International Journal of Control and Automation
Vol. 9, No. 8 (2016), pp.403-418
http//dx.doi.org/10.14257/ijca.2016.9.8.38

Correlation Coefficients of Intuitionistic Hesitant Fuzzy Sets and Their Applications to Clustering Analysis

Xiuming Chen [1,2] and Yezheng Liu [1]

[1] *School of Management, Hefei University of Technology, Hefei 230001, China*
[2] *School of Information Engineering, Anhui Xinhua University, Hefei 230088, China*
cxm9001@163.com

Abstract

An intuitionistic hesitant fuzzy set (IHFS) consists of hesitancy function intuitionistic fuzzy elements, which consists the membership degree and the non membership degree, supporting a more exemplary and flexible access to assign values for each element in the domain, and can handle this kind of hesitancy in this situation. It can be seen as a powerful tool to express uncertain information in the process of group recommendation. Therefore, we propose a correlation coefficient between IHFSs as a further extension of existing correlation coefficients for hesitant fuzzy sets and intuitionistic fuzzy sets and apply it to group recommendation under intuitionistic hesitant fuzzy environments. Through the weighted correlation coefficient between each item and the ideal item, the ranking order of all items can be identified and the best item can be easily identified as well. Finally, a practical example of a movie recommendation is given to demonstrate the practicality and effectiveness of the developed approach.

***Keywords**: Correlation coefficient, Intuitionistic hesitant fuzzy sets (IHFSs), Clustering analysis, Group recommendation*

1. Introduction

The concept of intuitionistic fuzzy set (IFS) was introduced by Atanassov [1] to generalize the concept of Zadeh's fuzzy set [2]. Each element in IFS is expressed by an ordered pair, and each ordered pair is distinguished by a membership degree and a non-membership degree. The sum of the membership degree and the non-membership degree of each ordered pair is lower than or equal to one. Since it was first introduced in 1986, the IFS theory has been widely investigated and applied to a variety of fields e.g. mathematical programming[3], decision-making[4], pattern recognition[5], image processing[6], grey relational analysis[7] , *etc*.

Hesitant fuzzy sets (HFSs), an extension of traditional fuzzy sets, can address these situations. HFSs were first introduced by Torra and Narukawa [8], and they permit the membership degrees of an element to be a set of several possible values between zero and one. HFSs are highly useful in resolving situations where people hesitate when providing their preferences in the decision-making process, and they have been a subject of great interest to researchers. Recently, Rodríguez *et al*. [9] presented an overview and discussed future trends for HFSs.

However, in the process of some practical group recommendation, sometimes, due to the time pressure and lack of knowledge or data or the recommenders' limited attention and information processing capacities, the recommenders cannot provide their recommendations with a single numerical value, a margin of error, some possibility distribution on the possible values, several possible numerical values, several possible interval numbers, but several possible intuitionistic fuzzy numbers. For example, to get a

ISSN: 2005-4297 IJCA

reasonable recommendation result, a recommendation organization, which contains a lot of recommenders, is required to estimate the degree that an item satisfies an attribute. Suppose there is a case: some recommenders provide, and the others provide, and these two parts cannot persuade each other to change their opinions. We can easily see that such case cannot be dealt with by fuzzy sets, hesitant fuzzy sets, and their extensions, such as interval-valued fuzzy sets, intuitionistic fuzzy sets, interval-valued intuitionistic fuzzy sets, type 2 fuzzy sets, interval-valued hesitant fuzzy sets [10]. Thus, it is very necessary to introduce a new extension of hesitant fuzzy sets to address this issue. The aim of this paper is to present the notion of intuitionistic hesitant fuzzy set (IFS), which extends the hesitant fuzzy set to intuitionistic fuzzy environments and permits the membership of an element to be a set of several possible intuitionistic fuzzy numbers. Thus, intuitionistic hesitant fuzzy set is a very useful tool to deal with the situations in which the recommenders hesitate between several possible intuitionistic fuzzy numbers to assess the degree to which an item satisfies an attribute. In the previous example, the degree to which the alternative satisfies the attribute can be represented by an intuitionistic hesitant fuzzy set.

In a real world, data used for clustering may be uncertain and fuzzy, to deal with various types of fuzzy data, a number of clustering algorithms corresponding to different fuzzy environments [11] have been proposed, e.g., intuitionistic fuzzy clustering algorithms [12] involving the correlation coefficient formulas for IFSs [13] and hesitant fuzzy sets clustering algorithms [14]. However, under the group recommendation situations, the recommendation information provided by different recommenders may have an obvious difference. These fuzzy clustering schemes mentioned above are unable to incorporate the differences in the opinions of different recommenders; that is, they are unsuitable to do clustering under intuitionistic hesitant fuzzy environments. IHFSs introduced here could resolve the issue, because they avoid performing data aggregation and can directly reflect the differences of the opinions of different recommenders. We will use the derived correlation coefficient formulas to calculate the degrees of correlation among IHFSs aiming at clustering different items.

The rest of the article is organized as follows. Section 2 reviews basic concepts related to IFSs and HFSs. In Section 3 we give the concept of IHFS, and some correlation coefficient formulas for IHFSs. Section 4 introduces an actual example is employed to illustrate the use of IHFSs in clustering and movie recommendations. Section 5 summarizes this study and presents future challenges.

2. Preliminaries

2.1. Intuitionistic Fuzzy Sets

Definition 1 ([1]) Let X be a universe of discourse, an IFS A in X is defined as:

$$A=\{<x,u_A(x),v_A(x)>|\, x\in X\} \tag{1}$$

where the functions $u_A(x)$ and $v_A(x)$ denote the degrees of membership and non-membership of the element $x\in X$ to the set A, respectively, with the condition:

$$0\le u_A(x)\le 1, 0\le v_A(x)\le 1, 0\le u_A(x)+v_A(x)\le 1 \tag{2}$$

and $\pi_A(x)=1-u_A(x)-v_A(x)$ is usually called the degree of hesitancy of x to $A, \alpha=(u_\alpha,v_\alpha)$ is named as an intuitionistic fuzzy value (IFV).

2.2. Hesitant fuzzy sets

Definition 2 ([8]) Let X be a reference set, a hesitant fuzzy set (HFS) A on X is defined in terms of a function $h_A(x)$ when applied to X returns a finite subset of [0, 1], *i.e.*,

$$A=\{<x,h_A(x)>|\, x\in X\} \tag{3}$$

where $h_A(x)$ is a set of some different values in [0, 1], representing the possible membership degrees of the element $x\in X$ to A. For convenience, we call $h_A(x)$ a hesitant fuzzy element (HFE) [15].

Definition 3 ([8]) Given a HFE h, its lower and upper bounds are defined as below:

Lowerbound : $h^-(x)=\min h(x)$; Upperbound : $h^+(x)=\max h(x)$.

Definition 4 ([8]) Given a HFE $h, A_{env}(h)$ is called the envelope of h which is represented by $(h^-, 1-h^+)$, with h^- and h^+ being its lower and upper bounds, respectively.

It should be mentioned that from Definitions 3 and 4, we can see that the envelop of a HFE h, denoted by $A_{env}(h)=(h^-, 1-h^+)$, is just an IFV.

2.3. Correlation coefficients of IFSs

Many approaches [16,17] have been introduced to compute the correlation coefficients of IFSs. Let $X=\{x_1,x_2,\cdots,x_n\}$ be a discrete universe of discourse, and let $IFS(X)$ denote the set of all the IFSs in X. For any $A,B\in IFS(X)$, Gerstenkorn and Man′ ko [16] extended the definition of informational energy given by Dumitrescu [18] to the case of IFSs, that is:

$$E_{IFS}(A)=\sum_{i=1}^{n}[(u_A^2(x_i)+v_A^2(x_i))] \tag{4}$$

The correlation of the IFSs A and B is defined as:

$$C_{IFS}(A,B)=\sum_{i=1}^{n}[(u_A(x_i)u_B(x_i)+v_A(x_i)v_B(x_i))] \tag{5}$$

They adopted the formula:

$$\rho_{IFS}(A,B)=\frac{C_{IFS}(A,B)}{[E_{IFS}(A)\cdot E_{IFS}(B)]^{\frac{1}{2}}}=\frac{\sum_{i=1}^{n}(u_A(x_i)u_B(x_i)+v_A(x_i)v_B(x_i))}{\{\sum_{i=1}^{n}[(u_A^2(x_i)+v_A^2(x_i))]\}^{\frac{1}{2}}\cdot\{\sum_{i=1}^{n}(u_B^2(x_i)+v_B^2(x_i))\}^{\frac{1}{2}}} \tag{6}$$

To define the correlation coefficient of A and B. Here, $\rho_{IFS}(A,B)$ satisfies the following conditions:

(1) $0\le\rho_{IFS}(A,B)\le 1$;

(2) $A=B\Rightarrow\rho_{IFS_1}(A,B)=1$;

(3) $\rho_{IFS}(A,B)=\rho_{IFS}(B,A)$

2.4. Correlation coefficients of HFSs

Let $X=\{x_1,x_2,\cdots,x_n\}$ be a discrete universe of discourse, A and B be two HFSs on X denoted as $A=\{<x_i,h_A(x_i)>|\, x_i\in X, i=1,2,\cdots,n\}$, $B=\{<x_i,h_B(x_i)>|\, x_i\in X, i=1,2,\cdots,n\}$ respectively.

The values of a HFE are usually given in a disorder, and for convenience, we arrange them in a decreasing order. For a HFE h, let $\sigma:(1,2,\cdots,n)\to(1,2,\cdots,n)$ be a permutation satisfying $h_{\sigma(i)}\ge h_{\sigma(i+1)}, i=1,2,\cdots,n-1$, and $h_{\sigma(j)}$ be the jth largest value in h.

Definition 5 ([14]) For a HFS $A=\{<x_i,h_A(x_i)>|\, x_i\in X, i=1,2,\cdots,n\}$, the informational energy of the set A is defined as:

$$E_{HFS}(A)=\sum_{i=1}^{n}(\frac{1}{l_i}\sum_{j=1}^{l_i}h_{A\sigma(j)}^2(x_i)) \tag{7}$$

Definition 6([14]) For two HFSs A and B, their correlation is defined by

$$C_{HFS}(A,B)=\sum_{i=1}^{n}(\frac{1}{l_i}\sum_{j=1}^{l_i}h_{A\sigma(j)}(x_i)h_{B\sigma(j)}(x_i)) \quad (8)$$

For $A,B\in HFSs$, the correlation (9) satisfies:

$C_{HFS}(A,A)=E_{HFS}(A);$

$C_{HFS}(A,B)=C_{HFS}(B,A)$

Using Definitions 5 and 6, we derive a correlation coefficient for $HFSs$:

Definition 7([14]) The correlation coefficient between two $HFSs$ A and B is given as:

$$\rho_{HFS}(A,B)=\frac{C_{HFS}(A,B)}{[C_{HFS}(A,A)]^{\frac{1}{2}}\cdot[C_{HFS}(B,B)]^{\frac{1}{2}}}=\frac{\sum_{i=1}^{n}(\frac{1}{l_i}\sum_{j=1}^{l_i}h_{A\sigma(j)}(x_i)h_{B\sigma(j)}(x_i))}{[\sum_{i=1}^{n}(\frac{1}{l_i}\sum_{j=1}^{l_i}h_{A\sigma(j)}^2(x_i)]^{\frac{1}{2}}\cdot[\sum_{i=1}^{n}(\frac{1}{l_i}\sum_{j=1}^{l_i}h_{B\sigma(j)}^2(x_i)]^{\frac{1}{2}}} \quad (9)$$

3. Correlation and Clustering Algorithm for Intuitionistic Hesitant Fuzzy Sets

It has been known that in the group recommendation systems, it is somewhat difficult for the recommenders to assign an exact value for a membership degree and a non-membership degree of a certain element to A, but maybe hesitant from different intuitionistic fuzzy values. It means that it is very necessary to introduce the concept of intuitionistic hesitant fuzzy set (IHFS).

Definition 8 Let be a fixed set; an intuitionistic hesitant fuzzy set (IHFS) on X is given in terms of a function that when applied to X returns a subset of D. To be easily understood, we express the IHFS by a mathematical symbol, and define as:

$$\tilde{A}=\{<x_i,h_{\tilde{A}}(x_i)>|\,x_i\in X, i=1,2,\cdots,n\} \quad (10)$$

where $h_{\tilde{A}}(x_i)$ is a set of some different intuitionistic fuzzy values in D, denoting the possible membership degree and nonmembership degree of the element $x\in X$ to the set $\tilde{A}$. where:

$$h_{\tilde{A}}(x_i)=\{(h_{\tilde{A}}^{\mu}(x_i),h_{\tilde{A}}^{\nu}(x_i))|(h_{\tilde{A}}^{\mu}(x_i),h_{\tilde{A}}^{\nu}(x_i))\in D)\}$$

For convenience, we call $(h_{\tilde{A}}^{\mu}(x_i),h_{\tilde{A}}^{\nu}(x_i))$ an intuitionistic hesitant fuzzy element (IHFE).

Example 1 Let $X=\{x_1,x_2\}$ be a reference set, $h_{\tilde{A}}(x_1)=\{(0.6,0.3),(0.7,0.2)\}$ and $h_{\tilde{A}}(x_2)=\{(0.5,0.4),(0.7,0.2),(0.8,0.1)\}$ be the IHFE of $x_i(i=1,2)$ to a set $\tilde{A}$, respectively. Then $\tilde{A}$ can be considered as an IHFS and given as:

$\tilde{A}=\{<x_1,(0.6,0.3),(0.7,0.2)\}>,<x_2,(0.5,0.4),(0.7,0.2),(0.8,0.1)>\}$ For an IHFE $h_{\tilde{A}}(x_i)$, we arrange the intuitionistic fuzzy numbers in $h_{\tilde{A}}(x_i)$ in a decreasing order.

Let $\sigma:(1,2,\cdots,n)\rightarrow(1,2,\cdots,n)$ be a permutation satisfying $h_{\tilde{A}\sigma(i)}\geq h_{\tilde{A}\sigma(i+1)}$, $i=1,2,\cdots,n-1$ and $h_{\tilde{A}\sigma(j)}(x_i)$ be the jth largest intuitionistic fuzzy numbers in $h_{\tilde{A}}(x_i)$, where $h_{\tilde{A}\sigma(j)}(x_i)=(h_{\tilde{A}\sigma(j)}^{\mu}(x_i),h_{\tilde{A}\sigma(j)}^{\nu}(x_i)), j=1,2,\cdots,l_i$, Similar to the previous definitions on the correlation coefficients of $HFSs$ A and B, we can define the correlation coefficients of $IHFSs$ $\tilde{A}$ and $\tilde{B}$ in X as:

$$\rho_{IHFS_1}(\tilde{A},\tilde{B})=\frac{C_{IHFS_1}(\tilde{A},\tilde{B})}{\left[C_{IHFS_1}(\tilde{A},\tilde{A})\right]^{\frac{1}{2}}\cdot\left[C_{IHFS_1}(\tilde{B},\tilde{B})\right]^{\frac{1}{2}}}$$
$$=\frac{\sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i)+h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)}{\left\{\sum_{i=1}^{n}\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2+\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right]\right\}^{\frac{1}{2}}\cdot\left\{\sum_{i=1}^{n}\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2+\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]\right\}^{\frac{1}{2}}} \tag{11}$$

Theorem 1. For two $IHFSs$ $\tilde{A}$ and $\tilde{B}$, the correlation coefficients defined by Eqs. (11) satisfy:

(1) $\rho_{IHFS_1}(\tilde{A},\tilde{B})=\rho_{IHFS_1}(\tilde{B},\tilde{A})$;

(2) $0\le\rho_{IHFS_1}(\tilde{A},\tilde{B})\le 1$;

(3) $\rho_{IHFS_1}(\tilde{A},\tilde{B})=1$,if $\tilde{A}=\tilde{B}$.

Proof 1

(1)It is straightforward.

(2)The inequality $\rho_{IHFS_1}(\tilde{A},\tilde{B})\ge 0$ is obvious. Below let us prove $\rho_{IHFS_1}(\tilde{A},\tilde{B})\le 1$:

$$\rho_{IHFS_1}(\tilde{A},\tilde{B})=\frac{C_{IHFS_1}(\tilde{A},\tilde{B})}{\left[C_{IHFS_1}(\tilde{A},\tilde{A})\right]^{\frac{1}{2}}\cdot\left[C_{IHFS_1}(\tilde{B},\tilde{B})\right]^{\frac{1}{2}}}$$

$$C_{IHFS_1}(\tilde{A},\tilde{B})=\sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i)+h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)$$
$$=\frac{1}{l_1}\sum_{j=1}^{l_1}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_1)h^{\mu}_{\tilde{B}\sigma(j)}(x_1)+h^{\nu}_{\tilde{A}\sigma(j)}(x_1)h^{\nu}_{\tilde{B}\sigma(j)}(x_1)\right)+\frac{1}{l_2}\sum_{j=1}^{l_2}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_2)h^{\mu}_{\tilde{B}\sigma(j)}(x_2)+h^{\nu}_{\tilde{A}\sigma(j)}(x_2)h^{\nu}_{\tilde{B}\sigma(j)}(x_2)\right)$$
$$+\cdots+\frac{1}{l_n}\sum_{j=1}^{l_n}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_n)h^{\mu}_{\tilde{B}\sigma(j)}(x_n)+h^{\nu}_{\tilde{A}\sigma(j)}(x_n)h^{\nu}_{\tilde{B}\sigma(j)}(x_n)\right)$$
$$=\sum_{j=1}^{l_1}\left(\frac{h^{\mu}_{\tilde{A}\sigma(j)}(x_1)}{\sqrt{l_1}}\cdot\frac{h^{\mu}_{\tilde{B}\sigma(j)}(x_1)}{\sqrt{l_1}}+\frac{h^{\nu}_{\tilde{A}\sigma(j)}(x_1)}{\sqrt{l_1}}\cdot\frac{h^{\nu}_{\tilde{B}\sigma(j)}(x_1)}{\sqrt{l_1}}\right)+\sum_{j=1}^{l_2}\left(\frac{h^{\mu}_{\tilde{A}\sigma(j)}(x_2)}{\sqrt{l_2}}\cdot\frac{h^{\mu}_{\tilde{B}\sigma(j)}(x_2)}{\sqrt{l_2}}+\frac{h^{\nu}_{\tilde{A}\sigma(j)}(x_2)}{\sqrt{l_2}}\cdot\frac{h^{\nu}_{\tilde{B}\sigma(j)}(x_2)}{\sqrt{l_2}}\right)$$
$$+\cdots+\sum_{j=1}^{l_n}\left(\frac{h^{\mu}_{\tilde{A}\sigma(j)}(x_n)}{\sqrt{l_n}}\cdot\frac{h^{\mu}_{\tilde{B}\sigma(j)}(x_n)}{\sqrt{l_n}}+\frac{h^{\nu}_{\tilde{A}\sigma(j)}(x_n)}{\sqrt{l_n}}\cdot\frac{h^{\nu}_{\tilde{B}\sigma(j)}(x_n)}{\sqrt{l_n}}\right)$$

Using the Cauchy–Schwarz inequality:

$$(x_1y_1+x_2y_2+\cdots+x_ny_n)^2\le(x_1^2+x_2^2+\cdots+x_n^2)\cdot(y_1^2+y_2^2+\cdots+y_n^2),$$

where $(x_1,x_2,\cdots,x_n)\in R^n,(y_1,y_2,\cdots,y_n)\in R^n$, we obtain:

$$(C_{IHFS_1}(\tilde{A},\tilde{B}))^2\le\left[\frac{1}{l_1}\sum_{j=1}^{l_1}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_1)\right)^2+\frac{1}{l_1}\sum_{j=1}^{l_1}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_1)\right)^2+\frac{1}{l_2}\sum_{j=1}^{l_2}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_2)\right)^2+\frac{1}{l_2}\sum_{j=1}^{l_2}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_2)\right)^2+\cdots+\frac{1}{l_n}\sum_{j=1}^{l_n}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_n)\right)^2+\frac{1}{l_n}\sum_{j=1}^{l_n}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_n)\right)^2\right]$$
$$\times\left[\frac{1}{l_1}\sum_{j=1}^{l_1}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_1)\right)^2+\frac{1}{l_1}\sum_{j=1}^{l_1}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_1)\right)^2+\frac{1}{l_2}\sum_{j=1}^{l_2}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_2)\right)^2+\frac{1}{l_2}\sum_{j=1}^{l_2}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_2)\right)^2+\cdots+\frac{1}{l_n}\sum_{j=1}^{l_n}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_n)\right)^2+\frac{1}{l_n}\sum_{j=1}^{l_n}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_n)\right)^2\right]$$
$$=\left[\sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2+\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right]\cdot\left[\sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2+\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]$$
$$=C_{IHFS_1}(\tilde{A},\tilde{A})\cdot C_{IHFS_1}(\tilde{B},\tilde{B})$$

(3) $\tilde{A}=\tilde{B}\ \Rightarrow h_{\tilde{A}\sigma(j)}(x_i)=h_{\tilde{B}\sigma(j)}(x_i),x_i\in X\ \Rightarrow\rho_{IHFS_1}(\tilde{A},\tilde{B})=1$

Example 2 Let $\tilde{A}$ and $\tilde{B}$ be two $IHFSs$ in $X=\{x_1,x_2,x_3\}$, and

$A=\{<x_1,\{(0.7,0.2)\}>,<x_2,\{(0.9,0.1),(0.8,0.1)\}>,<x_3,\{(0.7,0.2)\}>\}$,

$B=\{<x_1,\{(0.6,0.2)\}>,<x_2,\{(0.6,0.3),(0.5,0.2)\}>,<x_3,\{(0.6,0.3)\}>\}$

By using Eq. (11), we calculate:

$$C_{IHFS_1}(\tilde{A},\tilde{A})=\sum_{i=1}^{n}\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2+\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right]$$
$$=(0.7^2+0.2^2)+\tfrac{1}{2}(0.9^2+0.1^2+0.8^2+0.1^2)+(0.7^2+0.2^2)$$
$$=1.795$$
$$C_{IHFS_1}(\tilde{B},\tilde{B})=\sum_{i=1}^{n}\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2+\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]$$
$$=(0.6^2+0.2^2)+\tfrac{1}{2}(0.6^2+0.3^2+0.5^2+0.2^2)+(0.6^2+0.3^2)$$
$$=1.22$$
$$C_{IHFS_1}(\tilde{A},\tilde{B})=\sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i)+h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)$$
$$=(0.7\times0.6+0.2\times0.2)+\tfrac{1}{2}(0.9\times0.6+0.1\times0.3+0.8\times0.5+0.1\times0.2)+(0.7\times0.6+0.2\times0.3)$$
$$=1.435$$

Finally, we use Eq. (11) to calculate the correlation coefficient:

$$\rho_{IHFS_1}(\tilde{A},\tilde{B})=\frac{C_{IHFS_1}(\tilde{A},\tilde{B})}{\left[C_{IHFS_1}(\tilde{A},\tilde{A})\right]^{\frac{1}{2}}\cdot\left[C_{IHFS_1}(\tilde{B},\tilde{B})\right]^{\frac{1}{2}}}=\frac{1.435}{\sqrt{1.795}\cdot\sqrt{1.22}}=0.97$$

Obviously, $0<\rho_{IHFS_1}(\tilde{A},\tilde{B})<1$.

It is noted that the number of values in different IHFEs may be different. To compute the correlation coefficients between two IHFSs, let $l_i=\max\{l(h_{\tilde{A}}(x_i)),l(h_{\tilde{B}}(x_i))\}$ for each x_i in X, where $l(h_{\tilde{A}}(x_i))$ and $l(h_{\tilde{B}}(x_i))$ represent the number of values in $h_{\tilde{A}}(x_i)$ and $h_{\tilde{B}}(x_i)$, respectively. When $l(h_{\tilde{B}}(x_i))\neq l(h_{\tilde{A}}(x_i))$, one can make them having the same number of elements through adding some elements to the IHFE which has less number of elements. In terms of the pessimistic principle, the smallest element will be added while in the opposite case, the optimistic principle may be adopted. In the present work, we use the former case. Especially, if $l(h_{\tilde{A}}(x_i))<l(h_{\tilde{B}}(x_i))$, then $h_{\tilde{A}}(x_i)$ should be extended by adding the minimum value in it until it has the same length as $h_{\tilde{B}}(x_i)$. This idea has been successfully applied to distance and similarity measures for HFSs [19].

In what follows we give a new formula of calculating the correlation coefficient, which is similar to that used in IFSs [20]:

Definition 9 For two $IHFSs$ $\tilde{A}$ and $\tilde{B}$, their correlation coefficient is defined by

$$\rho_{IHFS_2}(\tilde{A},\tilde{B})=\frac{C_{IHFS_1}(\tilde{A},\tilde{B})}{\max\left\{C_{IHFS_1}(\tilde{A},\tilde{A}),C_{IHFS_1}(\tilde{B},\tilde{B})\right\}}$$
$$=\frac{\sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i)+h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)}{\max\left\{\sum_{i=1}^{n}\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2+\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right],\sum_{i=1}^{n}\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2+\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]\right\}}\quad(12)$$

Theorem 2 The correlation coefficient of two $IHFSs$ $\tilde{A}$ and $\tilde{B}$, $\rho_{IHFS_2}(\tilde{A},\tilde{B})$, follows the same properties listed in Theorem 1.

Proof 2 The process to prove the properties (1) and (3) is analogous to that in Theorem 1, we do not repeat it here.

(2) $\rho_{IHFS_2}(\tilde{A},\tilde{B})\geq0$ is obvious. We now only prove $\rho_{IHFS_2}(\tilde{A},\tilde{B})\leq1$.

Based on the proof process of Theorem 1, we have:

$C_{IHFS_1}(\tilde{A},\tilde{B})\leq\left[C_{IHFS_1}(\tilde{A},\tilde{A})\right]^{\frac{1}{2}}\cdot\left[C_{IHFS_1}(\tilde{B},\tilde{B})\right]^{\frac{1}{2}}$ and then

$C_{IHFS_1}(\tilde{A},\tilde{B}) \le \max\{C_{IHFS_1}(\tilde{A},\tilde{A}), C_{IHFS_1}(\tilde{B},\tilde{B})\}$ Thus, $0 \le \rho_{IHFS_2}(\tilde{A},\tilde{B}) \le 1$.

In practical applications, the elements $x_i (i=1,2,\cdots,n)$ in the universe X have different weights. Let $w=(w_1, w_2, \cdots, w_n)^T$ be the weight vector of $x_i (i=1,2,\cdots,n)$ with $w_i \ge 0, i=1,2,\cdots,n$ and $\sum_{i=1}^{n} w_i = 1$, we further extend the correlation coefficient formulas given in Eqs. (13) and (14) as: $\rho_{IHFS_3}(\tilde{A},\tilde{B}) = \dfrac{C_{IHFS_2}(\tilde{A},\tilde{B})}{[C_{IHFS_2}(\tilde{A},\tilde{A})]^{\frac{1}{2}} \cdot [C_{IHFS_2}(\tilde{B},\tilde{B})]^{\frac{1}{2}}}$

$$= \frac{\sum_{i=1}^{n} w_i \left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i) + h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)}{\left\{\sum_{i=1}^{n} w_i\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2 + \left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right]\right\}^{\frac{1}{2}} \cdot \left\{\sum_{i=1}^{n} w_i\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2 + \left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]\right\}^{\frac{1}{2}}} \quad (13)$$

$$\rho_{IHFS_4}(\tilde{A},\tilde{B}) = \frac{C_{IHFS_2}(\tilde{A},\tilde{B})}{\max\{C_{IHFS_2}(\tilde{A},\tilde{A}), C_{IHFS_2}(\tilde{B},\tilde{B})\}}$$

$$= \frac{\sum_{i=1}^{n} w_i \left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i) + h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)}{\max\left\{\sum_{i=1}^{n} w_i\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2 + \left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right], \sum_{i=1}^{n} w_i\left[\frac{1}{l_i}\sum_{j=1}^{l_i}\left(\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2 + \left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]\right\}} \quad (14)$$

It can be seen that if $w=(1/n, 1/n, \cdots, 1/n)^T$, then Eqs. (13) and (14) reduce to Eqs. (11) and (12), respectively. Note that both $\rho_{IHFS_3}(\tilde{A},\tilde{B})$ and $\rho_{IHFS_4}(\tilde{A},\tilde{B})$ also satisfy three properties of Theorem 1.

Theorem 3 Let $w=(w_1, w_2, \cdots, w_n)^T$ be the weight vector of $x_i (i=1,2,\cdots,n)$ with $w_i \ge 0, i=1,2,\cdots,n$ and $\sum_{i=1}^{n} w_i = 1$, the correlation coefficient $\rho_{IHFS_3}(\tilde{A},\tilde{B})$ between two *IHFSs* $\tilde{A}$ and $\tilde{B}$ defined in Eq. (13), which takes into account the weights, satisfies:

(1) $\rho_{IHFS_3}(\tilde{A},\tilde{B}) = \rho_{IHFS_3}(\tilde{B},\tilde{A})$;

(2) $0 \le \rho_{IHFS_3}(\tilde{A},\tilde{B}) \le 1$;

(3) $\rho_{IHFS_3}(\tilde{A},\tilde{B}) = 1$, if $\tilde{A} = \tilde{B}$.

Proof 3

(1) It is straightforward.

(2) $\rho_{IHFS_3}(\tilde{A},\tilde{B}) \ge 0$ is obvious. Below we prove $\rho_{IHFS_3}(\tilde{A},\tilde{B}) \le 1$. Since

$$C_{IHFS_2}(\tilde{A},\tilde{B}) = \sum_{i=1}^{n}\left(\frac{1}{l_i}\sum_{j=1}^{l_i} w_i\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)h^{\mu}_{\tilde{B}\sigma(j)}(x_i) + h^{\nu}_{\tilde{A}\sigma(j)}(x_i)h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)\right)$$

$$= \frac{w_1}{l_1}\sum_{j=1}^{l_1}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_1)h^{\mu}_{\tilde{B}\sigma(j)}(x_1) + h^{\nu}_{\tilde{A}\sigma(j)}(x_1)h^{\nu}_{\tilde{B}\sigma(j)}(x_1)\right) + \frac{w_2}{l_2}\sum_{j=1}^{l_2}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_2)h^{\mu}_{\tilde{B}\sigma(j)}(x_2) + h^{\nu}_{\tilde{A}\sigma(j)}(x_2)h^{\nu}_{\tilde{B}\sigma(j)}(x_2)\right)$$

$$+ \cdots + \frac{w_n}{l_n}\sum_{j=1}^{l_n}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_n)h^{\mu}_{\tilde{B}\sigma(j)}(x_n) + h^{\nu}_{\tilde{A}\sigma(j)}(x_n)h^{\nu}_{\tilde{B}\sigma(j)}(x_n)\right)$$

$$= \sum_{j=1}^{l_1}\left(\frac{\sqrt{w_1}h^{\mu}_{\tilde{A}\sigma(j)}(x_1)}{\sqrt{l_1}} \cdot \frac{\sqrt{w_1}h^{\mu}_{\tilde{B}\sigma(j)}(x_1)}{\sqrt{l_1}} + \frac{\sqrt{w_1}h^{\nu}_{\tilde{A}\sigma(j)}(x_1)}{\sqrt{l_1}} \cdot \frac{\sqrt{w_1}h^{\nu}_{\tilde{B}\sigma(j)}(x_1)}{\sqrt{l_1}}\right)$$

$$+\sum_{j=1}^{l_2}\left(\frac{\sqrt{w_2}h^{\mu}_{\tilde{A}\sigma(j)}(x_2)}{\sqrt{l_2}}\cdot\frac{\sqrt{w_2}h^{\mu}_{\tilde{B}\sigma(j)}(x_2)}{\sqrt{l_2}}+\frac{\sqrt{w_2}h^{\nu}_{\tilde{A}\sigma(j)}(x_2)}{\sqrt{l_2}}\cdot\frac{\sqrt{w_2}h^{\nu}_{\tilde{B}\sigma(j)}(x_2)}{\sqrt{l_2}}\right)$$

$$+\cdots+\sum_{j=1}^{l_n}\left(\frac{\sqrt{w_n}h^{\mu}_{\tilde{A}\sigma(j)}(x_n)}{\sqrt{l_n}}\cdot\frac{\sqrt{w_n}h^{\mu}_{\tilde{B}\sigma(j)}(x_n)}{\sqrt{l_n}}+\frac{\sqrt{w_n}h^{\nu}_{\tilde{A}\sigma(j)}(x_n)}{\sqrt{l_n}}\cdot\frac{\sqrt{w_n}h^{\nu}_{\tilde{B}\sigma(j)}(x_n)}{\sqrt{l_n}}\right)$$

Using the Cauchy–Schwarz inequality, we obtain:

$$(x_1y_1+x_2y_2+\cdots+x_ny_n)^2\le(x_1^2+x_2^2+\cdots+x_n^2)\cdot(y_1^2+y_2^2+\cdots+y_n^2),$$

where $(x_1,x_2,\cdots,x_n)\in R^n,(y_1,y_2,\cdots,y_n)\in R^n$, we obtain:

$$(C_{IHFS_2}(\tilde{A},\tilde{B}))^2$$

$$\le\left[\frac{w_1}{l_1}\sum_{j=1}^{l_1}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_1)\right)^2+\frac{w_1}{l_1}\sum_{j=1}^{l_1}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_1)\right)^2+\frac{w_2}{l_2}\sum_{j=1}^{l_2}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_2)\right)^2+\right.$$

$$\left.\frac{w_2}{l_2}\sum_{j=1}^{l_2}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_2)\right)^2+\cdots+\frac{w_n}{l_n}\sum_{j=1}^{l_n}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_n)\right)^2+\frac{w_n}{l_n}\sum_{j=1}^{l_n}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_n)\right)^2\right]\times$$

$$\left[\frac{w_1}{l_1}\sum_{j=1}^{l_1}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_1)\right)^2+\frac{w_1}{l_1}\sum_{j=1}^{l_1}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_1)\right)^2+\frac{w_2}{l_2}\sum_{j=1}^{l_2}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_2)\right)^2+\right.$$

$$\left.\frac{w_2}{l_2}\sum_{j=1}^{l_2}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_2)\right)^2+\cdots+\frac{w_n}{l_n}\sum_{j=1}^{l_n}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_n)\right)^2+\frac{w_n}{l_n}\sum_{j=1}^{l_n}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_n)\right)^2\right]$$

$$=\left[\sum_{i=1}^{n}\left(\frac{w_i}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2+\frac{w_i}{l_i}\sum_{j=1}^{l_i}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right]\cdot\left[\sum_{i=1}^{n}\left(\frac{w_i}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2+\frac{w_i}{l_i}\sum_{j=1}^{l_i}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]$$

$$=\left[\sum_{i=1}^{n}w_i\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{A}\sigma(j)}(x_i)\right)^2+\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\nu}_{\tilde{A}\sigma(j)}(x_i)\right)^2\right)\right]\cdot\left[\sum_{i=1}^{n}w_i\left(\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\mu}_{\tilde{B}\sigma(j)}(x_i)\right)^2+\frac{1}{l_i}\sum_{j=1}^{l_i}\left(h^{\nu}_{\tilde{B}\sigma(j)}(x_i)\right)^2\right)\right]$$

$$=C_{IHFS_2}(\tilde{A},\tilde{A})\cdot C_{IHFS_2}(\tilde{B},\tilde{B})$$

Therefore

$$C_{IHFS_2}(\tilde{A},\tilde{B})\le\left[C_{IHFS_2}(\tilde{A},\tilde{A})\right]^{\frac{1}{2}}\cdot\left[C_{IHFS_2}(\tilde{B},\tilde{B})\right]^{\frac{1}{2}}\text{ So, }0\le\rho_{IHFS_2}(\tilde{A},\tilde{B})\le 1.$$

(3) $\tilde{A}=\tilde{B}$

$$\Rightarrow h_{\tilde{A}\sigma(j)}(x_i)=h_{\tilde{B}\sigma(j)}(x_i),x_i\in X$$

$$\Rightarrow\rho_{IHFS_2}(\tilde{A},\tilde{B})=1.$$

Theorem 4 The correlation coefficient of two *IHFSs* $\tilde{A}$ and $\tilde{B}$ defined in Eq. (14), which accounts for the weights $\rho_{IHFS_4}(\tilde{A},\tilde{B})$, satisfies the same properties as those in Theorem 3.

Since the process to prove these properties is analogous to that in Theorem 2, we do not repeat it here.

Example 3 Let $\tilde{A}_1,\tilde{A}_2$,and $\tilde{A}_3$ be three *IHFSs* in $X=\{x_1,x_2,x_3\}$, $w=(0.3,0.3,0.4)^T$ be the weight vector of $x_i(i=1,2,3)$, and

$\tilde{A}_1=\{<x_1,\{(0.9,0.1)\}>,<x_2,\{(0.2,0.1)\}>,<x_3,\{(0.5,0.3),(0.2,0.1)\}>\}$;

$\tilde{A}_2=\{<x_1,\{(0.7,0.2)\}>,<x_2,\{(0.5,0.3)\}>,<x_3,\{(0.6,0.3),(0.4,0.1)\}>\}$;

$\tilde{A}_3=\{<x_1,\{(0.3,0.1)\}>,<x_2,\{(0.3,0.2)\}>,<x_3,\{(0.8,0.1),(0.5,0.3)>\}$;

From Eq. (13), we can obtain:

$\rho_{IHFS_3}(\tilde{A}_1, \tilde{A}_2) = 0.913; \rho_{IHFS_3}(\tilde{A}_1, \tilde{A}_3) = 0.731;\ \rho_{IHFS_3}(\tilde{A}_2, \tilde{A}_3) = 0.873$.

Obviously, $\rho_{IHFS_3}(\tilde{A}_1, \tilde{A}_2) > \rho_{IHFS_3}(\tilde{A}_2, \tilde{A}_3) > \rho_{IHFS_3}(\tilde{A}_1, \tilde{A}_3)$.

4. Correlation Coefficients of IHFSs' Applications

4.1. Clustering Algorithm for IHFSs

Based on the intuitionistic fuzzy clustering algorithm [20], and the correlation coefficient formulas developed previously for IHFSs, in what follows, we develop an algorithm to do clustering under intuitionistic hesitant fuzzy environments. Before doing this, some concepts are introduced firstly:

Definition 10 Let $\tilde{A}_j (j = 1,2,\cdots,m)$ be m $IHFSs$, and $C = (\rho_{ij})_{m\times m}$ be a correlation matrix, where $\rho_{ij} = \rho(\tilde{A}_i, \tilde{A}_j)$ denotes the correlation coefficient of two $IHFSs$ $\tilde{A}_i$ and $\tilde{A}_j$ and satisfies:

$0 \le \rho_{ij} \le 1, i, j = 1,2,\cdots,m;$

$\rho_{ii} = 1, i = 1,2,\cdots,m;$

$\rho_{ij} = \rho_{ji}, i, j = 1,2,\cdots,m.$

Definition 11 ([20]) Let $C = (\rho_{ij})_{m\times m}$ be a correlation matrix, if $C^2 = C \circ C = (\bar{\rho}_{ij})_{m\times m}$, then C^2 is called a composition matrix of C, where

$\bar{\rho}_{ij} = \max_k \{\min\{\rho_{ik}, \rho_{kj}\}\}, i, j = 1,2,\cdots,m$

Theorem 5 ([20]). Let $C = (\rho_{ij})_{m\times m}$ be a correlation matrix. Then the composition matrix $C^2 = C \circ C = (\bar{\rho}_{ij})_{m\times m}$ is also a correlation matrix.

Theorem 6 ([20]) Let C be a correlation matrix. Then for any nonnegative integers m_1 and m_2, the composition matrix $C^{m_1+m_2}$ derived from $C^{m_1+m_2} = C^{m_1} \circ {}^{m_2}$ is still a correlation matrix.

Definition 12 ([20]) Let $C = (\rho_{ij})_{m\times m}$ be a correlation matrix, if $C^2 \subseteq C, i.e$

$$\max_k \{\min\{\rho_{ik}, \rho_{kj}\}\} \le \rho_{ij}, i, j = 1,2,\cdots,m \tag{15}$$

then C is called an equivalent correlation matrix.

Theorem 7 ([20]) Let $C = (\rho_{ij})_{m\times m}$ be a correlation matrix. Then after the finite times of compositions: $C \to C^2 \to C^4 \to \cdots \to C^{2^k} \to \cdots$, there must exist a positive integer k such that $C^{2^k} = C^{2^{(k+1)}}$ and C^{2^k} is also an equivalent correlation matrix.

Definition 13 ([20]) Let $C = (\rho_{ij})_{m\times m}$ be an equivalent correlation matrix. Then we call $C_\lambda = ({}_\lambda\rho_{ij})_{m\times m}$ the λ-cutting matrix of C, where

$${}_\lambda\rho_{ij} = \begin{cases} 0 & if \quad \rho_{ij} < \lambda, \\ 1 & if \quad \rho_{ij} \ge \lambda, \end{cases} i, j = 1,2,\cdots,m \tag{16}$$

and λ is the confidence level with $\lambda \in [0,1]$.

We now propose an algorithm for clustering IHFSs as follows: Algorithm-IHFSC).

Step 1. Let $\{\tilde{A}_1, \tilde{A}_2, \cdots, \tilde{A}_m\}$ be a set of IHFSs in $X = \{x_1, x_2, \cdots, x_m\}$. Using Eq. (13) or Eq. (14), we can calculate the correlation coefficients of the IHFSs, and then construct a correlation matrix $C = (\rho_{ij})_{m\times m}$, where $\rho_{ij} = \rho(\tilde{A}_i, \tilde{A}_j)$.

Step 2. Check whether $C = (\rho_{ij})_{m\times m}$ is an equivalent correlation matrix, *i.e*. check whether it satisfies $C^2 \subseteq C$, where

$$C^2 = C \circ C = (\bar{\rho}_{ij})_{m\times m},\ \bar{\rho}_{ij} = \max_k\{\min\{\rho_{ik}, \rho_{kj}\}\}, i, j = 1,2,\cdots,m$$

If it does not hold, we construct the equivalent correlation matrix C^{2^k}:

$$C \to C^2 \to C^4 \to \cdots \to C^{2^k} \to \cdots, \text{ until } C^{2^k} = C^{2^{(k+1)}}.$$

Step 3. For a confidence level k, we construct a λ-cutting matrix $C_\lambda = ({}_\lambda\rho_{ij})_{m\times m}$ through Definition 13 in order to classify the IHFSs $\tilde{H}_i (i = 1,2,\cdots,m)$. If all elements of the ith line (column) in C_λ are the same as the corresponding elements of the jth line (column) in C_λ, then the IHFSs $\tilde{A}_i$ and $\tilde{A}_j$ are of the same type. By means of this principle, we can classify all these m IHFSs $\tilde{H}_i (i = 1,2,\cdots,m)$.

Below the real example is employed to illustrate the need of the clustering algorithm based on IHFSs:

Example. Consider a movie recommendation system. To better recommend different types of movies on the market, we perform clustering for them according to four attributes ： story (x_1), acting (x_2), visuals (x_3) and direction (x_4). The weighing vector of these four attributes is $w = (0.35, 0.3, 0.2, 0.15)$. Suppose that a group intends to give ratings on ten movies $M_1, M_2, \cdots, M_{10}$. Given the recommenders who make such a recommendation have different backgrounds and levels of knowledge, interests and hobbies, skills, experience and personality, *etc*., this could lead to a difference in the recommendation information. To clearly reflect the differences of the opinions of different recommenders, the data of recommendation information are represented by the IHFSs and listed in Table 1.

Table 1. The Intuitionistic Hesitant Fuzzy Judgment Matrix Provided by the Decision Group

	x_1	x_2	x_3	x_4
M_1	{(0.5,0.3),(0.4,0.2)}	{(0.6,0.2),(0.5,0.3)}	{(0.6,0.2),(0.4,0.4)}	{(0.6,0.4)}
M_2	{(0.8,0.1),(0.7,0.2)}	{(0.7,0.2),(0.6,0.3)}	{(0.7,0.2),(0.6,0.3)}	{(0.7,0.2)}
M_3	{(0.9,0.1),(0.8,0.1)}	{(0.8,0.1),(0.7,0.2)}	{(0.8,0.1),(0.7,0.2)}	{(0.9,0.1)}
M_4	{(0.4,0.5),(0.3,0.6)}	{(0.6,0.2),(0.5,0.4)}	{(0.6,0.2),(0.5,0.3)}	{(0.3,0.6)}
M_5	{(0.6,0.2),(0.5,0.3)}	{(0.3,0.5),(0.2,0.6)}	{(0.3,0.5),(0.2,0.6)}	{(0.1,0.8)}
M_6	{(0.6,0.2),(0.5,0.4)}	{(0.7,0.2),(0.6,0.3)}	{(0.2,0.2),(0.1,0.7)}	{(0.8,0.1)}
M_7	{(0.8,0.1),(0.6,0.2)}	{(0.6,0.2),(0.5,0.3)}	{(0.4,0.4),(0.3,0.5)}	{(0.5,0.4)}
M_8	{(0.7,0.1),(0.6,0.2)}	{(0.4,0.5),(0.3,0.6)}	{(0.6,0.3),(0.5,0.4)}	{(0.8,0.2)}
M_9	{(0.4,0.5),(0.3,0.6)}	{(0.4,0.5),(0.3,0.6)}	{(0.2,0.6),(0.1,0.8)}	{(0.2,0.6)}
M_{10}	{(0.2,0.6),(0.1,0.7)}	{(0.8,0.1),(0.6,0.2)}	{(0.6,0.2),(0.5,0.4)}	{(0.7,0.3)}

Step 1. Calculate the correlation coefficients of the IHFSs $M_i (i = 1,2,\cdots,10)$ by using Eq. (13). Then the derived correlation matrix is:

$$C=\begin{pmatrix}1.000 & 0.957 & 0.925 & 0.923 & 0.809 & 0.920 & 0.950 & 0.920 & 0.799 & 0.884\\ 0.957 & 1.000 & 0.988 & 0.830 & 0.740 & 0.910 & 0.961 & 0.939 & 0.676 & 0.781\\ 0.925 & 0.988 & 1.000 & 0.765 & 0.638 & 0.881 & 0.922 & 0.900 & 0.566 & 0.747\\ 0.923 & 0.830 & 0.765 & 1.000 & 0.830 & 0.809 & 0.845 & 0.796 & 0.877 & 0.924\\ 0.809 & 0.740 & 0.638 & 0.830 & 1.000 & 0.725 & 0.858 & 0.818 & 0.927 & 0.618\\ 0.920 & 0.910 & 0.881 & 0.809 & 0.725 & 1.000 & 0.927 & 0.881 & 0.776 & 0.836\\ 0.950 & 0.961 & 0.922 & 0.845 & 0.858 & 0.927 & 1.000 & 0.923 & 0.793 & 0.753\\ 0.920 & 0.939 & 0.900 & 0.796 & 0.818 & 0.881 & 0.923 & 1.000 & 0.758 & 0.724\\ 0.799 & 0.676 & 0.566 & 0.877 & 0.927 & 0.776 & 0.793 & 0.758 & 1.000 & 0.766\\ 0.884 & 0.781 & 0.747 & 0.924 & 0.618 & 0.836 & 0.753 & 0.724 & 0.766 & 1.000\end{pmatrix}$$

Step 2. Construct the equivalent correlation matrix and obtain:

$$C^2=C\circ C=\begin{pmatrix}1.000 & 0.957 & 0.957 & 0.923 & 0.858 & 0.927 & 0.957 & 0.939 & 0.877 & 0.923\\ 0.957 & 1.000 & 0.988 & 0.923 & 0.858 & 0.927 & 0.961 & 0.939 & 0.830 & 0.884\\ 0.957 & 0.988 & 1.000 & 0.923 & 0.858 & 0.922 & 0.961 & 0.939 & 0.799 & 0.884\\ 0.923 & 0.923 & 0.923 & 1.000 & 0.877 & 0.920 & 0.923 & 0.920 & 0.877 & 0.924\\ 0.858 & 0.858 & 0.858 & 0.877 & 1.000 & 0.858 & 0.858 & 0.858 & 0.927 & 0.830\\ 0.927 & 0.927 & 0.922 & 0.920 & 0.858 & 1.000 & 0.927 & 0.923 & 0.809 & 0.884\\ 0.957 & 0.961 & 0.961 & 0.923 & 0.858 & 0.927 & 1.000 & 0.939 & 0.858 & 0.884\\ 0.939 & 0.939 & 0.939 & 0.920 & 0.858 & 0.923 & 0.939 & 1.000 & 0.818 & 0.884\\ 0.877 & 0.830 & 0.799 & 0.877 & 0.927 & 0.809 & 0.858 & 0.818 & 1.000 & 0.877\\ 0.923 & 0.884 & 0.884 & 0.924 & 0.830 & 0.884 & 0.884 & 0.884 & 0.877 & 1.000\end{pmatrix}$$

It can be seen that $C^2 \subset C$ does not hold. That is to say, the correlation matrix C is not an equivalent correlation matrix. So, we further calculate:

$$C^4=C^2\circ C^2=\begin{pmatrix}1.000 & 0.957 & 0.957 & 0.923 & 0.877 & 0.927 & 0.957 & 0.939 & 0.877 & 0.923\\ 0.957 & 1.000 & 0.988 & 0.923 & 0.877 & 0.927 & 0.961 & 0.939 & 0.877 & 0.923\\ 0.957 & 0.988 & 1.000 & 0.923 & 0.877 & 0.927 & 0.961 & 0.939 & 0.877 & 0.923\\ 0.923 & 0.923 & 0.923 & 1.000 & 0.877 & 0.923 & 0.923 & 0.923 & 0.877 & 0.924\\ 0.877 & 0.877 & 0.877 & 0.877 & 1.000 & 0.877 & 0.877 & 0.877 & 0.927 & 0.877\\ 0.927 & 0.927 & 0.927 & 0.923 & 0.877 & 1.000 & 0.927 & 0.927 & 0.877 & 0.923\\ 0.957 & 0.961 & 0.961 & 0.923 & 0.877 & 0.927 & 1.000 & 0.939 & 0.877 & 0.923\\ 0.939 & 0.939 & 0.939 & 0.923 & 0.877 & 0.927 & 0.939 & 1.000 & 0.877 & 0.923\\ 0.877 & 0.877 & 0.877 & 0.877 & 0.927 & 0.877 & 0.877 & 0.877 & 1.000 & 0.877\\ 0.923 & 0.923 & 0.923 & 0.924 & 0.877 & 0.923 & 0.923 & 0.923 & 0.877 & 1.000\end{pmatrix}$$

and

$$C^8 = C^4 \circ C^4 = \begin{pmatrix} 1.000 & 0.957 & 0.957 & 0.923 & 0.877 & 0.927 & 0.957 & 0.939 & 0.877 & 0.923 \\ 0.957 & 1.000 & 0.988 & 0.923 & 0.877 & 0.927 & 0.961 & 0.939 & 0.877 & 0.923 \\ 0.957 & 0.988 & 1.000 & 0.923 & 0.877 & 0.927 & 0.961 & 0.939 & 0.877 & 0.923 \\ 0.923 & 0.923 & 0.923 & 1.000 & 0.877 & 0.923 & 0.923 & 0.923 & 0.877 & 0.924 \\ 0.877 & 0.877 & 0.877 & 0.877 & 1.000 & 0.877 & 0.877 & 0.877 & 0.927 & 0.877 \\ 0.927 & 0.927 & 0.927 & 0.923 & 0.877 & 1.000 & 0.927 & 0.927 & 0.877 & 0.923 \\ 0.957 & 0.961 & 0.961 & 0.923 & 0.877 & 0.927 & 1.000 & 0.939 & 0.877 & 0.923 \\ 0.939 & 0.939 & 0.939 & 0.923 & 0.877 & 0.927 & 0.939 & 1.000 & 0.877 & 0.923 \\ 0.877 & 0.877 & 0.877 & 0.877 & 0.927 & 0.877 & 0.877 & 0.877 & 1.000 & 0.877 \\ 0.923 & 0.923 & 0.923 & 0.924 & 0.877 & 0.923 & 0.923 & 0.923 & 0.877 & 1.000 \end{pmatrix} = C^4$$

Hence, C^4 is an equivalent correlation matrix.

Step 3. For a confidence level λ, to do clustering for IHFSs, we construct a λ-cutting matrix $C_\lambda = ({}_\lambda\rho_{ij})_{m\times m}$ by Definition 13, and based on which, we get all possible classifications of $M_i (i = 1,2,\cdots,10)$

(1)If $0 < \lambda \le 0.877$, then $M_i (i = 1,2,\cdots,10)$ are of the same type:

$\{M_1, M_2, M_3, M_4, M_5, M_6, M_7, M_8, M_9, M_{10}\}$

(2)If $0.877 < \lambda \le 0.923$, then $M_i (i = 1,2,\cdots,10)$ are classified into two types:

$\{M_1, M_2, M_3, M_4, M_6, M_7, M_8, M_{10}\}, \{M_5, M_9\}$

(3)If $0.923 < \lambda \le 0.924$, then $M_i (i = 1,2,\cdots,10)$ are classified into three types:

$\{M_1, M_2, M_3, M_6, M_7, M_8\}, \{M_5, M_9\}, \{M_4, M_{10}\}$

(4)If $0.924 < \lambda \le 0.927$, then $M_i (i = 1,2,\cdots,10)$ are classified into four types:

$\{M_1, M_2, M_3, M_6, M_7, M_8\}, \{M_4\}, \{M_5, M_9\}, \{M_{10}\}$

(5) If $0.927 < \lambda \le 0.939$, then $M_i (i = 1,2,\cdots,10)$ are classified into six types:

$\{M_1, M_2, M_3, M_7, M_8\}, \{M_4\}, \{M_5\}, \{M_6\}, \{M_9\}, \{M_{10}\}$

(6) If $0.939 < \lambda \le 0.957$, then $M_i (i = 1,2,\cdots,10)$ are classified into seven types:

$\{M_1, M_2, M_3, M_7\}, \{M_4\}, \{M_5\}, \{M_6\}, \{M_8\}, \{M_9\}, \{M_{10}\}$

(7) If $0.957 < \lambda \le 0.961$, then $M_i (i = 1,2,\cdots,10)$ are classified into eight types:

$\{M_1\}, \{M_2, M_3, M_7\}, \{M_4\}, \{M_5\}, \{M_6\}, \{M_8\}, \{M_9\}, \{M_{10}\}$

(8) If $0.961 < \lambda \le 0.988$, then $M_i (i = 1,2,\cdots,10)$ are classified into nine types:

$\{M_1\}, \{M_2, M_3\}, \{M_4\}, \{M_5\}, \{M_6\}, \{M_7\}, \{M_8\}, \{M_9\}, \{M_{10}\}$

(9) If $0.988 < \lambda \le 1$, then $M_i (i = 1,2,\cdots,10)$ are classified into ten types:

$\{M_1\}, \{M_2\}, \{M_3\}, \{M_4\}, \{M_5\}, \{M_6\}, \{M_7\}, \{M_8\}, \{M_9\}, \{M_{10}\}$

Under the group recommendation setting, the recommenders' information usually does not reach an agreement for the items that need to be classified. This example clearly shows that the clustering algorithm based on IHFSs provides a proper way to resolve this issue.

4.2. Correlation Coefficient for Group Recommendation

In this section, we shall utilize the correlation coefficient of IHFSs to group recommendation with intuitionistic hesitant fuzzy information.

For a group recommendation problem with intuitionistic hesitant fuzzy information, let $M = \{M_1, M_2, \cdots, M_m\}$ be a discrete set of items and $X = \{x_1, x_2, \cdots, x_n\}$ be a set of attributes. If the recommenders provide several values for the item $M_i (i = 1,2,\cdots,m)$ under the attribute $x_j (j = 1,2,\cdots,n)$, these values can be considered as a IHFE $ih_{ij} = (\mu_{ij}, \nu_{ij}), (j = 1,2,\cdots,n; i = 1,2,\cdots,m)$.

Therefore, we can elicit a intuitionistic hesitant fuzzy decision matrix $D = (ih_{ij})_{m\times n}$ where ih_{ij}, $(j = 1,2,\cdots,n; i = 1,2,\cdots,m)$ is in the form of intuitionistic hesitant fuzzy elements.

In group recommendation environments, the concept of ideal point has been used to help the identification of the best item in the recommendation set. Although the ideal item does not exist in real world, it does provide a useful theoretical construct to recommend items. Therefore, we define each ideal IHFE $ih^*_{ij} = (1,0), (j = 1,2,\cdots,n)$ in the ideal item $M^* = \{< x_j, ih^*_j >| x_j \in X\}$, $(j = 1,2,\cdots,n)$.

The weighting vector of attributes for the different importance of each attribute is given as $w = (w_1, w_2,\cdots,w_n)$, where $w_j \geq 0$, $(j = 1,2,\cdots,n)$, and $\sum_{j=1}^{n} w_j = 1$ .Then, we utilize the weighted correlation coefficient for group recommendation problems with intuitionistic hesitant fuzzy information, which can be described as follows:

Step1. Calculate the weighted correlation coefficient between an item $M_i (i = 1,2,\cdots,m)$ and the ideal item M^* by using the following Eq. (13):

Step2. Rank the items in accordance with the values of weighted correlation coefficients.

Step3. Select the best item according to the maximum value of weighted correlation coefficients.

Step4. End.

Then, we utilize the developed approach to obtain the ranking order of the movie recommendation and the most desirable one(s), which can be described as follows:

Step1. By applying Eq. (13), we can obtain the computing results of the weighted correlation coefficients between the items and the ideal item as follows:

$\rho_{IHFS_3}(M^*, M_1) = 0.860$, $\rho_{IHFS_3}(M^*, M_2) = 0.950$, $\rho_{IHFS_3}(M^*, M_3) = 0.983$, $\rho_{IHFS_3}(M^*, M_4) = 0.688$, $\rho_{IHFS_3}(M^*, M_5) = 0.518$, $\rho_{IHFS_3}(M^*, M_6) = 0.806$,
$\rho_{IHFS_3}(M^*, M_7) = 0.854$, $\rho_{IHFS_3}(M^*, M_8) = 0.817$, $\rho_{IHFS_3}(M^*, M_9) = 0.431$,
$\rho_{IHFS_3}(M^*, M_{10}) = 0.687$.

Step2. We can rank the items in accordance with the values of weighted correlation coefficients: $M_3 \succ M_2 \succ M_1 \succ M_7 \succ M_8 \succ M_6 \succ M_4 \succ M_{10} \succ M_5 \succ M_9$

Step3. The movie M_3 is the best choice according to the maximum value among ten weighted correlation coefficients.

The example clearly indicates that the proposed recommending method is simple and effective under intuitionistic hesitant fuzzy environments.

As mentioned above, the intuitionistic hesitant fuzzy set is a further generalization of hesitant fuzzy set and intuitionistic fuzzy set. So the intuitionistic hesitant fuzzy set contains more information than the hesitant fuzzy set (only membership hesitant degrees) and the intuitionistic fuzzy set (both membership degree and non-membership degree). Then, the proposed correlation coefficient of IHFSs is a further generalization of the correlation coefficients of hesitant fuzzy sets and intuitionistic fuzzy sets. Furthermore, correlation coefficients of hesitant fuzzy sets and intuitionistic fuzzy sets are special cases of the correlation coefficient of IHFSs proposed in this paper. Therefore, the correlation coefficient between IHFSs proposed in this paper can be used to solve not only correlation problems with IHFSs but also correlation problems of hesitant fuzzy sets and intuitionistic fuzzy sets, is only suitable for the correlation problem with hesitant fuzzy sets or intuitionistic fuzzy sets.

5. Conclusion

The IHFS is a comprehensive set encompassing several existing sets, and its membership degrees are represented by a set of possible intuitionistic fuzzy values. Therefore, it has the desirable characteristics and advantages of its own and appears to be a more flexible method to be valued in multifold ways according to the practical demands than the existing hesitant fuzzy sets, taking into account much more information given by recommenders. In this paper, we proposed the correlation coefficient of IHFSs as a new extension. Finally, a practical example of group recommendation was given to verify the developed approach and to demonstrate its practicality and effectiveness. The proposed method has a clear concept, simple calculation and provides a new idea for solving group recommendation and other clustering analysis problems.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable suggestions as to how to improve this paper. This work was supported by the National Natural Science Foundation of China (Grant Nos. 71490725, 71371062), the National Basic Research Development Program of China (973 Program) (Grant Nos.2013CB329603), the Natural Science Foundation of the Anhui Educational Committee of China (Grant Nos. KJ2015A300), the National Undergraduate Training Programs for Innovation and Entrepreneurship (Grant Nos. 201412216026, 201512216009), and Anhui Xinhua University Students' Quality education Project (Nos. IFQE201408).

References

[1] K.T. Atanassov, “Intuitionistic fuzzy sets”, Fuzzy sets and Systems, vol. 20, no.1, **(1986)**, pp. 87-96.
[2] L. A. Zadeh, “Fuzzy sets”, Information and control, vol. 8, no. 3, **(1965)**, pp. 338-353.
[3] D. F. Li, “Mathematical-programming approach to matrix games with payoffs represented by Atanassov's interval-valued intuitionistic fuzzy sets”, Fuzzy Systems, IEEE Transactions on, vol. 18, no. 6, **(2010)**, pp. 1112-1128.
[4] Z. Chen and W. Yang, “A new multiple attribute group decision making method in intuitionistic fuzzy setting”, Applied Mathematical Modelling, vol. 35, no.9, **(2011)**, pp. 4424-4437.
[5] D.F. Li and C.T. Cheng, “New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions”, Pattern Recognition Letters, vol. 23, no.1, **(2002)**, pp. 221–225.
[6] T. Chaira, “Intuitionistic fuzzy segmentation of medical images”, IEEE Transactions on, vol. 57, no.6, **(2010)**, pp. 1430-1436.
[7] G.W. Wei, “Gray relational analysis method for intuitionistic fuzzy multiple attribute decision making”, Expert systems with Applications, vol. 38, no.9, **(2011)**, pp. 11671–11677.
[8] V. Torra, “Hesitant fuzzy sets”, International Journal of Intelligent Systems, vol. 25, no.6 **(2010)**, pp. 529-539.
[9] R.M. Rodríguez, L. Martínez, V. Torra, Z.S. Xu and F. Herrera, “Hesitant fuzzy sets: state of the art and future directions”, International Journal of Intelligent Systems, vol. 29, no.6 **(2014)**, pp. 495-524.
[10] N. Chen, Z.S. Xu and M. M. Xia, “Interval-valued hesitant preference relations and their applications to group decision making”, Knowledge-Based Systems, vol. 37, no.1, **(2013)**, pp. 528-540.
[11] X.H. Wu, B. Wu, J. Sun and J.W. Zhao, “Mixed fuzzy inter-cluster separation clustering algorithm”, Applied Mathematical Modelling, vol. 35, no.10, **(2011)**, pp. 4790-4795.
[12] Z.S. Xu, “Intuitionistic fuzzy hierarchical clustering algorithms”, Journal of Systems Engineering and Electronics, vol. 20, no.1, **(2009)**, pp. 90-97.
[13] Z.S. Xu, J. Chen and J.J. Wu, “Clustering algorithm for intuitionistic fuzzy sets”, Information Sciences, vol. 178, no.19, **(2008)**, pp. 3775-3790.
[14] N. Chen, Z.S. Xu and M.M. Xia, “Correlation coefficients of hesitant fuzzy sets and their applications to clustering analysis”, Applied Mathematical Modelling, vol. 37, no.4, **(2013)**, pp. 2197–2211.

[15] M. M. Xia and Z.S. Xu, "Hesitant fuzzy information aggregation in decision making", International Journal of Approximate Reasoning, vol. 52, no.3, **(2011)**, pp. 395-407.
[16] T. Gerstenkorn and J. Man′ ko, "Correlation of intuitionistic fuzzy sets", *Fuzzy Sets and Systems*, vol. 44, no.1, **(1991)**, pp. 39-43.
[17] W.L. Hung and J.W. Wu, "Correlation of intuitionistic fuzzy sets by centroid method", Information Sciences, vol. 144, no.1, **(2002)**, pp. 219-225.
[18] D. Dumitrescu, "A definition of an informational energy in fuzzy sets theory", Babes-Bolyai Math, vol. 22, no.2, **(1977)**, pp. 57-59.
[19] Z.S. Xu and M.M. Xia, "Distance and similarity measures for hesitant fuzzy sets", Information Sciences, vol. 181, no.11, **(2011)**, pp. 2128-2138.
[20] Z.S. Xu, J. Chen and J.J. Wu, "Clustering algorithm for intuitionistic fuzzy sets", Information Sciences, vol. 178, no.19, **(2008)**, pp. 3775-3790.