Fuzzy Generalization and Comparisons for the Rand Index

Miin-Shen Yang,1,* Chiou-Cherng Yeh1,2
1Department of Applied Mathematics, Chung Yuan Christian University, Chung-Li 32023, Taiwan
2Teaching Center of Natural Science, Minghsin University of Science and Technology, Hsinchu, Taiwan

To generalize the Rand index ($RI$) from crisp partitions to fuzzy partitions, we first propose a graph method in which color edges in the graph for crisp partitions are used to determine the relation matrix between objects such that the matrix trace can be employed to calculate the $RI$. This approach is then introduced into fuzzy partitions to generalize the $RI$ to the fuzzy $RI$ ($FRI$). Compared with previous fuzzy generalizations, the most unique aspect of our method has the following important characteristics that for any two partition matrices $M^{(1)}$ and $M^{(2)}$, the result with $M^{(1)} = M^{(2)}$ is the necessary and sufficient condition for the result that the $FRI$ is equal to 1. This important characteristic renders our fuzzy generalization of the $RI$ is not only able to determine the similarities between fuzzy partitions and crisp reference partitions, but also to identify the similarity between fuzzy partitions and fuzzy reference partitions. The method can even be used to explore and compare the similarities between various data sets and the same fuzzy reference partition. Finally, we use synthetic data and real data to give more demonstrations, and further perform comparisons of our method with those existing fuzzy extensions of the $RI$. © 2017 Wiley Periodicals, Inc.

1. INTRODUCTION

In the literature, Jaccard1 seems to be the first one to define similarity measures between two clustering methods. In 1971, Rand2 proposed objective criteria for the evaluation of clustering methods. Afterwards, similarity indices between clustering methods (partitions) become popular. Up to now, different indices of similarities between two partitions have been widely studied and applied in various areas. For example, Jaccard1 defined a similarity as the ratio of the total number of pairs that any two objects from an object data set are sorted into the same cluster in two clusterings, and the total number of pairs that any two objects are sorted into the same cluster in the clusterings at least once. This result is known as the Jaccard index. Fowlkes and Mallows3 proposed dividing the total pairs, for which any two
objects are sorted into the same clusters, in two different clusterings by the geometric
mean of the number of pairs in which the objects are sorted into the same cluster in
the first clustering and the number of pairs in which the objects are sorted into the
same cluster in the second clustering.

Rand\textsuperscript{2} believed that any two objects sorted into different clusters in two clus-
terings could be considered a type of similarity. Therefore, Rand\textsuperscript{2} defined similarity
as the ratio of the total number of pairs of any two object sorted into either the same
cluster or different clusters in two clusterings to the total number of pairs of objects.
This similarity is also known as the Rand index (RI). Under the assumption of hyper-
geometric distribution, Hubert and Arabie\textsuperscript{4} used the expected and maximum values
of the RI to propose the adjusted RI. Triple of any three objects in the object data
set was used to define the concordant triple and discordant triple. Finally, Hubert
and Arabie\textsuperscripts{4} provided some opinions on measuring the similarity of two clusterings.
Anderson et al.\textsuperscript{5} helpfully collated the 14 methods developed by various researchers
in their study to provide a reference. Overall, various scholars differ regarding the
objective criteria for calculating the similarity of two clusterings, but most rely on
whether any two objects in the object data set are classified into the same or different
clusters.

Uncertainty always exists in nature and real systems. Measures of uncertainty
and uncertain decision making were studied.\textsuperscript{6,7} It is known that probability has been
used traditionally in modeling uncertainty. Since Zadeh\textsuperscript{8} proposed fuzzy set theory,
it has been widely used to handle another type of uncertainty called fuzziness. In
general, fuzziness commonly exists in real-world systems. Membership function is
a key idea in fuzzy set theory that produces partial membership grades of multiple
classes for the presentation of unsharp boundaries (see Tamir et al.\textsuperscript{9}). Membership
values are always ranged between 0 and 1 that represent the membership grades of
elements. In fuzzy set applications, fuzzy clustering had been widely studied,\textsuperscript{10–13}
where membership functions can be produced from most fuzzy clustering algo-

rithms. It can be well used as representation of human knowledge in complex
systems, and so using fuzzy set is a more suitable tool for handling uncertainty in
practical applications.\textsuperscript{14} According to actual applications in recent years, including
genetics, voice recognition, image recognition, and educational psychology, understand-
ing of how to generalize the RI from crisp partitions into fuzzy partitions is
urgently required. We mention that, although the adjusted RI\textsuperscript{4} is also commonly
employed, the premises required for hypergeometric distribution are incompatible
with fuzzy partition concept. Therefore, we will not consider the adjusted RI in this
paper. On the other hand, these fuzzy generalizations of the RI can be straightly used
as fuzzy generalizations of the Jaccard index. In this sense, we only focus on these
fuzzy generalizations of the RI.

In recent years, to widen the use of similarity between two clusterings,
many researchers\textsuperscript{5,15–19} had devoted considerable effort to generalizing the RI
into fuzzy partitions. In set theory, a binary relation between the object data
set \(O = \{o_1, o_2, \cdots, o_n\}\) and the crisp partition \(P = \{S_1, S_2, \cdots, S_k\}\), where
\(\bigcup_{h=1}^{k} S_h = P\) and \(S_h \cap S_{h'} = \emptyset\) for all \(h \neq h'\), is defined as
\(m(o_i, S_h) = \begin{cases} 1 & \text{if } o_i \in S_h \\ 0 & \text{if } o_i \notin S_h \end{cases}\)
for all \(i = 1, 2, \ldots, n\) and \(h = 1, 2, \ldots, k\). To generalize the binary relation, Zadeh\textsuperscript{8}
first extended \( m(o_i, S_h) \) to represent the strength of belonging for the object \( o_i \) in \( S_h \) as a fuzzy membership with \( 0 \leq m(o_i, S_h) \leq 1 \) and \( \sum_{h=1}^{k} m(o_i, S_h) = 1 \). Based on the idea of fuzzy memberships, there are several fuzzy extensions of the RI in the literature. In Section 2, we review these fuzzy extensions of the RI proposed by Anderson et al., Campello, Brouwer, Quere et al., and Hullermeier et al. In Section 3, we propose a new fuzzy generalization of the RI so that it can be suitable for finding not only similarities between fuzzy partitions and crisp reference partitions, but also similarities between fuzzy partitions and fuzzy reference partitions. In Section 4, we use several examples including synthetic data and real data sets to demonstrate these new concepts and also perform comparisons of our method with these fuzzy extensions of the RI. Finally, we make conclusions and our future studies in Section 5.

2. PREVIOUS METHODS

In this section, we give a brief review for these fuzzy extensions of the RI in the literature.

2.1. Rand Index

Rand\(^2\) proposed objective criteria for comparing the similarity between two clustering methods. These criteria have become an important tool that is consistently recommended in literature as the RI. The definition of RI is as follows.

Assume that we have the crisp partition matrix \( M_{C}^{(r)} \), \( r = 1, 2 \), between the object data set \( O \) and the partition \( P^{(r)} \). Let \( n_{uv} = \#(S_u^{(1)} \cap S_v^{(2)}) \), \( n_u = \sum_{v=1}^{k_2} n_{uv} \) and \( n_v = \sum_{u=1}^{k_1} n_{uv} \) for \( u = 1, 2, \cdots, k_1 \), \( v = 1, 2, \cdots, k_2 \). We get contingency table (see Hubert and Arabie\(^4\)) as

\[
N = [n_{uv}]_{k_1 \times k_2} = (M_{C}^{(1)}(M_{C}^{(2)})^T
\]

Thus, RI can be written as

\[
RI(M_{C}^{(1)}, M_{C}^{(2)}) = RI(P^{(1)}, P^{(2)}) = \frac{a + d}{a + b + c + d}
\]

where \( a = \#(o_i, o_j) \) belonging to the same cluster in \( P^{(1)} \) and to the same cluster in \( P^{(2)} \),
Figure 1. Two partitions of the object set $O = \{o_1, o_2, ..., o_6\}$.

\[
\forall i, j = 1, 2, \ldots, n, \text{ and } i \neq j \Rightarrow \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \left( \frac{n_{uv}}{2} \right)
\]

(3)

\[ b = \#(o_i, o_j)| (o_i, o_j) \text{ belonging to the same cluster in } P^{(1)} \text{ and to different clusters in } P^{(2)}, \]

\[
\forall i, j = 1, 2, \ldots, n, \text{ and } i \neq j = \frac{1}{2} \left[ \sum_{u=1}^{k_1} n_u^2 - \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} n_{uv} \right]
\]

(4)

\[ c = \#(o_i, o_j)| (o_i, o_j) \text{ belonging to different clusters in } P^{(1)} \text{ and to the same cluster in } P^{(2)}, \]

\[
\forall i, j = 1, 2, \ldots, n, \text{ and } i \neq j = \frac{1}{2} \left[ \sum_{v=1}^{k_2} n_v^2 - \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} n_{uv} \right]
\]

(5)

\[ d = \#(o_i, o_j)| (o_i, o_j) \text{ belonging to different clusters in } P^{(1)} \text{ and to different clusters in } P^{(2)}, \]

\[
\forall i, j = 1, 2, \ldots, n, \text{ and } i \neq j = \frac{1}{2} \left[ n^2 + \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} n_{uv}^2 - \left( \sum_{u=1}^{k_1} n_u^2 + \sum_{v=1}^{k_2} n_v^2 \right) \right]
\]

(6)

Obviously, $a + b + c + d = C_n^2 = n (n - 1)/2$.

**Example 1.** Assume that there is an object data set $O = \{o_1, o_2, \ldots, o_6\}$ with $P^{(1)} = \{S_1^{(1)}, S_2^{(1)}\}$ and $P^{(2)} = \{S_1^{(2)}, S_2^{(2)}, S_3^{(2)}, S_4^{(2)}\}$, where $S_1^{(1)} = \{o_1, o_2, o_4\}$, $S_2^{(1)} = \{o_3, o_5, o_6\}$, $S_1^{(2)} = \{o_1, o_2\}$, $S_2^{(2)} = \{o_3\}$, $S_3^{(2)} = \{o_4, o_5\}$, $S_4^{(2)} = \{o_6\}$ are shown in Figure 1.
Then we have the followings:

(i) Crisp partition matrices are

\[
M_C^{(1)} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}, \quad M_C^{(2)} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.
\]

(ii) The contingency table is

\[
N = [n_{ij}]_{2 \times 4} = (M_C^{(1)})(M_C^{(2)})^T = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}
\]

(iii) By Equations (3)–(6), we obtain that \( a = 1, b = 5, c = 1, \) and \( d = 8. \)

Thus, we have \( RI(M_C^{(1)}, M_C^{(2)}) = RI(P^{(1)}, P^{(2)}) = \frac{9}{15} = 0.6 \)

For fuzzy partitions, Anderson et al.,5 Campello,15 Brouwer,16 Hullermeier and Rifqi,17 Quere et al.,18 and Hullermeier et al.19 proposed different extensions according to various directions for evaluating \( RI \) to identify the similarity between two clusterings. We subsequently discuss these fuzzy extensions of \( RI \) as follows.

### 2.2. Campello’s Fuzzy \( RI \)

Campello15 used set-theoretic to formulate \( RI. \) Assume that \( M^{(r)} = [m_{hi}^{(r)}]_{k \times n}, \) \( r = 1, 2, \) is a fuzzy (or crisp) partition matrix between the set \( O \) and the partition \( P^{(r)}. \) Let \( G^{(r)} = [g_{ij}^{(r)}] \) where \( g_{ij}^{(r)} \) is the maximum of the minimum membership of \((o_i, o_j)\) belonging to the same cluster in \( P^{(r)}), \) and \( V^{(r)} = [v_{ij}^{(r)}] \) where \( v_{ij}^{(r)} \) be the maximum of the minimum membership of \((o_i, o_j)\) belonging to different clusters in \( P^{(r)}). \) That is,

\[
g_{ij}^{(r)} = \max \{ \min_{1 \leq h \leq k^r} \{ m_{hi}^{(r)} m_{hj}^{(r)} \} \} \quad \text{and} \quad v_{ij}^{(r)} = \max \{ \min_{1 \leq h, h' \leq k^r \text{ and } h \neq h'} \{ m_{hi}^{(r)} m_{h'j}^{(r)} \} \}.
\]

Then, we can obtain the followings:

\[
a = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \min(g_{ij}^{(1)}, g_{ij}^{(2)}) \quad \text{(7)}
\]

\[
b = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \min(g_{ij}^{(1)}, v_{ij}^{(2)}) \quad \text{(8)}
\]
\[ c = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \min \left( v_{ij}^{(1)}, g_{ij}^{(2)} \right) \]  (9)

\[ d = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \min \left( v_{ij}^{(1)}, v_{ij}^{(2)} \right) \]  (10)

Therefore, Campello\textsuperscript{15} defined the fuzzy RI (FRI) for the two fuzzy (or crisp) partition matrices \( M^{(1)} \) and \( M^{(2)} \) as \( FRC(M^{(1)}, M^{(2)}) = FRC(P^{(1)}, P^{(2)}) = \frac{a+d}{a+b+c+d} \).

Although Campello’s method can be used to find the FRI, they highlighted that \( M^{(1)} = M^{(2)} \) is no longer a necessary and sufficient condition for \( FRC(M^{(1)}, M^{(2)}) = 1 \). It is necessary, but sufficiency is only guaranteed by ensuring that the reference partition (for example, \( M^{(2)} \)) is crisp. In this sense, FRC cannot be seen as a general measure for comparing two fuzzy partitions, and so one of our research directions in this paper is to resolve this disadvantage of FRC.

### 2.3. Brouwer’s FRI

Brouwer\textsuperscript{16} used a vector to deal with the memberships of objects \( o_i \) belonging to each cluster. First, Brouwer\textsuperscript{16} defined bonding matrices in virtue of crisp partition matrices. Next, Brouwer\textsuperscript{16} used the inner product of two vectors to calculate cosine correlation, and to find RI according to the cosine correlation. Finally, Brouwer\textsuperscript{16} implied these results to fuzzy partitions. We briefly describe it as follows.

Assume that \( M^{(r)}_C, r = 1, 2, \) is the crisp partition matrix between the object data set \( O \) and the partition \( P^{(r)} \). Define \( B^{(r)} = [b_{ij}^{(r)}] = (M^{(r)}_C)^T (M^{(r)}_C). \) Then, \( B^{(r)} \) is called the bonding matrix of the partition \( P^{(r)} \). Let \( f(Z) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} z_{ij} \) and \( g(Z) = f(Z) - \frac{n}{2} \). Then, we have

\[ a = g\left( B^{(1)} \wedge (B^{(2)})^T \right) \]  (11)

\[ b = f\left( B^{(1)} \wedge (\neg B^{(2)})^T \right) \]  (12)

\[ c = f\left( (\neg B^{(1)}) \wedge (B^{(2)})^T \right) \]  (13)

\[ d = f\left( (\neg B^{(1)}) \wedge (\neg B^{(2)})^T \right) \]  (14)

where \( \neg B^{(r)} = J - B^{(r)}, r = 1, 2, \) and \( J \) is a \( n \times n \) ones-matrix. In the sequel, they generalized the above notion to fuzzy partitions.
Let the matrix \( U^{(r)} = [u_{hi}^{(r)}]_{k \times n} = [\frac{m_{hi}^{(r)}}{\|m_{hi}^{(r)}\|}] \), where \( \|m_{hi}^{(r)}\| \) is the length of \( i^{th} \) column vector of a fuzzy partition matrix \( M_{F}^{(r)} \), \( i = 1, 2, \ldots, n, r = 1, 2 \). Then, the bonding matrix \( B^{(r)} = (U^{(r)})^T(U^{(r)}) \) and the values of \( a, b, c, d \) can be obtained from Equation (11) to Equation (14). In fact, the entry \( b_{ij} \) of \( B^{(r)} \) is called the cosine correlation that is often used to measure the similarity between two vectors.\(^{16} \) For example, correlation-based similarities have been widely used in the micro-array literature.\(^{20,21} \) Nonetheless, \( x_{ij} \) denotes the multiplication of two cosine correlations in FRB.

### 2.4. Quere et al.’s FRI

For reinforcing the results of Borgelt and Kruse\(^{21} \) and Campello,\(^{15} \) Quere et al.\(^{18} \) used the cosine-correlation idea of Brouwer\(^{16} \) to generalize coincidence matrices by means of \( t \)-norms and normalization. Assume that \( M^{(r)} = [m_{hi}^{(r)}]_{k \times n}, r = 1, 2 \), is a fuzzy (or crisp) partition matrix between the set \( O \) and the partition \( P^{(r)} \). Borgelt and Kruse\(^{21} \) extended indices for similarities between clusterings by defining a coincidence matrix \( \phi_{M^{(r)}}^{T} \) of \( M^{(r)} \) based on a \( t \)-norm (triangular norm) \( T \) as \( \phi_{M^{(r)}}^{T} = [\phi_{M^{(r)},ij}^{T}]_{k \times n} \) with \( \phi_{M^{(r)},ij}^{T} = \sum_{h=1}^{k} T(m_{hi}^{(r)}, m_{hj}^{(r)}) \). However, Quere et al.\(^{18} \) pointed out that this extension may produce undesirable results. This is because, whatever the \( t \)-norm, diagonal terms \( \phi_{M^{(r)},ij}^{T} \), representing the degree with which each \( x_{ij} \) is as in the same class as itself, are no more equal to 1 so that the normalization is necessary. Quere et al.\(^{18} \) used the cosine-correlation idea of Brouwer\(^{16} \) to normalize coincidence matrices as follows. The normalized coincidence matrix is written as \( \Phi_{M^{(r)}}^{T} = [\Phi_{M^{(r)},ij}^{T}]_{n \times n} \) with \( \Phi_{M^{(r)},ij}^{T} = \frac{x}{T(K_{T}(\phi_{M^{(r)},ij}^{T}), K_{T}(\phi_{M^{(r)},ij}^{T})))} \), where \( K_{T}(x):[0, 1] \rightarrow [0, 1] \) is a normalizing function such that \( \frac{x}{T(K_{T}(x), K_{T}(x)} = 1 \). Let \( V^{(r)} = [v_{ij}^{(r)}]_{n \times n} = [1 - \Phi_{M^{(r)},ij}^{T}] \). Then,

\[
\begin{align*}
    a &= \sum_{i=2}^{n} \sum_{j=1}^{i-1} T(\Phi_{M^{(1)},ij}^{T}, \Phi_{M^{(2)},ij}^{T}) \\
    b &= \sum_{i=2}^{n} \sum_{j=1}^{i-1} T(\Phi_{M^{(1)},ij}^{T}, v_{ij}^{(2)}) \\
    c &= \sum_{i=2}^{n} \sum_{j=1}^{i-1} T(v_{ij}^{(1)}, \Phi_{M^{(2)},ij}^{T}) \\
    d &= \sum_{i=2}^{n} \sum_{j=1}^{i-1} T(v_{ij}^{(1)}, v_{ij}^{(2)})
\end{align*}
\]
As shown by Quere et al., they actually improved the undesirable results of coincidence matrices proposed by Borgelt and Kruse. Of course, the Quere et al. method has all diagonal terms $\Phi^T_{M^{(1)}M^{(2)}}$ equal to 1 for any given $t$-norm.

### 2.5. Anderson et al.’s FRI

To successfully generalize $RI$ from crisp partitions to fuzzy partitions, Anderson et al. recommended replacing the term $N$ of Equation (1) in $RI$ with $N^\ast$, as shown below. Let $M^{(r)} = [m^{(r)}_{hi}]_{k \times n}$ be any crisp or fuzzy membership matrix between the object data set $O$ and the partitions $P^{(r)}$. Then, we have $N = (M^{(1)})(M^{(2)})^T$ and $N^\ast = \phi N = [n^\ast_{uv}]_{k_1 \times k_2}$, where

\[
\phi = \begin{cases} 
1, & \text{if } M^{(1)}, M^{(2)} \text{ are crisp, fuzzy or probabilistic} \\
\frac{n}{\sum u n_u}, & \text{if } M^{(1)} \text{ and/or } M^{(2)} \text{ are possibilistic}
\end{cases}
\]

The FRA can be evaluated from Equation 2 and (3) to Equation (6). However, through the FRA calculation, one can determine that $a = \frac{1}{2} \sum_{h=1}^{k_1} \sum_{h=1}^{k_2} (n^\ast_{uv} - 1)$. If $0 < n^\ast_{uv} < 1$, then an illogical result of $(n^\ast_{uv})^2 > 0$ may occur. We may circumvent this disadvantage in our proposed method.

### 2.6. Hullermeier et al.’s FRI

Hullermeier et al. adopted a metric perspective and considered $RI$ as a distance function. Their approach is as follows. Let $M^{(r)} = [m^{(r)}_{hi}]_{k \times n}$, $r = 1, 2$, be any fuzzy (or crisp) partition matrix between the object data set $O$ and the partition $P^{(r)}$. Let

\[
E_{M^{(r)}}(o_i, o_j) = 1 - \frac{1}{2} \sum_{h=1}^{k_r} |m^{(r)}_{hi} - m^{(r)}_{hj}| = 1 - \|m(o_i) - m(o_j)\| \quad (19)
\]

\[
d(M^{(1)}, M^{(2)}) = \frac{1}{C_n^2} \sum_{o_i, o_j \in O} |E_{M^{(1)}}(o_i, o_j) - E_{M^{(2)}}(o_i, o_j)| \quad (20)
\]

Then, Hullermeier et al. proposed that the FRI is equal to $1 - d(M^{(1)}, M^{(2)})$. In response that similarity is not always the sum of $a + d$ advocated by other researchers, Hullermeier et al. divided concordance into two parts, $a$-concordance and $d$-concordance. The product $t$-norm was then used to define the following:
\[ a = \sum_{o_i, o_j \in O} \left[ (1 - |E_{M^{(1)}}(o_i, o_j) - E_{M^{(2)}}(o_i, o_j)|) \cdot E_{M^{(1)}}(o_i, o_j) \cdot E_{M^{(2)}}(o_i, o_j) \right] \]

\[ d = \sum_{o_i, o_j \in O} \left[ (1 - |E_{M^{(1)}}(o_i, o_j) - E_{M^{(2)}}(o_i, o_j)|) \right] \cdot (1 - E_{M^{(1)}}(o_i, o_j) \cdot E_{M^{(2)}}(o_i, o_j)) \]

They also divided discordance into \( b \)-discordance and \( c \)-discordance; therefore,

\[ b = \sum_{o_i, o_j \in O} \left[ \max\left( E_{M^{(1)}}(o_i, o_j) - E_{M^{(2)}}(o_i, o_j), 0 \right) \right] \]

\[ c = \sum_{o_i, o_j \in O} \left[ \max\left( E_{M^{(2)}}(o_i, o_j) - E_{M^{(1)}}(o_i, o_j), 0 \right) \right] \]

Hullermeier et al.\(^{19}\) clearly showed that the distance function (Equation (20)) between fuzzy partitions is a pseudo-metric. However, if \( M^{(1)} \) and \( M^{(2)} \) are restricted to normalized fuzzy partitions such as the following assumptions of Equations (21) and (22), then the distance function (Equation (20)) becomes a metric. In other words, if \( M^{(r)}, r = 1, 2 \), satisfies

\[(i) \quad \forall o_i \in O, \sum_{h=1}^{k_r} m_{hi}^{(r)} = 1 \quad (21)\]

\[(ii) \quad \forall S_h^{(r)} \in P^{(r)}, \exists o_i \in O \text{ such that } m_{hi}^{(r)} = 1, \quad (22)\]

then the distance function (Equation (20)) is a metric. Clearly, the requirements of Equation (21) can be satisfied by dividing the object data set with the fuzzy \( c \)-means (FCM) clustering algorithm. However, Equation (22) cannot be easily satisfied. Therefore, we aimed to overcome this limitation that required a fuzzy partition to satisfy Equation (22).

### 3. THE PROPOSED METHOD

In this section, we first construct \( RI \) via a graph and then define a color relation matrix. We next prove that \( RI \) can be obtained by the trace of the matrix. Finally, we generalize the color relation matrix from crisp partitions to fuzzy partitions.
3.1. RI via Graph

In recent years, the theory of multiple intelligence proposed by Gardner,22 as well as the graphical memory method in learning, adopts the clear overview of graph for learning. Graph theory not only provides deep impressions, but it is also simple and clear where it has wide applications, such as Li et al.23 Thus, we consider a new approach for calculating $RI$ through graph.

**Definition 1.** (Color edge) If the object pair $(o_i, o_j)$ in an object data set $O$ belongs to the same cluster in a partition $P = \{S_1, S_2, \ldots, S_k\}$, we connect $o_i$ and $o_j$ by a red edge. If the object pair $(o_i, o_j)$ in the set $O$ belongs to different clusters in the partition $P$, we connect $o_i$ and $o_j$ by a blue edge.

Based on the color edge definition, we define a color relation matrix as follows.

**Definition 2.** (Color relation matrix) Assume that there is a partition $P = \{S_1, S_2, \ldots, S_k\}$ in an object data set $O = \{o_1, o_2, \ldots, o_n\}$. The matrix $R = [r_{ij}]_{n \times n}$ is called a color relation matrix where

$$r_{ij} = \begin{cases} 
1 & \text{if the pair } (o_i, o_j) \text{ is connected by red edge} \\
-1 & \text{if the pair } (o_i, o_j) \text{ is connected by blue edge} \\
0 & \text{if } i = j.
\end{cases}$$

We define $r_{ij} = 0$ for all $i$ because any pair $(o_i, o_i)$ cannot draw a simple edge. This way may improve the trouble in Brouwer’s method15 which needs to minus $n/2$ when we use $g(B^{(1)} \wedge (B^{(2)})^T)$ to find $a$. Thus, we have the following theorems that $RI$ can be calculated via the trace of a color relation matrix.

**Theorem 1.** $R = R^T$.

*Proof.* It is clear that $(o_i, o_j)$ and $(o_j, o_i)$ have an identical color edge. \hfill $\Box$

**Example 2.** Those color edges of Example 1 can be referred as Figure 2.

**Figure 2.** Color edges of Example 1.
The color relation matrices \( R^{(1)} \) and \( R^{(2)} \) of Example 1 are symmetry as follows:

\[
R^{(1)} = \begin{bmatrix}
0 & 1 & -1 & 1 & -1 & 1 \\
1 & 0 & -1 & 1 & -1 & -1 \\
-1 & -1 & 0 & -1 & 1 & 1 \\
1 & 1 & -1 & 0 & -1 & -1 \\
-1 & -1 & 1 & -1 & 0 & 1 \\
-1 & -1 & 1 & -1 & 1 & 0
\end{bmatrix}
\]

\[
R^{(2)} = \begin{bmatrix}
0 & 1 & -1 & -1 & -1 & -1 \\
1 & 0 & -1 & -1 & -1 & -1 \\
-1 & -1 & 0 & -1 & -1 & -1 \\
-1 & -1 & -1 & 0 & -1 & -1 \\
-1 & -1 & -1 & 1 & -1 & -1 \\
-1 & -1 & -1 & -1 & -1 & 0
\end{bmatrix}
\]

**Theorem 2.** Let \( R^{(r)} = [r_{ij}^{(r)}]_{n \times n}, r = 1, 2, \) be the color relation matrix of the partition \( P^{(r)}. \) Let \( Y = [y_{ij}] = (R^{(1)})(R^{(2)})^T. \) Then, \( RI(M_C^{(1)}, M_C^{(2)}) = RI(P^{(1)}, P^{(2)}) = \frac{a+d}{a+b+c+d} = \frac{1}{2} + \frac{1}{4} tr(Y) \), where \( tr(Y) \) is the trace of the matrix \( Y. \)

**Proof.** Since \( C_2^a = a + b + c + d \) and \( tr(Y) = 2[(a + d) - (b + c)], \) \( a + d = \frac{1}{2} C_2^a + \frac{1}{4} tr(Y). \) Thus, \( RI(M_C^{(1)}, M_C^{(2)}) = RI(P^{(1)}, P^{(2)}) = \frac{a+d}{a+b+c+d} = \frac{1}{2} + \frac{1}{4} tr(Y) \).

According to Theorem 2, \( RI \) can be determined using only diagonal entries \( Y = (R^{(1)})(R^{(2)})^T. \) It is not necessary to calculate all entries \( y_{ij} \) of the matrix \( Y. \) In fact, let \( T R^{(r)} \) be the upper triangle matrix of \( R^{(r)}, r = 1, 2, \) and \( Y' = (T R^{(1)})(T R^{(2)})^T. \) Then \( tr(Y') = (a + d) - (b + c). \) Moreover, we can also separately calculate all terms of \( a, b, c, d \) according to the following proposition so that other related indices can be provided via the trace of a color relation matrix.

**Theorem 3.** Let \( R^{(r)} = [r_{ij}^{(r)}] \) be a color relation matrix.

(i) If \( r_{ij}^{(r)} = 1 \) if the pair \((o_i, o_j)\) is red edge in \( P^{(r)}, r = 1, 2, \) then \( a = \frac{1}{2} tr(Y) = \frac{1}{2} tr(Y); \)

(ii) If \( r_{ij}^{(1)} = 1 \) if the pair \((o_i, o_j)\) is red edge in \( P^{(1)} \) and \( r_{ij}^{(2)} = \frac{1}{2} tr(Y) = \frac{1}{2} tr(Y); \)

(iii) If \( r_{ij}^{(1)} = \frac{1}{2} tr(Y) = \frac{1}{2} tr(Y); \)

(iv) If \( r_{ij}^{(1)} = \frac{1}{2} tr(Y) = \frac{1}{2} tr(Y); \)

---

*International Journal of Intelligent Systems*  DOI 10.1002/int
Example 3. Refer to Example 1, we can find $a$, $b$, $c$, $d$ as follows.

\[
\begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0
\end{bmatrix}
\]

(i) If $Y = \begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0
\end{bmatrix}$, then the diagonal of $Y$ consists of the elements $1,1,0,0,0,0$. This implies $a = \frac{1}{2} |tr(Y)| = 1$.

(ii) If $Y = \begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0
\end{bmatrix}$, then the diagonal of $Y$ consists of the elements $-1,-1,-2,-2,-2,-2$. This implies $b = \frac{1}{2} |tr(Y)| = 5$.

(iii) If $Y = \begin{bmatrix}
0 & 0 & -1 & 0 & -1 & -1 \\
-1 & -1 & 0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & -1 & -1 \\
-1 & -1 & 0 & -1 & 0 & 0 \\
-1 & -1 & 0 & -1 & 0 & 0
\end{bmatrix}$, then the diagonal of $Y$ consists of the elements $0,0,0,-1,-1,0$. This implies $c = \frac{1}{2} |tr(Y)| = 1$.

(iv) If $Y = \begin{bmatrix}
0 & 0 & -1 & 0 & -1 & -1 \\
-1 & -1 & 0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & -1 & -1 \\
-1 & -1 & 0 & -1 & 0 & 0 \\
-1 & -1 & 0 & -1 & 0 & 0
\end{bmatrix}$, then the diagonal of $Y$ consists of the elements $3,3,2,2,3$. This implies $d = \frac{1}{2} |tr(Y)| = 8$.

In order to generalize a color relation matrix from crisp partitions to fuzzy partitions, we first propose the concept of a duplication factor $t$ as follows. For evaluating similarity measures between clusterings, $RI$ used the ratio to define similarity indices. Of course, the ratio is not influenced when an object is divided into $t$ equal parts. This is similar as that the ratio is not influenced when the unit is scaled. In this sense, we may use the duplication factor $t$ to represent that an object is divided into $t$ equal parts.

**Definition 3.** (Extended color relation matrix for crisp partition) Let $M_C = [m_{hi}]_{k \times n}$ be a crisp partition matrix between the object data set $O$ and the partition $P = \{S_1, S_2, \ldots, S_k\}$. For $i = 1, \ldots, n$, let each $o_i$ be divided into $t$ equal elements $o_{i1}, o_{i2}, \ldots, o_{it}$, where $t$ is a positive integer. Then

(i) The set $O = \{o_1, o_2, \ldots, o_n\}$ is extended to $O_E = \{o_{i1}, o_{i2}, \ldots, o_{it}, o_{i2}, \ldots, o_{it}\}$.
Example 4. Refer to Example 1, let the duplication factor \( t = 2 \). According to Definition 3, we have the following results.

(i) The set \( O = \{ o_1, \ldots, o_n \} \) is extended to \( \mathcal{O}_E = \{ o_1, o_{12}, o_{21}, o_{22}, o_3, o_4, o_5, o_6, o_7, o_8, o_9 \} \).
(ii) The extended partition matrices at the duplication factor \( t = 2 \) of \( M_C \) become

\[
\begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1
\end{bmatrix}
\]

\[\begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix}.
\]

(iii) The extended color relation matrices at the duplication factor \( t = 2 \) become

\[
\begin{bmatrix}
0 & 0 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
0 & 0 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & 0 & 0 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & 0 & 0 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
-1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1
\end{bmatrix}
\]

\[\begin{bmatrix}
0 & 0 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
0 & 0 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & 0 & 0 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & 1 & 0 & 0 & -1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\
-1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 0 & 0 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 & 1
\end{bmatrix}.
\]

For a simple expression, we can rewrite them as follows:

\[
\begin{bmatrix}
O_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & -J_{2 \times 2}
\end{bmatrix}
\]

\[\begin{bmatrix}
O_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & -J_{2 \times 2}
\end{bmatrix}
\]

\[
\begin{bmatrix}
O_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & -J_{2 \times 2}
\end{bmatrix}
\]

\[
\begin{bmatrix}
O_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & J_{2 \times 2} & -J_{2 \times 2} & -J_{2 \times 2}
\end{bmatrix}
\]
where every entry of $O_{2 \times 2}$ is 0.

(iv) $tr(2Y) = tr(Y) \times 2^2 = 24$ (i.e., $tr(Y) = tr(Y) \times t^2$).

(v) $RI(3M_{C}^{(1)}, 3M_{C}^{(2)}) = RI(M_{C}^{(1)}, M_{C}^{(2)}) = 0.6$.

(vi) Assume $t = 3$, we can also obtain $RI(3M_{C}^{(1)}, 3M_{C}^{(2)}) = 0.6$ according to the above method. Similarly, for $t = 4$, we have $RI(4M_{C}^{(1)}, 4M_{C}^{(2)}) = 0.6$. In other words, $RI(M_{C}^{(1)}, M_{C}^{(2)}) = RI(M_{C}^{(1)}, M_{C}^{(2)})$ for any positive integer $t$.

3.2. Color Relation Matrix Generalization from Crisp Partitions to Fuzzy Partitions

From Theorem 4, we find that $RI$ is invariant via any duplication factor $t$. Based on this invariant property, we next extend our defined color relation matrix from crisp partitions to fuzzy partitions.

**Definition 4.** (Extended color relation matrix for fuzzy partitions) Let $M_{F} = [m_{hi}]_{1 \times n}$ be a fuzzy partition matrix between the object data set $O$ and the partition $P = \{S_{1}, S_{2}, \ldots, S_{k}\}$. For $i = 1, 2, \ldots, n$, let each $o_{i}$ be divided into $t$ equal elements $o_{i1}, o_{i2}, \ldots, o_{it}$, where $t$ is a fixed number such that all $t \times m_{hi}$, $r = 1, 2$, become integers. Then

(i) The set $O = \{o_{1}, o_{2}, \ldots, o_{n}\}$ is extended to $O_{E} = \{o_{i1}, o_{i2}, \ldots, o_{it}, o_{i1}, o_{i2}, \ldots, o_{it}, \ldots, o_{nt}, \ldots, o_{nt}\}$.

(ii) There are $t \times m_{hi}$ elements of $o_{hi}$ belonging to the $h$th cluster $S_{h}$ of $P$.

(iii) The extended color relation matrix at the duplication factor $t$ for $M_{F}$ (or simply called as the extended color relation matrix, i.e., $R = [r_{u,v}]_{n \times nt}$) is defined according to the form of Definition 3. We can see that, for all $u, v$, $r_{u,v} = 0$ if $i = j$. This is because any pair $(o_{i}, o_{j})$ in the object data set $O$ cannot draw a simple edge (i.e., the same object cannot draw a simple edge).

This extension still reserves the original contents of $RI$ whenever the pair of $o_{ij}, o_{ji}$ belongs to the same cluster $S_{h}^{(r)}$ of the partition $P^{(r)}$, where $r = 1, 2$, $i, j = 1, 2, \ldots, n$, $v, w = 1, 2, \ldots, t$, $h = 1, 2, \ldots, k_{r}$. That is, we retain the relationship between $o_{ij}$ and $o_{ji}$ only with the two cases, 1 for the pair of $o_{ij}$, $o_{ji}$ belonging to the same cluster $S_{h}^{(r)}$ or $-1$ for the pair of $o_{ij}$, $o_{ji}$ not belonging to the same cluster $S_{h}^{(r)}$ when $o_{ij}$ and $o_{ji}$ are the two different objects, where $i \neq j$.

Note that the choice of a duplication factor $t$ in Definitions 3 and 4 is different. The duplication factor $t$ in Definition 3 can be any positive integer. However, the duplication factor $t$ in Definition 4 should be a specified number such that all $t \times m_{hi}$, $r = 1, 2$, become integers. That is, we choose a positive integer $t$ to make all $t \times m_{hi}$ be integers.

**Theorem 5.** Assume that $M_{F}^{(r)}$, $r = 1, 2$, is a fuzzy partition matrix between the object data set $O$ and the partition $P^{(r)}$. Let $t$ be a specified number such that all $t \times m_{hi}$, $r = 1, 2$, become integers. Let $R^{(r)} = [r_{ij}^{(r)}]_{nt \times nt}$, $r = 1, 2$, be the extended color relation matrices at the duplication factor $t$ for the fuzzy partition matrix $M_{F}^{(r)}$ based on Definition 4. Let $Y = (R^{(1)})^{T}(R^{(2)})^{T}$. Then, the FRI for fuzzy partition...
matrices $M_F^{(1)}$ and $M_F^{(2)}$ can be obtained by

$$FRI(M_F^{(1)}, M_F^{(2)}) = FRI(P^{(1)}, P^{(2)}) = \frac{1}{2} + \frac{1}{t^2} tr(\langle Y \rangle)$$

where $tr(\langle Y \rangle)$ is the trace of the matrix $Y$.

**Proof.** It is similar to the proof of Theorems 2 and 4. □

Note that, similar to Theorem 3, we can get all terms of $a, b, c, d$ for any two fuzzy partition matrices $M_F^{(1)}$ and $M_F^{(2)}$. This means that we can make fuzzy extension of other related indices, such as the Jaccard index, for any two fuzzy partition matrices $M_F^{(1)}$ and $M_F^{(2)}$.

**Example 5.** Assume that $M_F^{(1)}$ and $M_F^{(2)}$ are fuzzy partition matrices between the object data set $O$ and the partition $P^{(r)}$, defined as follows:

$$M_F^{(1)} = S_1^{(1)} \begin{bmatrix} 0.9 & 0.2 & 0.1 & 0.2 & 0.3 & 0.7 \\ 0.1 & 0.8 & 0.9 & 0.8 & 0.7 & 0.3 \end{bmatrix},$$

$$M_F^{(2)} = S_2^{(2)} \begin{bmatrix} 0.5 & 0.2 & 0.1 & 0.2 & 0.2 & 0.5 \\ 0.1 & 0.5 & 0.3 & 0.4 & 0.4 & 0.3 \\ 0.4 & 0.3 & 0.6 & 0.4 & 0.4 & 0.2 \end{bmatrix}.$$

It is trivial that we choose $t = 10$. This is because all $t \times m^{(r)}_{ij}$ become integers for all $i$ and $j$. Then, $R^{(r)}$, $r = 1, 2$, are $60 \times 60$ square matrices. Next, we partition $R^{(r)}$ into submatrices as follows:

$$R^{(r)} = \begin{bmatrix} R_{11}^{(r)} & R_{12}^{(r)} & \ldots & R_{16}^{(r)} \\ R_{21}^{(r)} & R_{22}^{(r)} & \ldots & R_{26}^{(r)} \\ \vdots & \vdots & \ddots & \vdots \\ R_{61}^{(r)} & \ldots & \ldots & R_{66}^{(r)} \end{bmatrix}$$

where each $R_{ij}^{(r)}$ is a $10 \times 10$ matrix for $r = 1, 2$, and $i, j = 1, 2, \ldots, 6$. Without loss of generality, we may assume $o_{11}, o_{12}, \ldots, o_{19} \in S_1^{(1)}$, $o_{110} \in S_2^{(1)}$, and $o_{21}, o_{22}, \ldots, o_{29} \in S_1^{(1)}$, $o_{210} \in S_2^{(1)}$, and so on. Let every entry of $J_{w \times t}$ (where $J_{w \times t}$ is $w \times t$ matrix) is 1 and every entry of $O_{w \times t}$ (where $O_{w \times t}$ is $w \times t$ matrix) is 0. Then, $R_{ij}^{(r)} = O_{10 \times 10}$, where $i = 1, 2, \ldots, 6, r=1,2$, $R_{12}^{(1)} = (R_{21}^{(1)})^T = [-J_{s \times 2} - J_{s \times 8}]$, and so forth.

*International Journal of Intelligent Systems*  DOI 10.1002/int
Using Theorem 5, we can obtain the FRI as 

$$FRI(M_F^{(1)}, M_F^{(2)}) = FRI(P^{(1)}, P^{(2)}) = 0.6873.$$ 

We next analyze some special points of our method as follows:

(i) The FRI of our method is between 0 and 1.

(ii) Because the selection of \( t \) renders \( t \times m_{hi}^{(r)} \) integers for all \( h, i \) and all entries of the extended color relation matrix belong to \{0, 1, −1\}, our method will not cause information missing for data and also provides the following result: \( M^{(1)} = M^{(2)} \) if and only if \( FRI(M^{(1)}, M^{(2)}) = 1 \) where the original data \( M^{(1)} \) and \( M^{(2)} \) may be a crisp or fuzzy partition matrix. This result is extremely important for a similarity.

(iii) If \( m_{hi}^{(r)} \) has multiple digits after the decimal point, we can round off \( m_{hi}^{(r)} \) to the nearest tenth or hundredth. In fact, a fuzzy clustering, such as FCM, is generally used to find \( m_{hi}^{(r)} \) with one or two digits after the decimal point. In this sense, setting \( t = 10 \) or \( t = 100 \) will be enough. If higher precision is needed, then \( m_{hi}^{(r)} \) may appear three digits (or more than three digits) after the decimal point. In this case, the extended color relation matrix will become larger and need more computational time. This is a limitation for our method. Besides, for all \( i \), \( \sum_{h=1}^{c} m_{hi}^{(r)} \) is restricted to 1 because the number of two data sets must be equal.

(iv) Assume that \( M_F^{(r)}, r = 1, 2 \), is a fuzzy partition matrix between the object data set \( O \) and the partition \( P^{(r)} \). Let \( t \) be a fixed number such that all \( t \times m_{hi}^{(r)} \), \( r = 1, 2 \), become integers. Then, the extended partition matrix at the duplication factor \( t \) for \( M_F^{(r)} \) is \( tM_F^{(r)} \). We choose another duplication factor \( s \), where \( s = kt \) (\( k \) is a positive integer). Equivalent to Theorem 4, \( FRI(tM_F^{(1)}, tM_F^{(2)}) = FRI(M_F^{(1)}, M_F^{(2)}) \).

4. NUMERICAL EXAMPLES AND COMPARISONS

In this section, we use some synthetic data and real data sets to make comparisons of our method with these fuzzy extensions of RI proposed by various researchers. We mention that Quere et al.\(^{18}\) proposed a generic solution to perform normalization of a fuzzy coincidence matrix, where Hamacher norm with parametric \( \gamma = 5 \) was chosen as a \( t \)-norm, that is,

$$T_{H_{\gamma}}(x_1, x_2) = \frac{x_1 x_2}{\gamma + (1 - \gamma)(x_1 + x_2 - x_1 x_2)}$$

and

$$K_{T}(x) = \frac{\gamma \sqrt{x}}{\sqrt{\gamma + (1 - \gamma)x} + (\gamma - 1)\sqrt{x}}.$$ 

These compared methods are denoted with abbreviations as shown in Table I to facilitate explanations.

In the following Example 6, we generate a data set from a mixture distribution

$$f(x; \alpha, \theta) = \sum_{k=1}^{c} \alpha_k f(x; \theta_k)$$

with \( f(x; \theta_k) \) being a bivariate Gaussian distribution \( N(\mu_k, \Sigma_k) \), where \( \mu_k \) and \( \Sigma_k \) indicate the mean and the covariance matrix of the \( k \)th component. We then implement the FCM\(^{10,11}\) on the data sets to get fuzzy partition matrices. Note that we use FCM as a clustering method because it is similar to
the examples in Anderson et al., who indicated that $RI$ and $FRI$ for the hard \(c\)-means and FCM behave in a similarly desirable fashion, but the possibilistic \(c\)-means (PCM) fails to reach a maximum at \(k = 4\) ($k$ indicates the true number of clusters). Therefore, we do not use PCM to divide the data sets. Finally, we use the obtained fuzzy partition matrix and the crisp reference partition to compare and analyze various methods.

**Example 6.** In this example, we use two data sets generated from Gaussian mixture distributions. The first data set, as shown in Figure 3a, is generated from a two-dimensional four-component Gaussian mixture distribution with the sample size \(n = 100\) and the parameters \(\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/4\); \(\mu_1 = \begin{pmatrix} 100 \\ 100 \end{pmatrix}, \mu_2 = \begin{pmatrix} 100 \\ 200 \end{pmatrix}, \mu_3 = \begin{pmatrix} 200 \\ 100 \end{pmatrix}, \mu_4 = \begin{pmatrix} 200 \\ 200 \end{pmatrix}\) and \(\Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma_4 = \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix}\). The second data set with the sample size \(n = 100\), as shown in Figure 3b, is generated from the same Gaussian mixture distribution except with the parameters \(\Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma_4 = \begin{pmatrix} 400 & 0 \\ 0 & 400 \end{pmatrix}\). Obviously, the crisp reference partition for the two data sets is with \(o_1, \ldots, o_{25} \in S_1, o_{26}, \ldots, o_{50} \in S_2, o_{51}, \ldots, o_{75} \in S_3, \text{ and } o_{76}, \ldots, o_{100} \in S_4\). Let \([m_{hi}]\) be written

<table>
<thead>
<tr>
<th>Abbreviations of different fuzzy extensions of $RI$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indices</strong></td>
</tr>
<tr>
<td><strong>Methods</strong></td>
</tr>
<tr>
<td>Fuzzy $RI$</td>
</tr>
<tr>
<td>Campello\textsuperscript{15}</td>
</tr>
<tr>
<td>Brouwer\textsuperscript{16}</td>
</tr>
<tr>
<td>Quere et al.\textsuperscript{18}</td>
</tr>
<tr>
<td>Anderson et al.\textsuperscript{5}</td>
</tr>
<tr>
<td>Hullermeier et al.\textsuperscript{19}</td>
</tr>
<tr>
<td>Our approach</td>
</tr>
</tbody>
</table>
Table II. Fuzzy RI between the data set 1 and the crisp reference partition $D_4$

<table>
<thead>
<tr>
<th>Methods</th>
<th>$FRC$</th>
<th>$FRB$</th>
<th>$FRQ$</th>
<th>$FRA$</th>
<th>$FRH$</th>
<th>$FRI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 2$</td>
<td>0.6455</td>
<td>0.5408</td>
<td>0.5882</td>
<td>0.6024</td>
<td>0.5743</td>
<td>0.6037</td>
</tr>
<tr>
<td>$k = 3$</td>
<td>0.7875</td>
<td>0.7525</td>
<td>0.7481</td>
<td>0.7850</td>
<td>0.7928</td>
<td>0.8063</td>
</tr>
<tr>
<td>$k = 4$</td>
<td>0.9675</td>
<td>0.9684</td>
<td>0.9730</td>
<td>0.9480</td>
<td>0.9475</td>
<td>0.9815</td>
</tr>
<tr>
<td>$k = 5$</td>
<td>0.9362</td>
<td>0.9508</td>
<td>0.9565</td>
<td>0.9198</td>
<td>0.9200</td>
<td>0.9524</td>
</tr>
<tr>
<td>$k = 6$</td>
<td>0.9049</td>
<td>0.9412</td>
<td>0.9453</td>
<td>0.8979</td>
<td>0.9055</td>
<td>0.9305</td>
</tr>
<tr>
<td>$k = 7$</td>
<td>0.8705</td>
<td>0.9249</td>
<td>0.9274</td>
<td>0.8750</td>
<td>0.8861</td>
<td>0.9020</td>
</tr>
<tr>
<td>$k = 8$</td>
<td>0.8314</td>
<td>0.9096</td>
<td>0.9103</td>
<td>0.8511</td>
<td>0.8690</td>
<td>0.8736</td>
</tr>
</tbody>
</table>

Table III. Fuzzy RI between the second data set and the crisp reference partition $D_4$

<table>
<thead>
<tr>
<th>Methods</th>
<th>$FRC$</th>
<th>$FRB$</th>
<th>$FRQ$</th>
<th>$FRA$</th>
<th>$FRH$</th>
<th>$FRI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 2$</td>
<td>0.6145</td>
<td>0.4966</td>
<td>0.5118</td>
<td>0.5897</td>
<td>0.5656</td>
<td>0.5939</td>
</tr>
<tr>
<td>$k = 3$</td>
<td>0.7534</td>
<td>0.7096</td>
<td>0.7249</td>
<td>0.7479</td>
<td>0.7305</td>
<td>0.7640</td>
</tr>
<tr>
<td>$k = 4$</td>
<td>0.9055</td>
<td>0.8951</td>
<td>0.9159</td>
<td>0.8699</td>
<td>0.8601</td>
<td>0.8969</td>
</tr>
<tr>
<td>$k = 5$</td>
<td>0.8761</td>
<td>0.8787</td>
<td>0.9001</td>
<td>0.8487</td>
<td>0.8290</td>
<td>0.8752</td>
</tr>
<tr>
<td>$k = 6$</td>
<td>0.8620</td>
<td>0.8723</td>
<td>0.8901</td>
<td>0.8403</td>
<td>0.8206</td>
<td>0.8716</td>
</tr>
<tr>
<td>$k = 7$</td>
<td>0.8378</td>
<td>0.8663</td>
<td>0.8822</td>
<td>0.8271</td>
<td>0.8108</td>
<td>0.8590</td>
</tr>
<tr>
<td>$k = 8$</td>
<td>0.8291</td>
<td>0.8616</td>
<td>0.8752</td>
<td>0.8191</td>
<td>0.8029</td>
<td>0.8516</td>
</tr>
</tbody>
</table>

as $D_4$, where $m_{hi}$ is equal to 1 if $o_i \in S_h$ and $m_{hi}$ is equal to 0 if $o_i \notin S_h$. For convenience, we use $M_k^{(r)}$ to express the fuzzy partition matrix of partitions of the data set $r$ in the situation of $k$ clusters, where $r = 1, 2$. Table II displays the results of different FRI between the first data set and the crisp reference partition $D_4$ after clustering into $k$ clusters by FCM (called compared partition). Table III shows the results of different FRI between the second data set and the crisp reference partition $D_4$ after clustering into $k$ clusters by FCM.

From Tables II and III, we can find the following conclusions:

(i) Regardless of methods, the first data set and second data set are both successfully calculated to be closest to the crisp reference partition when $k = 4$.
That is, $\forall k \neq 4$, $FRC(M_4^{(1)}, D_4) = 0.9675 > FRC(M_k^{(1)}, D_4)$

$\forall k \neq 4$, $FRB(M_4^{(1)}, D_4) = 0.9684 > FRB(M_k^{(1)}, D_4)$;

$\forall k \neq 4$, $FRQ(M_4^{(1)}, D_4) = 0.9730 > FRQ(M_k^{(1)}, D_4)$;

$\forall k \neq 4$, $FRA(M_4^{(1)}, D_4) = 0.9480 > FRA(M_k^{(1)}, D_4)$;

$\forall k \neq 4$, $FRH(M_4^{(1)}, D_4) = 0.9475 > FRH(M_k^{(1)}, D_4)$;
∀k \neq 4, FRI(M^{(1)}_k, D_4) = 0.9815 > FRI(M^{(1)}_4, D_4).

(ii) When k = 4, all methods show that if a crisp reference partition is used as the basis, the first data set is closer to the crisp reference partition compared with the second data set. This result actually corresponds with the expected results.

That is, \( FRC(M^{(1)}_4, D_4) = 0.9675 > FRC(M^{(2)}_4, D_4) = 0.9055; \)
\( FRB(M^{(1)}_4, D_4) = 0.9684 > FRB(M^{(2)}_4, D_4) = 0.8951; \)
\( FRQ(M^{(1)}_4, D_4) = 0.9730 > FRQ(M^{(2)}_4, D_4) = 0.9159; \)
\( FRA(M^{(1)}_4, D_4) = 0.9480 > FRA(M^{(2)}_4, D_4) = 0.8699; \)
\( FRH(M^{(1)}_4, D_4) = 0.9475 > FRH(M^{(2)}_4, D_4) = 0.8601; \)
\( FRI(M^{(1)}_4, D_4) = 0.9815 > FRI(M^{(2)}_4, D_4) = 0.8969. \)

We further consider the similarity between two fuzzy partition matrices. We then explore the issue of comparing the similarities between various data sets and fuzzy reference partitions. We use Examples 7 and 8 to show that these methods proposed by Anderson et al.,\(^5\) Campello,\(^15\) Brouwer,\(^16\) Quere et al.,\(^18\) and Hullermeier et al.\(^19\) are not suitable for assessing the similarity in the two conditions mentioned previously. However, our method can effectively handle similarity issues in these two situations.

Example 7. In this example, we explore the similarity issue between various data sets and the same fuzzy reference partition. We continue using Example 6 where we call the data set in Figure 3a as data set 1 and the data set in Figure 3b as data set 2. We also generate a data set from Gaussian mixture distributions with the sample size \( n = 100 \) and the parameters \( \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1/4; \mu_1 = (100, 100), \mu_2 = (100, 200), \mu_3 = (200, 100), \mu_4 = (200, 200) \) and \( \Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma_4 = (600 0 0 600). \) We call it data set 3, as shown in Figure 4.

Let \( M^{(r)}_4 \) be the fuzzy partition matrix obtained by using FCM for the data set \( r, r = 1, 2, 3. \) Furthermore, we use \( M^{(1)}_4 \) as the reference fuzzy partition matrix. The results of different \( FRI \) between the data set \( r, r = 1, 2, 3 \) and the reference fuzzy partition matrix \( M^{(1)}_4 \) are shown in Table IV.

From Tables II–IV, we can conclude the followings:

(i) When \( k = 4, \) if the fuzzy partition matrix of the data set 1 can be used as the fuzzy reference partition, then the results gained from all methods show that the data set 1 is closest to the fuzzy reference partition, followed by the data set 2 and the data set 3. It is reasonable because the data set 1 and the reference fuzzy partition matrix are actually the same data set. This result corresponds to the expected results.

(ii) In \( FRC, FRB, FRQ, \) and \( FRA, \) we observe an unusual situation where the fuzzy similarity index of the same data set under various dividing methods (regardless of
Figure 4. Data set 3 generated from Gaussian mixture distributions.

Table IV. Fuzzy $RI$ between data set $r$ and reference fuzzy partition $M_{4}^{(1)}$

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Methods</th>
<th>$FRC$</th>
<th>$FRB$</th>
<th>$FRQ$</th>
<th>$FRA$</th>
<th>$FRH$</th>
<th>$FRI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data set 1 and data set 1 (i.e., $M_{4}^{(1)}, M_{4}^{(1)}$)</td>
<td></td>
<td>0.9390</td>
<td>0.9419</td>
<td>0.9579</td>
<td>0.9073</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Data set 2 and data set 1 (i.e., $M_{4}^{(2)}, M_{4}^{(1)}$)</td>
<td></td>
<td>0.8764</td>
<td>0.8720</td>
<td>0.9058</td>
<td>0.8356</td>
<td>0.8944</td>
<td>0.8934</td>
</tr>
<tr>
<td>Data set 3 and data set 1 (i.e., $M_{4}^{(3)}, M_{4}^{(1)}$)</td>
<td></td>
<td>0.8348</td>
<td>0.8227</td>
<td>0.8544</td>
<td>0.8072</td>
<td>0.8570</td>
<td>0.8484</td>
</tr>
</tbody>
</table>

whether the number of division for the data set remains the same) possibly is higher than the fuzzy similarity index of the same data set divided under the same division method and the same number of divisions.

For instance, $FRC(M_{4}^{(1)}, M_{4}^{(1)}) = 0.9390 < FRC(M_{4}^{(1)}, D_{4}) = 0.9675$

$$FRB(M_{4}^{(1)}, M_{4}^{(1)}) = 0.9419 < FRB(M_{4}^{(1)}, D_{4}) = 0.9684$$

$$FRQ(M_{4}^{(1)}, M_{4}^{(1)}) = 0.9579 < FRQ(M_{4}^{(1)}, D_{4}) = 0.9730$$

$$FRA(M_{4}^{(1)}, M_{4}^{(1)}) = 0.9073 < FRA(M_{4}^{(1)}, D_{4}) = 0.9480$$

Quere et al.\textsuperscript{18} mentioned that normalization yielded to the loss of this property (i.e., $FRI$ between $M_{4}^{(1)}$ and $M_{4}^{(1)}$ should be larger than $FRI$ between $M_{4}^{(1)}$ and $D_{4}$ by making possible pertinent reinforcements. However, both the methods developed by Hullermeier et al.\textsuperscript{19} and our proposed method do not show this abnormality. This is because

*International Journal of Intelligent Systems* DOI 10.1002/int
FRH\( (M_F^{(1)}, M_F^{(2)}) = 1 > \) FRH\( (M_F^{(1)}, D_4) = 0.9475 \)

\[ FRI(M_F^{(1)}, M_F^{(2)}) = 1 > FRI(M_F^{(1)}, D_4) = 0.9815 \]

**Example 8.** Regarding learning strategies for general domains, Dansereau\textsuperscript{27} believed that learning strategies can be classified into primary strategies and supportive strategies. Pintrich and DeGroot\textsuperscript{28} believed that high-quality learning motives generate more intricate learning strategies and, therefore, advocated the classification of learning strategies into cognitive strategies, self-regulated strategies, and management strategies. Different studies divide them in various ways. Combined with the subjectivity, complexity, and ambiguity of human behavior, language, and thought, the application of learning strategies is intrinsically diverse, complex, and ambiguous. Therefore, encountering similarity problems that involve two fuzzy partitions during further investigation of these issues is highly probable. Another example would be when discussing changes in the teaching methodology of teachers, the similarity of learning strategies before and after changing the instruction method also raises the issue of similarity between two fuzzy partitions. To explore similarity issues regarding the various methods of partition proposed by Dansereau\textsuperscript{27} and Pintrich and DeGroot,\textsuperscript{28} we randomly sample six students for interviews, instructing them to answer the questionnaires using the membership method. In the questionnaire, \( S_1^{(1)} \) represents primary strategies, \( S_2^{(1)} \) represents supportive strategies, \( S_3^{(2)} \) represents cognitive strategies, \( S_4^{(2)} \) represents self-regulated strategies, and \( S_5^{(2)} \) represents management strategies. The objects \( o_1, o_2, \ldots, o_6 \) sequentially represent the six students in the sample investigation. In this case, we use the data in Example 5 as the results of the six student questionnaires. We introduce the data into the various methods of \( FRI \), the following results are obtained:

(i) \( FRC(M_F^{(1)}, M_F^{(2)}) = 0.5204, \) \( FRB(M_F^{(1)}, M_F^{(2)}) = 0.6963, \) \( FRQ(M_F^{(1)}, M_F^{(2)}) = 0.7178, \)

\( FRA(M_F^{(1)}, M_F^{(2)}) = 0.5935, \) \( FRH(M_F^{(1)}, M_F^{(2)}) = 0.8267, \) \( FRI(M_F^{(1)}, M_F^{(2)}) = 0.6873. \)

(ii) The memberships of \( o_3 \) in \( M_F^{(1)} \) are exactly the same. This means that these two students use the same learning strategies. If \( FRC \) and \( FRB \) are used, a situation where students in the same cluster have memberships that do not equal 1 will occur (same for \( o_4, o_5 \) in \( M_F^{(2)} \)).

(iii) In \( FRA, N^* = \{0.95, 0.63, 0.82, 0.75, 1.37, 1.43\} \). Therefore, all four items of \( n_{11}(n_{11} - 1), n_{12}(n_{12} - 1), n_{13}(n_{13} - 1) \) and \( n_{23}(n_{23} - 1) \) are less than 0. This renders the original meaning of \( a = \sum_{i=1}^{k_3} \sum_{e=1}^{k_2} \left( \frac{R_{i,e}}{2} \right) \) questionable.

(iv) Although symmetry exists between \( FRC, FRB, FRQ, \) and \( FRA \), no reflexivity is presented. For example, \( FRC(M_F^{(1)}, M_F^{(2)}) = 0.6637, \) \( FRC(M_F^{(2)}, M_F^{(1)}) = 0.5263, \)

\( FRB(M_F^{(1)}, M_F^{(1)}) = 0.7429, \) \( FRB(M_F^{(2)}, M_F^{(2)}) = 0.7517, \) \( FRQ(M_F^{(1)}, M_F^{(1)}) = 0.8313, \)

\( FRQ(M_F^{(2)}, M_F^{(2)}) = 0.7646, \) \( FRA(M_F^{(1)}, M_F^{(1)}) = 0.4897, \) \( FRA(M_F^{(2)}, M_F^{(2)}) = 0.4676. \) This indicates that results can be obtained regardless of whether we use \( FRC, FRB, FRQ, \) or \( FRA, M_F^{(1)} = M_F^{(2)} \) is not the necessary and sufficient condition for \( FRI \) to adopt the value 1.

(v) \( FRH \) and \( FRI \) both possess symmetry and reflexivity. However, \( M_F^{(1)} = M_F^{(2)} \) is still not the necessary and sufficient condition for \( FRH(M_F^{(1)}, M_F^{(2)}) = 1. \) For
Table V. Summary of sampled students

<table>
<thead>
<tr>
<th>Course department</th>
<th>Department of Information Management (represented by $O_A$)</th>
<th>Department of Finance (represented by $O_B$)</th>
<th>Department of Electrical Engineering (represented by $O_C$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High scoring</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Average scoring</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Low scoring</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

For example, assume $M'_F(1) = [0.5, 0.5, 0.2, 0.8]$ and $M'_F(2) = [0.4, 0.6, 0.1, 0.9]$. Visibly, $M'_F(1) \neq M'_F(2)$. However, unreasonable results such as $FRH(M'_F(1), M'_F(2)) = 1$ occurred, nevertheless, $FRI(M'_F(1), M'_F(2)) = 0.8200 \neq 1$.

Finally, we present a real data set by applying the proposed $FRI$ in the next Example 9.

Example 9. To educators, a deep understanding of the differences in learning motivations among different classes is necessary for facilitating teachers in improving teaching methods and creating effective learning circumstances. However, human behavior, thought, and language are subjective, complex, multiple, and unclear. Therefore, both Law29 and Yen30 agreed that applying fuzzy theory is suitable for researching educational and psychological topics.

In a Calculus course, a questionnaire regarding motivations for learning Calculus, which involved items regarding intrinsic motivations + self-expected, test anxiety, self-control, subject value, and extrinsic motivations (a list of the questionnaire items appears in Appendix A that was translated from Chinese in Hsin31) is modified to incorporate fuzzy memberships. A sample of first-year university students were asked to complete the questionnaire. The final data dimensions of each student $o_i$ were equal to 5 (i.e., $o_i \in R^5$). In an object $o_i = (o_{i1}, o_{i2}, o_{i3}, o_{i4}, o_{i5})$, $o_{i1}$ represents the intrinsic motivations + self-expected measurement number of $o_i$; $o_{i2}$ represents the test anxiety measurement number of $o_i$; $o_{i3}$ represents the self-control measurement number of $o_i$; $o_{i4}$ represents the subject value measurement number of $o_i$; and $o_{i5}$ represents the extrinsic motivations measurement number of $o_i$.

Because the questionnaire used a five-point Likert scale, $o_{ij} \in [1,5]$ for all $i$ and $j = 1, \ldots, 4$. The purpose of providing this example is to consider whether students in different calculus courses taught by the same professor have similar learning motivations. In each of three courses, a sample of 15 students was selected. The students were categorized as high scoring, average scoring, and low scoring based on whether the students scored between 70 and 100, 50 and 70, or less than 50 on the mid-term and final examinations during the first semester of 2013. Table V presents a summary of the selected sample students.

Let students fill with questionnaire by a membership form. Thus, we adopt a fuzzy partition matrix to divide the object data sets. In general, Euclidean distance is the most frequently used for the higher dimension. Therefore, the fuzzy clustering algorithm $FCM$ (using Euclidean distance) is used in this example. Because the
Table VI. Fuzzy $RI$ between two different object data sets

<table>
<thead>
<tr>
<th>(Data set, data set)</th>
<th>$FRC$</th>
<th>$FRB$</th>
<th>$FRQ$</th>
<th>$FRA$</th>
<th>$FRH$</th>
<th>$FRI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(O_A, O_B)$</td>
<td>0.5054</td>
<td>0.5780</td>
<td>0.5511</td>
<td>0.5182</td>
<td>0.7667</td>
<td>0.6309</td>
</tr>
<tr>
<td>$(O_A, O_C)$</td>
<td>0.5153</td>
<td>0.5940</td>
<td>0.5700</td>
<td>0.5216</td>
<td>0.7691</td>
<td>0.5681</td>
</tr>
<tr>
<td>$(O_B, O_C)$</td>
<td>0.5944</td>
<td>0.6854</td>
<td>0.7593</td>
<td>0.5513</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>$(O_A, O_B)$</td>
<td>0.5849</td>
<td>0.6900</td>
<td>0.7474</td>
<td>0.5421</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>$(O_C, O_C)$</td>
<td>0.5851</td>
<td>0.7056</td>
<td>0.7781</td>
<td>0.5514</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Students in each class are classified according to high, average, or low scores, regardless of the fuzzy partition used, the number of clusters is with $k = 3$. Let $O_t$ represent the object data set where the $O_t$ is partitioned by FCM, $t = A, B, C$. The results are provided in Table VI.

From Table VI, we can conclude the following results:

(i) Regardless of methods, $FRI$ of two data sets $t$ and $t'(t', t = A, B, C)$ are successfully calculated.

(ii) Assume that two students $o_i$ and $o_j$ in the same data set have the same learning motivations in calculus, then we possess the result that the memberships of $o_i$ and $o_j$ in the data set are exactly the same. However, inappropriate results where students in the same cluster have memberships not equal 1 will occur when $FRC$ and $FRB$ are used. Of course, this is not good.

(iii) In $FRA(O_A, O_B)$, $N^* = [1.2622 1.6326 0.9294 \ 1.4333 2.6258 2.0859 \ 1.3685 2.0162 1.6461]$. Therefore, $n^*_{13}(n^*_{13} - 1)$ is less than 0.

This violates the original meaning because each term $(n_{uv})^2 = n_{uv}(n_{uv} - 1)$ in Equation (3) should be greater than or equal to zero. And, this situation also occurred in $FRA(O_A, O_C)$ and $FRA(O_B, O_C)$.

(iv) Although $FRC$, $FRB$, $FRQ$, and $FRA$ possess symmetry, they do not possess reflexivity. For example, $FRC(O_A, O_A) = 0.5944$, $FRB(O_A, O_A) = 0.6854$, $FRQ(O_A, O_A) = 0.7593$, and $FRA(O_A, O_A) = 0.5513$. In other words, $M^{(1)}_F = M^{(2)}_F$ is not the necessary and sufficient condition for $FRI$ to adopt the value 1 when we use $FRC$, $FRB$, $FRQ$, or $FRA$.

(v) In spite that $FRH$ possesses symmetry and reflexivity (i.e., $FRH(O_A, O_A) = 1$, $FRH(O_B, O_B) = 1$, and $FRH(O_C, O_C) = 1$), $M^{(1)}_F = M^{(2)}_F$ is still not the necessary and sufficient condition for $FRH(M^{(1)}_F, M^{(2)}_F) = 1$ by the explanation (v) of Example 8. Fortunately, our proposed method $FRI$ not only holds symmetry and reflexivity, but also owns the following important result:

$$M^{(1)}_F = M^{(2)}_F$$

if and only if $FRI(M^{(1)}_F, M^{(2)}_F) = 1$.

In reality, the color relation matrix used by the proposed $FRI$ divides the object unit of fuzzy partition matrix into smaller units. The fuzzy partition matrix for the processed smaller units is the crisp partition matrix. The fundamental essence is the same as $RI$.

(vi) In $FRI$, the calculus learning motivations of Information Management students are similar to those of Department of Finance students. This result is logical and reflects actual conditions. Because these two departments are within the School of Business (the Department of Electrical Engineering is in the School of Engineering), in which more
female than male students are enrolled (the reverse is true in the School of Engineering). In addition, business students do not possess the same subject value of Calculus as do engineering students. Therefore, in studying the mathematics class calculus, business students differ slightly from electrical engineering students in subject value and other learning motivations. However, other methods cannot get this result.

As a summary, this real example shows that the proposed FRI method is successfully generalized to treat similarities between two different object data sets with the same cardinal number and the same partition matrices. Overall, FRI actually provides a useful evaluation method that can be applied in future practical studies in various fields according to actual demands.

5. CONCLUSIONS AND DISCUSSION

In this paper, we propose a color relation matrix via a graph. We also verify that RI (or JI, etc.) can be found easily using a color relation matrix and the trace of the matrix. In addition, to consider the similarity between fuzzy partitions and a fuzzy reference partition, we extend a color relation matrix from crisp partitions to fuzzy partitions. We know that the RI (or JI, etc.) actually satisfies the important property that, for any two crisp partition matrices $M^{(1)}_C$ and $M^{(2)}_C$ of an object data set $O$, $M^{(1)}_C = M^{(2)}_C$ if and only if $RI(M^{(1)}_C, M^{(2)}_C) = 1$. Although Anderson et al.

Campello,15 Brouwer,16 Hullermeier and Rifqi,17 Quere et al.,18 and Hullermeier et al.19 had extended the RI to the FRI, the important property that $M^{(1)}_F = M^{(2)}_F$ if and only if $FRI$ between $M^{(1)}_F$ and $M^{(2)}_F$ equals 1 is ultimately unreachable. This led to some of results that should not have occurred. By defining a fuzzy color relation matrix and the trace matrix, we propose that FRI not only possess symmetry and reflexivity, but also possess the important property originally satisfied by RI. In other words, regardless of whether $M^{(1)}$ and $M^{(2)}$ are crisp or fuzzy partition matrices, our proposed FRI still provides that $M^{(1)} = M^{(2)}$ is the necessary and sufficient condition for $FRI(M^{(1)}, M^{(2)}) = 1$. This is the reason that FRI is not only appropriate for determining the similarities between fuzzy partitions and crisp reference partition, but also the similarity between fuzzy partition and fuzzy reference partition. They can even be used to investigate problems concerning the similarities between various data sets and the same fuzzy reference partitions. Thus, this provides a useful method that can be applied in future practical studies in various fields according to actual demands.

Acknowledgment

This work was supported in part by the Ministry of Science and Technology (MOST) of Taiwan under Grant MOST-103-2118-M-033-001-MY2.

References


International Journal of Intelligent Systems DOI 10.1002/int

APPENDIX A: CALCULUS LEARNING MOTIVATIONS QUESTIONNAIRE ITEMS AND CATEGORIES (HSIN31)

<table>
<thead>
<tr>
<th>Category</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic motivation + self-expected</td>
<td>I study calculus because I liked mathematics before.</td>
</tr>
<tr>
<td>Intrinsic motivation + self-expected</td>
<td>I study calculus because I like to solve problems.</td>
</tr>
<tr>
<td>Intrinsic motivation + self-expected</td>
<td>I study calculus because I want to challenge myself.</td>
</tr>
<tr>
<td>Intrinsic motivation + self-expected</td>
<td>I study calculus because of the feeling of achievement.</td>
</tr>
<tr>
<td>Intrinsic motivation + self-expected</td>
<td>I am confident of self-urging to learn calculus.</td>
</tr>
<tr>
<td>Intrinsic motivation + self-expected</td>
<td>I am confident to finish calculus homework independently.</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>I become nervous because of insufficient time before the test.</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>I stress for not finishing the book before examination.</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>During the test, I get nervous and cannot solve the easy problems if I have several questions unsolved.</td>
</tr>
<tr>
<td>Self-control</td>
<td>I cannot understand the content of calculus because I do not study hard.</td>
</tr>
<tr>
<td>Self-control</td>
<td>The grades of calculus can progress if I am diligent in calculus.</td>
</tr>
<tr>
<td>Subject value</td>
<td>I feel that calculus is helpful for the next courses, such as engineering mathematics, management mathematics, statistics, and so on.</td>
</tr>
<tr>
<td>Subject value</td>
<td>I feel that calculus is helpful for next grade school and employment in the future.</td>
</tr>
<tr>
<td>Subject value</td>
<td>I feel that calculus is important for other subjects (for example, physics, economics).</td>
</tr>
<tr>
<td>Extrinsic motivation</td>
<td>I study calculus because I am afraid of failure in class.</td>
</tr>
<tr>
<td>Extrinsic motivation</td>
<td>I study calculus in order to get good grades.</td>
</tr>
<tr>
<td>Extrinsic motivation</td>
<td>I study calculus because I do not want to waste time in makeup courses.</td>
</tr>
</tbody>
</table>