Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 9975983, 9 pages
https://doi.org/10.1155/2021/9975983

Research Article

Consistency Fuzzy Sets and a Cosine Similarity Measure in Fuzzy Multiset Setting and Application to Medical Diagnosis

Ezgi Türkarslan,[1] **Jun Ye**,[2] **Mehmet Ünver**,[3] **and Murat Olgun**[3]

[1]*TED University, Faculty of Arts and Science, Department of Mathematics, Ankara 06420, Turkey*
[2]*Ningbo University, School of Civil and Environmental Engineering, 818 Fenghua Road, Jiangbei District, Ningbo, Zhejiang Provice, China*
[3]*Ankara University, Faculty of Science, Department of Mathematics, Ankara 06100, Turkey*

Correspondence should be addressed to Murat Olgun; olgun@ankara.edu.tr

Received 11 March 2021; Revised 5 May 2021; Accepted 24 May 2021; Published 11 June 2021

Academic Editor: Kamal Kumar

The main purpose of this study is to construct a base for a new fuzzy set concept that is called consistency fuzzy set (CFS) which expresses the multidimensional uncertain data quite successfully. Our motive is to reduce the complexity and difficulty caused by the information contained in the truth sequence in a fuzzy multiset (FMS) and to present the data of the truth sequence in a more understandable and compact manner. Therefore, this paper introduces the concept of CFS that is characterized with a truth function defined on a universal set $[0, 1]^2$. The first component of the truth pair of a CFS is the average value of the truth sequence of a FMS and the second component is the consistency degree, that is, the fuzzy complement of the standard deviation of the truth sequence of the same FMS. The main contribution of a CFS is the reflection of both the level of the average of the data that can be expressed with the different sequence lengths and the degree of the reasonable information in data via consistency degree. To develop this new concept, this paper also presents a correlation coefficient and a cosine similarity measure between CFSs. Furthermore, the proposed correlation coefficient and cosine similarity measure are applied to a multiperiod medical diagnosis problem. Finally, a comparison analysis is given between the obtained results and the existing results in literature to show the efficiency and rationality of the proposed correlation coefficient and cosine similarity measure.

1. Introduction

Fuzzy set theory was introduced by Zadeh [1] in 1965 with the help of the concept of membership (truth) function that is used as an effective tool to overcome uncertainty in science, and it has applications in many different fields such as economics, engineering, decision-making, management, and medicine [2–4]. There are many generalizations of the concept of the fuzzy set in the literature, and their applications to several areas such as decision-making and medical diagnosis are studied to model uncertain data that is encountered in science often. For example, Akram et al. [5] have proposed a new decision-making method in complex spherical fuzzy environment and Das et al. [6] have introduced a medical diagnosis model by using fuzzy logic and intuitionistic fuzzy logic. Moreover, a decision-making method, for the selection of an effective sanitizer to reduce COVID-19 which is one of the most up-to-date problems of recent times, has been presented in [7]. One of the generalizations of fuzzy sets is the concept of hesitant fuzzy set (HFS) [8], which is characterized by a membership (truth) function that is a set of crisp values in $[0, 1]$. A HFS can model uncertain data better than a fuzzy set, thanks to its handy structure, so it has been frequently preferred by researchers to solve multicriteria (group) decision-making or multiperiod medical diagnosis problems [9–12]. However, the concept of HFS eliminates and ignores repetitive information because of the nature of the crisp sets. For example, suppose that a doctor evaluates a target patient's symptoms at four different times with membership degrees

0.1, 0.3, 0.7, and 0.3, respectively. If the result of this evaluation is expressed as a HFS, then the repetitive 0.3 assessment is lost due to the formation structure of the HFS. In such a situation, the concept of fuzzy multiset (FMS) is a useful method to express the ambiguous information which is lost.

The concept of FMS was proposed by Yager in 1986 [13, 14] with the help of a count function. In a fuzzy multiset setting, the membership degrees of elements in a universal set are presented as a sequence having different sequence lengths/cardinalities with the same or different fuzzy values. Therefore, more accurate results can be obtained by preventing the loss of the repetitive information. Moreover, it is more appropriate to use this fuzzy set in solving multicriteria group decision-making problems and multiperiod medical diagnosis problems. Although FMSs have the property of saving repetitive information, the uncertainty increases as the length of the sequences in the FMSs increases. This situation causes a difficulty while expressing reasonable information and complicates the selection of the alternative in a decision-making problem. To make the information carried by the sequence in the FMS more understandable and to reduce the dependence of this information on the length of the sequence, some statistical methods such as arithmetic mean and standard deviation for the elements of this sequence can be used. Recently, Ye et al. [15] have used this idea in neutrosophic environment. Motivated from this, we propose a new concept which is called consistency fuzzy set (CFS). This concept is expressed as an ordered pair whose components are the average value and the consistency degree of the sequence, respectively. Later, we propose a correlation coefficient and a cosine similarity measure between CFSs.

Correlation analysis is an important research issue in the fuzzy set theory and in its generalizations because it can measure the relationship between two fuzzy sets. Therefore, they have gained attention from researchers and their wide applications in various fields have been considered. For instance, Ye [16] has proposed a weighted correlation coefficient between intuitionistic fuzzy sets. Moreover, Guan et al. [17] have put forward a synthetic correlation coefficient between HFSs. Recently, Lin et al. [18] have developed the directional correlation coefficient measures for Pythagorean fuzzy information and have applied them to the medical diagnosis and the cluster analysis. Also, several researchers have proposed some correlation coefficients in various fuzzy environments (see, e.g., [19, 20]).

The concept of similarity measure plays an important role to determine the degree of similarity between two fuzzy sets. There are several types of similarity measures in the literature (see, e.g., [21–25]). The concept of cosine similarity measure is one of them, and it is defined as the inner product of two vectors divided by the product of their lengths, that is, the cosine of the angle between the vector representations of fuzzy sets [26]. In this paper, we introduce a correlation coefficient and a cosine similarity measure between CFSs, and we give the multiperiod medical diagnosis approaches by using the proposed correlation coefficient and cosine similarity measure to show the efficiency of these new concepts.

The important contributions of the paper are listed below:

(i) The concept of CFS reduces the dependence of information on the length of the sequence in FMS and presents the information carried by the sequence in FMS in a more compact form.

(ii) A CFS that is based on the average values and the consistency degree can give reasonable information about sequences in a FMS.

(iii) A CFS contains both the level of the average of the data that can be expressed with different sequence lengths and the degree of consistency of the data via fuzzy complement of standard deviation of a sequence in FMS.

(iv) A CFS facilitates the understanding of the problem, so the decision-making process has compact information due to the ability of CFSs.

(v) The proposed correlation coefficient and cosine similarity measure between CFSs provide useful ranking method, and they are beneficial mathematical tools for multiperiod medical diagnosis and multicriteria group decision-making problems in the FMS environment.

(vi) The developed medical diagnosis approach not only improves the decision-making reliability but also supplies a new influential way for multiperiod medical diagnosis problems in the FMS environment. The remainder of this paper is set out as follows. In Section 2, we introduce the concept of CFS and we give a correlation coefficient between CFSs. Later, we apply it to a multiperiod medical diagnosis problem to demonstrate the efficiency of the proposed correlation coefficient. In Section 3, we propose a cosine similarity measure between CFSs. Then, we apply it to the same multiperiod medical diagnosis problem. Moreover, we compare the results of the proposed correlation coefficient and the proposed cosine similarity measure with each other and the existing results in literature. In Section 4, we give a conclusion with some remarks.

2. CFSs and a Correlation Coefficient between CFSs

In this section, we recall the concepts of FMS and a correlation coefficient between FMSs. Then, we introduce the concept of CFS and a correlation coefficient between CFSs. Next, we apply it to a multiperiod medical diagnosis problem.

2.1. The Concept of CFS

Definition 1 (see [14]). Let $X = \{x_1, \ldots, x_m\}$ be a finite set. A FMS A in X is characterized by a count membership

function count_A such that $\text{count}_A\colon X \longrightarrow Q$, where Q is the set of all crisp multisets in $[0,1]$. The membership (truth) sequence is defined as $(\mu_A^1(x_j), \mu_A^2(x_j), \ldots, \mu_A^{n_j}(x_j))$ such that $\mu_A^1(x_j) \geq \mu_A^2(x_j) \geq \cdots \geq \mu_A^{n_j}(x_j)$, for $j = 1, \ldots, m$. Therefore, a FMS A is given by

$$A = \left\{\left\langle x_j, \left(\mu_A^1\left(x_j\right), \mu_A^2\left(x_j\right), \ldots, \mu_A^{n_j}\left(x_j\right)\right)\right\rangle\colon\ j = 1, \ldots, m\right\}, \tag{1}$$

where n_j is the length of the sequence for jth element. Obviously, a FMS reduces to a fuzzy set when $n_j = 1$.

Now, we define the concept of CFS which reduces the dependence of information on the length of the sequence in a FMS and to present the information carried by the sequence in a FMS in a more compact form.

Definition 2. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let A be a FMS in X. Average values and consistency degrees of the membership (truth) sequences in A are defined by

$$m_A^{(j)} = \frac{1}{n_j}\sum_{k=1}^{n_j}\mu_A^k\left(x_j\right), \tag{2}$$

$$c_A^{(j)} = 1 - \sigma_A^{(j)}, \tag{3}$$

for each $x_j \in X$ $(j = 1, \ldots, m)$, respectively, where $\sigma_A^{(j)}$ is the standard deviation of the jth membership (truth) sequence in FMS A. A CFS C_A is defined by

$$C_A = \left\{\left\langle x_j, \left(m_A^{(j)}, c_A^{(j)}\right)\right\rangle\colon\ x_j \in X\right\}. \tag{4}$$

Moreover, the consistency fuzzy element (CFE) in CFS C_A is simply denoted as $a_j = (m_A^{(j)}, c_A^{(j)})$, for each $j = 1, \ldots, m$.

Example 1. Let $X = \{x_1, x_2, x_3\}$ be a finite set and let A be the FMS in X defined by

$$A = \left\{\left\langle x_1, (0.7, 0.3)\right\rangle, \left\langle x_2, (0.5, 0.2, 0.1)\right\rangle, \left\langle x_3, (0.9)\right\rangle\right\}. \tag{5}$$

Then, we construct the corresponding CFS C_A to FMS A by

$$\begin{aligned} C_A &= \left\{\left\langle x_1, (0.5, 1 - 0.2828)\right\rangle, \left\langle x_2, (0.4, 1 - 0.2345)\right\rangle, \left\langle x_3, (0.9, 1 - 0)\right\rangle\right\} \\ &= \left\{\left\langle x_1, (0.5, 0.7172)\right\rangle, \left\langle x_2, (0.4, 0.7655)\right\rangle, \left\langle x_3, (0.9, 1)\right\rangle\right\}, \end{aligned} \tag{6}$$

by using (2) and (3).

By using CFSs, we make a statistical inference for the information carried by the truth sequences in a FMS, and we express the information presented in these sequences as a compact and understandable way. Thus, we simplify the decision-making process by reducing the complexity created by the length of the truth sequences in a FMS. We also eliminate the dependence of the information on the length of these truth sequences in a FMS.

The fuzzy set theory has been often preferred by researchers especially to solve real-life problems such as medical diagnosis and decision-making, since it can model uncertain information very well. While solving these problems, the optimal choice is usually determined by using an aggregation functions or information measures such as similarity measures, entropy measures, and divergence measures, after the uncertainty in the environment is modeled with fuzzy sets. The concept of correlation coefficient is a crucial measure that determines the relationship between two fuzzy sets. Now, we recall a correlation coefficient for FMSs.

Definition 3 (see [27]). Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let

$$\begin{aligned} A &= \left\{\left\langle x_j, \left(\mu_A^1\left(x_j\right), \mu_A^2\left(x_j\right), \ldots, \mu_A^{n_j}\left(x_j\right)\right)\right\rangle\colon\ j = 1, \ldots, m\right\}, \\ B &= \left\{\left\langle x_j, \left(\mu_B^1\left(x_j\right), \mu_B^2\left(x_j\right), \ldots, \mu_B^{n_j}\left(x_j\right)\right)\right\rangle\colon\ j = 1, \ldots, m\right\}, \end{aligned} \tag{7}$$

be two FMSs in X. A correlation coefficient between A and B is given with

$$\rho_{\text{FMS}}(A, B) \coloneqq \frac{C_{\text{FMS}}(A, B)}{\sqrt{C_{\text{FMS}}(A, A) C_{\text{FMS}}(B, B)}}, \tag{8}$$

where

$$C_{\text{FMS}}(A, B) = \frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{n_j}\mu_A^k\left(x_j\right)\mu_B^k\left(x_j\right). \tag{9}$$

Proposition 1 (see [27]). *The correlation coefficient ρ_{FMS} satisfies the following properties:*

($\mathbf{P}_1$) $0 \leq \rho_{FMS}(A, B) \leq 1$

($\mathbf{P}_2$) $\rho_{FMS}(A, B) = \rho_{FMS}(B, A)$

($\mathbf{P}_3$) *If $A = B$, then* $\rho_{FMS}(A, B) = 1$

Now, we propose a correlation coefficient between CFSs by motivating from the definition of the correlation coefficients between FMSs.

Definition 4. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let A and B be two FMSs in X. The correlation coefficient between CFSs C_A and C_B is given with

$$\rho_{\mathrm{CFS}}(A,B) \coloneqq \frac{C_{\mathrm{CFS}}(A,B)}{\sqrt{C_{\mathrm{CFS}}(A,A)C_{\mathrm{CFS}}(B,B)}}, \tag{10}$$

where

$$C_{\mathrm{CFS}}(A,B) = \sum_{j=1}^{m}\left(m_A^{(j)}m_B^{(j)} + c_A^{(j)}c_B^{(j)}\right). \tag{11}$$

Proposition 2. *The correlation coefficient ρ_{CFS} satisfies the following properties:*

$(\mathbf{P}_1)$ $0 \le \rho_{CFS}(A,B) \le 1$

$(\mathbf{P}_2)$ $\rho_{CFS}(A,B) = \rho_{CFS}(B,A)$

$(\mathbf{P}_3)$ *If $A = B$, then* $\rho_{CFS}(A,B) = 1$

Proof. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let $a_j = (m_A^{(j)}, c_A^{(j)})$ and $b_j = (m_B^{(j)}, c_B^{(j)})$ be two CFEs in X for a fixed $j \in \{1, \ldots, m\}$. From Schwarz inequality, we obtain

$$m_A^{(j)}m_B^{(j)} + c_A^{(j)}c_B^{(j)} \le \sqrt{\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2}\sqrt{\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2}. \tag{12}$$

Thus, we have

$$\sum_{j=1}^{m}\left(m_A^{(j)}m_B^{(j)} + c_A^{(j)}c_B^{(j)}\right) \le \sum_{j=1}^{m}\sqrt{\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2}\sqrt{\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2}. \tag{13}$$

Now, using Cauchy Schwarz inequality, we have

$$\sum_{j=1}^{m}\left(m_A^{(j)}m_B^{(j)} + c_A^{(j)}c_B^{(j)}\right) \le \sqrt{\sum_{j=1}^{m}\left(\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2\right)}\sqrt{\sum_{j=1}^{m}\left(\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2\right)}. \tag{14}$$

Then, we obtain

$$\rho_{\mathrm{CFS}}(A,B) = \frac{\sum_{j=1}^{m}\left(m_A^{(j)}m_B^{(j)} + c_A^{(j)}c_B^{(j)}\right)}{\sqrt{\sum_{j=1}^{m}\left(\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2\right)}\sqrt{\sum_{j=1}^{m}\left(\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2\right)}} \le 1. \tag{15}$$

The proofs of $(\mathbf{P}_2)$ and $(\mathbf{P}_3)$ are straightforward. □

Now, we propose a weighted version of the correlation coefficient ρ_{CFS} for CFSs as follows.

Definition 5. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let A and B be two FMSs in X. A weighted correlation coefficient between CFSs C_A and C_B is given with

$$W\rho_{\mathrm{CFS}}(A,B) \coloneqq \frac{\mathrm{WC}_{\mathrm{CFS}}(A,B)}{\sqrt{\mathrm{WC}_{\mathrm{CFS}}(A,A)\mathrm{WC}_{\mathrm{CFS}}(B,B)}}, \tag{16}$$

where

$$\mathrm{WC}_{\mathrm{CFS}}(A,B) = \sum_{j=1}^{m}\omega_j\left(m_A^{(j)}m_B^{(j)} + c_A^{(j)}c_B^{(j)}\right), w = (w_1, \ldots, w_m). \tag{17}$$

w is the weight vector with $w_j \in [0,1]$, for all $j = 1, \ldots, m$, such that $\sum_{j=1}^{m} w_j = 1$.

2.2. An Application. A multiperiod medical diagnosis is a process of decision-making on a disease which has a target patient. In this process, the decision maker evaluates the effect of symptoms on the target patient several different times. The most important factor that discriminates this process from other medical diagnosis processes is the presentation of the solution algorithm that pays attention to the time variable [24]. Therefore, it can be convenient to present the patient's symptoms and diseases with the help of a sequence of fuzzy values.

Now, we adopt an illustrative example from [27] to show the applicability and effectiveness of the proposed correlation coefficient under FMS setting.

Example 2. Let $P = \{P_1, P_2, P_3, P_4\}$ be a set of patients and let

$$\begin{aligned} Q &= \{Q_1(\mathrm{Viralfever}), Q_2(\mathrm{Tuberculosis}), Q_3(\mathrm{Typhoid}), Q_4(\mathrm{ThroatDisease})\}, \\ S &= \{s_1(\mathrm{Temperature}), s_2(\mathrm{Cough}), s_3(\mathrm{Throatpain}), s_4(\mathrm{Headache}), s_5(\mathrm{Bodypain})\}, \end{aligned} \tag{18}$$

be sets of disease and symptoms, respectively. Suppose that all patients are examined at different time intervals with respect to all the symptoms and they are represented by the following FMSs:

$$\begin{aligned}
P_1 &= \{\langle s_1, (0.6, 0.7, 0.5)\rangle, \langle s_2, (0.4, 0.3, 0.4)\rangle, \langle s_3, (0.1, 0.2, 0.0)\rangle, \langle s_4, (0.5, 0.6, 0.7)\rangle, \langle s_5, (0.2, 0.3, 0.4)\rangle\},\\
P_2 &= \{\langle s_1, (0.4, 0.3, 0.5)\rangle, \langle s_2, (0.7, 0.6, 0.8)\rangle, \langle s_3, (0.6, 0.5, 0.4)\rangle, \langle s_4, (0.3, 0.6, 0.2)\rangle, \langle s_5, (0.8, 0.7, 0.5)\rangle\},\\
P_3 &= \{\langle s_1, (0.1, 0.2, 0.1)\rangle, \langle s_2, (0.3, 0.2, 0.1)\rangle, \langle s_3, (0.8, 0.7, 0.8)\rangle, \langle s_4, (0.3, 0.2, 0.2)\rangle, \langle s_5, (0.4, 0.3, 0.2)\rangle\},\\
P_4 &= \{\langle s_1, (0.3, 0.2, 0.2)\rangle, \langle s_2, (0.4, 0.3, 0.1)\rangle, \langle s_3, (0.2, 0.1, 0.0)\rangle, \langle s_4, (0.5, 0.6, 0.3)\rangle, \langle s_5, (0.4, 0.5, 0.4)\rangle\}.
\end{aligned} \tag{19}$$

Moreover, assume that each disease Q_i, for $i = 1, 2, 3, 4$, is given as a FMS with respect to all of the symptoms as follows:

$$\begin{aligned}
Q_1 &= \{\langle s_1, (0.8, 0.9, 0.85)\rangle, \langle s_2, (0.2, 0.3, 0.25)\rangle, \langle s_3, (0.3, 0.5, 0.4)\rangle, \langle s_4, (0.5, 0.7, 0.6)\rangle, \langle s_5, (0.5, 0.6, 0.55)\rangle\},\\
Q_2 &= \{\langle s_1, (0.2, 0.3, 0.25)\rangle, \langle s_2, (0.9, 0.1, 0.95)\rangle, \langle s_3, (0.7, 0.8, 0.75)\rangle, \langle s_4, (0.6, 0.7, 0.65)\rangle, \langle s_5, (0.7, 0.8, 0.75)\rangle\},\\
Q_3 &= \{\langle s_1, (0.5, 0.7, 0.6)\rangle, \langle s_2, (0.3, 0.5, 0.4)\rangle, \langle s_3, (0.2, 0.3, 0.25)\rangle, \langle s_4, (0.2, 0.4, 0.3)\rangle, \langle s_5, (0.4, 0.6, 0.5)\rangle\},\\
Q_4 &= \{\langle s_1, (0.1, 0.3, 0.2)\rangle, \langle s_2, (0.3, 0.4, 0.35)\rangle, \langle s_3, (0.8, 0.9, 0.85)\rangle, \langle s_4, (0.1, 0.2, 0.15)\rangle, \langle s_5, (0.1, 0.2, 0.15)\rangle\}.
\end{aligned} \tag{20}$$

Now, we construct CFSs. Firstly, all patients in P are expressed as CFSs P_1^*, P_2^*, P_3^*, and P_4^* as follows:

$$\begin{aligned}
P_1^* &= \{\langle s_1, (0.6, 0.9)\rangle, \langle s_2, (0.36, 0.9423)\rangle, \langle s_3, (0.1, 0.9)\rangle, \langle s_4, (0.6, 0.9)\rangle, \langle s_5, (0.3, 0.9)\rangle\},\\
P_2^* &= \{\langle s_1, (0.4, 0.9)\rangle, \langle s_2, (0.7, 0.9)\rangle, \langle s_3, (0.5, 0.9)\rangle, \langle s_4, (0.36, 0.7919)\rangle, \langle s_5, (0.66, 0.8473)\rangle\},\\
P_3^* &= \{\langle s_1, (0.13, 0.9423)\rangle, \langle s_2, (0.2, 0.9)\rangle, \langle s_3, (0.76, 0.9423)\rangle, \langle s_4, (0.23, 0.9423)\rangle, \langle s_5, (0.3, 0.9)\rangle\},\\
P_4^* &= \{\langle s_1, (0.23, 0.9423)\rangle, \langle s_2, (0.26, 0.8473)\rangle, \langle s_3, (0.1, 0.9)\rangle, \langle s_4, (0.46, 0.8473)\rangle, \langle s_5, (0.43, 0.9423)\rangle\},
\end{aligned} \tag{21}$$

respectively, and all diseases in Q are expressed as CFSs Q_1^*, Q_2^*, Q_3^*, and Q_4^* as follows:

$$\begin{aligned}
Q_1^* &= \{\langle s_1, (0.85, 0.95)\rangle, \langle s_2, (0.25, 0.95)\rangle, \langle s_3, (0.4, 0.9)\rangle, \langle s_4, (0.6, 0.9)\rangle, \langle s_5, (0.55, 0.95)\rangle\},\\
Q_2^* &= \{\langle s_1, (0.25, 0.95)\rangle, \langle s_2, (0.65, 0.5231)\rangle, \langle s_3, (0.75, 0.95)\rangle, \langle s_4, (0.65, 0.95)\rangle, \langle s_5, (0.75, 0.95)\rangle\},\\
Q_3^* &= \{\langle s_1, (0.6, 0.9)\rangle, \langle s_2, (0.4, 0.9)\rangle, \langle s_3, (0.25, 0.95)\rangle, \langle s_4, (0.3, 0.9)\rangle, \langle s_5, (0.5, 0.9)\rangle\},\\
Q_4^* &= \{\langle s_1, (0.2, 0.9)\rangle, \langle s_2, (0.35, 0.95)\rangle, \langle s_3, (0.85, 0.95)\rangle, \langle s_4, (0.15, 0.95)\rangle, \langle s_5, (0.15, 0.95)\rangle\},
\end{aligned} \tag{22}$$

respectively. Let the weight of each symptom be $\omega_j = 0.2$, for $j = 1, 2, 3, 4, 5$. Now, we apply the proposed weighted correlation coefficient $W\rho_{\mathrm{CFS}}$ to determine the optimal disease for each patient. New results obtained in this study and some existing results in [27] are given in Table 1.

The process of assigning each patient P_k^* to a disease Q_i^* is described by

$$t = \arg\max_{1\le i\le 4}\{W\rho_{\mathrm{CFS}}(P_k^*, Q_i^*)\}, \tag{23}$$

for fixed $k \in \{1, 2, 3, 4\}$.

The numerical results in Table 1 show that third and fourth patients suffer from throat disease and typhoid, respectively, according to both correlation coefficients for FMSs [27] and the proposed correlation coefficient for CFSs. The rest of Table 1 is different for two approaches. The novelty of the approach used in this study may cause this difference.

3. A Cosine Similarity Measure for CFSs

3.1. A Cosine Similarity Measure. The concept of cosine similarity measure is defined as the inner product of two vectors divided by the product of their lengths. In other words, a cosine similarity measure is the cosine of the angle between the vector representations of the two fuzzy sets.

Table 1: Comparison analysis of the results of ρ_{FMS} and $W\rho_{\mathrm{CFS}}$.

	Correlation coefficient	Viral fever	Tuberculosis disease	Typhoid disease	Throat disease
The results in [27]	$\rho_{\mathrm{FMS}}(P_1, Q_i)$	0.929	0.757	0.899	0.403
	$\rho_{\mathrm{FMS}}(P_2, Q_i)$	0.822	0.951	0.897	0.747
	$\rho_{\mathrm{FMS}}(P_3, Q_i)$	0.644	0.872	0.641	0.943
	$\rho_{\mathrm{FMS}}(P_4, Q_i)$	0.855	0.835	0.895	0.480
The results of the proposed weighted correlation coefficient	$W\rho_{\mathrm{CFS}}(P_1^*, Q_i^*)$	0.982	0.910	0.984	0.909
	$W\rho_{\mathrm{CFS}}(P_2^*, Q_i^*)$	0.956	0.969	0.977	0.941
	$W\rho_{\mathrm{CFS}}(P_3^*, Q_i^*)$	0.925	0.935	0.944	0.994
	$W\rho_{\mathrm{CFS}}(P_4^*, Q_i^*)$	0.959	0.928	0.978	0.925

Now, we introduce a cosine similarity measure and its weighted version for CFSs by motivating from [26] as follows.

Definition 6. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let A and B be two FMSs in X. A cosine similarity measure between CFSs C_A and C_B is given with

$$\delta_{\mathrm{CFS}}(A, B) := \frac{1}{m}\sum_{j=1}^{m} \frac{m_A^{(j)} m_B^{(j)} + c_A^{(j)} c_B^{(j)}}{\sqrt{\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2}\sqrt{\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2}}. \tag{24}$$

If we take $m = 1$, the cosine similarity measure δ_{CFS} reduces the correlation coefficient ρ_{CFS}, i.e., $\delta_{\mathrm{CFS}}(A, B) = \rho_{\mathrm{CFS}}(A, B)$.

Proposition 3. *The cosine similarity measure δ_{CFS} satisfies the following properties:*

($\mathbf{P}_1$) $0 \le \delta_{CFS}(A, B) \le 1$

($\mathbf{P}_2$) $\delta_{CFS}(A, B) = \delta_{CFS}(B, A)$

($\mathbf{P}_3$) *If $A = B$, then* $\delta_{CFS}(A, B) = 1$

Proof. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let $(m_A^{(j)}, c_A^{(j)})$ and $(m_B^{(j)}, c_B^{(j)})$ be two CFEs in X. Then, we have

$$\frac{m_A^{(j)} m_B^{(j)} + c_A^{(j)} c_B^{(j)}}{\sqrt{\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2}\sqrt{\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2}} = \cos\,\theta_j, \tag{25}$$

where θ_j be the radian measure of the angle between $(m_A^{(j)}, c_A^{(j)})$ and $(m_B^{(j)}, c_B^{(j)})$. Therefore, ($\mathbf{P}_1$) is true. ($\mathbf{P}_2$) and ($\mathbf{P}_3$) are trivial. □

Now, we introduce the weighted version of the proposed cosine similarity measure δ_{CFS} between CFSs.

Definition 7. Let $X = \{x_1, \ldots, x_m\}$ be a finite set and let A and B be two FMSs in X. A weighted cosine similarity measure between CFSs C_A and C_B is given with

$$W\delta_{\mathrm{CFS}}(A, B) := \sum_{j=1}^{m} \omega_j \frac{m_A^{(j)} m_B^{(j)} + c_A^{(j)} c_B^{(j)}}{\sqrt{\left(m_A^{(j)}\right)^2 + \left(c_A^{(j)}\right)^2}\sqrt{\left(m_B^{(j)}\right)^2 + \left(c_B^{(j)}\right)^2}}, \tag{26}$$

where $w = (w_1, \ldots, w_m)$ is the weight vector with $w_j \in [0, 1]$, for all $j = 1, \ldots, m$, such that $\sum_{j=1}^{m} w_j = 1$.

It is clear that if we take $w_j = 1/m$, for any $j = 1, \ldots, m$, then $\delta_{\mathrm{CFS}}(A, B) = W\delta_{\mathrm{CFS}}(A, B)$. Obviously, the proposed weighted cosine similarity measure $W\delta_{\mathrm{CFS}}(A, B)$ also satisfies the properties $P_1 - P_3$.

3.2. An Application. Now, we examine the same multiperiod medical diagnosis problem which is adapted from [27] to illustrate the applicability and effectiveness of the proposed cosine similarity measure for CFSs under the FMS setting. For this aim, we use CFSs for all of the patients and all diseases in Example 2.

Example 3. Let the weight of each symptom be $\omega_j = 0.2$ for each $j = 1, 2, 3, 4, 5$. Now, we apply the proposed weighted cosine similarity measure $W\rho_{\mathrm{CFS}}$ to determine the optimal disease for all patients. New results obtained in this study and some existing results in [27] are given in Table 2.

The process of assigning each patients P_k^* to a disease Q_i^* is described by

$$t = \arg\max_{1 \le i \le 4}\{W\delta_{\mathrm{CFS}}(P_k^*, Q_i^*)\}, \tag{27}$$

for fixed $k \in \{1, 2, 3, 4\}$.

The results in Table 2 show that second, third, and fourth patients suffer from tuberculosis, throat disease, and typhoid, respectively, according to both the correlation coefficient in [27] and the proposed cosine similarity measure in this study. The rest of Table 2 is different for two approaches. The novelty of the approach used in this study may cause this difference.

The results in Table 3 show that first and fourth patients suffer from typhoid whereas the third patient suffers from throat disease according to both the proposed correlation coefficient and the proposed cosine similarity measure.

Table 2: Comparison analysis of the results of ρ_{FMS} and $W\delta_{\mathrm{CFS}}$.

	Correlation coefficient	Viral fever	Tuberculosis disease	Typhoid disease	Throat disease
The results in [27]	$\rho_{\mathrm{FMS}}(P_1, Q_i)$	0.929	0.757	0.899	0.403
	$\rho_{\mathrm{FMS}}(P_2, Q_i)$	0.822	0.951	0.897	0.747
	$\rho_{\mathrm{FMS}}(P_3, Q_i)$	0.644	0.872	0.641	0.943
	$\rho_{\mathrm{FMS}}(P_4, Q_i)$	0.855	0.835	0.895	0.480
	Similarity measure	Viral fever	Tuberculosis disease	Typhoid disease	Throat disease
The results of the proposed weighted cosine similarity measure	$W\delta_{\mathrm{CFS}}(P_1^*, Q_i^*)$	0.9833	0.9196	0.9870	0.9282
	$W\delta_{\mathrm{CFS}}(P_2^*, Q_i^*)$	0.969	0.9864	0.9815	0.9493
	$W\delta_{\mathrm{CFS}}(P_3^*, Q_i^*)$	0.9427	0.9299	0.9544	0.9937
	$W\delta_{\mathrm{CFS}}(P_4^*, Q_i^*)$	0.9650	0.9284	0.9806	0.9436

Table 3: Comparison analysis of the results of $W\rho_{\mathrm{FMS}}$ and $W\delta_{\mathrm{CFS}}$.

	Correlation coefficient	Viral fever	Tuberculosis disease	Typhoid disease	Throat disease
The results of the proposed weighted correlation coefficient	$W\rho_{\mathrm{CFS}}(P_1^*, Q_i^*)$	0.982	0.910	0.984	0.909
	$W\rho_{\mathrm{CFS}}(P_2^*, Q_i^*)$	0.956	0.969	0.977	0.941
	$W\rho_{\mathrm{CFS}}(P_3^*, Q_i^*)$	0.925	0.935	0.944	0.994
	$W\rho_{\mathrm{CFS}}(P_4^*, Q_i^*)$	0.959	0.928	0.978	0.925
	Similarity measure	Viral fever	Tuberculosis disease	Typhoid disease	Throat disease
The results of the proposed weighted cosine similarity measure	$W\delta_{\mathrm{CFS}}(P_1^*, Q_i^*)$	0.9833	0.9196	0.9870	0.9282
	$W\delta_{\mathrm{CFS}}(P_2^*, Q_i^*)$	0.969	0.9864	0.9815	0.9493
	$W\delta_{\mathrm{CFS}}(P_3^*, Q_i^*)$	0.9427	0.9299	0.9544	0.9937
	$W\delta_{\mathrm{CFS}}(P_4^*, Q_i^*)$	0.9650	0.9284	0.9806	0.9436

Table 4: Decision results and standard deviations for the proposed weighted correlation coefficient and weighted cosine similarity measure.

	Patients	Standard deviation	Ranking order	Best selection
ρ_{CFS}	P_1^*	0.04244	$Q_3^* \succ Q_1^* \succ Q_2^* \succ Q_4^*$	Q_3^*
	P_2^*	0.01575	$Q_3^* \succ Q_2^* \succ Q_1^* \succ Q_4^*$	Q_3^*
	P_3^*	0.03066	$Q_4^* \succ Q_3^* \succ Q_2^* \succ Q_1^*$	Q_4^*
	P_4^*	0.02548	$Q_3^* \succ Q_1^* \succ Q_2^* \succ Q_4^*$	Q_3^*
δ_{CFS}	P_1^*	0.03556	$Q_3^* \succ Q_1^* \succ Q_4^* \succ Q_2^*$	Q_3^*
	P_2^*	0.01654	$Q_2^* \succ Q_3^* \succ Q_1^* \succ Q_4^*$	Q_2^*
	P_3^*	0.02756	$Q_4^* \succ Q_3^* \succ Q_1^* \succ Q_2^*$	Q_4^*
	P_4^*	0.02303	$Q_3^* \succ Q_1^* \succ Q_4^* \succ Q_2^*$	Q_3^*

3.3. Comparison Analysis of the Proposed Two Approaches. In this section, firstly, we compare the results of the proposed correlation coefficient ρ_{CFS} with the results of the proposed cosine similarity measure δ_{CFS} by using standard deviation of the obtained results. Then, we explain the advantage of two approaches. The numerical results in Table 4 show that the best selections of these two approaches are consistent with each other for patients P_1^*, P_3^*, and P_4^*. However, we know that larger standard deviations show higher determination due to larger difference in calculation values, while smaller standard deviations show smaller determination. Therefore, we look at the standard deviations for patients P_1^*, P_3^*, and P_4^* in both approaches. The standard deviation of the results of ρ_{CFS} is greater than the standard deviation of the results of δ_{CFS} except for patient P_2^*. In this case, ρ_{CFS} has higher ability to determine the disease of the patients P_1^*, P_3^*, and P_4^* than δ_{CFS} under FMS setting.

The (weighted) correlation coefficient and (weighted) cosine similarity measure given in this paper provide the useful ranking method and they are beneficial mathematical tools for multiperiod medical diagnosis in the FMS environment because the new concepts simplify the decision-making process. Therefore, developed medical diagnosis approaches not only improve the decision-making reliability but also supply a new influential way for multiperiod medical diagnosis problems in the FMS environment.

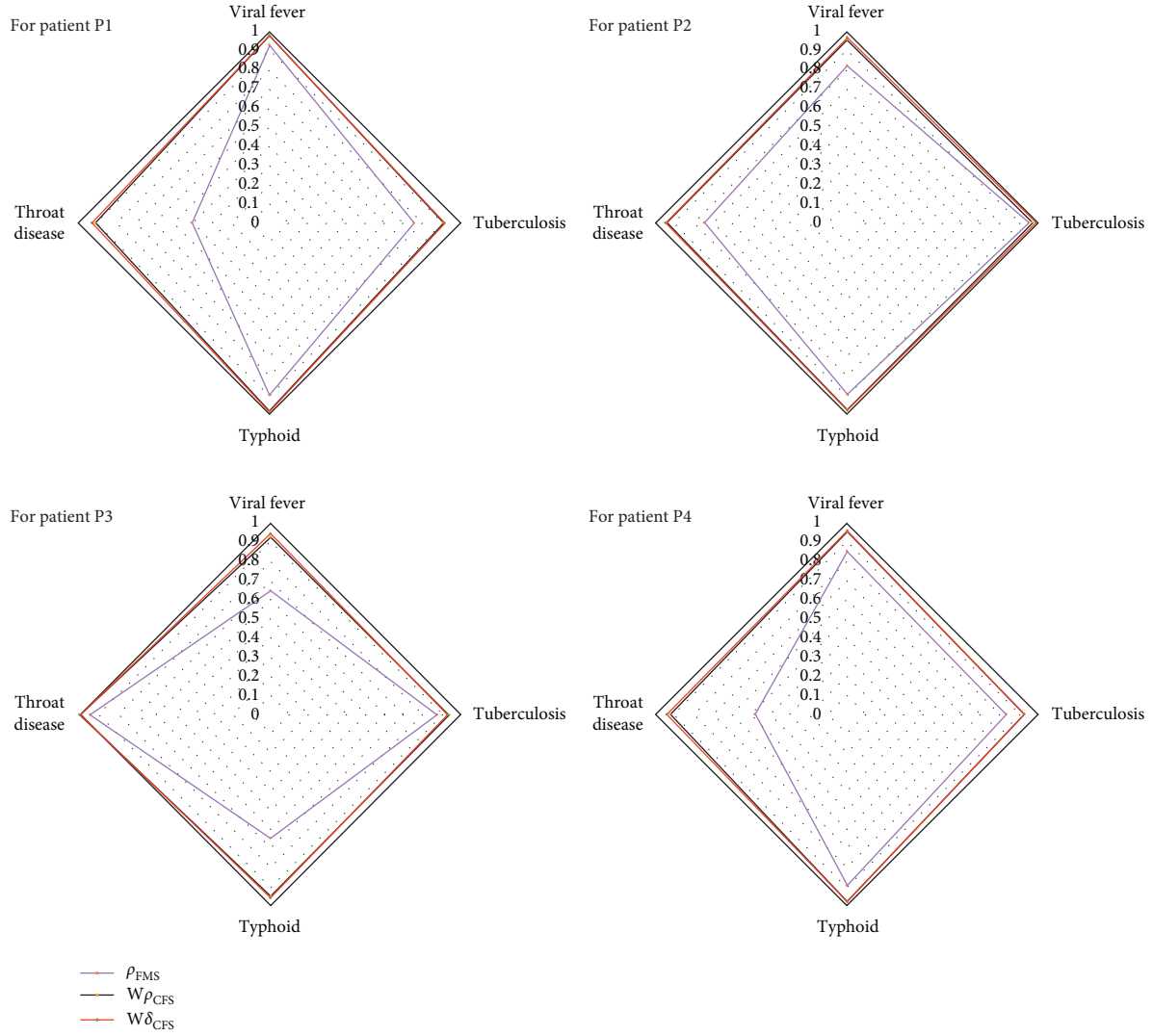

Figure 1: The comparison of given diseases.

Figure 1 shows the comparison of the results of the present paper and the results of [27].

4. Conclusion

In this paper, we introduce a new fuzzy set that is called consistency fuzzy set (CFS). The difference of this new concept from other existing multivalued fuzzy sets is that it uses not only the information from fuzzy multiset (FMS) but also the information provided by both the consistency degree and average of the sequences (truth sequences) in FMSs. Therefore, CFSs contain more useful information than other multivalued fuzzy sets because they use two statistical comparison methods. The aim of this new fuzzy set is to obtain more reasonable results by facilitating the decision-making process and to offer more understandable methods. Since other methods cannot take the consistency degree and average into account, their results may be unreasonable in the decision-making process. Moreover, we also propose a correlation coefficient and a cosine similarity measure between CFSs by taking the advantages of CFSs to solve a multiperiod medical diagnosis problem. Then, we compare them with some existing methods to show the usefulness of CFSs. These proposed approaches can give more detailed information and valuable results to the decision makers as compared to the other existing ones. In the future, we focus on extending the theory under q- rung fuzzy information or we shall develop new aggregation operators and some information measures algorithms in FMS setting.

Data Availability

No data were used to support the findings of the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[2] V. U. Nguyen, "Tender evaluation by fuzzy sets," *Journal of Construction Engineering and Management*, vol. 111, no. 3, pp. 231–243, 1985.

[3] S. K. Pal and R. A. King, "On edge detection of X-ray images using fuzzy sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, no. 1, pp. 69–77, 1983.

[4] R. R. Yager, "Database discovery using fuzzy sets," *International Journal of Intelligent Systems*, vol. 11, no. 9, pp. 691–712, 1996.

[5] M. Akram, C. Kahraman, and K. Zahid, "Group decision-making based on complex spherical fuzzy VIKOR approach," *Knowledge-Based Systems*, vol. 216, Article ID 106793, 2021.

[6] S. Das, D. Guha, and B. Dutta, "Medical diagnosis with the aid of using fuzzy logic and intuitionistic fuzzy logic," *Applied Intelligence*, vol. 45, no. 3, pp. 850–867, 2016.

[7] M. Akram, G. Shahzadi, and A. A. H. Ahmadini, "Decision-making framework for an effective sanitizer to reduce COVID-19 under fermatean fuzzy environment," *Journal of Mathematics*, vol. 2020, Article ID 3263407, 19 pages, 2020.

[8] V. Torra, "Hesitant fuzzy sets," *International Journal of Intelligent Systems*, vol. 25, no. 6, pp. 529–539, 2010.

[9] S. Faizi, T. Rashid, W. Sałabun, S. Zafar, and J. Wątróbski, "Decision making with uncertainty using hesitant fuzzy sets," *International Journal of Fuzzy Systems*, vol. 20, no. 1, pp. 93–103, 2018.

[10] S. Faizi, W. Sałabun, T. Rashid, J. Wątróbski, and S. Zafar, "Group decision-making for hesitant fuzzy sets based on characteristic objects method," *Symmetry*, vol. 9, no. 8, p. 136, 2017.

[11] B. Farhadinia, "A hesitant fuzzy based medical diagnosis problem," *International Journal on Data Science and Technology*, vol. 3, no. 1, pp. 1–7, 2017.

[12] J. Lan, M. Yang, M. Hu, and F. Liu, "Multi-attribute group decision making based on hesitant fuzzy sets TOPSIS method and fuzzy preference relations," *Technological and Economic Development of Economy*, vol. 24, no. 6, pp. 2295–2317, 2018.

[13] S. Miyamoto, "Fuzzy multisets and their generalizations," pp. 225–235, Springer, Berlin, Germany, 2000.

[14] R. R. Yager, "On the theory of bags," *International Journal of General Systems*, vol. 13, no. 1, pp. 23–37, 1986.

[15] J. Ye, J. Song, and S. Du, "Correlation coefficients of consistency neutrosophic sets regarding neutrosophic multivalued sets and their multi-attribute decision-making method," *International Journal of Fuzzy Systems*, 2020.

[16] J. Ye, "Fuzzy decision-making method based on the weighted correlation coefficient under intuitionistic fuzzy environment," *European Journal of Operational Research*, vol. 205, no. 1, pp. 202–204, 2010.

[17] X. Guan, G. Sun, X. Yi, and Z. Zhou, "Synthetic correlation coefficient between hesitant fuzzy sets with applications," *International Journal of Fuzzy Systems*, vol. 20, no. 6, pp. 1968–1985, 2018.

[18] M. Lin, C. Huang, R. Chen, H. Fujita, and X. Wang, "Directional correlation coefficient measures for Pythagorean fuzzy sets: their applications to medical diagnosis and cluster analysis," *Complex & Intelligent Systems*, vol. 7, no. 2, pp. 1025–1043, 2021.

[19] W. S. Du, "Correlation and correlation coefficient of generalized orthopair fuzzy sets," *International Journal of Intelligent Systems*, vol. 34, no. 4, pp. 564–583, 2019.

[20] T. Zheng, M. Zhang, L. Li, Q. Wu, and L. Zhou, "Correlation coefficients of interval-valued pythagorean hesitant fuzzy sets and their applications," *IEEE Access*, vol. 8, no. 1, pp. 9271–9286, 2020.

[21] S.-C. Ngan, "An activation detection based similarity measure for intuitionistic fuzzy sets," *Expert Systems with Applications*, vol. 60, pp. 62–80, 2016.

[22] X. T. Nguyen, V. D. Nguyen, V. H. Nguyen, and H. Garg, "Exponential similarity measures for Pythagorean fuzzy sets and their applications to pattern recognition and decision-making process," *Complex & Intelligent Systems*, vol. 5, no. 2, pp. 217–228, 2019.

[23] Y. Song, X. Wang, L. Lei, and A. Xue, "A new similarity measure between intuitionistic fuzzy sets and its application to pattern recognition," *Abstract and Applied Analysis*, vol. 2014, Article ID 384241, 11 pages, 2014.

[24] J. Ye and J. Fu, "Multi-period medical diagnosis method using a single valued neutrosophic similarity measure based on tangent function," *Computer Methods and Programs in Biomedicine*, vol. 123, pp. 142–149, 2015.

[25] J. Ye, "Vector similarity measures of simplified neutrosophic sets and their application in multicriteria decision making," *International Journal of Fuzzy Systems*, vol. 16, no. 2, pp. 204–211, 2014.

[26] J. Ye, "Cosine similarity measures for intuitionistic fuzzy sets and their applications," *Mathematical and Computer Modelling*, vol. 53, no. 1-2, pp. 91–97, 2011.

[27] M. S. El-Azab, M. Shokry, and R. A. Abo khadra, "Correlation measure for fuzzy multisets," *Journal of the Egyptian Mathematical Society*, vol. 25, no. 3, pp. 263–267, 2017.