Measures of Uncertainty for an Incomplete Set-Valued Information System With the Optimal Selection of Subsystems: Gaussian Kernel Method

LIJUN CHEN1, SHIMIN LIAO2, NINGXIN XIE2, ZHAOWEN LI3, GANGQIANG ZHANG2, AND CHING-FENG WEN4,5

1School of Mathematics and Statistics, Yulin Normal University, Yulin 537000, China
2School of Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, China
3Key Laboratory of Complex System Optimization and Big Data Processing, Department of Guangxi Education, Yulin Normal University, Yulin 537000, China
4Center for Fundamental Science, Kaohsiung Medical University, Kaohsiung 80708, Taiwan
5Research Center for Nonlinear Analysis and Optimization, Kaohsiung Medical University, Kaohsiung 80708, Taiwan

Corresponding authors: Ningxin Xie (ningxinxie100@126.com) and Ching-Feng Wen (cfwen@kmu.edu.tw).

This work was supported in part by the National Natural Science Foundation of China under Grant 11971420, in part by the Project of Improving the Basic Scientific Research Ability of Young and Middle-aged Teachers in Guangxi Universities under Grant 2020KY14013, and in part by the Natural Science Foundation of Guangxi under Grant 2019JJA110027 and Grant 2018GXNSFDA294003.

ABSTRACT A set-valued information system (SVIS) with missing values is known as an incomplete set-valued information system (ISVIS). This article focuses on studying uncertainty measurement for an ISVIS and the optimal selection of subsystems by means of Gaussian kernel. First, the distance between two information values on each attribute in an ISVIS is put forward. Second, the fuzzy $T_{cos}$-equivalence relation induced by a given subsystem is proposed based on Gaussian kernel. Next, some tools are used to measure the uncertainty of an ISVIS. Moreover, effectiveness analysis is done from a statistical point of view. In the end, the optimal selection of subsystems based on $\delta$-information granulation and $\delta$-information amount is given. These results will help us comprehend nature of uncertainty in an ISVIS.

INDEX TERMS ISVIS, distance, Gaussian kernel, $T_{cos}$-equivalence relation, measure, effectiveness analysis, optimal selection.

I. INTRODUCTION Rough set theory as a mathematical tool for dealing with inaccuracy and uncertainty in data analysis has been successfully applied to many fields [17]–[21], [25], [26]. From philosophical point of view, rough set theory is established on the assumption that each object in the universe is connected with some information, expressed by means of some attributes used for object description [18]. Accordingly, an information system (IS) is a database that represents relationships between objects and attributes. If the information values of each object in an IS are sets, then this IS is called a set-valued information system (SVIS). Some scholars have studied SVISs. For instance, Yao [30] presented a set model for SVISs with upper and lower approximations, moreover, studied generalized decision logic; On the basis of knowledge induction process, Leung et al. [10] discussed a rough set approach for selecting decision rules with minimum feature sets in SVISs; Qian et al. [22] proposed a dominance relation for SVISs.

Uncertainty is caused by the limited resolution and incomplete description of the data. Measures of uncertainty have gradually become a significant research topic and given rise to a large number of people’s attentions. Aiming at uncertainty of IS, Shannon [24] introduced the concept of entropy and discussed the uncertainty with entropy. Later, Liang et al. [11] studied information granules and entropy theory in ISs; Liang et al. [12] investigated several kinds of entropy in incomplete ISs; Dai et al. [2] thought about entropy measures in SVISs; Qian et al. [23] considered fuzzy information entropy and granularity; Xu et al. [28] investigated rough entropy in ordered ISs; Dai et al. [4] proposed an extended conditional entropy in interval-valued decision systems; Dai et al. [3] put forward $\theta$-rough degree in IVISs on the foundation of $\theta$-similarity entropy; Dai et al. [5] explored entropy and granularity measures in SVISs; Huang et al. [8]
investigated uncertainty measures for intuitionistic fuzzy approximation space; Huang et al. [9] gave uncertainty measures in interval-valued intuitionistic fuzzy ISs; Xie et al. [27] took into account new method to measure the uncertainty of interval-valued ISs; Zhang et al. [37] measured the uncertainty of fully fuzzy ISs; Li et al. [13], [14] considered uncertainty measurements in fuzzy relation ISs and covering ISs.

An incomplete set-valued information system (ISVIS) is a SVIS with missing values. An ISVIS itself has uncertainty. How to measure its uncertainty is a crucial research topic. This article will study this issue. The similarity degree between two information values on a given attribute in an ISVIS is constructed and the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.

The rest of the article is arranged as follows. In the second section, we review some basic concepts about fuzzy sets and fuzzy relations. In the third section, we construct the distance degree between two information values on a given attribute in an ISVIS and give the distance between two objects is given. Fuzzy $T_{cos}$-equivalence relation is induced by a given subsystem of an ISVIS by means of Gaussian kernel. The uncertainty of a SVIS is measured. Effectiveness analysis is done from the angle of statistics. Based on them, the optimal selection of subsystems is given. The work process of the article is shown in FIGURE 1.
where * is a missing value.

Let \((U,A)\) be an IIS. For each \(a \in A\),

\[ V_a^* = V_a - \{a(u) : a(u) = *\}. \]

**Definition 7** [29]: Suppose that \((U,A)\) is an IS. Then \((U,A)\) is referred to as a set-valued information system (SVIS), if for any \(a \in A\) and \(u \in U\), \(a(u)\) is a set.

If \((U,A)\) is a SVIS. Given \(P \subseteq A\) and \(\theta \in [0, 1]\). Then a tolerance relation on \(U\) can be defined as

\[ R_P^\theta = \{(u,v) \in U \times U : \forall a \in P, s(a(u),a(v)) \leq \theta\}, \]

where \(s(a(u),a(v)) = \frac{|a(u) \cap a(v)|}{|a(u)| + |a(v)|}\) means the similarity degree between \(a(u)\) and \(a(v)\).

**Definition 8** [29]: Given that \((U,A)\) is an IS. Then \((U,A)\) is called an incomplete set-valued information system (ISVIS), if \((U,A)\) is both incomplete and set-valued.

If \(P \subseteq A\), then \((U,P)\) is referred to as the subsystem of \((U,A)\).

**Example 9**: TABLE 1 depicts an ISVIS \((U,A)\) with \(U = \{u_1, u_2, \ldots, u_{10}\}\) and \(A = \{a_1, a_2, \ldots, a_6\}\).

\[
\begin{align*}
V_{a_1}^* = & \{L,M,N\}, V_{a_2}^* = \{L,M\}, V_{a_3}^* = \{L,M\}, V_{a_4}^* = \{L,N\}, V_{a_5}^* = \{L,M,N\}, V_{a_6}^* = \{L,M\},
\end{align*}
\]

**III. DISTANCE BETWEEN TWO OBJECTS IN AN ISVIS**

**Definition 9**: Let \((U,A)\) be an ISVIS. Then \(\forall u,v \in U\), \(a \in A\), the distance between \(a(u)\) and \(a(v)\) is defined as

\[
d(a(u), a(v)) =
\begin{align*}
0, & \quad u = v; \\
1 - \frac{1}{|V_a^*|^2}, & \quad u \neq v, a(u) = *, a(v) = *; \\
1 - \frac{1}{|V_a^*|}, & \quad u \neq v, a(u) \neq *, a(v) = *; \\
1 - \frac{1}{|V_a^*|}, & \quad u \neq v, a(u) = *, a(v) \neq *; \\
0, & \quad u \neq v, a(u) \neq *, a(v) \neq *, a(u) \neq a(v); \\
1 - \frac{|a(u) \cap a(v)|}{|a(u) \cup a(v)|}, & \quad u \neq v, a(u) \neq *, a(v) \neq *, a(u) \neq a(v).
\end{align*}
\]

According to the above definition, the distance between two objects in an ISVIS is defined as follows.

**Definition 10**: Suppose that \((U,A)\) is an ISVIS. Given \(P \subseteq A\). \(\forall u,v \in U\), the distance between \(u\) and \(v\) in the subsystem \((U,P)\) is defined as

\[
d_P(u,v) = \sqrt{\sum_{a \in P} d^2(a(u),a(v))},
\]

where \(d(u,v) = d(a(u),a(v))\), \(a\) is a set-valued attribute.

**Proposition 12**: Assume that \((U,A)\) is an ISVIS. Given \(P \subseteq A\). Then \(\forall u,v \in U\),

\[
0 \leq d_P(u,v) \leq \sqrt{|P|}.
\]

**Proof**: Obviously. \(\square\)

**Example 13 (Continued From Example 9)**: Given \(P = \{a_1, a_2, a_3, a_4\}\). Calculate \(d_P(u_1,u_3)\) in TABLE 1.

By Definition 10, we have

\[
d(a_1(u_1), a_1(u_3)) = 1 - \frac{|a_1(u_1) \cap a_1(u_3)|}{|a_1(u_1) \cup a_1(u_3)|} = 1 - \frac{2}{3} \approx 0.3333;
\]

\[
d_P(u_1,u_3) = \sqrt{\sum_{a \in P} d^2(a(u_1),a(u_3))} = \sqrt{d(a_1(u_1),a_1(u_3))^2 + \ldots + d(a_4(u_1),a_4(u_3))^2}.
\]
TABLE 1. An ISVIS.

|   | a₁ | a₂ | a₃ | a₄ | a₅ | a₆ |
|---|----|----|----|----|----|----|
| u₁ | \{L, M, N\} | \{L, M\} | *  | \{L, N\} | \{L, M\} | \{L, M\} |
| u₂ | \{L, M, N\} | \{M, N\} | \{L, M\} | \{L, N\} | \{L, M\} | \{L, M\} |
| u₃ | \{L, M, N\} | \{M, N\} | \{L, N\} | \{M, N\} | \{L, M\} | \{L, M\} |
| u₄ | \{L, M, N\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} | \{L, N\} |
| u₅ | \{M, N\} | \{L, M\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} |
| u₆ | \{M, N\} | \{L, M\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} |
| u₇ | \{L, M, N\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} | \{L, M\} |
| u₈ | \{L, M, N\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} | \{L, M\} |
| u₉ | \{M, N\} | \{L, M\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} |
| u₁₀| \{L, M, N\} | \{L, N\} | \{L, M\} | \{M, N\} | \{L, M\} | \{L, M\} |

\[
d(a₂(u₁), a₂(u₃)) = 1 - \frac{|a₂(u₁) \cap a₂(u₃)|}{|a₂(u₁) \cup a₂(u₃)|} = 1 - \frac{1}{3} \approx 0.6667;
\]
\[
d(a₃(u₁), a₃(u₃)) = 1 - \frac{|a₃(u₁) \cap a₃(u₃)|}{|a₃(u₁) \cup a₃(u₃)|} = 1 - \frac{1}{2} = 0.5;
\]
\[
d(a₄(u₁), a₄(u₃)) = 1 - \frac{|a₄(u₁) \cap a₄(u₃)|}{|a₄(u₁) \cup a₄(u₃)|} = 1 - 1 = 0.
\]

Then
\[
d_A(u₁, u₃) = \sqrt{\sum_{a \in A} d²(a(u₁), a(u₃))}
\]
\[
\approx \sqrt{0.3333² + 0.6667² + 0.5² + 0²} \approx 0.8975.
\]

IV. FUZZY T₄-EQUIVALENCE RELATION BASED ON GAUSSIAN KERNEL IN AN ISVIS

In this section, the fuzzy T₄-equivalence relation induced by a given subsystem of an ISVIS is given by means of Gaussian kernel.

Gaussian kernel \(G(u, v) = \exp(-\frac{\|u-v\|^2}{2\delta^2})\) is used to compute the similarity between two objects \(u\) and \(v\), where \(\|u-v\|\) is the Euclidean distance between two objects \(u\) and \(v\), \(\delta\) is a threshold. In this article, pick \(\delta \in (0, 1]\).

Obviously, \(G(u, v)\) satisfies:

1. \(G(u, v) \in [0, 1]\);
2. \(G(u, v) = G(v, u)\);
3. \(G(u, u) = 1\).

Definition 14: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\), denote
\[
\begin{align*}
R_{P}^G(\delta)(u₁, u₂) &= \exp(-\frac{d_P^2(u₁, u₂)}{2\delta^2}), \\
M(R_{P}^G(\delta)) &= (R_{P}^G(\delta)(u₁, u₂))_{n \times n}.
\end{align*}
\]

Then \(M(R_{P}^G(\delta))\) is called the Gaussian kernel matrix of the subsystem \((U, P)\) with respect to \(\delta\).

Theorem 15: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then \(R_{P}^G(\delta)\) is a T₄-equivalence relation on \(U\).

Proof: This holds by Corollary 4.

Definition 16: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then \(R_{P}^G(\delta)\) is called the T₄-equivalence relation induced by the subsystem \((U, P)\) with respect to \(\delta\).

For any \(u \in U\), a fuzzy set \(\overline{[u]}_{R_{P}^G(\delta)}\) is defined as follows:
\[
\overline{[u]}_{R_{P}^G(\delta)}(v) = R_{P}^G(\delta)(u, v), \quad \forall v \in U.
\]

Algorithm 1 The T₄-Equivalence Relation

Input: An ISVIS \((U, A)\), \(P \subseteq A\) and \(\delta \in (0, 1]\).
Output: A T₄-equivalence relation \(R_{P}^G(\delta)\).

1. for \(i = 1; i \leq |U|; i++\) do
2. 
3. for \(j = |U| - 1; j > i; j--\) do
4. 
5.  \(d(a(u_i), a(u_j)) = 0;\)
6. 
7. for each \(a \in P\) do
8. 
9. 
10. end
11. end
12. end

Then \([u]_{R_{P}^G(\delta)}\) can be viewed as the fuzzy neighborhood of the point \(u\) on \(U\) with respect to \(\delta\) in the subsystem \((U, P)\).

Example 17: (Continued from Example 9) In TABLE 1, pick \(\delta = \sqrt{0.8}\), we have

Similarly,

Then \(R_{P}^G(\delta)\) is the T₄-equivalence relation induced by the system \((U, A)\) with respect to \(\delta\).

Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then an algorithm on a T₄-equivalence relation \(R_{P}^G(\delta)\) is designed as follows.

V. RELATIONSHIPS BETWEEN TWO ISVISs

In this section, we investigate relationships between two ISVISs and display inclusion degree of IVISs below.

Definition 18: Let \((U, P)\) and \((U, Q)\) be two ISVISs. Given \(\delta \in (0, 1]\). If for any \(u \in U\), \([u]_{R_{P}^G(\delta)} \subseteq [u]_{R_{Q}^G(\delta)}\), then \((U, P)\) and \((U, Q)\) are called to be equivalent with respect to \(\delta\). We write \((U, P) \approx_{\delta} (U, Q)\).

Obviously,
\[
(U, P) \approx_{\delta} (U, Q) \Leftrightarrow R_{P}^G(\delta) = R_{Q}^G(\delta).
\]

Definition 19: Assuming that \((U, P)\) and \((U, Q)\) are two ISVISs. Given \(\delta \in (0, 1]\).
(1) \((U, Q)\) is called to depend on \((U, P)\) with respect to \(\delta\), if for any \(u \in U\), \([u]^{R^G_{\{u\}}(\delta)} \subseteq [u]^{R^G_{\{u\}}(\delta)}\), we write \((U, P) \preceq_\delta (U, Q)\); \((U, Q)\) is known as to depend strictly on \((U, P)\) with respect to \(\delta\), if \((U, P) \preceq_\delta (U, Q)\) and \((U, P) \not\sim_\delta (U, Q)\), we write \((U, P) \prec_\delta (U, Q)\).

(2) \((U, Q)\) is called to depend partially on \((U, P)\) with respect to \(\delta\), if exists \(u \in U\), \([u]^{R^G_{\{u\}}(\delta)} \subseteq [u]^{R^G_{\{u\}}(\delta)}\), we write \((U, P) \sqsubseteq_\delta (U, Q)\); \((U, P)\) is known as to depend strictly on \((U, Q)\), if \((U, Q) \sqsubseteq_\delta (U, P)\) and \((U, Q) \not\approx_\delta (U, P)\), we can write \((U, Q) \sqsubset_\delta (U, P)\).

(3) \((U, P)\) is referred to be independent of \((U, Q)\), if for each \(u \in U\), \([u]^{R^G_{\{u\}}(\delta)} \subseteq [u]^{R^G_{\{u\}}(\delta)}\), we write \((U, Q) \gtrsim_\delta (U, P)\).

Clearly, the following conclusions can be obtained.

\[(U, Q) \preceq_\delta (U, P) \iff (U, Q) \preceq_\delta (U, P) \text{ and } (U, Q) \approx_\delta (U, P)\]

\begin{equation}
M(R^G_{\{u_1\}}(\delta)) = \begin{bmatrix}
1.00 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 & 1.00 \\
1.00 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 & 1.00 \\
0.993 & 0.993 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 \\
1.00 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 & 1.00 \\
0.993 & 0.993 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 \\
0.993 & 0.993 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 \\
1.00 & 1.00 & 0.993 & 1.00 & 0.993 & 1.00 & 1.00 & 1.00 & 0.993 & 0.993 & 1.00 \\
\end{bmatrix}
\end{equation}

\begin{equation}
M(R^G_{\{u_2\}}(\delta)) = \begin{bmatrix}
1.00 & 0.966 & 1.00 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.966 & 0.966 & 1.00 \\
0.966 & 1.00 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.895 & 0.966 & 0.966 & 0.966 \\
1.00 & 0.966 & 1.00 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.966 & 0.966 & 1.00 \\
0.895 & 0.966 & 0.895 & 1.00 & 1.00 & 0.895 & 1.00 & 1.00 & 0.966 & 0.966 & 0.895 \\
0.895 & 0.966 & 0.895 & 1.00 & 1.00 & 0.895 & 1.00 & 1.00 & 0.966 & 0.966 & 0.895 \\
0.966 & 0.838 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 1.00 \\
1.00 & 0.966 & 1.00 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.966 & 0.966 & 1.00 \\
\end{bmatrix}
\end{equation}

\begin{equation}
M(R^G_{\{u_3\}}(\delta)) = \begin{bmatrix}
1.00 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 \\
0.966 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 1.00 & 0.966 & 0.966 & 0.966 & 0.966 \\
0.966 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 1.00 & 0.966 & 0.966 & 0.966 \\
0.966 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 1.00 & 0.966 & 0.966 \\
0.966 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 1.00 & 0.966 \\
0.838 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 1.00 \\
0.966 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 \\
0.838 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 1.00 \\
\end{bmatrix}
\end{equation}

\begin{equation}
M(R^G_{\{u_4\}}(\delta)) = \begin{bmatrix}
1.00 & 0.966 & 1.00 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 1.00 & 0.895 & 1.00 \\
0.966 & 1.00 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.895 & 0.966 & 0.966 & 0.966 \\
0.966 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 0.895 & 1.00 & 0.966 & 0.966 & 0.966 \\
0.895 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 \\
0.895 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 \\
0.895 & 0.895 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 \\
0.895 & 0.895 & 0.895 & 0.895 & 0.895 & 1.00 & 0.895 & 0.895 & 0.895 & 0.895 & 0.895 \\
\end{bmatrix}
\end{equation}
\[(U, P) \preceq_\delta (U, Q), \]

\[(U, Q) \preceq_\delta (U, P) \Rightarrow (U, Q) \subseteq_\delta (U, P), \]

\[(U, Q) \nsubseteq_\delta (U, P) \Rightarrow (U, Q) \supseteq_\delta (U, P). \]

**Theorem 20:** Suppose that \((U, P)\) and \((U, Q)\) are two ISVISs. If \(P \subseteq Q\), then for any \(\delta \in (0, 1]\), \((U, P) \preceq_\delta (U, Q)\).

**Proof:** Obviously.

Suppose that \((U, A)\) is an ISVIS. Denote

\[\Sigma_{(U, A)} = \{(U, P) : P \subseteq A\}. \]

Given \(\delta \in (0, 1]\). It is obvious that \((\Sigma_{(U, A)}, \preceq_\delta)\) is a partial order set.

**Definition 21** [39]: Let \((U, A)\) be an ISVIS. Given \(\delta \in (0, 1]\). Assuming that a mapping \(D_\delta : \Sigma_{(U, A)} \times \Sigma_{(U, A)} \rightarrow [0, 1]\) is said to be the inclusion degree on \(\Sigma_{(U, A)}\) with respect to \(\delta\), if it satisfies the following conditions: for any \((U, O), (U, P), (U, Q) \in \Sigma_{(U, A)}\),

1. \(0 \leq D_\delta((U, P)/(U, O)) \leq 1;\)
2. \((U, O) \preceq_\delta (U, P)\) implies \(D_\delta((U, P)/(U, O)) = 1;\)
3. \((U, O) \preceq_\delta (U, P)\) implies \(D_\delta((U, O)/(U, Q)) \leq D_\delta((U, O)/(U, P)).\)

**Definition 22:** Assuming that \((U, P)\) and \((U, Q)\) are two ISVISs. Given \(\delta \in (0, 1]\), define

\[D_\delta((U, Q)/(U, P)) = \sum_{i=1}^{n} \frac{[u_i]^{R_G^\delta}(\delta)}{\sum_{i=1}^{n} [u_i]^{R_G^\delta}(\delta)} \cdot \chi_{[u_i]^{R_G^\delta}(\delta)}([u_i]^{R_G^\delta}(\delta)), \]

where

\[\chi_{[u_i]^{R_G^\delta}(\delta)}([u_i]^{R_G^\delta}(\delta)) = \begin{cases} 1, & \text{if } [u_i]^{R_G^\delta}(\delta) \subseteq [u_i]^{R_G^\delta}(\delta), \\ 0, & \text{if } [u_i]^{R_G^\delta}(\delta) \nsubseteq [u_i]^{R_G^\delta}(\delta). \end{cases} \]

**Proposition 23:** \(D_\delta\) in Definition 22 is the inclusion degree under Definition 21.

**Proof:** Suppose \(O, P, Q \subseteq A\) and \(\delta \in (0, 1]\).

1. Obviously, \(0 \leq D_\delta((U, Q)/(U, P)) \leq 1.\)
2. Suppose \((U, P) \preceq_\delta (U, Q)\). Then, by Definition 19, \([u_i]^{R_G^\delta}(\delta) \subseteq [u_i]^{R_G^\delta}(\delta)\). Thus, for each \(l\), \([u_i]^{R_G^\delta}(\delta) \subseteq [u_i]^{R_G^\delta}(\delta)\).

This result implies that

\[\chi_{[u_i]^{R_G^\delta}(\delta)}([u_i]^{R_G^\delta}(\delta)) = 1. \]

Thus, \(D_\delta((U, Q)/(U, P)) = 1.\)

3. Suppose \((U, P) \nsubseteq_\delta (U, Q)\). Then, by Definition 19, \([u_i]^{R_G^\delta}(\delta) \subseteq [u_i]^{R_G^\delta}(\delta) \nsubseteq [u_i]^{R_G^\delta}(\delta)\). Thus, for each \(l\), \([u_i]^{R_G^\delta}(\delta) \nsubseteq [u_i]^{R_G^\delta}(\delta)\).

By Definition 22,

\[D_\delta((U, P)/(U, Q)) = \sum_{i=1}^{n} \frac{[u_i]^{R_G^\delta}(\delta)}{\sum_{i=1}^{n} [u_i]^{R_G^\delta}(\delta)} \cdot \chi_{[u_i]^{R_G^\delta}(\delta)}([u_i]^{R_G^\delta}(\delta)). \]

If \([u_i]^{R_G^\delta}(\delta) \nsubseteq [u_i]^{R_G^\delta}(\delta)\), then \([u_i]^{R_G^\delta}(\delta) \nsubseteq [u_i]^{R_G^\delta}(\delta)\). This result implies that

\[\chi_{[u_i]^{R_G^\delta}(\delta)}([u_i]^{R_G^\delta}(\delta)) = 0 \implies \chi_{[u_i]^{R_G^\delta}(\delta)}([u_i]^{R_G^\delta}(\delta)) = 0. \]

Thus,

\[D_\delta((U, P)/(U, Q)) \leq D_\delta((U, P)/(U, Q)). \]
From the above, we know that $D_δ$ is the inclusion degree.

It can be obtained that the inclusion degree has the ability to quantify relationships by the theorem below.

**Theorem 24:** Assuming that $(U, P)$ and $(U, Q)$ are two ISVISs. Given $δ ∈ (0, 1]$, (1) $(U, P) ≤_δ (U, Q) ⇔ D_δ((U, Q)/(U, P)) = 1$; (2) $(U, P) ≈_δ (U, Q) ⇔ D_δ((U, Q)/(U, P)) = 0$; (3) $(U, P) ≥_δ (U, Q) ⇔ 0 < D_δ((U, Q)/(U, P)) ≤ 1$.

Proof: (1) “$⇒$” is evident. We prove “$⇐$”. Suppose $|[u_i]|_{RG_δ} = q_i$, $\sum_{i=1}^{n} |[u_i]|_{RG_δ} = q$.

Then,

$$q = \sum_{i=1}^{n} q_i.$$

Owing to $D_δ((U, Q)/(U, P)) = 1$, it can be obtained that $n \sum_{i=1}^{n} q_i \chi_{[u_i]_{RG_δ}}([[u_i]|_{RG_δ})) = \sum_{i=1}^{n} q_i = q$.

Then,

$$q(1 - \chi_{[u_i]_{RG_δ}}([[u_i]|_{RG_δ})) = 0.$$

Consequently, $\forall l$,

$$1 - \chi_{[u_i]_{RG_δ}}([[u_i]|_{RG_δ}) = 0.$$

Thus, it can be obtained that $\forall l, [u_i]|_{RG_δ} ⊆ [u_i]|_{RG_δ}$. By Definition 19, $(U, P) ≤_δ (U, Q)$.

(2) “$⇒$”. Owing to $(U, P) ≈_δ (U, Q)$, it can be obtained that $[u_i]|_{RG_δ} ≼ [u_i]|_{RG_δ}$. Then $\forall l$,

$$\chi_{[u_i]_{RG_δ}}([[u_i]|_{RG_δ})) = 0.$$

By Definition 22, $D_δ((U, Q)/(U, P)) = 0$.

“$⇐$”. Owing to $D_δ((U, Q)/(U, P)) = 0$, it can be obtained that $\forall l, \chi_{[u_i]_{RG_δ}}([[u_i]|_{RG_δ})) = 0$.

Then, $\forall l, [u_i]|_{RG_δ} ≼ [u_i]|_{RG_δ}$. By Definition 19, $(U, P) ≈_δ (U, Q)$.

(3) The result can be obtained from (1) and (2). □

**VI. MEASURING UNCERTAINTY IN AN ISVIS**

Uncertainty of a given ISVIS is derived from uncertainty of fuzzy relations. In this section, we put forward some tools to measure uncertainty.

**A. GRANULATION MEASUREMENT FOR AN ISVIS**

**Definition 25:** Let $(U, A)$ be an ISVIS. Given $δ ∈ (0, 1]$. Suppose that $G_δ : 2^A \to (-∞, +∞)$ is a function. Then $G$ is called a $δ$-information granulation function in $(U, A)$ with respect to $δ$, if $G$ satisfies the following conditions:

(1) $\forall P \in 2^A$, $G_δ(P) ≥ 0$ (Non-negativity);
(2) $\forall P, Q \in 2^A$, if $(U, P) ≈_δ (U, Q)$, then $G_δ(P) = G_δ(Q)$ (Invarianlity);
(3) $\forall P, Q \in 2^A$, if $(U, P) ≤_δ (U, Q)$, then $G_δ(P) < G_δ(Q)$ (Monotonicity).

**Definition 26:** Suppose that $(U, A)$ is an ISVIS. Given $P \subseteq A$ and $δ ∈ (0, 1]$. Then $δ$-information granulation of $(U, P)$ with respect to $δ$ is defined as

$$G_δ(P) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} [u_i]|_{RG_δ}.$$
Theorem 29: $G_\delta$ in Definition 26 is an $\delta$-information granulation function under Definition 25.

Proof: (1) Obviously, “Non-negativity” holds.

(2) Given $P, Q \subseteq A$. If $(U, P) \approx_\delta (U, Q)$, then $\forall i$, $[u_i]^{G_P(\delta)} = [u_i]^{G_Q(\delta)}$.

By Definition 26, $G_\delta(P) = G_\delta(Q)$.

(3) “Monotonicity” follows from Theorem 28. \hfill \Box

B. ENTROPY MEASUREMENTS FOR AN ISVIS

Definition 30: Let $(U, A)$ be an ISVIS. Given $P \subseteq A$ and $\delta \in (0, 1]$. Then $\delta$-rough entropy of $(U, P)$ with respect to $\delta$ is defined as

$$(E_\delta)_s(P) = -\sum_{i=1}^{n} \frac{|[u_i]^{G_P(\delta)}|}{n} \log_2 \frac{1}{|[u_i]^{G_P(\delta)}}.$$

Proposition 31: Let $(U, A)$ be an ISVIS. Given $P \subseteq A$ and $\delta \in (0, 1]$. Then

$$-\infty < (E_\delta)_s(P) \leq \log_2 n.$$

Furthermore, if $R_P^G(\delta) = \omega$, then $E_\delta$ reaches the maximum value $\log_2 n$; if $R_P^G(\delta)$ is reflexive, then

$$0 \leq (E_\delta)_s(P) \leq \log_2 n.$$

Proof: (1) By Definition 30,

$$(E_\delta)_s(P) = \frac{1}{n} \sum_{i=1}^{n} \log_2 \left( \frac{1}{|[u_i]^{G_P(\delta)}} \right) \sum_{j=1}^{n} R_P^G(\delta)(u_j) \forall i, j,$$

$$0 \leq R_P^G(\delta)(u_i, u_j) \leq 1.\,\text{Then}\,\forall i,$$

$$-\infty < \log_2\left( \sum_{j=1}^{n} R_P^G(u_i, u_j) \right) \leq \log_2 n.$$

This means that

$$-\infty < \sum_{j=1}^{n} \log_2 \left( \sum_{j=1}^{n} R_P^G(u_i, u_j) \right) \leq n \log_2 n.$$

Thus

$$-\infty < (E_\delta)_s(P) \leq \log_2 n.$$

(2) Suppose $R_P^G(\delta) = \omega$. Then $\forall i, j, R_P^G(\delta)(u_i, u_j) = 1$. Thus

$$(E_\delta)_s(P) = \log_2 n.$$

(3) Suppose that $R_P^G(\delta)$ is reflexive. Then $\forall i, R(u_i, u_i) = 1$. So $\forall i,$

$$1 \leq \sum_{j=1}^{n} R_P^G(\delta)(u_i, u_j) \leq n.$$

Thus $\forall i,$

$$0 \leq \log_2\left( \sum_{j=1}^{n} R_P^G(\delta)(u_i, u_j) \right) \leq \log_2 n.$$

Hence

$$0 \leq (E_\delta)_s(P) \leq \log_2 n.$$

\hfill \Box

Proposition 32: Let $(U, A)$ be an ISVIS. Given $P, Q \subseteq A$ and $\delta \in (0, 1]$. If $(U, P) \nsim_\delta (U, Q)$, then $(E_\delta)_s(P) < (E_\delta)_s(Q)$.

Proof: (1) Similar to the proof of Theorem 28, we obtain that $\forall i, j,$

$$R_P^G(\delta)(u_i, u_j) \leq R_Q^G(\delta)(u_i, u_j),$$

and $\exists i', j'$,

$$R_P^G(\delta)(u_{i'}, u_{j'}) < R_Q^G(\delta)(u_{i'}, u_{j'}).$$

Then

$$\sum_{j=1}^{n} R_P^G(u_{i'}, u_{j'}) < \sum_{j=1}^{n} R_Q^G(u_{i'}, u_{j'}).$$

So

$$\log_2\left( \sum_{j=1}^{n} R_P^G(u_{i'}, u_{j'}) \right) < \log_2\left( \sum_{j=1}^{n} R_Q^G(u_{i'}, u_{j'}) \right).$$

Thus

$$(E_\delta)_s(P) < (E_\delta)_s(Q).$$

\hfill \Box

Theorem 33: $(E_\delta)_s$ in Definition 30 is an $\delta$-information granulation function under Definition 25.

Proof: (1) Obviously, “Non-negativity” holds.

(2) Given $P, Q \subseteq A$. If $(U, P) \approx_\delta (U, Q)$, then $\forall i, [u_i]^{G_P(\delta)} = [u_i]^{G_Q(\delta)}$. 

VOLUME 8, 2020 212029
(3) By Definition 30, \((E_r)_\delta(P) = (E_r)_\delta(Q)\). “Monotonicity” follows from Proposition 32.

C. INFORMATION ENTROPY FOR AN ISVIS

Definition 34: Suppose that \((U, A)\) is an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then \(\delta\)-information entropy of \((U, P)\) with respect to \(\delta\) is defined as

\[
H_\delta(P) = -\sum_{i=1}^{n} \frac{|u_i|^\delta \log_2 \frac{|u_i|^\delta}{n}}{n}.
\]

Theorem 35: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then

\[
(E_r)_\delta(P) + H_\delta(P) = \log_2 n.
\]

Proof: By Definitions 30 and 34,

\[
(E_r)_\delta(P) = -\sum_{i=1}^{n} \frac{1}{|u_i|^\delta} \frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{|u_i|^\delta}{n}
\]

\[
H_\delta(P) = -\frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{|u_i|^\delta}{n}.
\]

Then

\[
(E_r)_\delta(P) + H_\delta(P) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{|u_i|^\delta} \frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{|u_i|^\delta}{n}
\]

\[
= -\frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{1}{|u_i|^\delta} + \frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{|u_i|^\delta}{n}
\]

\[
= -\frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{1}{n} = \log_2 n.
\]

Thus

\[
(E_r)_\delta(P) + H_\delta(P) = \log_2 n.
\]

Corollary 36: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then

\[
0 \leq H_\delta(P) < +\infty.
\]

Besides, if \(R_G^\delta(\delta) = \infty\), then \(H\) reaches the minimum value 0; if \(R_G^\delta(\delta)\) is reflective, then

\[
0 \leq H_\delta(P) \leq \log_2 n.
\]

Proof: This holds by Proposition 31 and Theorem 35.

D. INFORMATION AMOUNT OF AN ISVIS

Definition 37: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then \(\delta\)-information amount of \((U, P)\) with respect to \(\delta\) is defined as

\[
E_\delta(P) = \sum_{i=1}^{n} \frac{|u_i|^\delta \log_2 \frac{|u_i|^\delta}{n}}{n}.
\]

Theorem 38: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta = 0, 1\). Then

\[
G_\delta(P) + E_\delta(P) = 1.
\]

Proof: By Definition 26, \(G_\delta(P) = \frac{1}{n} \sum_{i=1}^{n} |u_i|^\delta\).

By Definition 37, \(E_\delta(P) = \frac{1}{n} \sum_{i=1}^{n} |u_i|^\delta \log_2 \frac{|u_i|^\delta}{n}.

\[
= \frac{1}{n} \sum_{i=1}^{n} (|u_i|^\delta \log_2 \frac{|u_i|^\delta}{n} - |u_i|^\delta + n - 1)
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} (|u_i|^\delta \log_2 \frac{|u_i|^\delta}{n} + (n - \frac{|u_i|^\delta}{n}))
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} n = 1.
\]

Thus

\[
G_\delta(P) + E_\delta(P) = 1.
\]

Corollary 39: Let \((U, A)\) be an ISVIS. Given \(P \subseteq A\) and \(\delta \in (0, 1]\). Then

\[
0 \leq E_\delta(P) \leq 1.
\]

Furthermore, if \(R_G^\delta(\delta) = \infty\), then \(E\) reaches the minimum value 0; if \(R_G^\delta(\delta) = 0\), then \(E\) reaches the maximum value 1.

Proof: This holds by Proposition 27 and Theorem 38.

Example 40: (Continued from Example 17) Pick \(B_1 = \{a_1, a_2, \ldots, a_6\}\) and \(\delta = \sqrt{0.8}\).

By Definition 26,

\[
G_\delta(B_1) = \frac{1}{10} \sum_{i=1}^{10} |u_i|^\delta \approx 0.9966,
\]

\[
G_\delta(B_2) = \frac{1}{10} \sum_{i=1}^{10} |u_i|^\delta \approx 0.9356,
\]

\[
G_\delta(B_3) = \frac{1}{10} \sum_{i=1}^{10} |u_i|^\delta \approx 0.8235,
\]

\[
G_\delta(B_4) = \frac{1}{10} \sum_{i=1}^{10} |u_i|^\delta \approx 0.6652,
\]

\[
G_\delta(B_5) = \frac{1}{10} \sum_{i=1}^{10} |u_i|^\delta \approx 0.4450,
\]

\[
G_\delta(B_6) = \frac{1}{10} \sum_{i=1}^{10} |u_i|^\delta \approx 0.3383.
\]

By Definition 30,

\[
(E_r)_\delta(B_1) = -\frac{1}{10} \sum_{i=1}^{10} \frac{|u_i|^\delta}{|u_i|^\delta} + \frac{1}{n} \log_2 \frac{1}{|u_i|^\delta} \approx 33.0555,
\]
By Definition 34,

\[
H_3(B_1) = - \sum_{i=1}^{10} \frac{[u_i]^G_{B_1(3)}}{10} \log_2 \left( \frac{[u_i]^G_{B_1(3)}}{10} \right) \approx 0.0495,
\]

\[
H_3(B_2) = - \sum_{i=1}^{10} \frac{[u_i]^G_{B_2(3)}}{10} \log_2 \left( \frac{[u_i]^G_{B_2(3)}}{10} \right) \approx 0.8986,
\]

\[
H_3(B_3) = - \sum_{i=1}^{10} \frac{[u_i]^G_{B_3(3)}}{10} \log_2 \left( \frac{[u_i]^G_{B_3(3)}}{10} \right) \approx 2.3055,
\]

\[
H_3(B_4) = - \sum_{i=1}^{10} \frac{[u_i]^G_{B_4(3)}}{10} \log_2 \left( \frac{[u_i]^G_{B_4(3)}}{10} \right) \approx 3.9846,
\]

\[
H_3(B_5) = - \sum_{i=1}^{10} \frac{[u_i]^G_{B_5(3)}}{10} \log_2 \left( \frac{[u_i]^G_{B_5(3)}}{10} \right) \approx 5.1579,
\]

\[
H_3(B_6) = - \sum_{i=1}^{10} \frac{[u_i]^G_{B_6(3)}}{10} \log_2 \left( \frac{[u_i]^G_{B_6(3)}}{10} \right) \approx 5.2473.
\]

By Definition 37,

\[
E_3(B_1) = \sum_{i=1}^{10} \frac{[u_i]^G_{B_1(3)} - [u_i]^G_{B_1(3)}}{10} \approx 0.0343,
\]

\[
E_3(B_2) = \sum_{i=1}^{10} \frac{[u_i]^G_{B_2(3)} - [u_i]^G_{B_2(3)}}{10} \approx 0.6026,
\]

\[
E_3(B_3) = \sum_{i=1}^{10} \frac{[u_i]^G_{B_3(3)} - [u_i]^G_{B_3(3)}}{10} \approx 1.4513,
\]

\[
E_3(B_4) = \sum_{i=1}^{10} \frac{[u_i]^G_{B_4(3)} - [u_i]^G_{B_4(3)}}{10} \approx 2.2483,
\]

\[
E_3(B_5) = \sum_{i=1}^{10} \frac{[u_i]^G_{B_5(3)} - [u_i]^G_{B_5(3)}}{10} \approx 2.4441,
\]

\[
E_3(B_6) = \sum_{i=1}^{10} \frac{[u_i]^G_{B_6(3)} - [u_i]^G_{B_6(3)}}{10} \approx 2.2183.
\]

The results of these experiments are shown in FIGURE 2.

It can be seen the truth that with the attribute subset $B \subseteq A$, uncertainty measures of the ISVIS ($U$, $A$) show certain regularity, which are reflected by the following truths:

1) $G_3$ and $(E_r)_3$ monotonically decrease with the increase of number of attributes;

2) $(E_r)_3$ and $H_3$ are more sensitive than $G_3$;

3) $(E_r)_3$ and $H_3$ are more sensitive than $E_3$;

4) The difference among $G_3$ and $E_3$ is almost the same.

Thus, $\delta$-rough entropy and $\delta$-information amount and $\delta$-information granularity are more suitable than $\delta$-information amount and $\delta$-information granularity for an ISVIS.

VII. EFFECTIVENESS ANALYSIS

In this section, effectiveness analysis is put forward from three aspects.

A. DISPERSION ANALYSIS

Assume that $X = \{x_1, \ldots, x_n\}$ is a data set. Then its arithmetic average value (resp. standard deviation, standard deviation coefficient) is regarded as $\bar{x}$ ($\sigma(X)$, $CV(X)$), they are defined as follows:

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x, \quad \sigma(X) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x - \bar{x})^2}, \quad CV(X) = \frac{\sigma(X)}{\bar{x}}.
\]

Example 41 (Continued From Example 40): Denote

\[
X_G = \{G_3(B_1), \ldots, G_3(B_6)\}, \quad X_E = \{(E_r)_3(B_1), \ldots, (E_r)_3(B_6)\},
\]

\[
X_H = \{H_3(B_1), \ldots, H_3(B_6)\}, \quad X_E = \{E_3(B_1), \ldots, E_3(B_6)\}.
\]

Then

\[
CV(X_G) = 0.3487, \quad CV(X_E) = 0.4963, \quad CV(X_H) = 0.6838, \quad CV(X_E) = 0.6040.
\]

The results are shown in FIGURE 3.

So

\[
CV(X_H) > CV(X_E) > CV(X_E) > CV(X_G).
\]

Then dispersion degree of $G$ reaches minimum.
L. Chen et al.: Measures of Uncertainty for an ISVIS With the Optimal Selection of Subsystems

From FIGUREs 2 and 3, the following results can be obtained:

1. \( (E_r)_h \) and \( H_h \) have better performance to measure uncertainty of an ISVIS if the monotonicity is only considered;
2. \( (E_r)_h \) has better performance to measure uncertainty of an ISVIS if the monotonicity and dispersion degree are both considered.

B. ASSOCIATION ANALYSIS

In statistics, Pearson correlation coefficient is a measure of the strength of a linear correlation between two data sets. Suppose that \( X = \{x_1, x_2, \ldots, x_n\} \) and \( Y = \{y_1, y_2, \ldots, y_n\} \) are two data sets. Pearson correlation coefficient between \( X \) and \( Y \), denoted by \( r(X, Y) \), is defined as

\[
r(X, Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}.
\]

where \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \), \( \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \).

Obviously,

\[-1 \leq r(X, Y) \leq 1.\]

The correlation between \( X \) and \( Y \) can be obtained according to TABLE 2.

Example 42 (Continued From Example 40): Pearson correlation coefficients are calculated as follows (see TABLE 3).

From TABLE 3, the following results are obtained (see TABLE 4):

C. FRIEDMAN TEST AND BONFERRONI-DUNN TEST

To further explore whether the performance of each uncertainty measurement with the six subsystems are significantly different, Friedman test [6] and Bonferroni-Dunn test [1] are given in this subsection.

Friedman test is a statistical test that uses the rank of algorithms. Friedman statistic is defined as

\[
\chi^2_F = \frac{12N}{k(k+1)} \left( \sum_{i=1}^{k} r_i^2 - \frac{k(k+1)^2}{4} \right)
\]

where \( k \) is the number of algorithms, \( N \) is the number of data sets, \( r_i \) is the average ranking of the \( i \)-th algorithm. When \( k \) and \( N \) are large enough, Friedman statistic follows the chi-square distribution with \( k - 1 \) degrees of freedom. However, such Friedman test is too conservation, and is usually replaced by the next statistic

\[
F_F = \frac{(N-1)\chi^2_F}{N(k-1)-\chi^2_F}.
\]

The statistic \( F_F \) follows the Fisher distribution with \( k - 1 \) and \( (k - 1)(N - 1) \) degrees of freedom. If the value of the
TABLE 5. The ranking of uncertainty measurements for ISVIS with different subsystems.

| Vectors | $G_3$ | $(E_r)_3$ | $H_3$ | $E_3$ |
|---------|-------|-----------|-------|-------|
| $B_1$   | 3     | 4         | 2     | 1     |
| $B_2$   | 3     | 4         | 1     | 2     |
| $B_3$   | 1     | 4         | 3     | 2     |
| $B_4$   | 1     | 4         | 3     | 2     |
| $B_5$   | 1     | 4         | 3     | 2     |
| Average | 1.67  | 4.0       | 2.5   | 1.83  |

The statistic $F_F$ is larger than the critical value of $F_\alpha(k-1, N-1)$, it means the null hypothesis is rejected under the Friedman test. Then the Bonferroni-Dunn test can be used to further explore which algorithm is better in the statistical term. If the average level of distance exceeds the critical distance $CD_\alpha$, then the performance of the two algorithms will be significantly different. The critical distance $CD_\alpha$ is denoted as

$$CD_\alpha = q_\alpha \sqrt{\frac{k(k+1)}{6N}},$$

where $q_\alpha$ is a critical value calculated by the qtukey function in r and $\alpha$ is the significance level.

Example 43 (Continued From Example 40): We have

$$G_3(B_1) \approx 0.9966, \ G_3(B_2) \approx 0.9356, \ G_3(B_3) \approx 0.8235, \ G_3(B_4) \approx 0.6652, \ G_3(B_5) \approx 0.4450, \ G_3(B_6) \approx 0.3383;$$

$$(E_r)_3(B_1) \approx 33.0555, \ (E_r)_3(B_2) \approx 30.1812, \ (E_r)_3(B_3) \approx 25.0493, \ (E_r)_3(B_4) \approx 17.7807, \ (E_r)_3(B_5) \approx 9.6253, \ (E_r)_3(B_6) \approx 5.9907;$$

$$H_3(B_1) \approx 0.0495, \ H_3(B_2) \approx 0.8986, \ H_3(B_3) \approx 2.3055, \ H_3(B_4) \approx 3.9846, \ H_3(B_5) \approx 5.1579, \ H_3(B_6) \approx 5.2473;$$

$$E_3(B_1) \approx 0.0343, \ E_3(B_2) \approx 0.6026, \ E_3(B_3) \approx 1.4513, \ E_3(B_4) \approx 2.2483, \ E_3(B_5) \approx 2.4441, \ E_3(B_6) \approx 2.2183.$$

Below, we view the four uncertainty measurements for ISVIS as four algorithms and demonstrate the statistical significance by using Friedman test and Bonferroni-Dunn test.

1. We give the ranking of the four measurements with six subsystems, respectively (see TABLE 5).
2. We conduct Friedman test to investigate whether the performance of the four measurements are significantly different. Under the four measurements and the 6 subsystems, $F_F$ follows the Fisher distribution with 3 and 15 degrees of freedom. Note that the critical value $F_{0.05}(3, 15)$ is 3.287, and $F_F = 10.517$. Obviously, the value of $F_F$ is larger than the value of $F_{0.05}(3, 15)$. This means that at the significant level $\alpha = 0.05$, it is evidence to reject the null hypothesis, which means that the four uncertainty measurements are different in the statistical significance.
3. To further show the significant differences of the four measurements, Bonferroni-Dunn test is introduced. For $\alpha = 0.05$, we can easily calculate the corresponding critical distance $CD_\alpha = 2.569 \sqrt{\frac{3(3+1)}{6 \times 6}} = 1.915$. FIGURE 4 shows the results with $\alpha = 0.05$ on the four measurements. The line segments in FIGURE 4 indicate the average ranking of the four measurements. The line segments in FIGURE 4 carves out the scope of $CD_\alpha$. If the two roots partially overlap on the y-axis, then there is no significant difference between these two uncertainty measurements.

4. From FIGURE 4, the following results are obtained:
   1. a) The performance of $G_3$ is statistically different from the performance of $(E_r)_3$;
   b) The performance of $E_3$ is statistically different from the performance of $(E_r)_3$.
   2. a) There is no significant difference among $G_3$, $E_3$ and $H_3$;
   b) There is no significant difference between $H_3$ and $E_3$.

VIII. OPTIMAL SELECTION OF SUBSYSTEMS BASED ON UNCERTAINTY MEASURES

In the above section, we use relationships between two ISVIS to study uncertainty measures, which naturally causes a problem. When uncertainty measure reaches the optimal value (i.e. the maximum or minimum value)? How to determine the corresponding subsystem (we call it the optimal system)? In this section, the optimal selection of subsystems based on $\delta$-information granulation and $\delta$-information amount is obtained.

Definition 44: Let $(U, A)$ be an ISVIS. Given $\delta \in (0, 1]$. (1) If there exists $B_1 \subseteq A$ such that $G_3(B_1) = \max\{G_3(B) : B \subseteq A\}$, then $(U, B_1)$ is called a maximum subsystem in $(U, A)$ based on $\delta$-information granulation;
   (2) If there exists $B_2 \subseteq A$ such that $G_3(B_2) = \min\{G_3(B) : B \subseteq A\}$, then $(U, B_2)$ is called a minimum subsystem in $(U, A)$ based on $\delta$-information granulation.

The maximum subsystem and minimum subsystem in $(U, A)$ are collectively called the optimal subsystems based on $\delta$-information granulation.

Theorem 45: Let $(U, A)$ be an ISVIS. Given $\delta \in (0, 1]$. (1) If there exists $a_0 \in A$ such that $G_3([a_0]) = \max\{G_3([a]) : a \in A\}$, then $(U, [a_0])$ is a maximum subsystem in $(U, A)$ based on $\delta$-information granulation;
(2) \((U, A)\) is a minimum subsystem in \((U, A)\) based on \(\delta\)-information granulation.

Proof: (1) By Theorem 28,
\[ \max \{G_\delta(B) : B \subseteq A\} = \max \{G_\delta(\{a\}) : a \in A\}. \]

Note that \(G_\delta(\{a_0\}) = \max \{G_\delta(\{a\}) : a \in A\}. \) Then
\[ \max \{G_\delta(B) : B \subseteq A\} = G_\delta(\{a_0\}). \]

Thus \((U, \{a_0\})\) is a maximum subsystem in \((U, A)\) based on \(\delta\)-information granulation.

(2) By Theorem 28, \(\forall B \subseteq A,\)
\[ G_\delta(B) \leq G_\delta(A). \]

This shows that
\[ G_\delta(A) = \min \{G_\delta(B) : B \subseteq A\}. \]

By Definition 44, \((U, A)\) is a minimum subsystem in \((U, A)\) based on \(\delta\)-information granulation.

Example 46: (Continued from Example 40) Pick \(\delta = \sqrt{0.8}\). Then
\[ G_\delta(\{a_1\}) = 0.9966, G_\delta(\{a_2\}) = 0.9523, G_\delta(\{a_3\}) = 0.9508, \]
\[ G_\delta(\{a_4\}) = 0.9336, G_\delta(\{a_5\}) = 0.8225, G_\delta(\{a_6\}) = 0.8573, \]
\[ G_\delta(\{a_7\}) = 0.3383. \]

Thus, \((U, \{a_1\})\) is a maximum subsystem in \((U, A)\) based on \(\delta\)-information granulation, \((U, A)\) is a minimum subsystem in \((U, A)\) based on \(\delta\)-information granulation.

Definition 47: Let \((U, A)\) be an ISVIS. Given \(\delta \in (0, 1]\),
(1) If there exists \(B_1 \subseteq A\) such that \(E_\delta(B_1) = \max \{E_\delta(B) : B \subseteq A\}\), then \((U, B_1)\) is called a maximum subsystem in \((U, A)\) based on \(\delta\)-information amount;
(2) If there exists \(B_2 \subseteq A\) such that \(E_\delta(B_2) = \min \{E_\delta(B) : B \subseteq A\}\), then \((U, B_2)\) is called a minimum subsystem in \((U, A)\) based on \(\delta\)-information amount.

The maximum subsystem and minimum subsystem in \((U, A)\) based on \(\delta\)-information amount are collectively called the optimal subsystems based on \(\delta\)-information amount.

Example 48: (Continued from Example 40) Pick \(\delta = \sqrt{0.8}\). Then the results are obtained by calculating as follows:
\[ E_\delta(\{a_1\}) = \min \{E_\delta(B) : B \subseteq A\} = 0.0343, \]
\[ E_\delta(\{a_1, a_3, a_5, a_6\}) = \max \{E_\delta(B) : B \subseteq A\} = 2.4664. \]

Thus, \((U, \{a_1\})\) is a minimum subsystem in \((U, A)\) based on \(\delta\)-information amount, \((U, \{a_1, a_3, a_5, a_6\})\) is a maximum subsystem in \((U, A)\) based on \(\delta\)-information amount.

IX. CONCLUSION

This article has measured the uncertainty of an ISVIS by means of Gaussian kernel and given the optimal selection of subsystems. Relationships between ISVISs have been investigated. Four tools of measuring the uncertainty of an ISVIS have been proposed. Effectiveness analysis about the proposed measures has been done from the angle of statistics. Based on \(\delta\)-information granulation and \(\delta\)-information amount, the optimal selection of subsystems has been given. In the future, we will examine applications of the proposed measures for an ISVIS.

REFERENCES

[1] O. J. Dunn, “Multiple comparisons among means,” J. Amer. Stat. Assoc., vol. 56, no. 293, pp. 52–64, Mar. 1961.
[2] J. Dai and H. Tian, “Entropy measures and granularity measures for set-valued information systems,” Inf. Sci., vol. 240, pp. 72–82, Aug. 2013.
[3] J. Dai, W. Wang, and J.-S. Mi, “Uncertainty measurement for interval-valued information systems,” Inf. Sci., vol. 251, pp. 63–78, Dec. 2013.
[4] J. Dai, W. Wang, Q. Xu, and H. Tian, “Uncertainty measurement for interval-valued decision systems based on extended conditional entropy,” Knowl.-Based Syst., vol. 27, pp. 443–450, Mar. 2012.
[5] J. Dai, B. Wei, X. Zhang, and Q. Zhang, “Uncertainty measurement for incomplete interval-valued information systems based on \(\alpha\)-weak similarity,” Knowl.-Based Syst., vol. 136, pp. 159–171, Nov. 2017.
[6] M. Friedman, “A comparison of alternative tests of significance for the problem of \(m\) rankings,” Ann. Math. Statist., vol. 11, no. 1, pp. 86–92, Mar. 1940.
[7] G. Facchinetti, R. G. Ricci, and S. Muzzio, “Note on ranking fuzzy triangular numbers,” Int. J. Intell. Syst., vol. 13, no. 7, pp. 613–622, Jul. 1998.
[8] B. Huang, C.-X. Guo, H.-X. Li, G.-F. Feng, and X.-Z. Zhou, “Hierarchical structures and uncertainty measures for intuitionistic fuzzy approximation space,” Inf. Sci., vol. 336, pp. 92–114, Apr. 2016.
[9] B. Huang, Y.-L. Zhuang, and H.-X. Li, “Information granulation and uncertainty measures in interval-valued intuitionistic fuzzy information systems,” Eur. J. Oper. Res., vol. 231, no. 1, pp. 162–170, Nov. 2013.
[10] Y. Leung, M. M. Fischer, W. Z. Wu, and J. S. Mi, “A rough set approach for the discovery of classification rules in set-valued information systems,” Int. J. Approx. Reasoning, vol. 47, pp. 233–246, Feb. 2008.
[11] J. Liang and Y. Qian, “Information granules and entropy theory in information systems,” Sci. China Ser. F, Inf. Sci., vol. 51, no. 10, pp. 1427–1444, Oct. 2008.
[12] J. Liang, Z. Shi, D. Li, and M. J. Wierman, “Entropy information, rough entropy and knowledge granulation in incomplete information systems,” Int. J. Gen. Syst., vol. 35, no. 6, pp. 641–654, Dec. 2006.
[13] Z. Li, P. Zhang, X. Ge, N. Xie, and G. Zhang, “Uncertainty measurement for a covering information system,” Soft Comput., vol. 23, no. 14, pp. 5307–5325, Jul. 2019.
[14] Z. Li, P. Zhang, X. Ge, N. Xie, G. Zhang, and C. Wen, “Uncertainty measurement for a fuzzy relation information system,” IEEE Trans. Fuzzy Syst., vol. 27, no. 12, pp. 2338–2352, Dec. 2019.
[15] B. Moser, “On the \(T\)-transitivity of kernels,” Fuzzy Sets Syst., vol. 157, no. 13, pp. 1787–1796, Jul. 2006.
[16] B. Moser, “On representing and generating kernels by fuzzy equivalence relations,” J. Mach. Learn. Res., vol. 7, pp. 2603–2630, Dec. 2006.
[17] Z. Pawlak, “Rough sets,” Int. J. Comput. Inf. Sci., vol. 11, pp. 341–356, Oct. 1982.
[18] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data. Norwell, MA, USA: Kluwer, 1991.
[19] Z. Pawlak and A. Skowron, “Rudiments of rough sets,” Int. J. Approx. Reason., vol. 177, no. 1, pp. 3–27, Jan. 2007.
[20] Z. Pawlak and A. Skowron, “Rough sets: Some extensions,” Int. J. Approx. Reason., vol. 177, no. 1, pp. 28–40, Jan. 2007.
[21] Z. Pawlak and A. Skowron, “Rough sets and Boolean reasoning,” Int. J. Approx. Reason., vol. 177, no. 1, pp. 41–73, Jan. 2007.
[22] Y. H. Qian, J. Y. Liang, and C. Y. Dang, “On the \(T\)-transitivity of kernels,” Fuzzy Sets Syst., vol. 157, no. 6, pp. 1591–1712, Jun. 2006.
[23] Y. Qian, J. Liang, W.-Z.-Z. Wu, and C. Dang, “Information granularity in fuzzy binary GrC model,” IEEE Trans. Fuzzy Syst., vol. 19, no. 2, pp. 253–264, Apr. 2011.
[24] C. E. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., vol. 27, pp. 737–842, Jan. 1948.
[25] C. Wang, Y. Huang, M. Shao, Q. Hu, and D. Chen, “Feature selection based on neighborhood self-information,” IEEE Trans. Cybern., vol. 50, no. 9, pp. 4031–4042, Sep. 2020.
[26] C. Wang, Y. Wang, M. Shao, Y. Qian, and D. Chen, “Fuzzy rough attribute reduction for categorical data,” IEEE Trans. Fuzzy Syst., vol. 28, no. 5, pp. 818–830, May 2020.
[27] N. Xie, M. Liu, Z. Li, and G. Zhang, “New measures of uncertainty for an interval-valued information system,” Inf. Sci., vol. 470, pp. 156–174, Jan. 2019.

[28] X. Wei-hua, Z. Xiao-yan, and Z. Wen-xiu, “Knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems,” Appl. Soft Comput., vol. 9, no. 4, pp. 1244–1251, Sep. 2009.

[29] Y. Y. Yao and N. Noroozi, “A unified framework for set-based computations,” in Proc. 3rd Int. Workshop Rough Sets Soft Comput., 1994, pp. 10–12.

[30] Y. Y. Yao and X. Li, “Comparison of rough-set and interval-set models for uncertain reasoning,” Fundamenta Informaticae, vol. 27, no. 2, pp. 289–298, 1996.

[31] X. B. Yang, D. J. Yu, J. Y. Yang, and L. H. Wei, “Dominance-based rough set approach to incomplete set-valued information system,” Data Knowl. Eng., vol. 68, pp. 1331–1347, Nov. 2009.

[32] C. Virgil Negoita, “Fuzzy sets,” Fuzzy Sets Syst., vol. 133, no. 2, p. 275, Jan. 2003.

[33] L. A. Zadeh, “Fuzzy logic equals computing with words,” IEEE Trans. Fuzzy Syst., vol. 4, no. 2, pp. 103–111, May 1996.

[34] L. A. Zadeh, “Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic,” Fuzzy Sets Syst., vol. 90, no. 2, pp. 111–127, Sep. 1997.

[35] L. A. Zadeh, “A new direction in AI-toward a computational theory of perceptions,” AI Mag., vol. 22, no. 1, pp. 73–84, 2001.

[36] L. A. Zadeh, “Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems,” Soft Comput. A. Fusion Found., Methodologies Appl., vol. 2, no. 1, pp. 23–25, Apr. 1998.

[37] G. Zhang, Z. Li, W.-Z. Wu, X. Liu, and N. Xie, “Information structures and uncertainty measures in a fully fuzzy information system,” Int. J. Approx. Reasoning, vol. 101, pp. 119–149, Oct. 2018.

[38] A. Zeng, T. Li, D. Liu, J. Zhang, and H. Chen, “A fuzzy rough set approach for incremental feature selection on hybrid information systems,” Fuzzy Sets Syst., vol. 258, pp. 39–60, Jan. 2015.

[39] W. X. Zhang and G. F. Qiu, Uncertain Decision Making Based on Rough Sets. Beijing, China: Tsinghua Univ., 2005.

LIJUN CHEN received the M.Sc. degree in statistics from Hohai University, Nanjing, China, in 2016. She is currently an Instructor with the School of Mathematics and Statistics, Yulin Normal University. Her main research interests include rough set theory and information systems.

SHIMIN LIAO received the M.Sc. degree in mathematics from Guangxi University for Nationalities, Nanning, China, in 2020. His main research interests include rough set theory and information systems.

NINGXIN XIE received the M.Sc. degree in computer from Guangxi University, Nanning, China, in 2001. He is currently a Professor with the School of Artificial Intelligence, Guangxi University for Nationalities, Nanning. His research interests include rough set theory, fuzzy set theory, and information systems.

ZHAOWEN LI received the M.Sc. degree in mathematics from Guangxi University, Nanning, China, in 1988, and the Ph.D. degree in mathematics from Hunan University, Changsha, China, in 2008. He is currently a Professor with the Key Laboratory of Complex System Optimization and Big Data Processing, Department of Guangxi Education, Yulin Normal University. His research interests include topology and its applications, rough set theory, fuzzy set theory, and information systems.

GANGQIANG ZHANG received the M.Sc. degree in software engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2006. He is currently an Associate Professor with the School of Artificial Intelligence, Guangxi University for Nationalities. His main research interests include rough set theory, fuzzy set theory, and information systems.

CHING-FENG WEN received the M.Sc. degree in mathematics from Kaohsiung Normal University, Kaohsiung, Taiwan, in 1995, and the Ph.D. degree in mathematics from National Cheng Kung University, Tainan, Taiwan, in 2003. He is currently a Professor and the Director with the Center for Fundamental Science, Kaohsiung Medical University, Kaohsiung, where he is also the Research Center for Nonlinear Analysis and Optimization. He is also a Research Fellow with the Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung. His research interests include nonlinear analysis, optimization, fuzzy set theory, and information systems.

***