IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 24, 2020, accepted November 16, 2020, date of publication November 23, 2020, date of current version December 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3039778

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

LIJUN CHEN[1], SHIMIN LIAO[2], NINGXIN XIE[2], ZHAOWEN LI[3], GANGQIANG ZHANG[2], AND CHING-FENG WEN[4,5]

[1]School of Mathematics and Statistics, Yulin Normal University, Yulin 537000, China
[2]School of Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, China
[3]Key Laboratory of Complex System Optimization and Big Data Processing, Department of Guangxi Education, Yulin Normal University, Yulin 537000, China
[4]Center for Fundamental Science, Kaohsiung Medical University, Kaohsiung 80708, Taiwan
[5]Research 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 δ-information granulation and δ-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 θ-rough degree in IVISs on the foundation of θ-similarity entropy; Dai *et al.* [5] explored entropy and granularity measures in SVISs; Huang *et al.* [8]

The associate editor coordinating the review of this manuscript and approving it for publication was Farhana Jabeen Jabeen.

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 in an ISVIS. In the fourth section, we study fuzzy T_{cos}-equivalence relation by means of Gaussian kernel. In the fifth section, we research relationships between two ISVISs and display inclusion degree of IIVISs. In the sixth section, we measure of uncertainty for a given ISVIS. In the seventh section, we do effectiveness analysis from three aspects. In the eighth section, we obtain the optimal selection of subsystems based on the proposed measures. In the ninth section, we summarize the article.

II. PRELIMINARIES

In this section, we briefly recall some concepts about fuzzy sets, fuzzy relations and ISVISs.

U denotes a non-empty finite set and I expresses [0, 1] in this article.

Put

$$U = \{u_1, u_2, \ldots, u_n\}.$$

A. FUZZY SETS AND FUZZY RELATIONS

If F is a mapping defined by $F : U \to I$, then F is a fuzzy set on U.

In this article, I^U indicates the family of all fuzzy sets on U. $\bar{a}$ denotes the constant fuzzy set on U for each $a \in I$. $|F| = \sum_{u\in U} F(u)$ means the cardinality of $F \in I^U$.

If R is a fuzzy set in $U \times U$, then R is called a fuzzy relation on U.

In this article, $I^{U\times U}$ denotes the family of all fuzzy relations on U.

Given $F \in I^U$ and $u \in U$. Then $F(u)$ indicates the degree that u belongs to F. Similarly, given $R \in I^{U\times U}$ and $u, v \in U$. Then $R(u, v)$ indicates the degree that (u, v) belongs to R. Thus $R(u, v)$ can be regarded as the degree of similarity between u and v. In general, R is denoted by the following matrix:

$$M(R) = (R(u_i, u_j))_{n\times n}.$$

Suppose $R \in I^{U\times U}$. R is said to be a fuzzy identity relation on U if $M(R)$ is an identity matrix, we indicate $R = \triangle$; R is said to be a fuzzy zero relation on U if $M(R) = 0$, we indicate $R = o$; R is said to be a fuzzy universal relation on U if $R(u_i, u_j) = 1$, we indicate $R = \omega$.

Suppose $R \in I^{U\times U}$ and $u \in U$. Then a fuzzy set $[u]^R$ is defined as follows:

$$[u]^R(v) = R(u, v), \quad \forall`v \in U.$$

$[u]^R$ can be viewed as the fuzzy neighborhood of the point u on U under R.

Definition 1 [16]: A function $T : I^2 \to I$ is called a t-norm, if meets the conditions as follows:

(1) $T(a, b) = T(b, a)$ *(Commutativity);*

(2) $T(T(a, b), c) = T(a, T(b, c))$ *(Associativity);*

(3) $a \leq c, b \leqslant d = T(a, b) \leqslant T(c, d)$ *(Monotonicity);*

(4) $T(a, 1) = a$ *(Boundary condition).*

Definition 2 [38]: Let T be the t-norm. Suppose $R \in I^{U\times U}$. Then R is a T-fuzzy equivalence relation on U if it meets the following conditions:

(1) $R(u, u) = 1$ *(Reflexivity);*

(2) $R(u, v) = R(v, u)$ *(Symmetry);*

(3) $T(R(u, v), R(v, w)) \leqslant R(u, w)$ *(T-transitivity).*

Proposition 3 [15]: Assume that $f : U \times U \to I$ satisfies $f(u, u) = 1$ for all $u \in U$. Then $u, v, w \in U$,

$$T_{cos}(f(u, v), f(v, w)) \leq f(u, w).$$

Corollary 4: Given $R \in I^{U\times U}$. If R is reflexive, then R is T_{cos}-transitive.

B. ISVISs

Definition 5 [18]: Consider that U is an object set, A is an attribute set, U and A are finite sets. Then the pair (U, A) is called an information system (IS), if each attribute $a \in A$ determines an information function $a : U \to V_a$, where $V_a = \{a(u) : u \in U\}$ is the set of information function values of the attribute a.

Let (U, A) be an IS, given $P \subseteq A$, then an equivalence relation on U can be defined as

$$ind(P) = \{(u, v) \in U \times U : \quad \forall a \in P,\ a(u) = a(v)\}.$$

Definition 6 [18]: Assume that (U, A) is an IS. Then the pair (U, A) is said to be an incomplete information system (IIS), if there are $u \in U$ and $a \in A$ then $a(u)$ is missing.

If (U, A) is an IIS. Given $P \subseteq A$. Then a tolerance relation on U can be defined as

$$sim(P) = \{(u, v) \in U \times U : \quad \forall a \in P,\ a(u) = a(v) or a(u) = * or a(v) = *\},$$

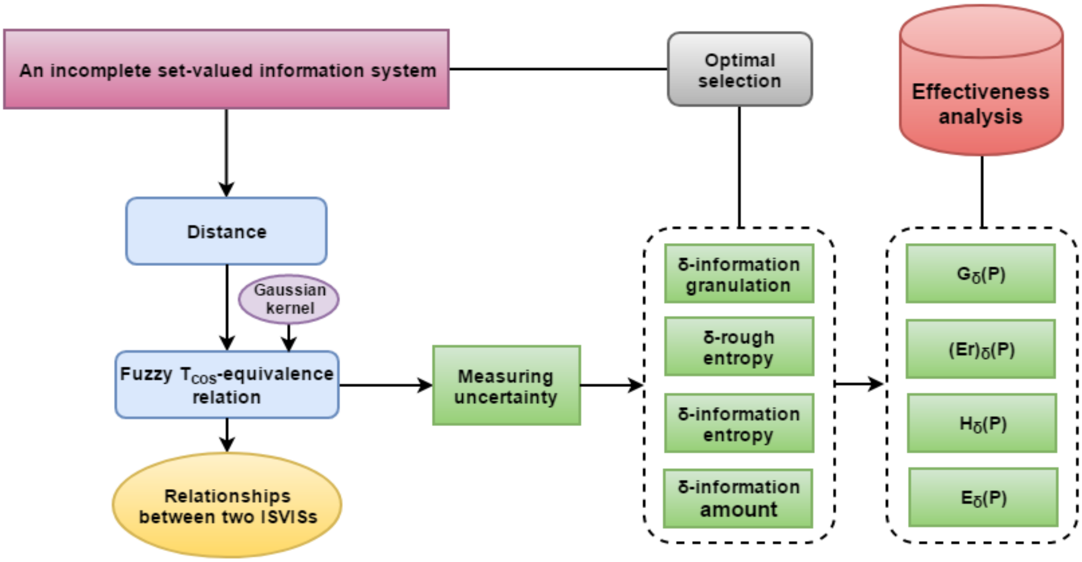

FIGURE 1. The work process of the article.

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

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

where $s(a(u), a(v)) = \frac{|a(u) \bigcap a(v)|}{|a(u) \bigcup a(v)|}$ 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 $A = \{a_1, a_2, \ldots, a_6\}$.

$$V_{a_1}^* = V_{a_1} = \{\{L, M, N\}, \{M, N\}\},\ V_{a_2}^* = \{\{L, M\}, \{L, N\}\},$$
$$V_{a_3}^* = \{\{L, M\}, \{M, N\}\},\ V_{a_4}^* = \{\{L, N\}, \{M, N\}, \{N\}\},$$
$$V_{a_5}^* = \{\emptyset, \{L, M, N\}\},\ V_{a_6}^* = \{\{L, M\}, \{M\}\},$$

III. DISTANCE BETWEEN TWO OBJECTS IN AN ISVIS

Definition 10: 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

$$d(a(u), a(v)) = \begin{cases} 0, & u = v; \\ 1 - \dfrac{1}{|V_a^*|^2}, & u \neq v,\ a(u) = *,\ a(v) = *; \\ 1 - \dfrac{1}{|V_a^*|}, & u \neq v,\ a(u) \neq *,\ a(v) = *; \\ 1 - \dfrac{1}{|V_a^*|}, & u \neq v,\ a(u) = *,\ a(v) \neq *; \\ 0, & u \neq v,\ a(u) \neq *, \\ & a(v) \neq *,\ a(u) = a(v); \\ 1 - \dfrac{|a(u) \bigcap a(v)|}{|a(u) \bigcup a(v)|}, & u \neq v,\ a(u) \neq *,\ a(v) \neq *, \\ & a(u) \neq a(v). \end{cases}$$

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

Definition 11: 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. □

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

$$d(a_1(u_1), a_1(u_3)) = 1 - \frac{|a_1(u) \bigcap a_1(v)|}{|a_1(u) \bigcup a_1(v)|} = 1 - \frac{2}{3} \approx 0.3333;$$

TABLE 1. An ISVIS.

	a_1	a_2	a_3	a_4	a_5	a_6
u_1	$\{L,M,N\}$	$\{L,M\}$	$*$	$\{L,N\}$	$\emptyset$	$\{L,M\}$
u_2	$\{L,M,N\}$	$*$	$\{M,N\}$	$\{N\}$	$\{L,M,N\}$	$\{N\}$
u_3	$\{M,N\}$	$\{L,N\}$	$\{L,M\}$	$\{L,N\}$	$\emptyset$	$\{N\}$
u_4	$\{L,M,N\}$	$\{L,M\}$	$\{M,N\}$	$*$	$\emptyset$	$\{L,M\}$
u_5	$\{M,N\}$	$\{L,M\}$	$\{L,M\}$	$\{M,N\}$	$\{L,M,N\}$	$*$
u_6	$\{M,N\}$	$\{L,N\}$	$\{M,N\}$	$*$	$\{L,M,N\}$	$\{N\}$
u_7	$\{L,M,N\}$	$\{L,M\}$	$\{M,N\}$	$\{M,N\}$	$*$	$\{N\}$
u_8	$\{M,N\}$	$\{L,M\}$	$*$	$\{N\}$	$\emptyset$	$*$
u_9	$\{M,N\}$	$*$	$\{L,M\}$	$\{L,N\}$	$\emptyset$	$\{L,M\}$
u_{10}	$\{L,M,N\}$	$\{L,N\}$	$*$	$\{M,N\}$	$\{L,M,N\}$	$\{N\}$

$$d(a_2(u_1), a_2(u_3)) = 1 - \frac{|a_2(u)\bigcap a_2(v)|}{|a_2(u)\bigcup a_2(v)|} = 1 - \frac{1}{3} \approx 0.6667;$$
$$d(a_3(u_1), a_3(u_3)) = 1 - \frac{|a_3(u)\bigcap a_3(v)|}{|a_3(u)\bigcup a_3(v)|} = 1 - \frac{1}{2} = 0.5;$$
$$d(a_4(u_1), a_4(u_3)) = 1 - \frac{|a_4(u)\bigcap a_4(v)|}{|a_4(u)\bigcup a_4(v)|} = 1 - 1 = 0.$$

Then

$$d_A(u_1, u_3) = \sqrt{\sum_{a\in A} d^2(a(u_1), a(u_3))} \approx \sqrt{0.3333^2 + 0.6667^2 + 0.5^2 + 0^2} \approx 0.8975.$$

IV. FUZZY T_{cos}-EQUIVALENCE RELATION BASED ON GAUSSIAN KERNEL IN AN ISVIS

In this section, the fuzzy T_{cos}-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, δ 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

$$R_P^G(\delta)(u_i, u_j) = exp(-\frac{d_P^2(u_i, u_j)}{2\delta^2}),$$
$$M(R_P^G(\delta)) = (R_P^G(\delta)(u_i, u_j))_{n\times n}.$$

Then $M(R_P^G(\delta))$ is called the Gaussian kernel matric of the subsystem (U, P) with respect to δ.

Theorem 15: Let (U, A) be an ISVIS. Given $P \subseteq A$ and $\delta \in (0, 1]$. Then $R_P^G(\delta)$ is a T_{cos}-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 T_{cos}-equivalence relation induced by the subsystem (U, P) with respect to δ.

For any $u \in U$, a fuzzy set $[u]^{R_P^G(\delta)}$ is defined as follows:

$$[u]^{R_P^G(\delta)}(v) = R_P^G(\delta)(u, v), \quad \forall v \in U.$$

Algorithm 1 The T_{cos}-Equivalence Relation

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

for $i = 1$; $i \leq |U|$; $i++$ **do**
  for $j = |U| - 1$; $j > i$; $j--$ **do**
    $d(a(u_i), a(u_j)) = 0$;
    for *each* $a \in P$ **do**
      $d(a(u_i), a(u_j) += d^2(a(u_i), a(u_j))$;
    end
    Compute $d_P(u_i, u_j) = \sqrt{\sum_{a\in P} d(a(u_i), a(u_j))}$;
    $R_P^G(\delta)(u_i, u_j) = exp(-\frac{d_P^2(u_i,u_j)}{2\times\delta^2})$;
    $M(R_P^G(\delta)) = (R_P^G(\delta)(u_i, u_j))_{n\times n}$;
    Obtain $R_P^G(\delta)$.
  end
end

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

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

Similarly,

Then $R_A^G(\delta)$ is the T_{cos}-equivalence relation induced by the system (U, A) with respect to δ.

Given $P \subseteq A$ and $\delta \in (0, 1]$. Then an algorithm on a T_{cos}-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 IIVISs below.

Definition 18: Let (U, P) and (U, Q) be two ISVISs. Given $\delta \in (0, 1]$. If for any $u \in U$, $[u]^{R_P^G(\delta)} = [u]^{R_Q^G(\delta)}$, then (U, P) and (U, Q) are called to be equivalent with respect to δ. 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 δ, if for any $u \in U$, $[u]^{R_P^G(\delta)} \subseteq [u]^{R_Q^G(\delta)}$, we write $(U, P) \preceq_\delta (U, Q)$; (U, Q) is known as to depend strictly on (U, P) with respect to δ, if $(U, P) \preceq_\delta (U, Q)$ and $(U, P) \not\approx_\delta (U, Q)$, we write $(U, P) \prec_\delta (U, Q)$.

(2) (U, Q) is called to depend partially on (U, P) with respect to δ, if exists $u \in U$, $[u]^{R_P^G(\delta)} \subseteq [u]^{R_Q^G(\delta)}$, 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 $u \in U$, $[u]^{R_P^G(\delta)} \not\subseteq [u]^{R_Q^G(\delta)}$, we write $(U, Q) \bowtie_\delta (U, P)$.

Clearly, the following conclusions can be obtained.

$$(U, Q) \approx_\delta (U, P) \Leftrightarrow (U, Q) \preceq_\delta (U, P) \text{ and}$$

$$M(R^G_{\{a_1\}}(\delta)) = \begin{pmatrix}
1.000 & 1.000 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 \\
1.000 & 1.000 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 \\
0.993 & 0.993 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 \\
1.000 & 1.000 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 \\
0.993 & 0.993 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 \\
0.993 & 0.993 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 \\
1.000 & 1.000 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 \\
0.993 & 0.993 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 \\
0.993 & 0.993 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 & 1.000 & 1.000 & 0.993 \\
1.000 & 1.000 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000 & 0.993 & 0.993 & 1.000
\end{pmatrix},$$

$$M(R^G_{\{a_2\}}(\delta)) = \begin{pmatrix}
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000 \\
0.966 & 1.000 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.838 & 0.966 \\
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000 \\
0.895 & 0.966 & 0.895 & 1.000 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 \\
0.895 & 0.966 & 0.895 & 1.000 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 \\
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000 \\
0.895 & 0.966 & 0.895 & 1.000 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 \\
0.895 & 0.966 & 0.895 & 1.000 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 \\
0.966 & 0.838 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 1.000 & 0.966 \\
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000
\end{pmatrix},$$

$$M(R^G_{\{a_3\}}(\delta)) = \begin{pmatrix}
1.000 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.838 & 0.966 & 0.838 \\
0.966 & 1.000 & 0.895 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 & 0.966 \\
0.966 & 0.895 & 1.000 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000 & 0.966 \\
0.966 & 1.000 & 0.895 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 & 0.966 \\
0.966 & 0.895 & 1.000 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000 & 0.966 \\
0.966 & 1.000 & 0.895 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 & 0.966 \\
0.966 & 1.000 & 0.895 & 1.000 & 0.895 & 1.000 & 1.000 & 0.966 & 0.895 & 0.966 \\
0.838 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 1.000 & 0.966 & 0.838 \\
0.966 & 0.895 & 1.000 & 0.895 & 1.000 & 0.895 & 0.895 & 0.966 & 1.000 & 0.966 \\
0.838 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.838 & 0.966 & 1.000
\end{pmatrix},$$

$$M(R^G_{\{a_4\}}(\delta)) = \begin{pmatrix}
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 0.895 & 0.895 & 0.966 & 1.000 & 0.895 \\
0.966 & 1.000 & 0.966 & 0.895 & 0.966 & 0.895 & 0.966 & 1.000 & 0.966 & 0.966 \\
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 0.895 & 0.895 & 0.966 & 1.000 & 0.895 \\
0.895 & 0.895 & 0.895 & 1.000 & 0.895 & 0.705 & 0.895 & 0.895 & 0.895 & 0.895 \\
0.895 & 0.966 & 0.895 & 0.895 & 1.000 & 0.895 & 1.000 & 0.966 & 0.895 & 1.000 \\
0.895 & 0.895 & 0.895 & 0.705 & 0.895 & 1.000 & 0.895 & 0.895 & 0.895 & 0.895 \\
0.895 & 0.966 & 0.895 & 0.895 & 1.000 & 0.895 & 1.000 & 0.966 & 0.895 & 1.000 \\
0.966 & 1.000 & 0.966 & 0.895 & 0.966 & 0.895 & 0.966 & 1.000 & 0.966 & 0.966 \\
1.000 & 0.966 & 1.000 & 0.895 & 0.895 & 0.895 & 0.895 & 0.966 & 1.000 & 0.895 \\
0.895 & 0.966 & 0.895 & 0.895 & 1.000 & 0.895 & 1.000 & 0.966 & 0.895 & 1.000
\end{pmatrix},$$

$$(U,P) \preceq_\delta (U,Q),$$
$$(U,Q) \preceq_\delta (U,P) \Rightarrow (U,Q) \sqsubseteq_\delta (U,P),$$
$$(U,Q) \prec_\delta (U,P) \Rightarrow (U,Q) \sqsubset_\delta (U,P).$$

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 $D_\delta : \Sigma_{(U,A)} \times \Sigma_{(U,A)} \to [0,1]$ is said to be the inclusion degree on $\Sigma_{(U,A)}$ with respect to δ, if it satisfies the following conditions: for any $(U,O), (U,P), (U,Q) \in \Sigma_{(U,A)}$,

(1) $0 \le D_\delta((U,P)/(U,O)) \le 1$;

(2) $(U,O) \preceq_\delta (U,P)$ implies $D_\delta((U,P)/(U,O)) = 1$;

(3) $(U,O) \preceq_\delta (U,P) \preceq_\delta (U,Q)$ implies $D_\delta((U,O)/(U,Q)) \le D_\delta((U,O)/(U,P))$.

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

$$D_\delta((U,Q)/(U,P)) = \sum_{l=1}^{n} \frac{|[u_l]^{R_Q^G(\delta)}|}{\sum_{i=1}^{n} |[u_i]^{R_Q^G(\delta)}|} \chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}),$$

where

$$\chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}) = \begin{cases} 1, & \text{if } [u_l]^{R_P^G(\delta)} \subseteq [u_l]^{R_Q^G(\delta)}, \\ 0, & \text{if } [u_l]^{R_P^G(\delta)} \not\subseteq [u_l]^{R_Q^G(\delta)}. \end{cases}$$

Proposition 23: D_δ in Definition 22 is the inclusion degree under Definition 21.

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

(1) Obviously, $0 \le D_\delta((U,Q)/(U,P)) \le 1$.

(2) Suppose $(U,P) \preceq_\delta (U,Q)$. Then, by Definition 19, $[u]^{R_P^G(\delta)} \subseteq [u]^{R_Q^G(\delta)}$. Thus, for each l, $[u_l]^{R_P^G(\delta)} \subseteq [u_l]^{R_Q^G(\delta)}$. This result implies that

$$\textit{for each } l,\ \chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}) = 1.$$

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

(3) Suppose $(U,P) \preceq_\delta (U,Q) \preceq_\delta (U,O)$. Then, by Definition 19, $[u]^{R_P^G(\delta)} \subseteq [u]^{R_Q^G(\delta)} \subseteq [u]^{R_O^G(\delta)}$. Thus, for each l, $[u_l]^{R_P^G(\delta)} \subseteq [u_l]^{R_Q^G(\delta)} \subseteq [u_l]^{R_O^G(\delta)}$.

By Definition 22,

$$D_\delta((U,P)/(U,O)) = \sum_{l=1}^{n} \frac{|[u_l]^{R_P^G(\delta)}|}{\sum_{i=1}^{n} |[u]^{R_P^G(\delta)}(u_i)|} \chi_{[u_l]^{R_P^G(\delta)}}([u_l]^{R_O^G(\delta)}),$$
$$D_\delta((U,P)/(U,Q)) = \sum_{l=1}^{n} \frac{|[u_l]^{R_P^G(\delta)}|}{\sum_{i=1}^{n} |[u]^{R_P^G(\delta)}(u_i)|} \chi_{[u_l]^{R_P^G(\delta)}}([u_l]^{R_Q^G(\delta)}).$$

If $[u_l]^{R_Q^G(\delta)} \not\subseteq [u_l]^{R_P^G(\delta)}$, then $[u_l]^{R_O^G(\delta)} \not\subseteq [u_l]^{R_P^G(\delta)}$. This result illustrates that

$$\chi_{[u_l]^{R_P^G(\delta)}}([u_l]^{R_Q^G(\delta)}) = 0 \textit{ implies } \chi_{[u_l]^{R_P^G(\delta)}}([u_l]^{R_O^G(\delta)}) = 0.$$

Thus,

$$D_\delta((U,P)/(U,O)) \le D_\delta((U,P)/(U,Q)).$$

$$M(R_{\{a_5\}}^G(\delta)) = \begin{pmatrix} 1.000 & 0.572 & 1.000 & 1.000 & 0.572 & 0.572 & 0.966 & 1.000 & 1.000 & 0.572 \\ 0.572 & 1.000 & 0.572 & 0.572 & 1.000 & 1.000 & 0.966 & 0.572 & 0.572 & 1.000 \\ 1.000 & 0.572 & 1.000 & 1.000 & 0.572 & 0.572 & 0.966 & 1.000 & 1.000 & 0.572 \\ 1.000 & 0.572 & 1.000 & 1.000 & 0.572 & 0.572 & 0.966 & 1.000 & 1.000 & 0.572 \\ 0.572 & 1.000 & 0.572 & 0.572 & 1.000 & 1.000 & 0.966 & 0.572 & 0.572 & 1.000 \\ 0.572 & 1.000 & 0.572 & 0.572 & 1.000 & 1.000 & 0.966 & 0.572 & 0.572 & 1.000 \\ 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 0.966 & 1.000 & 0.966 & 0.966 & 0.966 \\ 1.000 & 0.572 & 1.000 & 1.000 & 0.572 & 0.572 & 0.966 & 1.000 & 1.000 & 0.572 \\ 1.000 & 0.572 & 1.000 & 1.000 & 0.572 & 0.572 & 0.966 & 1.000 & 1.000 & 0.572 \\ 0.572 & 1.000 & 0.572 & 0.572 & 1.000 & 1.000 & 0.966 & 0.572 & 0.572 & 1.000 \end{pmatrix},$$

$$M(R_{\{a_6\}}^G(\delta)) = \begin{pmatrix} 1.000 & 0.572 & 0.572 & 1.000 & 0.966 & 0.572 & 0.572 & 0.966 & 1.000 & 0.572 \\ 0.572 & 1.000 & 1.000 & 0.572 & 0.966 & 1.000 & 1.000 & 0.966 & 0.572 & 1.000 \\ 0.572 & 1.000 & 1.000 & 0.572 & 0.966 & 1.000 & 1.000 & 0.966 & 0.572 & 1.000 \\ 1.000 & 0.572 & 0.572 & 1.000 & 0.966 & 0.572 & 0.572 & 0.966 & 1.000 & 0.572 \\ 0.966 & 0.966 & 0.966 & 0.966 & 1.000 & 0.966 & 0.966 & 0.838 & 0.966 & 0.966 \\ 0.572 & 1.000 & 1.000 & 0.572 & 0.966 & 1.000 & 1.000 & 0.966 & 0.572 & 1.000 \\ 0.572 & 1.000 & 1.000 & 0.572 & 0.966 & 1.000 & 1.000 & 0.966 & 0.572 & 1.000 \\ 0.966 & 0.966 & 0.966 & 0.966 & 0.838 & 0.966 & 0.966 & 1.000 & 0.966 & 0.966 \\ 1.000 & 0.572 & 0.572 & 1.000 & 0.966 & 0.572 & 0.572 & 0.966 & 1.000 & 0.572 \\ 0.572 & 1.000 & 1.000 & 0.572 & 0.966 & 1.000 & 1.000 & 0.966 & 0.572 & 1.000 \end{pmatrix},$$

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* $\delta \in (0,1]$,

(1) $(U,P) \preceq_\delta (U,Q) \Leftrightarrow D_\delta((U,Q)/(U,P)) = 1$;

(2) $(U,P) \bowtie_\delta (U,Q) \Leftrightarrow D_\delta((U,Q)/(U,P)) = 0$;

(3) $(U,P) \sqsubseteq_\delta (U,Q) \Leftrightarrow 0 < D_\delta((U,Q)/(U,P)) \leq 1$.

Proof: (1) "$\Longrightarrow$" is evident. We prove "$\Longleftarrow$". Suppose

$$|[u_l]^{R_Q^G(\delta)}| = q_l, \quad \sum_{l=1}^{n} |[u_l]^{R_Q^G(\delta)}| = q.$$

Then,

$$q = \sum_{l=1}^{n} q_l.$$

Owing to $D_\delta((U,Q)/(U,P)) = 1$, it can be obtained that $\sum_{l=1}^{n} q_l \chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}) = \sum_{l=1}^{n} q_l = q$.

Then,

$$q(1 - \chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)})) = 0.$$

Consequently, $\forall\, l$,

$$1 - \chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}) = 0.$$

Thus, it can be obtained that $\forall\, l$, $[u_l]^{R_P^G(\delta)} \subseteq [u_l]^{R_Q^G(\delta)}$.

By Definition 19, $(U,P) \preceq_\delta (U,Q)$.

(2) "$\Longrightarrow$". Owing to $(U,P) \bowtie_\delta (U,Q)$, it can be obtained that $[u_l]^{R_P^G(\delta)} \nsubseteq [u_l]^{R_Q^G(\delta)}$. Then $\forall\, l$,

$$\chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}) = 0.$$

By Definition 22, $D_\delta((U,Q)/(U,P)) = 0$.

"$\Longleftarrow$". Owing to $D_\delta((U,Q)/(U,P)) = 0$, it can be obtained that $\forall\, l$, $\chi_{[u_l]^{R_Q^G(\delta)}}([u_l]^{R_P^G(\delta)}) = 0$.

Then, $\forall\, l$, $[u_l]^{R_P^G(\delta)} \nsubseteq [u_l]^{R_Q^G(\delta)}$. By Definition 19, $(U,P) \bowtie_\delta (U,Q)$.

(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* $\delta \in (0,1]$. *Suppose that* $G_\delta : 2^A \to (-\infty, +\infty)$ *is a function. Then G is called an* δ*-information granulation function in* (U,A) *with respect to* δ*, if G satisfies the following conditions:*

(1) $\forall\, P \in 2^A$, $G_\delta(P) \geq 0$ *(Non-negativity);*

(2) $\forall\, P, Q \in 2^A$, *if* $(U,P) \approx_\delta (U,Q)$, *then* $G_\delta(P) = G_\delta(Q)$ *(Invariability);*

(3) $\forall\, P, Q \in 2^A$, *if* $(U,P) \prec_\delta (U,Q)$, *then* $G_\delta(P) < G_\delta(Q)$ *(Monotonicity).*

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

$$G_\delta(P) = \frac{1}{n^2}\sum_{i=1}^{n} |[u_i]^{R_P^G(\delta)}|.$$

Proposition 27: Let (U,A) *be an ISVIS. Given* $P \subseteq A$ *and* $\delta \in (0,1]$. *Then*

$$0 \leq G_\delta(P) \leq 1.$$

If $R_P^G(\delta) = o$, *then G reaches the minimum value* 0; *if* $R_P^G(\delta) = \omega$, *then G reaches the maximum value* 1.

Proof: By Definition 26,

$$G_\delta(P) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} R_P^G(u_i, u_j).$$

Obviously, $\forall\, i, j$, $0 \leq R_P^G(u_i, u_j) \leq 1$.

Then

$$0 \leq \sum_{i=1}^{n}\sum_{j=1}^{n} R_P^G(u_i, u_j) \leq n^2.$$

Thus

$$0 \leq G_\delta(P) \leq 1.$$

If $R_P^G(\delta) = o$, then $\forall\, i, j$, $R_P^G(\delta)(u_i, u_j) = 0$ and so $G_\delta(P) = 0$.

If $R_P^G(\delta) = \omega$, then $\forall\, i, j$, $R_P^G(\delta)(u_i, u_j) = 1$ and so $G_\delta(P) = 1$. □

Theorem 28: Let (U,A) *be an ISVIS. Given* $P, Q \subseteq A$ *and* $\delta \in (0,1]$. *If* $(U,P) \prec_\delta (U,Q)$, *then* $G_\delta(P) < G_\delta(Q)$.

Proof: Since $(U,P) \prec_\delta (U,Q)$, we have $(U,P) \preceq_\delta (U,Q)$ and $(U,P) \not\approx_\delta (U,Q)$.

So $\forall\, i$, $[u_i]^{R_P^G(\delta)} \subseteq [u_i]^{R_Q^G(\delta)}$ and $\exists\, i'$, $G_P(u_{i'}) \subsetneq G_Q(u_{i'})$.

Thus $\forall\, i, j$, $G_P(x)(u_j) \leq G_Q(x)(u_j)$ and $\exists\, i', j'$, $[u_{i'}]^{R_P^G(\delta)}(u_{j'}) < [u_{i'}]^{R_Q^G(\delta)}(u_{j'})$.

Hence $\forall\, i, j$,

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

and $\exists\, i', j'$,

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

By Definition 26,

$$G_\delta(P) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} R_P^G(\delta)(u_i, u_j),$$

$$G_\delta(Q) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} R_Q^G(\delta)(u_i, u_j).$$

Hence $G_\delta(P) < G_\delta(Q)$. □

Theorem 29: G_δ *in Definition 26 is an* δ*-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 $\forall\, i$, $[u_i]^{R_P^G(\delta)} = [u_i]^{R_Q^G(\delta)}$.

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

(3) "Monotonicity" follows from Theorem 28. □

B. ENTROPY MEASUREMENTS FOR AN ISVIS

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

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

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

$$-\infty < (E_r)_\delta(P) \le \log_2\ n.$$

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

$$0 \le (E_r)_\delta(P) \le \log_2\ n.$$

Proof: (1) By Definition 30,

$$(E_r)_\delta(P) = \sum_{i=1}^{n} \frac{1}{n} \log_2 \sum_{j=1}^{n} [u_i]^{R_P^G(\delta)}(u_j)$$
$$= \frac{1}{n} \sum_{i=1}^{n} \log_2 (\sum_{j=1}^{n} R_P^G(u_i, u_j)). \quad \forall\, i, j,$$
$$0 \le R_P^G(\delta)(u_i, u_j) \le 1.$$

Then $\forall\, i$,

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

This means that

$$-\infty < \sum_{j=1}^{n} \log_2(\sum_{j=1}^{n} R_P^G(u_i, u_j)) \le nlog_2\ n.$$

Thus

$$-\infty < (E_r)_\delta(P) \le \log_2\ n.$$

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

$$(E_r)_\delta(P) = \log_2\ n.$$

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

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

Thus $\forall\, i$,

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

Hence

$$0 \le (E_r)_\delta(P) \le \log_2\ n.$$

□

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

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

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

and $\exists\, i', j'$,

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

Then

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

So

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

Thus

$$\begin{aligned}
&(E_r)_\delta(P)\\
&= \frac{1}{n}\sum_{i=1}^{n} \log_2 |[u_i]^{R_P^G(\delta)}| = \frac{1}{n}\sum_{i=1}^{n} \log_2(\sum_{j=1}^{n} R_P^G(\delta)(u_i, u_j))\\
&= \frac{1}{n}\sum_{i\neq i'} \log_2(\sum_{j=1}^{n} R_P^G(\delta)(u_i, u_j)) + \frac{1}{n}\log_2(\sum_{j=1}^{n} R_P^G(\delta)(u_{i'}, u_j))\\
&\le \frac{1}{n}\sum_{i\neq i'} \log_2(\sum_{j=1}^{n} R_Q^G(\delta)(u_i, u_j)) + \frac{1}{n}\log_2(\sum_{j=1}^{n} R_P^G(\delta)(u_{i'}, u_j))\\
&< \frac{1}{n}\sum_{i\neq i'} \log_2(\sum_{j=1}^{n} R_Q^G(\delta)(u_i, u_j)) + \frac{1}{n}\log_2(\sum_{j=1}^{n} R_Q^G(\delta)(u_{i'}, u_j))\\
&= \frac{1}{n}\sum_{i=1}^{n} \log_2(\sum_{j=1}^{n} R_Q^G(\delta)(u_i, u_j))\\
&= \frac{1}{n}\sum_{i=1}^{n} \log_2 |[u_i]^{R_Q^G(\delta)}| = (E_r)_\delta(Q).
\end{aligned}$$

Hence

$$(E_r)_\delta(P) < (E_r)_\delta(Q).$$

□

Theorem 33: $(E_r)_\delta$ *in Definition 30 is an* δ*-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 $\forall\, i$, $[u_i]^{R_P^G(\delta)} = [u_i]^{R_Q^G(\delta)}$.

(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 δ-information entropy of (U, P) with respect to δ is defined as

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

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,

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

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

Then

$$\begin{aligned}
&(E_r)_\delta(P) + H_\delta(P)\\
&= -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{|[u_i]^{R_P^G(\delta)}|} - \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{|[u_i]^{R_P^G(\delta)}|}{n}\\
&= -\frac{1}{n}\sum_{i=1}^{n}(\log_2 \frac{1}{|[u_i]^{R_P^G(\delta)}|} + \log_2 \frac{|[u_i]^{R_P^G(\delta)}|}{n})\\
&= -\frac{1}{n}\sum_{j=1}^{n} \log_2 \frac{1}{n}\\
&= \log_2\ n.
\end{aligned}$$

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

$$0 \le H_\delta(P) < +\infty.$$

Besides, if $R_P^G(\delta) = \omega$, then H reaches the minimum value 0; if $R_P^G(\delta)$ is reflexive, then

$$0 \le H_\delta(P) \le \log_2\ n.$$

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 δ-information amount of (U, P) with respect to δ is defined as

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

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, $G_\delta(P) = \frac{1}{n^2}\sum_{i=1}^{n} |[u_i]^{R_P^G(\delta)}|$.

By Definition 37, $E_\delta(P) = \sum_{i=1}^{n} \frac{1}{n}(1 - \frac{|[u_i]^{R_P^G(\delta)}|}{n})$.

Then

$$\begin{aligned}
G_\delta(P) + E_\delta(P) &= \frac{1}{n^2}\sum_{i=1}^{n} |[u_i]^{R_P^G(\delta)}| + \sum_{i=1}^{n} \frac{1}{n}(1 - \frac{|[u_i]^{R_P^G(\delta)}|}{n})\\
&= \frac{1}{n}\sum_{i=1}^{n}(\frac{|[u_i]^{R_P^G(\delta)}|}{n} + (n - \frac{|[u_i]^{R_P^G(\delta)}|}{n}))\\
&= \frac{1}{n}\sum_{j=1}^{n} n\\
&= 1.
\end{aligned}$$

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

$$0 \le E_\delta(P) \le 1.$$

Furthermore, if $R_P^G(\delta) = \omega$, then E reaches the minimum value 0; if $R_P^G(\delta) = o$, then E reaches the maximum value 1.

Proof: This holds by Proposition 27 and Theorem 38. □

Example 40: (Continued from Example 17) Pick $B_i = \{a_1, a_2, \ldots, a_i\}$ $(i = 1, 2, \ldots, 6)$ and $\delta = \sqrt{0.8}$.

By Definition 26,

$$G_\delta(B_1) = \frac{1}{10^2}\sum_{i=1}^{10} |[u_i]^{R_{B_1}^G(\delta)}| \approx 0.9966,$$

$$G_\delta(B_2) = \frac{1}{10^2}\sum_{i=1}^{10} |[u_i]^{R_{B_2}^G(\delta)}| \approx 0.9356,$$

$$G_\delta(B_3) = \frac{1}{10^2}\sum_{i=1}^{10} |[u_i]^{R_{B_3}^G(\delta)}| \approx 0.8235,$$

$$G_\delta(B_4) = \frac{1}{10^2}\sum_{i=1}^{10} |[u_i]^{R_{B_4}^G(\delta)}| \approx 0.6652,$$

$$G_\delta(B_5) = \frac{1}{10^2}\sum_{i=1}^{10} |[u_i]^{R_{B_5}^G(\delta)}| \approx 0.4450,$$

$$G_\delta(B_6) = \frac{1}{10^2}\sum_{i=1}^{10} |[u_i]^{R_{B_6}^G(\delta)}| \approx 0.3383.$$

By Definition 30,

$$(E_r)_\delta(B_1) = -\sum_{i=1}^{10} \frac{|[u_i]^{R_P^G(\delta)}|}{10} log_2 \frac{1}{|[u_i]^{R_{B_1}^G(\delta)}|} \approx 33.0555,$$

$$(E_r)_\delta(B_2) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{1}{|[u_i]^{R_{B_2}^G(\delta)}|} \approx 30.1812,$$

$$(E_r)_\delta(B_3) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{1}{|[u_i]^{R_{B_3}^G(\delta)}|} \approx 25.0493,$$

$$(E_r)_\delta(B_4) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{1}{|[u_i]^{R_{B_4}^G(\delta)}|} \approx 17.7807,$$

$$(E_r)_\delta(B_5) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{1}{|[u_i]^{R_{B_5}^G(\delta)}|} \approx 9.6253,$$

$$(E_r)_\delta(B_6) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{1}{|[u_i]^{R_{B_6}^G(\delta)}|} \approx 5.9907.$$

By Definition 34,

$$H_\delta(B_1) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{|[u_i]^{R_{B_1}^G(\delta)}|}{10} \approx 0.0495,$$

$$H_\delta(B_2) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{|[u_i]^{R_{B_2}^G(\delta)}|}{10} \approx 0.8986,$$

$$H_\delta(B_3) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{|[u_i]^{R_{B_3}^G(\delta)}|}{10} \approx 2.3055,$$

$$H_\delta(B_4) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{|[u_i]^{R_{B_4}^G(\delta)}|}{10} \approx 3.9846,$$

$$H_\delta(B_5) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{|[u_i]^{R_{B_5}^G(\delta)}|}{10} \approx 5.1579,$$

$$H_\delta(B_6) = -\sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}log_2\frac{|[u_i]^{R_{B_6}^G(\delta)}|}{10} \approx 5.2473.$$

By Definition 37,

$$E_\delta(B_1) = \sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}(1-\frac{|[u_i]^{R_{B_1}^G(\delta)}|}{10}) \approx 0.0343,$$

$$E_\delta(B_2) = \sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}(1-\frac{|[u_i]^{R_{B_2}^G(\delta)}|}{10}) \approx 0.6026,$$

$$E_\delta(B_3) = \sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}(1-\frac{|[u_i]^{R_{B_3}^G(\delta)}|}{10}) \approx 1.4513,$$

$$E_\delta(B_4) = \sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}(1-\frac{|[u_i]^{R_{B_4}^G(\delta)}|}{10}) \approx 2.2483,$$

$$E_\delta(B_5) = \sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}(1-\frac{|[u_i]^{R_{B_5}^G(\delta)}|}{10}) \approx 2.4441,$$

$$E_\delta(B_6) = \sum_{i=1}^{10}\frac{|[u_i]^{R_P^G(\delta)}|}{10}(1-\frac{|[u_i]^{R_{B_6}^G(\delta)}|}{10}) \approx 2.2183.$$

The results of these experiments are shown in FIGURE 2.

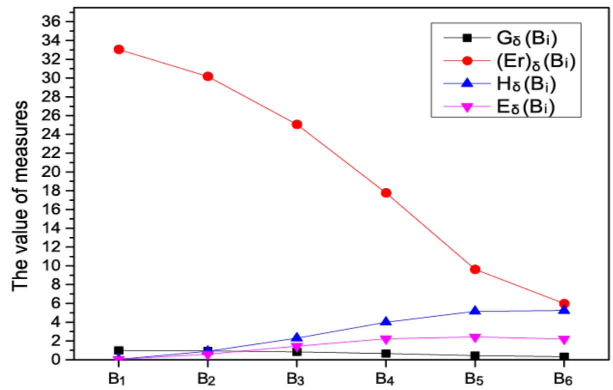

FIGURE 2. Uncertainty measurement for ISVIS with different subsystems.

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

1) *G_δ and $(E_r)_\delta$ monotonically decrease with the increase of number of attributes;*

2) *$(E_r)_\delta$ and H_δ are more sensitive than G_δ;*

3) *$(E_r)_\delta$ and H_δ are more sensitive than E_δ;*

4) *The difference among G_δ and E_δ are almost the same.*

Thus, δ-rough entropy and δ-information entropy are more suitable than δ-information amount and δ-information granulation 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 $\overline{x}$ ($\sigma(X)$, $CV(X)$), they are defined as follows:

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

Example 41 (Continued From Example 40): Denote

$$X_G = \{G_\delta(B_1), \ldots, G_\delta(B_6)\},\ X_{E_r} = \{(E_r)_\delta B_1), \ldots, E_r(B_6)\},$$
$$X_H = \{H_\delta(B_1), \ldots, H_\delta(B_6)\},\ X_E = \{E_\delta(B_1), \ldots, E_\delta(B_6)\}.$$

Then

$$CV(X_G) = 0.3487,\ CV(X_{E_r}) = 0.4963,$$
$$CV(X_H) = 0.6838,\ CV(X_E) = 0.6040.$$

The results are shown in FIGURE 3.
So

$$CV(X_H) > CV(X_E) > CV(X_{E_r}) > CV(X_G).$$

Then dispersion degree of G reaches minimum.

TABLE 2. The corresponding correlation between *X* and *Y*.

$r(X,Y)$	Correlation between X and Y	Abbreviation
$r(X,Y)=1$	Completely positive correlation	CPC
$0.7 \leq r(X,Y) < 1$	Height positive correlation	HPC
$0.4 \leq r(X,Y) < 0.7$	Moderate positive correlation	MPC
$0 < r(X,Y) < 0.4$	Low positive correlation	LPC
$r(X,Y)=0$	No correlation	NC
$-0.4 < r(X,Y) < 0$	Low negative correlation	LNC
$-0.7 \leq r(X,Y) < -0.4$	Moderate negative correlation	MNC
$-1 \leq r(X,Y) < -0.7$	Height negative correlation	HNC
$r(X,Y)=-1$	Completely negative correlation	CNC

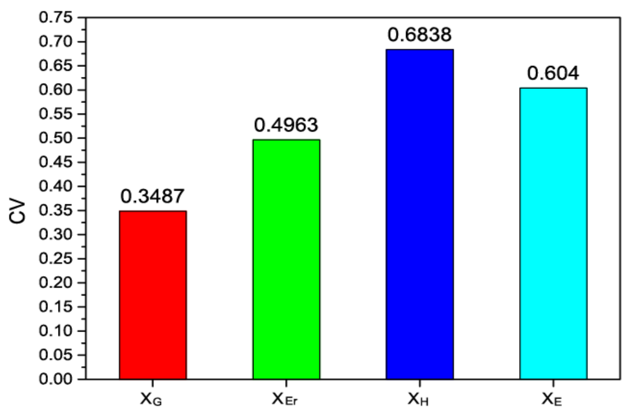

FIGURE 3. *CV*-values for measuring uncertainty of the subsystems.

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

(1) $(E_r)_\delta$ and H_δ have better performance to measure uncertainty of an ISVIS if the monotonicity is only considered;

(2) $(E_r)_\delta$ 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

$$r(X,Y) = \frac{\sum\limits_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}},$$

where $\overline{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$, $\overline{y} = \frac{1}{n}\sum\limits_{i=1}^{n} y_i$.

Obviously,

$$-1 \leq r(X,Y) \leq 1.$$

TABLE 3. *r*-values of measure values sets.

r	X_G	X_{E_r}	X_H	X_E
X_G	1			
X_H	0.9990	1		
X_{E_r}	-0.9736	-0.9830	1	
X_E	-0.8953	-0.9146	0.9729	1

TABLE 4. The correlation between two measures.

	G	E	E_r	H
G	CPC			
E	HPC	CPC		
E_r	HNC	HNC	CPC	
H	HNC	HNC	HPC	CPC

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

$$\chi_F^2 = \frac{12N}{k(k+1)}\left(\sum_{i=1}^{k} r_i^2 - \frac{k(k+1)^2}{4}\right)$$

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

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1)-\chi_F^2}.$$

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.

$Vector$	G_δ	$(E_r)_\delta$	H_δ	E_δ
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
B_6	1	4	3	2
Average	1.67	4.0	2.5	1.83

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_α, then the performance of the two algorithms will be significantly different. The critical distance CD_α is denoted as

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

where q_α is a critical value calculated by the qtukey function in r and α is the significance level.

Example 43 (Continued From Example 40): *We have*

$$G_\delta(B_1) \approx 0.9966,\ G_\delta(B_2) \approx 0.9356,\ G_\delta(B_3) \approx 0.8235,$$
$$G_\delta(B_4) \approx 0.6652,\ G_\delta(B_5) \approx 0.4450,\ G_\delta(B_6) \approx 0.3383;$$
$$(E_r)_\delta(B_1) \approx 33.0555,\ (E_r)_\delta(B_2) \approx 30.1812,\ (E_r)_\delta(B_3) \approx 25.0493,$$
$$(E_r)_\delta(B_4) \approx 17.7807,\ (E_r)_\delta(B_5) \approx 9.6253,\ (E_r)_\delta(B_6) \approx 5.9907;$$
$$H_\delta(B_1) \approx 0.0495,\ H_\delta(B_2) \approx 0.8986,\ H_\delta(B_3) \approx 2.3055,$$
$$H_\delta(B_4) \approx 3.9846,\ H_\delta(B_5) \approx 5.1579,\ H_\delta(B_6) \approx 5.2473;$$
$$E_\delta(B_1) \approx 0.0343,\ E_\delta(B_2) \approx 0.6026,\ E_\delta(B_3) \approx 1.4513,$$
$$E_\delta(B_4) \approx 2.2483,\ E_\delta(B_5) \approx 2.4441,\ E_\delta(B_6) \approx 2.2183.$$

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* $CD_\alpha = 2.569 \times \sqrt{\frac{4\times(4+1)}{6\times 6}} = 1.915$*. FIGURE* 4 *shows the results with* $\alpha = 0.05$ *on the four measurements. The dots in FIGURE* 4 *indicate the average ranking of the four measurements. The line segments in FIGURE* 4 *carves out the scope of* CD_α*. If the two roots partially overlap on the y-axis, then there is no significant difference between these two uncertainty measurements.*

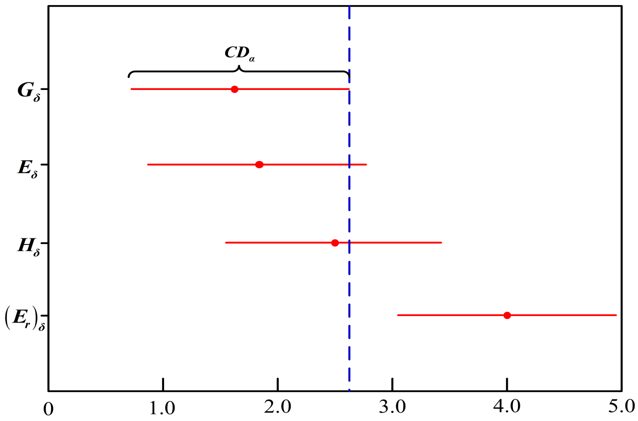

FIGURE 4. The performance of the four measurements under Bonferroni-Dunn test.

(4) *From FIGURE 4, the following results are obtained:*

1) *a) The performance of* G_δ *is statistically different from the performance of* $(E_r)_\delta$*;*

b) The performance of E_δ *is statistically different from the performance of* $(E_r)_\delta$*.*

2) *a) There is no significant difference among* G_δ*,* E_δ *and* H_δ*;*

b) There is no significant difference between H_δ *and* E_δ*.*

VIII. OPTIMAL SELECTION OF SUBSYSTEMS BASED ON UNCERTAINTY MEASURES

In the above section, we use relationships between two ISVISs 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 δ-information granulation and δ-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* $G_\delta(B_1) = max\{G_\delta(B) : B \subseteq A\}$*, then* (U, B_1) *is called a maximum subsystem in* (U, A) *based on* δ*-information granulation;*

(2) *If there exists* $B_2 \subseteq A$ *such that* $G_\delta(B_2) = min\{G_\delta(B) : B \subseteq A\}$*, then* (U, B_2) *is called a minimum subsystem in* (U, A) *based on* δ*-information granulation.*

The maximum subsystem and minimum subsystem in (U, A) are collectively called the optimal subsystems based on δ-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* $G_\delta(\{a_0\}) = max\{G_\delta(\{a\}) : a \in A\}$*, then* $(U, \{a_0\})$ *is a maximum subsystem in* (U, A) *based on* δ*-information granulation;*

(2) (U, A) *is a minimum subsystem in* (U, A) *based on* δ-*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 δ-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 δ-information granulation.

□

Example 46: (Continued from Example 40) Pick $\delta = \sqrt{0.8}$. *Then*

$$G_\delta(\{a_1\}) = 0.9966, G_\delta(\{a_2\}) = 0.9523, G_\delta(\{a_3\}) = 0.9508,$$
$$G_\delta(\{a_4\}) = 0.9336, G_\delta(\{a_5\}) = 0.8225, G_\delta(\{a_6\}) = 0.8573.$$
$$G_\delta(A) = 0.3383.$$

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

The maximum subsystem and minimum subsystem in (U, A) based on δ-information amount are collectively called the optimal subsystems based on δ-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* δ-*information amount,* $(U, \{a_1, a_3, a_5, a_6\})$ *is a maximum subsystem in* (U, A) *based on* δ-*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 δ-information granulation and δ-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 α-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. Muzzioli, "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, "Information entropy, 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," *Inf. Sci.*, vol. 177, no. 1, pp. 3–27, Jan. 2007.

[20] Z. Pawlak and A. Skowron, "Rough sets: Some extensions," *Inf. Sci.*, vol. 177, no. 1, pp. 28–40, Jan. 2007.

[21] Z. Pawlak and A. Skowron, "Rough sets and Boolean reasoning," *Inf. Sci.*, vol. 177, no. 1, pp. 41–73, Jan. 2007.

[22] Y. H. Qian, J. Y. Liang, and C. Y. Dang, "Set ordered information systems," *Comput. Math. Appl.*, vol. 56, pp. 1994–2009, Jul. 2008.

[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. 379–423, 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. 2,3, 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.