The Algorithm for Evaluation of Landscape Water Quality based on intuitionistic fuzzy sets of emphasizing hesitancy degree

HaiFeng Wang¹,², a, Kun Zhang¹,², b, Zhuang Li¹,², c, HongXu Wang¹, d

¹College of Computer Engineering, Hainan Tropical Ocean University, Sanya, Hainan, 572022, China
²Sanya Key Laboratory of Computer Vision, Sanya, Hainan, 572022, China
awyxzxz@163.com, bzk0588@163.com, chitlz@163.com, dwhx16233@126.com

Keywords: Intuitionistic fuzzy sets; Water quality analysis; Distance; New definition; New formula.

Abstract: Distance measure of intuitionistic fuzzy sets (IFS) is a measure of the difference sizes between two IFS. The original distance measure definition is modified in this paper, the new distance measure definition and a new distance formula of IFS are provided. Pattern recognition method based on distance of intuitionistic fuzzy sets is used for researching on water quality analysis problem. That provides a new way for the study of the same problem. The method is simpler than the current method and the results are in line with the actual.

Introduction

Intuitionistic fuzzy sets theory was proposed by Atanassov in 1986[1], and vague sets theory was proposed by Gao and Buehrer in 1993[2]. They are all extension of fuzzy sets theory[3]. At the same time, Literature [4] have proved that vague sets are IFS. In addition, Deschrijver and Kerre have also proved that interval-valued fuzzy sets and IFS are essentially same[5]. Compared with the fuzzy sets, IFS consider simultaneously membership degree, non-membership degree and hesitancy degree. IFS are more comprehensive, meticulous and intuitive for the fuzzy information description. Distance is a measure of IFS and an important part of IFS theory. Especially it is applied in the intuitionistic fuzzy reasoning, intuition fuzzy pattern recognition, intuitionistic fuzzy multiple attribute decision making, and many other aspects. So distance formula of IFS is constantly proposed. Such as literature [6, 7, 8, 9], it is a pity that many distance formula of IFS are flawed. In this paper a new distance measure definition and distance formula of IFS are put forward. And water quality assessment algorithm based on the distance formula is put forward. The experimental results show that the algorithm is simple and feasible, and the calculation result is correct.

Propaedeutics and Existing Distance Formula of IFS

Definition 1 [1]: Let \( X = \{x\} \) be nonempty set, \( A = \{< x, \mu_A(x), \nu_A(x) >| x \in X\} \) be named Intuitionistic Fuzzy Sets (IFS) in \( X \). Among \( x \in X \), \( \mu_A : x \to [0,1] \) and \( \nu_A : x \to [0,1] \) are respectively named membership degree and non-membership degree of IFS \( A \), and they satisfy constraint condition \( \forall x \in X \) \( 0 \leq \mu_A(x) + \nu_A(x) \leq 1 \). IFS \( (X) \) is defined IFS in domain of discourse \( X \).

Definition 2 For \( x \in X \), \( S_A(x) = \mu_A(x) - \nu_A(x) \) is named score values of \( x \) in \( A \) and \( \pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \) is named hesitancy degree of element \( x \) in \( A \) by intuitionistic fuzzy value \( \langle x, \mu_A(x), \nu_A(x) \rangle \). \( \langle x, 0, 1 \rangle \) and \( \langle x, 1, 0 \rangle \) are named ordinary intuitionistic fuzzy values. \( \langle x, \beta, 1 - \beta \rangle \) \( (0 < \beta < 1) \) is named Fuzzy value. \( \langle x, \alpha, \beta \rangle \) \( (\alpha = 1 - \beta, \alpha, \beta \in [0,1]) \) is named true intuitionistic fuzzy value. \( \langle x, 0, 0 \rangle \) is named special intuitionistic fuzzy value.

Let \( A, B \in IFS(X) \), \( A = \{< x, \mu_A(x), \nu_A(x) >| x \in X\} \) and \( B = \{< x, \mu_B(x), \nu_B(x) >| x \in X\} \), where \( x \in X \), \( A \subseteq B \) satisfy \( \mu_A(x) \leq \mu_B(x), \nu_A(x) \geq \nu_B(x) \).

In order to structure new distance measure formula, we introduce a data mining methods about intuitionistic fuzzy values on the basis of literature [6].
Definition 3 A kind of data mining methods about intuitionistic fuzzy values as follows: let $x = (x, \mu(x), \nu(x))$, then $\mu^{(0)}(x) = \mu(x), \nu^{(0)}(x) = \nu(x), \pi^{(0)}(x) = \pi(x), \mu^{(m)}(x) = \mu(x)(1 + \pi(x) + \pi^2(x) + \cdots + \pi^m(x)), \nu^{(m)}(x) = \nu(x)(1 + \pi(x) + \pi^2(x) + \cdots + \pi^m(x)), \pi^{(m)}(x) = \pi^{m+1}(x), m = 1, 2, \cdots$.

For the Definition 3, when the parameter $m$ is valued differently, we will get a series of intuitionistic fuzzy values $x^{(m)} = (x^{(m)}, \mu^{(m)}, \nu^{(m)})$ by intuitionistic fuzzy values $x = (x, \mu(x), \nu(x))$.

Definition 4 Let $A, B, C \in IFS(X)$, $D: IFS(X) \times IFS(X) \to [0, 1]$ be named distance measure between IFS $A$ and $B$. If it satisfies the following natures:

(P1) $D(A, B) = D(B, A)$;
(P2) $D(A, B) = 0$ if and only if $A = B$;
(P3) $0 \leq D(A, B) \leq 1$;
(P4) if $A \subseteq B \subseteq C$, $D(A, C) \leq D(A, B) \lor D(B, C)$.

Formula (1)-(4) all satisfy (P1)-(P1).

A new Distance Measure Formula of IFS

Based on the above, a new distance measure definition of IFS should be highlighted problem that is: when hesitate degrees is not zero, two equal distance measures of intuitionistic fuzzy value should be not zero. And (P4) is recovered into classic triangle inequality of distance measures, we can obtain:

Definition 5 Let $A, B, C \in IFS(X)$, then $D: IFS(X) \times IFS(X) \to [0, 1]$ is said to be distance measures of IFS $A$ and $B$. If it satisfies the following criterion:

(C1) $D(A, B) = D(B, A)$;
(C2) $D(A, B) = 0$, if and only if $A = B$ and $\pi_a(x_i) = 0, \pi_b(x_i) = 0, i = 1, 2, \cdots, n$;
(C3) $0 \leq D(A, B) \leq 1$;
(C4) $D(A, C) \leq D(A, B) + D(B, C)$.

Theorem 1 Let $x = (x, \mu(x), \nu(x)), y = (y, \mu(y), \nu(y))$ is intuitionistic fuzzy value, then the following formula is the distance measure between intuitionistic fuzzy values $x$ and $y$ ($m = 0, 1, 2, \cdots$).

$$D^{(m)}(x, y) = \frac{3[\mu^{(m)}(x) - \mu^{(m)}(y)] + 3[\nu^{(m)}(x) - \nu^{(m)}(y)] + 3[S^{(m)}(x) - S^{(m)}(y)] + 2\pi^{(m)}(x) + 2\pi^{(m)}(y)}{12}$$

(6)

Analogous, the distance definition of IFS is given and the weighted distance of IFS is defined as:

Theorem 2 Let $A = \{(x_i, \mu_a(x_i), \nu_a(x_i) : x_i \in X)\}$ and $B = \{(x_i, \mu_b(x_i), \nu_b(x_i) : x_i \in X)\}$ are shown as IFS in limited theory domain $X$. The following formula $D^{(m)}(E, C)$ is the distance ($m = 0, 1, 2, \cdots$) between IFS $E$ and $C$.

$$D^{(m)}(E, C) = \frac{1}{12n} \sum_{i=1}^{n} \left[3[\mu^{(m)}_e(x_i) - \mu^{(m)}_c(x_i)] + 3[\nu^{(m)}_e(x_i) - \nu^{(m)}_c(x_i)] + 3[S^{(m)}_e(x_i) - S^{(m)}_c(x_i)] + 2\pi^{(m)}_e(x_i) + 2\pi^{(m)}_c(x_i)\right]$$

(7)

Theorem 3 The following formula $WD^{(m)}(E, C)$ is the weighted distance($m = 0, 1, 2, \cdots$) between IFS $E$ and $C$.

$$WD^{(m)}(E, C) = \frac{1}{12} \sum_{i=1}^{n} \omega_i \left[3[\mu^{(m)}_e(x_i) - \mu^{(m)}_c(x_i)] + 3[\nu^{(m)}_e(x_i) - \nu^{(m)}_c(x_i)] + 3[S^{(m)}_e(x_i) - S^{(m)}_c(x_i)] + 2\pi^{(m)}_e(x_i) + 2\pi^{(m)}_c(x_i)\right]$$

(8)

Among the weight of one element $x_i$ is $\omega_i (0 \leq \omega_i \leq 1)$ and $\sum_{i=1}^{n} \omega_i = 1$.

Water quality analysis algorithm based on IFS distance and Application Example

Pattern recognition methods based on IFS are widely used in many fields, such as literature [9]. Aim at water quality analysis problem, the specific steps that are suggested by new IFS distance formula for water quality analysis algorithm are as follows:

Step 1: Set up index set $X = \{x_1, x_2, \cdots, x_n\}$, when $n = (1, 2, 3, 4)$, among $x_i$ be chemical oxygen demand (COD); $x_2$ be ammonia nitrogen(NH3-N); $x_3$ be total phosphorus(TP); $x_4$ be total nitrogen(TN). The unit be mg/L.
Step 2: Set up the standard sample water set \( E = \{ E_1, E_2, E_3, E_4, E_5 \} \) in index set \( X \), among water quality \( E_i \) be Class i. Pollution levels corresponding to clean, no pollution, light pollution, pollution and heavy pollution.

Set up sample water set \( C = \{ C_1, C_2, C_3 \} \), among \( C_1 \) be center sampling points in Ruyi lake, \( C_2 \) be the southern tip of sampling point, \( C_3 \) be the north end of sampling points.

Step 3: Creating an environment of IFS. The original data is changed into IFS data. The conversion formula that changed the nonnegative single value data into IFS data was given. In this case, the form as follow [10]:

\[
x_j = \langle x_j, \mu(x_j), v(x_j) \rangle = \left( x_j, \frac{x_j}{x_{j_{\max}}}, \left( \frac{x_j}{x_{j_{\max}}} \right)^2 \right)
\]  

(9)

Among \( x_{j_{\max}} = \max \{ x_{j_1}, x_{j_2}, x_{j_3}, x_{j_4}, x_{j_5}, x_{j_6}, x_{j_7}, x_{j_8} \} \), \( j = 1, 2, 3, 4 \).

Step 4: Making use of formula (7), we counts the distance of IFS between awaiting assessment sample \( C_j (j = 1, 2, \ldots, k) \) and standard sample \( E_i (i = 1, 2, \ldots, h) \).

Step 5: Pattern recognition based on distance of IFS. If the distance is numerical order as follow:

\[
D(C_j, E_i) \leq D(C_j, E_{i_2}) \leq \cdots \leq D(C_j, E_{i_k}) (j = 1, 2, \ldots, k),
\]

awaiting assessment sample \( C_j \) belongs to standard sample \( E_{i_k} \), among \( i_{k_1} \ldots i_k \) be a no duplicate permutations about \( 1, 2, \ldots, h \).

To study the Ruyi lake center in ZhengZhou, Monitoring data of the southern tip and the north end are acted as awaiting assessment sample. National standards are acted as the standard sample. Data took from literature [11],[12]. Using the distance pattern recognition method of IFS based on the distance, we proceed to analysis water quality. The original standard sample set data \( E \) and awaiting assessment sample set \( C \) data are shown in Table 1. they are changed into IFS making use of formula (9) in Table 2.

Table 1 Chinese Environmental quality standards and awaiting assessment samples for surface water

<table>
<thead>
<tr>
<th>( x_1 ) (COD)</th>
<th>( x_2 ) (NH3-N)</th>
<th>( x_3 ) (TP)</th>
<th>( x_4 ) (TN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_1 )</td>
<td>17</td>
<td>0.28</td>
<td>0.12</td>
</tr>
<tr>
<td>( C_2 )</td>
<td>20</td>
<td>0.20</td>
<td>0.12</td>
</tr>
<tr>
<td>( C_3 )</td>
<td>14</td>
<td>0.17</td>
<td>0.10</td>
</tr>
<tr>
<td>( E_1 )</td>
<td>15</td>
<td>0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>( E_2 )</td>
<td>15</td>
<td>0.50</td>
<td>0.10</td>
</tr>
<tr>
<td>( E_3 )</td>
<td>20</td>
<td>1.00</td>
<td>0.20</td>
</tr>
<tr>
<td>( E_4 )</td>
<td>30</td>
<td>1.50</td>
<td>0.30</td>
</tr>
<tr>
<td>( E_5 )</td>
<td>40</td>
<td>2.00</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Table 2 Chinese Environmental quality standards and awaiting assessment samples for surface water

<table>
<thead>
<tr>
<th>( x_1 ) (COD)</th>
<th>( x_2 ) (NH3-N)</th>
<th>( x_3 ) (TP)</th>
<th>( x_4 ) (TN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_1 )</td>
<td>( x_1, 0.425, 0.562 )</td>
<td>( x_2, 0.140, 0.374 )</td>
<td>( x_3, 0.300, 0.548 )</td>
</tr>
<tr>
<td>( C_2 )</td>
<td>( x_1, 0.500, 0.707 )</td>
<td>( x_2, 0.100, 0.316 )</td>
<td>( x_3, 0.300, 0.548 )</td>
</tr>
<tr>
<td>( C_3 )</td>
<td>( x_1, 0.350, 0.592 )</td>
<td>( x_2, 0.085, 0.292 )</td>
<td>( x_3, 0.250, 0.500 )</td>
</tr>
<tr>
<td>( E_1 )</td>
<td>( x_1, 0.375, 0.612 )</td>
<td>( x_2, 0.075, 0.274 )</td>
<td>( x_3, 0.050, 0.224 )</td>
</tr>
<tr>
<td>( E_2 )</td>
<td>( x_1, 0.375, 0.612 )</td>
<td>( x_2, 0.250, 0.500 )</td>
<td>( x_3, 0.250, 0.500 )</td>
</tr>
<tr>
<td>( E_3 )</td>
<td>( x_1, 0.500, 0.707 )</td>
<td>( x_2, 0.500, 0.707 )</td>
<td>( x_3, 0.500, 0.707 )</td>
</tr>
<tr>
<td>( E_4 )</td>
<td>( x_1, 0.750, 0.866 )</td>
<td>( x_2, 0.750, 0.866 )</td>
<td>( x_3, 0.750, 0.866 )</td>
</tr>
<tr>
<td>( E_5 )</td>
<td>( x_1, 1.000, 1.000 )</td>
<td>( x_2, 1.000, 1.000 )</td>
<td>( x_3, 1.000, 1.000 )</td>
</tr>
</tbody>
</table>
Making use of formula (7), when parameter be $m = 2$, the distance of IFS is calculated between awaiting assessment sample $C_i (i=1,2,3)$ and the standard sample $E_i (i=1,2,3,4,5)$. The results as follows:

$$
D^{(2)}(C_1, E_1) = 0.207, D^{(2)}(C_1, E_2) = 0.104, D^{(2)}(C_1, E_3) = 0.25, D^{(2)}(C_1, E_4) = 0.451, D^{(2)}(C_1, E_5) = 0.603.
$$

$$
D^{(2)}(C_2, E_1) = 0.212, D^{(2)}(C_2, E_2) = 0.14, D^{(2)}(C_2, E_3) = 0.229, D^{(2)}(C_2, E_4) = 0.461, D^{(2)}(C_2, E_5) = 0.599.
$$

$$
D^{(2)}(C_3, E_1) = 0.225, D^{(2)}(C_3, E_2) = 0.16, D^{(2)}(C_3, E_3) = 0.267, D^{(2)}(C_3, E_4) = 0.448, D^{(2)}(C_3, E_5) = 0.60.
$$

And IFS distance sorted by size: $D^{(2)}(C_1, E_1) < D^{(2)}(C_1, E_2) < D^{(2)}(C_1, E_3) < D^{(2)}(C_1, E_4) < D^{(2)}(C_1, E_5)$, $D^{(2)}(C_2, E_1) < D^{(2)}(C_2, E_2) < D^{(2)}(C_2, E_3) < D^{(2)}(C_2, E_4) < D^{(2)}(C_2, E_5)$, $D^{(2)}(C_3, E_1) < D^{(2)}(C_3, E_2) < D^{(2)}(C_3, E_3) < D^{(2)}(C_3, E_4) < D^{(2)}(C_3, E_5)$. We can obtain that the following result. $C_1$, $C_2$ and $C_3$ all belong to the standard sample $E_2$ and are Class II uncontaminated water quality.

**Conclusion**

About monitoring data, we make use of the vague set pattern recognition method based on distance to obtain that the evaluation results of awaiting assessment samples are all Class II uncontaminated water quality. The results are in accordance with the evaluation in literature [12]. In this paper, we give the vague set pattern recognition method based on distance and use a new distance formula of intuitionist fuzzy sets which give full play to their role as the hesitation degree proposed to research the water quality analysis problem. Meanwhile we come up with the water quality assessment algorithm and provide a new way for the study of the same problem. The method is simple based on the current method and the results conform to the actual.

**Acknowledgements**

This work was financially supported by the key project of scientific research of Hainan Province (No.Hnky2015ZD-14), the Hainan Planning projects of philosophy and Social Sciences (No.HNSK(QN)15-60), the scientific and technological cooperative project for college and region of Sanya(2015YD16,2015YD43,2014YD30), the key Laboratory of Sanya Project(No.L1410).

**References**


