , disbelief dis_A as follows:

$$A = \{bel_A(H_i), u_A(H_i), dis_A(H_i) | H_i \in 2^X\} \quad (11)$$

where $bel_A(H_i), u_A(H_i), dis_A(H_i) \in [0, 1]$ for each $H_i \in 2^X$ and $0 \leq bel_A(H_i) + u_A(H_i) + dis_A(H_i) \leq 1$.

Example

Let $X = \{x_1, x_2, x_3\}$ be the belief set of hypotheses $H_4 = \{x_1, x_2\}$ with respect to the given mass assignment in Table 1, the belief set of hypotheses is determined as follows:

$$bel(H_4) = bel(\{x_1, x_2\}) = m(\{x_1\}) + m(\{x_2\}) + m(\{x_1, x_2\}) = 0.1 + 0.2 + 0.1 = 0.4$$

$$pl(H_4) = pl(\{x_1, x_2\}) = m(\{x_1\}) + m(\{x_2\}) + m(\{x_1, x_2\}) + m(\{x_1, x_3\}) + m(\{x_2, x_3\}) + m(\{x_1, x_2, x_3\}) \\ = 0.1 + 0.2 + 0.1 + 0.3 + 0.1 = 0.9$$

$$dis(H_4) = dis(\{x_1, x_2\}) = 1 - pl(\{x_1, x_2\}) \\ = 1 - 0.9 = 0.1$$

$$u(H_4) = u(\{x_1, x_2\}) = 1 - bel(\{x_1, x_2\}) - dis(\{x_1, x_2\}) = 1 - bel(H_4) - dis(H_4) = 1 - 0.4 - 0.1 = 0.5$$

Therefore, the belief set of hypotheses $H_4 = \{x_1, x_2\}$ is defined as $A = (0.4, 0.5, 0.1)$. Furthermore, assume that belief set is $B = (0.5, 0.3, 0.2)$, then aggregate the belief sets A and B as a new belief set C that is calculated by Eqs. (8–10) as follows:

$$bel_C(H_4) = bel_A(H_4) \oplus bel_B(H_4) = \frac{0.4 \times 0.5 + 0.4 \times 0.3 + 0.5 \times 0.5}{1 - 0.4 \times 0.2 - 0.1 \times 0.5} = 0.655172$$

$$dis_C(H_4) = dis_A(H_4) \oplus dis_B(H_4) = \frac{0.1 \times 0.2 + 0.1 \times 0.3 + 0.2 \times 0.5}{1 - 0.4 \times 0.2 - 0.1 \times 0.5} = 0.172414$$

$$u_C(H_4) = u_1(H_4) \oplus u_2(H_4) = \frac{0.3 \times 0.5}{1 - 0.4 \times 0.2 - 0.1 \times 0.5} = 0.172414$$

Thus, the aggregated belief set is $C = (0.655172, 0.0.172414, 0.0.172414)$

Definition 2.2.2. The complement of belief set A is denoted by A^c , and can be defined as $bel_{A^c}(H_i) = dis_A(H_i)$, $u_{A^c}(H_i) = 1 - u_A(H_i)$, $dis_{A^c}(H_i) = bel_A(H_i)$, for any $H_i \in 2^X$.

$bel_A(H_i) \leq bel_B(H_i)$, $u_A(H_i) \geq u_B(H_i)$, $dis_A(H_i) \geq dis_B(H_i)$ for any $H_i \in 2^X$.

Definition 2.2.3. A belief set A is contained in the other belief set B , or $A \subseteq B$ if and only if $bel_A(H_i) \leq bel_B(H_i)$, $u_A(H_i) \geq u_B(H_i)$, $dis_A(H_i) \geq dis_B(H_i)$ for any $H_i \in 2^X$.

3. SIMILARITY METHOD

In this section, three important similarity measures are considered, and then their application is shown in belief sets. Accordingly, the basic mathematical preliminaries of these measures are first presented. Then, it is used for defining similarity measures between two belief sets.

Definition 3.1. The Jaccard distance is an index to show the relationship between two or more vectors or random variables and defined by [34]

$$J(X, Y) = \frac{XY}{\|X\|_2 + \|Y\|_2 - XY} = \frac{\sum X_i Y_i}{\sum x_i^2 + \sum y_i^2 - \sum x_i y_i} \quad (12)$$

where $\|X\|_2 = \sqrt{\sum x_i^2}$, $\|Y\|_2 = \sqrt{\sum y_i^2}$ ($i = 1, 2, \dots, n$) are the Euclidean norm of X and Y , and $XY = \sum x_i y_i$ is the inner product of variables or vectors X and Y .

Table 1 Basic Probability Assignments (BPA) to the focal element or hypotheses.

H_i	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\{x_1, x_2, x_3\}$
$m(H_i)$	0.1	0.2	0.1	0.1	0.1	0.3	0.1

Definition 3.2. Dice similarity measures the similarity value between two variable or sets, in order to assess the relation between them, as follows [18]:

$$D(X, Y) = \frac{XY}{\|X\|_2 + \|Y\|_2} = \frac{2 \sum X_i Y_i}{\sqrt{\sum x_i^2} + \sqrt{\sum y_i^2}} \quad (13)$$

Definition 3.3. Cosine similarity index measures the similarity between two vectors and defined based on the cosine of the angle between them, by [40]

$$C(X, Y) = \frac{XY}{\|X\|_2 \|Y\|_2} = \frac{\sum X_i Y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}} \quad (14)$$

4. SIMILARITY MEASURE BETWEEN TWO BELIEF SET

In this section, using the definition of belief values set as a three-dimensional vector and similarity measure, improved methods of similarity measure between belief sets are introduced. Thus, three commonly and famous similarity measures (i.e., Jaccard, Dic and Cosine) as introduced in the previous section are calculated based on the belief set in order to use in a real scientific and engineering application. Then numerical example and application of this method in the decision-making process are presented.

Definition 4.1. In order to extend and generalize of the similarity coefficient of belief sets, we first define the informational energy of belief sets, based on the informational energy measure, which is introduced by Onicescu [41,42]. Accordingly, let $A = \{bel_A(H_i), u_A(H_i), dis_A(H_i) | H_i \in 2^X\}$ is a belief set, where $bel_A(H_i), u_A(H_i), dis_A(H_i) \in [0, 1]$ for every $x \in 2^X$. The informational energy of belief sets can be expressed by

$$I = bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i) \quad (15)$$

The informational energy is criterion to describe and compute the relevances of input features, which is bounded by 0 and 1. When it increases, the entropy decreases that is an index to evaluate uncertainty.

Definition 4.2. Let A and B are two belief sets in the frame of discernment X , $A = \{bel_A(H_i), u_A(H_i), dis_A(H_i) | H_i \in 2^X\}$ and $B = \{bel_B(H_i), u_B(H_i), dis_B(H_i) | x \in 2^X\}$, where $bel_A(H_i), u_A(H_i), dis_A(H_i) \in [0, 1]$ and $bel_B(H_i), u_B(H_i), dis_B(H_i) \in [0, 1]$ for every $H_i \in 2^X$. We define, the inner products of three-dimensional belief vectors, as a correlation of belief functions A and B by

$$K(A, B) = \sum_{i=1}^n [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)] \quad (16)$$

Obviously, the correlation of belief sets satisfies the following properties:

- $K(A, A) = I$
- $K(A, B) = K(B, A)$

Definition 4.3. A similarity measure on a frame of discernment 2^X is a real function; $S : A \times B \rightarrow [0, 1]$, where A and B are belief sets and the distance of $S(A, B)$ is a similarity measure that satisfies the following conditions:

- $0 \leq S(A, B) \leq 1$; nonnegatively
- $S(A, B) = S(B, A)$; symmetry
- $S(A, B) = 0$ if only if $A = B$ definiteness
- If $A \subseteq B \subseteq C$ then $S(A, B) \leq S(A, C)$ and $S(B, C) \leq S(A, C)$
- $S(A, C) \leq S(A, B) + S(B, C)$; triangle inequality

Definition 4.4. Let A and B are two belief sets in the frame of discernment, 2^X , $A = \{(bel_A(H_i), u_A(H_i), dis_A(H_i) | H_i \in 2^X)\}$ and $B = \{(bel_B(H_i), u_B(H_i), dis_B(H_i) | x \in 2^X)\}$. Then, the Jaccard similarity measure between A and B is defined by

$$J(A, B) = \frac{\sum_{i=1}^n [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)]}{\sqrt{bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i)} + \sqrt{bel_B^2(H_i) + u_B^2(H_i) + dis_B^2(H_i)}} - \sum_{i=1}^n [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)] \quad (17)$$

Definition 4.5. Let A and B are two belief sets, the Dice similarity between A and B is defined by

$$D(A, B) = \frac{2 \sum_{i=1}^n [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)]}{\sqrt{bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i)} + \sqrt{bel_B^2(H_i) + u_B^2(H_i) + dis_B^2(H_i)}} \quad (18)$$

Definition 4.6. Let A and B are two belief sets, the cosine similarity between A and B is defined as follows:

$$C(A, B) = \frac{\sum_{i=1}^n [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)]}{\sqrt{bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i)} \sqrt{bel_B^2(H_i) + u_B^2(H_i) + dis_B^2(H_i)}} \quad (19)$$

It is easy to see that the similarity measures in Eqs. (17–19) have the properties, which are presented in Definition 4.3.

In order to consider and account the weights of criteria or index, the weighted similarity measure can be developed as the following definition:

Definition 4.7. Let $w = \{w_1, w_2, \dots, w_n\}$ is a weight vector of criteria, thus the weighted similarity measure of the proposed similarity measure (i.e., Eqs. (17–19)) is determined by respectively:

$$J(X, Y) = \frac{\sum_{i=1}^n w_i [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)]}{\sqrt{\sum_{i=1}^n w_i [bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i)]} + \sqrt{\sum_{i=1}^n w_i [bel_B^2(H_i) + u_B^2(H_i) + dis_B^2(H_i)]}} - \sum_{i=1}^n w_i [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)] \quad (20)$$

$$D(X, Y) = \frac{2 \sum_{i=1}^n w_i [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)]}{\sqrt{\sum_{i=1}^n w_i [bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i)]} + \sqrt{\sum_{i=1}^n w_i [bel_B^2(H_i) + u_B^2(H_i) + dis_B^2(H_i)]}} \quad (21)$$

$$C(X, Y) = \frac{\sum_{i=1}^n [bel_A(H_i) bel_B(H_i) + u_A(H_i) u_B(H_i) + dis_A(H_i) dis_B(H_i)]}{\sqrt{\sum_{i=1}^n w_i [bel_A^2(H_i) + u_A^2(H_i) + dis_A^2(H_i)]} \sqrt{\sum_{i=1}^n w_i [bel_B^2(H_i) + u_B^2(H_i) + dis_B^2(H_i)]}} \quad (22)$$

If $w = \{\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\}$, then Eqs. (20–22) is reduced to Eqs. (17–19), respectively. It is easy to see that three weighted similarity measures between two belief sets A and B is belong to $[0, 1]$.

5. NUMERICAL IMPLEMENTATION AND DECISION-MAKING APPLICATION

5.1. Numerical Examples

In this paper, the meaning of the similarity/dissimilarity is a measurement deal with the strength of the relation or divergence between two belief sets. The type of the function to evaluate similarity measures depends on the intended application. Generally, several possible techniques to find a proper result in similarity computational methods are considered with respect to contents of data and its type. Then, using comparison results, a desirable measure can be selected to provide a utilized similarity function and discover a better and effective one. On the other hand, recognizing a similarity computation function is one of the fundamental problems upon different types of applications (e.g., clustering, classification and decision-making problem). However, the focuses of this section are to present a capability of the defined similarity measure as a numerical example and it is not to infer that. Thus, it is only assumed to evaluate general impact of similarity measures and their trends of three different similarity measure functions are compared.

Accordingly, suppose that a belief value set of alternatives is evaluated by experts with respect to criteria as listed in Table 2. Alternatives are denoted by A_i ($i = 1, 2, \dots, 15$) and evaluated by using the belief value set to define the belief degrees (i.e., truth, uncertainty and falsity) on A_i with respect to the criteria based on expert or multiple expert opinions. In this example, from the expert's point of view, the positive ideal value or ideal belief set for investment is alternative $A_7 = (0.8, 0.1, 0.1)$ and the negative ideal belief set is $A_{13} = (0.3, 0.2, 0.5)$. The similarities between each alternative and ideal point ($A^* = (0.8, 0.1, 0.1)$) can lead to identify the best selection of alternatives in the decision items. It provides a useful theoretical construct against which to evaluate alternatives.

Hence, by applying Eqs. (17–19), the correlation or similarity measure is evaluated between alternative A_i and ideal alternative A^* , which is represented as belief value set. Accordingly, the correlation or dependence between each alternative and ideal value is calculated. The results value is between 0 and 1, where 1 is total positive correlation and indicates a perfect positive correlation, as well as 0 is no correlation and it means that there is no relationship between two belief sets. Furthermore, the ranking in ascending order of similarity results (i.e., Jaccard, Dice and cosine) are determined in Table 1. The bigger similarity measure is shown the better alternative rather than the best one.

From the comparison of three similarity measures, it is easy to see that the influence of three correlations is similar, and the ranking order of the result is the same. Furthermore, the curve of the variation coefficient is drowned in Figure 1, the trend of the results for different similarity measures shows that the influence of three similarity measures is same. Also, comparing between the numerical results and the variation curve of three similarities shows a greater effect of the cosine coefficient.

According to analysis results using numerical examples and its application in a decision-making process, it is logically reasonable as an approximate reasoning which can be applied in many fields such as decision-making problems. It should be noted that the developmental approach of the similarity measure between two belief sets based on different implication similarity functions and comparison of its results shows that the proposed method is applicable and effective. Additionally, the practical example of decision-making problems, the weighted similarity measure between each alternative and the positive ideal value is used to rank the alternatives and to select desirable one. The decision numerical consequence using the different similarity measure shows the same results as $A_3 < A_5 < A_4 < A_1 < A_2$.

Table 2 | Comparison of three similarity measures (i.e., Jaccard, Dice and Cosine).

Alternatine	Jaccard Index	Ranking	Dice Index	Ranking	Cosine Index	Ranking
$A_1 = (0.6, 0.2, 0.2)$	0.89655	0.42465 (A_{13})	0.94545	0.59615 (A_{13})	0.96495	0.61901 (A_{13})
$A_2 = (0.5, 0.2, 0.3)$	0.76271	0.59385 (A_8)	0.86538	0.74509 (A_8)	0.89856	0.77957 (A_8)
$A_3 = (0.7, 0.1, 0.2)$	0.96721	0.61290 (A_6)	0.98333	0.76 (A_6)	0.98828	0.80218 (A_6)
$A_4 = (0.6, 0.2, 0.2)$	0.89655	0.71428 (A_5)	0.94545	0.83333 (A_5)	0.96495	0.85470 (A_5)
$A_5 = (0.5, 0.4, 0.1)$	0.71428	0.71428 (A_9)	0.83333	0.83333 (A_9)	0.85470	0.85470 (A_9)
$A_6 = (0.4, 0.3, 0.3)$	0.61290	0.76271 (A_{10})	0.76	0.86538 (A_{10})	0.80218	0.89856 (A_{10})
$A_7 = (0.8, 0.1, 0.1)$	1	0.76271 (A_{14})	1	0.86538 (A_{14})	1	0.89856 (A_{14})
$A_8 = (0.4, 0.2, 0.4)$	0.59375	0.76271 (A_2)	0.74509	0.86538 (A_2)	0.77957	0.89856 (A_2)
$A_9 = (0.5, 0.4, 0.1)$	0.71428	0.89655 (A_1)	0.83333	0.94545 (A_1)	0.85470	0.96495 (A_1)
$A_{10} = (0.5, 0.3, 0.2)$	0.76271	0.89655 (A_{12})	0.86538	0.94545 (A_{12})	0.89856	0.96495 (A_{12})
$A_{11} = (0.7, 0.2, 0.1)$	0.96721	0.89655 (A_{15})	0.98333	0.94545 (A_{15})	0.98828	0.96495 (A_{15})
$A_{12} = (0.6, 0.2, 0.2)$	0.89655	0.89655 (A_4)	0.94545	0.94545 (A_4)	0.96495	0.96495 (A_4)
$A_{13} = (0.3, 0.2, 0.5)$	0.42465	0.96721 (A_{11})	0.59615	0.98333 (A_{11})	0.61901	0.98828 (A_{11})
$A_{14} = (0.5, 0.2, 0.3)$	0.76271	0.96721 (A_3)	0.86538	0.98333 (A_3)	0.89856	0.98828 (A_3)
$A_{15} = (0.6, 0.2, 0.2)$	0.89655	1 (A_7)	0.94545	1 (A_7)	0.96495	1 (A_7)

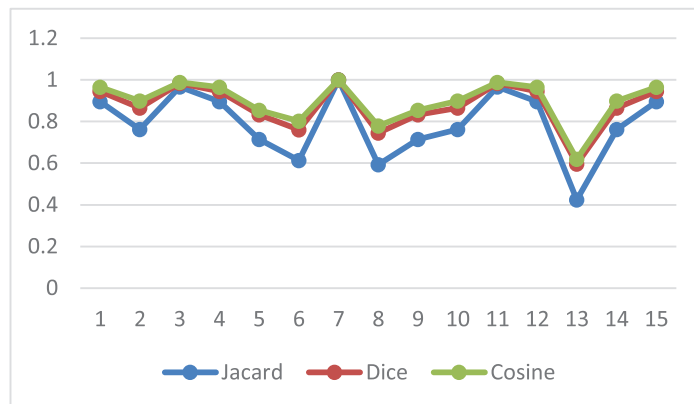


Figure 1 | Curve of three similarity measures (i.e., Jaccard, Dice and Cosine).

5.2. Application in Decision-Making

In the multi-decision-making environment, various term-term similarity measures based on the collocation have been suggested to select the additional search terms. The Jaccard, Dice and cosine similarity measures (Dice [43]; Salton and McGill [40]; Wang [44]; Kima and Choi [45]) are often used for this purpose. Ye [46,47] proposed a MCDM group method using vector similarity measure for trapezoidal fuzzy numbers; he extended the Jaccard, Dice and cosine similarity measures. In this section, a MCDM problem from a practical point of view is considered for the proposed method. Assuming a decision-making problem to select the best company to invest the sum of money with five alternatives, which is evaluated by three criteria given as profit analysis (C_1), loss analysis (C_2) and environment impact (C_3). The decision makers evaluate the five possible alternatives for investment under the three criteria. They approximate and calculate the relative belief degrees of each alternative with respect to criteria through express, acceptable, uncertain and falsity amounts for investment separately. Furthermore, suppose that the weights of criteria are $w = (0.4, 0.35, 0.25)$. The results are listed in the following matrix:

$$D_{5 \times 3} = \begin{bmatrix} (0.6, 0.2, 0.2) & (0.4, 0.3, 0.3) & (0.7, 0.2, 0.1) \\ (0.5, 0.2, 0.3) & (0.8, 0.1, 0.1) & (0.6, 0.2, 0.2) \\ (0.7, 0.1, 0.2) & (0.4, 0.2, 0.4) & (0.3, 0.2, 0.5) \\ (0.6, 0.2, 0.2) & (0.5, 0.4, 0.1) & (0.5, 0.2, 0.3) \\ (0.5, 0.4, 0.1) & (0.5, 0.3, 0.2) & (0.6, 0.2, 0.2) \end{bmatrix}$$

Generally, in order to obtain the most desirable alternatives using its ranking, a greater similarity between each value and positive ideal value can be helped to find the best alternative among a set of feasible alternatives. Although the ideal value or solution does not exist in the real world, it can be provided a useful construct to make a preference decision using prioritization the alternatives. Hence, alternative selection A_i with respect to criteria is based on an ideal belief set, which is $a_j^* = (bel^*, unc^*, dis^*) = (1, 0, 0)$. The weighted similarity measures of each belief set with respect to the positive ideal belief set (a_j^*) are calculated by using Eqs. (20–22) separately. The evaluated results are shown in Table 3.

It is easy to see that, the ranking orders of three similarity measures (i.e., Cosine, Jaccard and Dice) are the same and comparison of them is shown that the most desirable alternative for investing with respect to three criteria is the company A_2 . In addition, the decision results of different similarity measures based on new definition of the belief set are demonstrated that the proposed method under uncertainty and insufficient information is applicable and effective.

6. CONCLUSION

In the belief function theory, thanks advantages of the similarity measure in handling of existence conflict under incomplete, imprecise and uncertain information, a considerable number of methods were proposed about this issue. In this paper, based on the extension of the new definition of the belief set using the belief function theory as a three-dimensional vector, the informational energy of the belief set and different similarity correlation between two belief sets was proposed. Similarity measurements are elemental concepts in this paper and information retrieval that helps us to quantify the similarity between belief sets, which is effective in the improvement of the performance of MCDM and ranking alternatives. Nowadays, different types of similarity measures are available, but it is not explain that which similarity measure is more effective in detection the similarity of uncertain information's. The aim of this paper is to provide a comparative analysis of various belief sets based similarity measures such as Cosine similarity, Jaccard and Dice coefficient in order to evaluate. Therefore the Cosine, Jaccard and Dice similarity measures, as a special case of the correlation coefficient were demonstrated in this filed. Then, the simple and effective application of this method was established in decision-making problems. The results were illustrated by making comparisons between various types of the similarity measure. This paper could not be asserted, in which the proposed similarity measures were better or could be replaced by others. Contradicting, it was said that the existing similarity measures were practical and valuable. However, since requirements specification during the application procedure of the similarity measure was led to select different indices. Up to now, the conflict of evidence body was not directly evaluated in terms of a coefficient similarity measure in the classical belief function theory based on a belief degree that is presented by using available evidence, document or expert's opinions as truth, falsity and uncertain belief degrees. In the other words, no study does not pay attention to this drawback. Therefore, there is not only set up a new viewpoint that focuses

Table 3 | Numerical results of three similarity measures

Alternatives	$C(X, Y)$	$D(X, Y)$	$J(X, Y)$
A_1	0.748798	0.688935	0.525478
A_2	0.818914	0.744153	0.592551
A_3	0.835229	0.761636	0.615034
A_4	0.846367	0.776224	0.634286
A_5	0.897257	0.843938	0.730012
Ranking	$A_3 < A_5 <$ $A_4 < A_1 < A_2$	$A_3 < A_5 <$ $A_4 < A_1 < A_2$	$A_3 < A_5 <$ $A_4 < A_1 < A_2$

on a belief degree using a belief function, we consider specific applications and needs into decision-making practice. Furthermore, in the conducted several results using similarity coefficient between two belief sets from the Table 3 and Figure 1 it is clearly visible that similar behavior based on fitness values were obtained using the Cosine, Dice and Jaccard similarity.

CONFLICT OF INTEREST

There is no conflict of interest in this article.

AUTHORS' CONTRIBUTIONS

Authors wrote the initial draft of the paper and did the analysis. Then, both of them helped in analyzing the finding and supervised overall work.

Funding Statement

We have solely funded the research by ourself.

ACKNOWLEDGMENTS

The authors would like to sincerely thank the Editor-in-Chief, Associate Editor and referees for their useful comments.

REFERENCES

1. E. Lefevre, Z. Elouedi, *Decis. Support Syst.* 56 (2013), 326–333.
2. J.B. Yang, D.L. Xu, *Artif. Intell.* 205 (2013), 1–29.
3. J. Wang, F. Xiao, X. Deng, L. Fei, Y. Deng, *Int. J. Distrib. Sensor Netw.* 12 (2016), 1–10.
4. K.S. Chin, C. Fu, *Inf. Sci.* 325 (2015), 70–86.
5. L.G. de Oliveira Silva, A.T. de Almeida-Filho, *Inf. Sci.* 346 (2016), 275–285.
6. W. Jiang, M. Zhuang, X. Qin, Y. Tang, *SpringerPlus.* 5 (2016), 1–11.
7. X. Deng, D. Han, J. Dezert, Y. Deng, Y. Shyr, *IEEE Trans. Cybern.* 46 (2016), 2070–2082.
8. Y. Lin, C. Wang, C. Ma, Z. Dou, X. Ma, J. Supercomput. 72 (2016), 2874–2890.
9. Y. Yang, D. Han, *Knowl. Based Syst.* 94 (2016), 114–123.
10. Y. Zhao, R. Jia, P. Shi, *Inf. Sci.* 367 (2016), 125–142.
11. J. Schubert, *Int. J. Approx. Reason.* 48 (2008), 466–480.
12. Q. Ye, X. Wu, Z. Chen, *J. Electron.* 26 (2009), 18–23.
13. T. Denoeux, M.H. Masson, *IEEE Trans. Syst. Man Cybern. Part B Cybern.* 34 (2004), 95–109.
14. Y. Peng, H. Shen, Z. Hu, Y. Ma, *Int. J. Image Graph. Signal Process.* 3 (2011), 31–37.
15. D. Fixsen, R.P.S. Mahler, *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* 27 (1997), 96–104.
16. H. Guo, W. Shi, Y. Deng, *IEEE Trans. Syst. Man Cybern. Part B Cybern.* 36 (2006), 970–981.
17. C. Zhang, W. Zhu, S. Yang, in: *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing (WiCom 2007)*, Shanghai, China, 2007, pp. 4640–4644.
18. M.C. Florea, E. Bosse, A.L. Joussetme, in: *Cognitive Systems with Interactive Sensors (COGIS'09)*, Paris, France, Nov 2009, 16–18.
19. M.C. Florea, E. Bosse, in: *C3I in Crisis, Emergency and Consequence Management*, RTO-MP-IST-086, Bucharest, Romania, 2009.
20. P. Étude, *Bull. Soc. Vaudoise Sci. Nat.* 37 (1901), 547–579.
21. H. Zhang, Z. Liu, G. Chen, Y. Zhang, in: *Chinese Control and Decision Conference*, Yinchuan, China, 2016, pp. 2794–2798.
22. M. Ma, J. An, *J. Sensors.* 2015 (2015), 1–9.
23. Y. Song, X. Wang, L. Lei, A. Xue, in: *International Conference on Signal Processing*, Hangzhou, China, 2014, pp. 1392–1396.
24. P.Z. Wang, *Fuzzy Sets and its Applications*, Shanghai Science and Technology Press, Shanghai, China, 1983.
25. L. Yingfanga, Q. Keyunb, H. Xingxingb, M. Dana, *J. Intell. Fuzzy Syst.* 28 (2015), 2113–2125.
26. W. Zeng, Y. Zhao, Y. Gu, *Knowl. Based Syst.* 94 (2016), 124–131.
27. Ş. Ridvan, K. Ahmet, *J. Intell. Fuzzy Syst.* 27 (2014), 2417–2430.
28. B. Ismata, R. Tabasamb, *J. Intell. Fuzzy Syst.* 30 (2016), 821–829.
29. S.C. Ngan, *Expert Syst. Appl.* 60 (2016), 62–80.
30. F. Qinrong, G. Xiao, *Intell. Data Anal.* 21 (2017), 77–95.
31. L. Fei, J. Xia, Y. Feng, L. Liu, *IEEE Access.* 7 (2019), 84701–84716.
32. L. Fei, Y. Deng, Y. Hu, *Int. J. Fuzzy Syst.* 21 (2019), 157–175.

33. A.L. Jousselme, P. Maupin, *Int. J. Approx. Reason.* 53 (2012), 118–145.
34. M. Bouchard, A.-L. Jousselme, P.-E. Dore, *Int. J. Approx. Reason.* 54 (2013), 615–626.
35. Y. Song, X. Wang, H. Zhang, *Knowl. Based Syst.* 86 (2015), 288–298.
36. T. Feng, J.S. Mi, S.P. Zhang, *Inform. Sci.* 271 (2014), 143–158.
37. A.P. Dempster, *J. R. Stat. Soc.* 30 (1968), 205–247.
38. G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, USA, 1976.
39. D.T. Bui, B. Pradhan, O. Lofman, I. Revhaug, O.B. Dick, *CATENA*. 96 (2012), 28–40.
40. G. Salton, M.J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, New York, NY, USA, 1987.
41. O. Onicescu, *St. Cerc.* 18 (1966), 1419–1430.
42. O. Onicescu, *C. R. Acad. Sci. Paris Ser.* 263 (1966), 841–842.
43. L.R. Dice, *Ecology*. 26 (1945), 297–302.
44. W.J. Wang, *Fuzzy Sets Syst.* 85 (1997), 305–309.
45. M.C. Kima, K.S. Choi, *Inf. Process. Manag.* 35 (1999), 19–30.
46. J. Ye, *Math. Comput. Model.* 53 (2011), 91–97.
47. J. Ye, *Group Decis. Negat.* 21 (2012), 519–530.