Information Sciences 470 (2019) 156–174

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

New measures of uncertainty for an interval-valued information system

Ningxin Xie[a,*], Meng Liu[b], Zhaowen Li[c], Gangqiang Zhang[d]

[a] School of Software and Information Security, Guangxi University for Nationalities, Nanning, Guangxi, 530006, PR China
[b] School of Science, Guangxi University for Nationalities, Nanning, Guangxi 530006, PR China
[c] Key Laboratory of Complex System Optimization and Big Data Processing in Department of Guangxi Education, Yulin Normal University, Yulin, Guangxi, 537000, PR China
[d] School of Software and Information Security, Guangxi University for Nationalities, Nanning, Guangxi, 530006, PR China

A R T I C L E I N F O

Article history:
Received 14 November 2017
Revised 22 August 2018
Accepted 24 August 2018
Available online 25 August 2018

Keywords:
Interval-valued information system
Rough set theory
Information granule
Information structure
Dependence
Measure
Uncertainty
Effectiveness

A B S T R A C T

An information system as a database that represents relationships between objects and attributes is an important mathematical model. An interval-valued information system is a generalized model of single-valued information systems. As important evaluation tools in the field of machine learning, measures of uncertainty can quantify the dependence and similarity between two targets. However, the existing measures of uncertainty for interval-valued information systems have not been thoroughly researched. This paper is devoted to the study of new measures of uncertainty for an interval-valued information system. Information structures are first introduced in a given interval-valued information system. Then, the dependence between two information structures is depicted. Next, new measures of uncertainty for an interval-valued information system are investigated by using the information structures. As an application of the proposed measures, the rough entropy of a rough set is proposed by means of information granulation. Finally, a numerical experiment on the Face recognition dataset is presented to demonstrate the feasibility of the proposed measures, and a statistical effectiveness analysis is conducted. The results are helpful for understanding the essence of uncertainty in interval-valued information systems.

1. Introduction

Rough set theory, a mathematical tool to address imprecision and uncertainty in data analysis, has been successfully applied to intelligent systems, machine learning, knowledge discovery, expert systems, decision analysis, inductive reasoning, pattern recognition, split theory and signal analysis [17,18,22–24].

An information system based on rough sets was introduced by Pawlak [18]. An interval-valued information system is a generalization of the classic information system. To address interval data, scholars have used the methods for managing classic information systems to manage interval-valued information systems. For example, Yao and Li [34] proposed an interval set model for interval-valued information systems with upper and lower approximations and introduced generalized

* Corresponding author.
E-mail addresses: xieningxin@mail.gxun.cn (N. Xie), mengliu8846@gxun.edu.cn (M. Liu), 20110019@gxun.edu.cn (Z. Li), 402503@gxun.edu.cn (G. Zhang).

https://doi.org/10.1016/j.ins.2018.08.047
0020-0255/

decision logic. Dai et al. [3] studied the algebraic structures of interval-set-valued rough sets generated from an approximation space. Leung et al. [10] investigated a rough set approach on the basis of a knowledge induction process for selecting decision rules with minimum feature sets in interval-valued information systems. Qian et al. [25] presented a dominance relation for interval-valued information systems. Yang et al. [35] proposed a dominance relation and generated the optimal decision rules in incomplete interval-valued information systems. Wu et al. [30] considered the real formal concept analysis of grey-rough set theory by using grey numbers and proposed a grey-rough set approach to Galois lattice reductions. Sakai et al. [29] developed a rule generation prototype system for incomplete information databases in Lipski that can process interval-valued information systems.

To measure the uncertainty of a system, Shannon [28] introduced the concept of entropy, which is a very useful mechanism for characterizing information content in various modes and has been used in diverse fields. Some scholars have applied the extension of entropy and its variants to information systems or rough sets. For example, Liang and Qian [12] studied the theory of information particles and entropy in information systems. Liang et al. [13] researched information entropy, rough entropy and knowledge granularity in incomplete information systems. Dai and Tian [4] considered the entropy measure and granularity measure of set-valued information systems. Wang and Yue [32] discussed the entropy measure and granularity measure of interval-valued and set-valued information systems. Qian et al. [27] researched fuzzy information entropy and the granularity of fuzzy information in fuzzy granular structures. Qian et al. [26] proposed fuzzy granularity structure distance. Xu et al. [33] presented rough entropy of rough sets in ordered information systems. Dai et al. [7] studied uncertainty measurement based on the α-weak similarity for incomplete interval-valued information systems. Dai et al. [5] introduced θ-similarity entropy and proposed the θ-rough degree to measure the uncertainty of concepts or rough sets in an interval-valued information system. Dai et al. [6] constructed an extended conditional entropy for interval-valued decision information systems. Zhang et al. [41] investigated uncertainty measures in a fully fuzzy information system.

Granular computing, proposed by Zadeh [37], is a vital topic in data mining and knowledge representation and is also a crucial tool in artificial intelligence [37–40]. The aim of granular computing is to find an approximation scheme that enables us to observe phenomena of different sizes and solve complex problems effectively. Granularity and granular structures are two important concepts in granular computing. Information granules are a collection of objects that are drawn together by certain constraints, such as ambiguity, similarity, or function. The construction process of information granules is called granulation. A granular structure is a set of information particles in which the inner structure of each information particle can be considered to be a substructure. The construction, interpretation and representation of management information granules is an important problem in granular computing. Four main approaches to granular computing have been reported: rough set theory [17,18,22–24], fuzzy set theory [36], concept lattices [14,31] and quotient spaces [43]. Given an information system, the information structures are granular structures by definition of granular computing. Information structures are the basic structures of an information system. Some scholars have done some good work in granular computing. For example, Al-Hmouz et al. [1] established a general framework of granular computing for description and prediction of time series. Pedrycz [19] studied granular fuzzy data analysis. Hu et al. [9] considered fuzzy classifiers with information granules in feature space and logic-based computing. Loia et al. [11] investigated a granular computing approach based on formal concept analysis for discovering periodicities in data. Pedrycz and Homenda [20] build the fundamentals of granular computing for a principle of justifiable granularity. Zhong et al. [42] gave a framework of granular computing on granular data imputation. Pedrycz et al. [21] proposed granular representation and granular computing with fuzzy sets.

So far, however, measuring the uncertainty in an interval-valued information system itself has not been thoroughly studied. The aim of this paper is to address this topic. We first propose the concept of information structures in an interval-valued information system and then present four types of measuring tools to measure uncertainty in this system by using its information structures. Why do we study information structures and uncertainty measures together? Because information structures are very helpful for knowledge discovery from an interval-valued information system, an interval-valued information system has uncertainty, and it is difficult to compare the magnitude of the uncertainty measures of two interval-valued information systems. If the dependence between the information structures of two interval-valued information systems is given, then the magnitudes of the uncertainty measures can be compared by means of the dependence.

The remainder of this paper is organized as follows. In Section 2, we recall some basic concepts on binary relations, interval-valued numbers and tolerance relations in an interval-valued information system. In Section 3, we introduce the information structures in the given interval-valued information system and study the dependence between two information structures. In Section 4, we propose tools for measuring uncertainty in interval-valued information systems. In Section 5, we define the rough entropy of a rough set in a given interval-valued information system by means of information granulation and illustrate that the rough entropy is more accurate than the roughness. In Section 6, a numerical experiment is presented, and a statistical effectiveness analysis is conducted. Section 7 summarizes this paper.

2. Preliminaries

In this section, we recall some basic notions about binary relations, interval-valued numbers and tolerance relations in interval-valued information systems.

Throughout this paper, U denotes a non-empty finite set, 2^U denotes the family of all subsets of U, and $|X|$ denotes the cardinality of $X \in 2^U$.

In this paper, denote

$$U = \{x_1, x_2, \cdots, x_n\},$$

$$\delta = U \times U,\ \ \triangle = \{(x, x) \mid x \in U\}.$$

2.1. Interval-valued numbers

Let

$$[R] = \{a = [a^-, a^+] \mid a^-, a^+ \in R,\ \ a^- \leq a^+\}.$$

For any $a \in R$, denote $\overline{a} = [a, a]$.
For any a, $b \in [R]$, define
(1) $a = b \Longleftrightarrow a^- = b^-, a^+ = b^+$.
(2) $a \leq b \Longleftrightarrow a^- \leq b^-, a^+ \leq b^+$; $\ a < b \Longleftrightarrow a \leq b, a \neq b$.

Definition 2.1 [15,16]**.** Let a, $b \in [R]$. Then, the possible degree of a relative to b is defined as

$$P(a, b) = min\left\{1, max\left\{\frac{a^+ - b^-}{(a^+ - a^-) + (b^+ - b^-)}, 0\right\}\right\}.$$

Proposition 2.2 [8,16]**.** *The following properties hold:*
(1) $\forall\, a, b \in [R]$, $0 \leq P(a, b) \leq 1$*;*
(2) $\forall\, a \in [R]$, $P(a, a) = 0.5$*;*
(3) $\forall\, a, b \in [R]$, $P(a, b) + P(b, a) = 1$.

Definition 2.3 [6]**.** Let a, $b \in [R]$. Then, the similarity degree of a and b is defined as

$$S(a, b)\ or\ S_{ab} = 1 - |P(a, b) - P(b, a)|.$$

Proposition 2.4 [6]**.** *The following properties hold:*
(1) $\forall\, a, b \in [R]$, $S(a, b) = S(b, a)$*;*
(2) $\forall\, a, b \in [R]$, $0 \leq S(a, b) \leq 1$*;*
(3) $\forall\, a, b \in [R]$, $S(a, b) = 1 \ \Leftrightarrow\ a = b$.

Example 2.5 [5]**.** Select $a = [1, 3]$ and $b = [2, 5]$. Then,

$$P(a, b) = min\left\{1, max\left\{\frac{3 - 2}{(3 - 1) + (5 - 2)}, 0\right\}\right\} = \frac{1}{5},$$

$$P(b, a) = min\left\{1, max\left\{\frac{5 - 1}{(5 - 2) + (3 - 1)}, 0\right\}\right\} = \frac{4}{5},$$

$$S(a, b) = 1 - |P(a, b) - P(b, a)| = 1 - |\frac{1}{5} - \frac{4}{5}| = 0.4.$$

2.2. Tolerance relations in interval-valued information systems

Recall that R is a binary relation on U whenever $R \subseteq U \times U$.

Let R be a binary relation on U. Then, R is called an equivalence relation on U if R is reflexive, symmetric and transitive; R is called a tolerance relation on U if R is reflexive and symmetric; R is called a universal relation on U if $R = \delta$; and R is called an identity relation on U if $R = \triangle$.

Let U be an object set, and let A be an attribute set. Suppose that U and A are finite sets. Then, the pair (U, A) is called an information system if each attribute $a \in A$ determines an information function a: $U \rightarrow V_a$, where V_a is the set of all information function values with respect to attribute a.

Given that (U, A) is an information system, (U, A) is referred to as an interval-valued information system if for each $a \in A$, $V_a \subseteq [R]$. If $P \subseteq A$, then (U, P) is called the subsystem of (U, A).

Definition 2.6 [5]**.** Let (U, A) be an interval-valued information system. Given $\theta \in (0, 1]$ and $P \subseteq A$, a binary relation on U can be defined as

$$S_P^\theta = \{(x, y) \in U \times U \mid S(a(x), a(y)) \geq \theta,\ \ \forall\ a \in P\}.$$

Clearly, S_P^θ is a tolerance relation on U, and $S_P^\theta = \bigcap_{a \in P} S_{\{a\}}^\theta$.

Denote

$$S_P^\theta(x) = \{y \in U \mid (x, y) \in S_P^\theta\}.$$

Table 1
An interval-valued information system.

	a_1	a_2	a_3	a_4	a_5
x_1	[2.17,2.86]	[2.45,2.96]	[5.32,7.23]	[3.21,3.95]	[2.54,3.12]
x_2	[3.37,4.75]	[3.43,4.85]	[7.24,10.47]	[4.00,5.77]	[3.24,4.70]
x_3	[1.83,2.70]	[1.78,2.98]	[7.23,10.27]	[2.96,4.07]	[2.06,2.79]
x_4	[1.35,2.12]	[1.42,2.09]	[2.59,3.93]	[1.87,2.62]	[1.67,2.32]
x_5	[3.46,5.35]	[3.37,5.11]	[6.37,10.28]	[3.76,5.70]	[3.41,5.28]
x_6	[2.29,3.43]	[2.60,3.48]	[6.71,8.81]	[3.30,4.23]	[3.01,3.84]
x_7	[2.22,3.07]	[2.43,3.32]	[4.37,7.05]	[2.66,3.68]	[2.39,3.20]
x_8	[2.51,4.04]	[2.52,4.12]	[7.12,11.26]	[4.44,6.91]	[3.06,4.65]
x_9	[1.24,2.00]	[1.35,1.91]	[3.83,5.31]	[2.13,3.01]	[1.72,2.34]
x_{10}	[1.00,1.72]	[1.10,1.82]	[3.58,5.65]	[1.67,2.53]	[1.10,1.84]

Then, $S_P^\theta(x)$ is called the tolerance class of the point x under the tolerance relation S_P^θ.

Proposition 2.7 [5]**.** *Let (U, A) be an interval-valued information system. Then, the following properties hold:*

(1) If $P_1 \subseteq P_2$, then for any $\theta \in (0, 1]$ and $x \in U$,

$$S_{P_2}^\theta(x) \subseteq S_{P_1}^\theta(x);$$

(2) If $0 \le \theta_1 \le \theta_2 \le 1$, then for any $\theta \in (0, 1]$ and $P \subseteq A$,

$$S_P^{\theta_2}(x) \subseteq S_P^{\theta_1}(x).$$

3. Information structures in interval-valued information systems

In this section, we propose some concepts of information structures in a given interval-valued information system.

3.1. Information granules and information structures

Let (U, A) be an interval-valued information system. Given $\theta \in (0, 1]$ and $P \subseteq A$, we obtain a tolerance relation S_P^θ on U. For each $x \in U$, $S_P^\theta(x)$ is the tolerance class of point x under the tolerance relation S_P^θ. This tolerance relation S_P^θ divides the object set U into tolerance classes. All tolerance classes form a covering of U. If $y_1, y_2 \in S_P^\theta(x)$, then we may say that y_1, y_2 cannot be distinguished under the tolerance relation S_P^θ. Thus, each tolerance class is seen to be an information granule consisting of indistinguishable objects. The family of all these information granules constitutes a vector; this vector can be seen as the information structure of the subsystem (U, P). This type of information structure will be helpful for establishing the framework of granular computing in interval-valued information systems.

Definition 3.1. Let (U, A) be an interval-valued information system. Given $\theta \in (0, 1]$ and $P \subseteq A$, then

$$S^\theta(P) = (S_P^\theta(x_1), S_P^\theta(x_2), \cdots\cdots, S_P^\theta(x_n))$$

is called the information structure of the subsystem (U, P) with respect to θ.

Example 3.2. Table 1 depicts an interval-valued information system (U, A) (see Table 1 in [10]), where $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ is an object set and $A = \{a_1, a_2, a_3, a_4, a_5\}$ is an attribute set.

Select $\theta = 0.4$ and $P = \{a_1, a_2\}$. Then,

$S^\theta(P) = (\{x_1, x_3, x_6, x_7\}, \{x_2, x_5, x_8\}, \{x_1, x_3, x_7\}, \{x_4, x_9, x_{10}\}, \{x_2, x_5\}, \{x_1, x_6, x_7, x_8\}, \{x_1, x_3, x_6, x_7, x_8\}, \{x_2, x_6, x_7, x_8\}, \{x_4, x_9, x_{10}\}, \{x_4, x_9, x_{10}\})$.

Definition 3.3. Let (U, A) be an interval-valued information system, given θ_1, $\theta_2 \in (0, 1]$ and P, $Q \subseteq A$. Suppose that $S^{\theta_1}(P)$ and $S^{\theta_2}(Q)$ are the information structures of P and Q with respect to θ_1 and θ_2, respectively. If for each i, $S_P^{\theta_1}(x_i) = S_Q^{\theta_2}(x_i)$, then $S^{\theta_1}(P)$ and $S^{\theta_2}(Q)$ are considered to be the same. We write $S^{\theta_1}(P) = S^{\theta_2}(Q)$.

3.2. Dependence between information structures

In this subsection, we propose the dependence between information structures in interval-valued information systems.

The following definition depicts three aspects of the dependence between information structures in interval-valued information systems.

Definition 3.4. Let (U, A) be an interval-valued information system. Given θ_1, $\theta_2 \in (0, 1]$ and P, $Q \subseteq A$, suppose that $S^{\theta_1}(P)$ and $S^{\theta_2}(Q)$ are the information structures of P and Q with respect to θ_1 and θ_1, respectively. Then,

(1) $S^{\theta_2}(Q)$ is considered to be dependent on $S^{\theta_1}(P)$ if for each i, $S_P^{\theta_1}(x_i) \subseteq S_Q^{\theta_2}(x_i)$. We write $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$.

(2) $S^{\theta_2}(Q)$ is considered to be strictly dependent on $S^{\theta_1}(P)$ if $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$ and $S^{\theta_1}(P) \neq S^{\theta_2}(Q)$. We write $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$.

Clearly,

$$S^{\theta_1}(P) = S^{\theta_2}(Q) \iff S^{\theta_1}(P) \preceq S^{\theta_2}(Q) \text{ and } S^{\theta_2}(Q) \preceq S^{\theta_1}(P);$$

$$S^{\theta_1}(P) \prec S^{\theta_2}(Q) \Rightarrow S^{\theta_1}(P) \preceq S^{\theta_2}(Q).$$

Theorem 3.5. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$, then*

$$S^{\theta_1}(P) = S^{\theta_2}(Q) \iff S_P^{\theta_1} = S_Q^{\theta_2}.$$

Proof. The proof is trivial. □

Theorem 3.6. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$, then*

$$S^{\theta_1}(P) \preceq S^{\theta_2}(Q) \iff S_P^{\theta_1} \subseteq S_Q^{\theta_2}.$$

Proof. The proof is trivial. □

Corollary 3.7. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$, then*

$$S^{\theta_1}(P) \prec S^{\theta_2}(Q) \iff S_P^{\theta_1} \subset S_Q^{\theta_2}.$$

Proof. This holds by Theorems 3.5 and 3.6. □

Theorem 3.8. *Let (U, A) be an interval-valued information system.*
(1) If $P \subseteq Q$ with $P, Q \subseteq A$, then for any $\theta \in (0, 1]$, $S^{\theta}(Q) \preceq S^{\theta}(P)$.
(2) If $0 < \theta_1 \leq \theta_2 \leq 1$, then for any $P \subseteq A$, $S^{\theta_2}(P) \preceq S^{\theta_1}(P)$.

Proof. (1) Since $P \subseteq Q$, by Proposition 2.8(1), for any i, we have

$$S_Q^{\theta}(x_i) \subseteq S_P^{\theta}(x_i).$$

Then, $S^{\theta}(Q) \preceq S^{\theta}(P)$.
(1) Since $\theta_1 \leq \theta_2$, by Proposition 2.8(2), for any $P \subseteq A$, we have

$$S_P^{\theta_2}(x) \subseteq S_P^{\theta_1}(x).$$

Then, $S^{\theta_2}(P) \preceq S^{\theta_1}(P)$. □

Corollary 3.9. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$, if $P \subseteq Q$ and $\theta_1 \leq \theta_2$, then*

$$S^{\theta_2}(Q) \preceq S^{\theta_1}(Q) \preceq S^{\theta_1}(P), \;\; S^{\theta_2}(Q) \preceq S^{\theta_2}(P) \preceq S^{\theta_1}(P).$$

Proof. This holds by Theorem 3.8. □

4. New measures of uncertainty for interval-valued information systems

In this section, we propose four tools to measure uncertainty in interval-valued information systems.

4.1. Granulation measures for interval-valued information systems

We first give an axiomatic definition of the information granulation of an interval-valued information system.

Definition 4.1. Let (U, A) be an interval-valued information system. Given $\theta \in (0, 1]$. Suppose that $G^{\theta} : 2^A \to (-\infty, +\infty)$ is a function. G^{θ} is called an information granulation function in (U, A) with respect to θ if G^{θ} satisfies the following conditions:

(1) Non-negativity: $\forall P \subseteq A$, $G^{\theta}(P) \geq 0$;
(2) Invariability: $\forall P, Q \subseteq A$, if $S^{\theta}(P) = S^{\theta}(Q)$, then $G^{\theta}(P) = G^{\theta}(Q)$;
(3) Monotonicity: $\forall P, Q \subseteq A$, if $S^{\theta}(P) \prec S^{\theta}(Q)$, then $G^{\theta}(P) < G^{\theta}(Q)$.
Here, $G^{\theta}(P)$ is called the information granulation of subsystem (U, P) with respect to θ.

Definition 4.2. Suppose that (U, A) is an interval-valued information system. Given $\theta \in (0, 1]$, then the θ-information granulation of subsystem (U, P) with respect to θ is defined as

$$G^{\theta}(P) = \frac{1}{n^2} \sum_{i=1}^{n} |S_P^{\theta}(x_i)|^2.$$

Proposition 4.3. *Assume that* (U, A) *is an interval-valued information system. Given* $P, Q \subseteq A$ *and* $\theta \in (0, 1]$, *then*

$$\frac{1}{n} \leq G^{\theta}(P) \leq n.$$

Moreover, if S_P^{θ} *is an identity relation on* U, *then* G^{θ} *achieves the minimum value* $\frac{1}{n}$; *if* S_P^{θ} *is a universal relation on* U, *then* G^{θ} *achieves the maximum value* n.

Proof. Since for each i, $1 \leq |S_P^{\theta}(x_i)| \leq n$, we have $n \leq \sum_{i=1}^{n} |S_P^{\theta}(x_i)|^2 \leq n^3$.

By Definition 4.2,

$$\frac{1}{n} \leq G^{\theta}(P) \leq n.$$

If S_P^{θ} is an identity relation on U, then $\forall i$, $|S_P^{\theta}(x_i)| = 1$, so $G^{\theta}(P) = \frac{1}{n}$.
If S_P^{θ} is a universal relation on U, then $\forall i$, $|S_P^{\theta}(x_i)| = n$, so $G^{\theta}(P) = n$. □

Theorem 4.4. *Assume that* (U, A) *is an interval-valued information system. Given* $P, Q \subseteq A$ *and* $\theta_1, \theta_2 \in (0, 1]$,
(1) If $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$, *then* $G^{\theta_1}(P) \leq G^{\theta_2}(Q)$.
(2) If $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$, *then* $G^{\theta_1}(P) < G^{\theta_2}(Q)$.

Proof. (1) The proof is trivial.
(2) By Definition 4.2,

$$G^{\theta_1}(P) = \frac{1}{n^2} \sum_{i=1}^{n} |S_P^{\theta_1}(x_i)|^2, \quad G^{\theta_2}(Q) = \frac{1}{n^2} \sum_{i=1}^{n} |S_Q^{\theta_2}(x_i)|^2.$$

Note that $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$. Then, $\forall i$, $S_P^{\theta_1}(x_i) \subseteq S_Q^{\theta_2}(x_i)$ and $\exists j$, $S_P^{\theta_1}(x_j) \subsetneq S_Q^{\theta_2}(x_j)$. Thus, $\forall i$, $|S_P^{\theta_1}(x_i)| \leq |S_Q^{\theta_2}(x_i)|$ and $\exists j$, $|S_P^{\theta_1}(x_j)| < |S_Q^{\theta_2}(x_j)|$.

Hence, $G^{\theta_1}(P) < G^{\theta_2}(Q)$. □

This theorem shows that when the available information becomes coarse, the θ-information granulation increases, and when the available information becomes fine, the θ-information granulation decreases. In other words, the greater the uncertainty of the existing information, the greater the value of the θ-information granulation. Therefore, we can conclude that the θ-information granulation introduced in definition 4.2 can be used to evaluate the degree of an interval-valued information system.

Proposition 4.5. *Assume that* (U, A) *is an interval-valued information system. Then,*
(1) If $P \subseteq Q$ *with* $P, Q \subseteq A$, *then for any* $\theta \in (0, 1]$, $G^{\theta}(Q) \preceq G^{\theta}(P)$.
(2) If $0 < \theta_1 \leq \theta_2 \leq 1$, *then for any* $P \subseteq A$, $G^{\theta_2}(P) \preceq G^{\theta_1}(P)$.

Proof. This holds by Theorem 3.8 and Proposition 4.4(1). □

Corollary 4.6. *Let* (U, A) *be an interval-valued information system. Given* $\theta_1, \theta_2 \in (0, 1]$ *and* $P, Q \subseteq A$, *if* $P \subseteq Q$, $\theta_1 \leq \theta_2$, *then*

$$G^{\theta_2}(Q) \preceq G^{\theta_1}(Q) \preceq G^{\theta_1}(P), \quad G^{\theta_2}(Q) \preceq G^{\theta_2}(P) \preceq G^{\theta_1}(P).$$

Proof. This holds by Proposition 4.5. □

Theorem 4.7. G^{θ} *in Definition 4.2 is an information granulation function under Definition 4.1.*

Proof. (1) Clearly, "non-negativity" holds.
(2) Given $P, Q \subseteq A$ and θ, if $S^{\theta}(P) = S^{\theta}(Q)$, then $\forall i$, $S_P^{\theta}(x_i) = S_Q^{\theta}(x_i)$.
By Definition 4.2, $G^{\theta}(P) = G^{\theta}(Q)$.
(3) "Monotonicity" follows from Theorem 4.4. □

4.2. Information amounts in interval-valued information systems

Definition 4.8. Suppose that (U, A) is an interval-valued information system. Given $\theta \in (0, 1]$, then the θ-information amount of subsystem (U, P) with respect to θ is defined as

$$E^{\theta}(P) = \sum_{i=1}^{n} \frac{|S_P^{\theta}(x_i)|}{n} \left(1 - \frac{|S_P^{\theta}(x_i)|}{n}\right).$$

Theorem 4.9. *Assume that* (U, A) *is an interval-valued information system. Given* $P, Q \subseteq A$ *and* $\theta_1, \theta_2 \in (0, 1]$,
(1) If $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$, *then* $E^{\theta_1}(P) \leq E^{\theta_2}(Q)$.

(2) If $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$, then $E^{\theta_1}(P) < E^{\theta_2}(Q)$.

Proof. (1) The proof is trivial.

(2) By Definition 4.8,

$$E^{\theta_1}(P) = \sum_{i=1}^{n} \frac{|S_P^{\theta_1}(x_i)|}{n}\left(1 - \frac{|S_P^{\theta_1}(x_i)|}{n}\right), \ E^{\theta_2}(Q) = \sum_{i=1}^{n} \frac{|S_Q^{\theta_2}(x_i)|}{n}\left(1 - \frac{|S_Q^{\theta_2}(x_i)|}{n}\right).$$

Note that $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$. Then, $\forall i$, $S_P^{\theta_1}(x_i) \subseteq S_Q^{\theta_2}(x_i)$ and $\exists\ j$, $S_P^{\theta_1}(x_j) \subsetneq S_Q^{\theta_2}(x_j)$. Thus, $\forall i$, $|S_P^{\theta_1}(x_i)| \leq |S_Q^{\theta_2}(x_i)|$ and $\exists\ j$, $|S_P^{\theta_1}(x_j)| < |S_Q^{\theta_2}(x_j)|$.

Hence, $E^{\theta_1}(P) < E^{\theta_2}(Q)$. □

This theorem shows that as the structure of interval-valued information becomes finer, the θ-information amount decreases, and when the interval-valued information structure becomes rougher, the θ-information amount increases.

Proposition 4.10. *Assume that (U, A) is an interval-valued information system.*

(1) If $P \subseteq Q$ with $P, Q \subseteq A$, then for any $\theta \in (0, 1]$, $E^{\theta}(Q) \preceq E^{\theta}(P)$.

(2) If $0 < \theta_1 \leq \theta_2 \leq 1$, then for any $P \subseteq A$, $E^{\theta_2}(P) \preceq E^{\theta_1}(P)$.

Proof. This holds by Theorems 3.8 and 4.9. □

Corollary 4.11. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$, if $P \subseteq Q$, $\theta_1 \leq \theta_2$, then*

$$E^{\theta_2}(Q) \preceq E^{\theta_1}(Q) \preceq E^{\theta_1}(P), \ E^{\theta_2}(Q) \preceq E^{\theta_2}(P) \preceq E^{\theta_1}(P).$$

Proof. This follows from Proposition 4.10. □

Theorem 4.12. *E^{θ} in Definition 4.8 is an information granulation function under Definition 4.1.*

Proof. (1) Clearly, "non-negativity" holds.

(2) Given $P, Q \subseteq A$ and θ, if $S^{\theta}(P) = S^{\theta}(Q)$, then $\forall i$, $S_P^{\theta}(x_i) = S_Q^{\theta}(x_i)$.

By Definition 4.8, $E^{\theta}(P) = E^{\theta}(Q)$.

(3) "Monotonicity" follows from Theorem 4.9. □

Theorem 4.13. *Suppose that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta \in (0, 1]$, then*

$$1 \leq G^{\theta}(P) + E^{\theta}(P) \leq n.$$

Proof.

$$\begin{aligned} G^{\theta}(P) + E^{\theta}(P) &= \frac{1}{n^2}\sum_{i=1}^{n}|S_P^{\theta}(x_i)|^2 + \sum_{i=1}^{n}\frac{|S_P^{\theta}(x_i)|}{n}\left(1 - \frac{|S_P^{\theta}(x_i)|}{n}\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}|S_P^{\theta}(x_i)|(|S_P^{\theta}(x_i)| + n - |S_P^{\theta}(x_i)|) \\ &= \frac{1}{n}\sum_{i=1}^{n}|S_P^{\theta}(x_i)|. \end{aligned}$$

Since $1 \leq |S_P^{\theta}(x_i)| \leq n$,

$$1 \leq G^{\theta}(P) + E^{\theta}(P) \leq n.$$

□

Corollary 4.14. *Assume that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta \in (0, 1]$, then*

$$0 \leq E^{\theta}(P) \leq n - \frac{1}{n}.$$

Proof. Clearly, $0 \leq E^{\theta}(P)$.

By Proposition 4.3, $\frac{1}{n} \leq G^{\theta}(P) \leq n$. Then,

$$-n \leq -G^{\theta}(P) \leq -\frac{1}{n}.$$

By Theorem 4.13, $1 \leq G^{\theta}(P) + E^{\theta}(P) \leq n$.

Thus,

$$E^{\theta}(P) \leq n - \frac{1}{n}.$$

□

4.3. Entropy measure of interval-valued information systems

Definition 4.15. Suppose that (U, A) is an interval-valued information system. Given $\theta \in (0, 1]$, then the θ-rough entropy of subsystem (U, P) with respect to θ is defined as

$$E_r^{\theta}(P) = -\sum_{i=1}^{n} \frac{|S_P^{\theta}(x_i)|}{n} \log_2 \frac{1}{|S_P^{\theta}(x_i)|}.$$

Proposition 4.16. *Assume that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta \in (0, 1]$, then*

$$0 \le E_r^{\theta}(P) \le n \log_2 n.$$

Moreover, if S_P^{θ} is an identity relation on U, then E_r^{θ} achieves the minimum value 0; if S_P^{θ} is a universal relation on U, then E_r^{θ} achieves the maximum value $n \log_2 n$.

Proof. Since $\forall i,\ 1 \le |S_P^{\theta}(x_i)| \le n,$ we have

$$0 \le -\log_2 \frac{1}{|S_P^{\theta}(x_i)|} = \log_2 |S_P^{\theta}(x_i)| \le \log_2 n, \quad 0 \le \frac{|S_P^{\theta}(x_i)|}{n} \le 1.$$

Then,

$$0 \le \frac{|S_P^{\theta}(x_i)|}{n} \log_2 \frac{1}{|S_P^{\theta}(x_i)|} \le \log_2 n.$$

By Definition 4.15,

$$0 \le E_r^{\theta}(P) \le n \log_2 n.$$

If S_P^{θ} is an identity relation on U, then $\forall i,\ |S_P^{\theta}(x_i)| = 1$. Therefore, $E_r^{\theta}(P) = 0$.
If S_P^{θ} is a universal relation on U, then $\forall i,\ |S_P^{\theta}(x_i)| = n$. Therefore, $E_r^{\theta}(P) = n \log_2 n$.

□

Theorem 4.17. *Let (U, A) be an interval-valued information system. Given $P, Q \subseteq A$ and $\theta_1, \theta_2 \in (0, 1]$:*
(1) If $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$, then $E_r^{\theta_1}(P) \le E_r^{\theta_2}(Q)$.
(2) If $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$, then $E_r^{\theta_1}(P) < E_r^{\theta_2}(Q)$.

Proof. (1) The proof is trivial.
(2) By Definition 4.15,

$$E_r^{\theta_1}(P) = -\sum_{i=1}^{n} \frac{|S_P^{\theta_1}(x_i)|}{n} \log_2 \frac{1}{|S_P^{\theta_1}(x_i)|}, \quad E_r^{\theta_2}(Q) = -\sum_{i=1}^{n} \frac{|S_Q^{\theta_2}(x_i)|}{n} \log_2 \frac{1}{|S_Q^{\theta_2}(x_i)|}.$$

Note that $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$. Then, $\forall i,\ S_P^{\theta_1}(x_i) \subseteq S_Q^{\theta_2}(x_i)$ and $\exists\ j,\ S_P^{\theta_1}(x_j) \subsetneq S_Q^{\theta_2}(x_j)$. Thus, $\forall i,\ |S_P^{\theta_1}(x_i)| \le |S_Q^{\theta_2}(x_i)|$ and $\exists\ j$, $|S_P^{\theta_1}(x_j)| < |S_Q^{\theta_2}(x_j)|$.

Hence, $\forall i$,

$$-|S_P^{\theta_1}(x_i)| \log_2 \frac{1}{|S_P^{\theta_1}(x_i)|} = |S_P^{\theta_1}(x_i)| \log_2 |S_P^{\theta_1}(x_i)|$$
$$\le |S_Q^{\theta_2}(x_i)| \log_2 |S_Q^{\theta_2}(x_i)| = -|S_Q^{\theta_2}(x_i)| \log_2 \frac{1}{|S_Q^{\theta_2}(x_i)|},$$

$\exists\ j$,

$$-|S_P^{\theta_1}(x_j)| \log_2 \frac{1}{|S_P^{\theta_1}(x_j)|} = |S_P^{\theta_1}(x_j)| \log_2 |S_P^{\theta_1}(x_j)|$$
$$< |S_Q^{\theta_2}(x_j)| \log_2 |S_Q^{\theta_2}(x_j)| = -|S_Q^{\theta_2}(x_j)| \log_2 \frac{1}{|S_Q^{\theta_2}(x_j)|}.$$

This result shows that $E_r^{\theta_1}(P) < E_r^{\theta_2}(Q)$. □

This theorem indicates that the greater the uncertainty of the available information, the greater the θ-rough entropy. Therefore, we can conclude that the θ-rough entropy proposed in definition 4.15 can be used to evaluate the degree of determination of interval-valued information systems.

Proposition 4.18. *Assume that (U, A) is an interval-valued information system.*

(1) If $P \subseteq Q$ with $P, Q \subseteq A$, then for any $\theta \in (0, 1]$, $E_r^\theta(Q) \preceq E_r^\theta(P)$.
(2) If $0 < \theta_1 \le \theta_2 \le 1$, then for any $P \subseteq A$, $E_r^{\theta_2}(P) \preceq E_r^{\theta_1}(P)$.

Proof. This holds by Theorems 3.8 and 4.17. □

Corollary 4.19. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$. If $P \subseteq Q$, $\theta_1 \le \theta_2$, then*

$$E_r^{\theta_2}(Q) \preceq E_r^{\theta_1}(Q) \preceq E_r^{\theta_1}(P), \quad E_r^{\theta_2}(Q) \preceq E_r^{\theta_2}(P) \preceq E_r^{\theta_1}(P).$$

Proof. This follows from Proposition 4.18. □

Theorem 4.20. *E_r^θ in Definition 4.15 is also an information granulation function under Definition 4.1.*

Proof. (1) Clearly, "non-negativity" holds.
(2) Given $P, Q \subseteq A$ and θ, if $S^\theta(P) = S^\theta(Q)$, then $\forall i$, $S_P^\theta(x_i) = S_Q^\theta(x_i)$.
By Definition 4.15, $E_r^\theta(P) = E_r^\theta(Q)$.
(3) "Monotonicity" follows from Theorem 4.17. □

Definition 4.21. Suppose that (U, A) is an interval-valued information system. Given $\theta \in (0, 1]$, then the θ-information entropy of subsystem (U, P) with respect to θ is defined as

$$H^\theta(P) = -\sum_{i=1}^{n} \frac{|S_P^\theta(x_i)|}{n} \log_2 \frac{|S_P^\theta(x_i)|}{n}.$$

Theorem 4.22. *Assume that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta \in (0, 1]$, then*

$$\log_2 n \le E_r^\theta(P) + H^\theta(P) \le n \log_2 n.$$

Proof.

$$\begin{aligned} E_r^\theta(P) + H^\theta(P) &= -\sum_{i=1}^{n} \frac{|S_P^\theta(x_i)|}{n} \log_2 \frac{1}{|S_P^\theta(x_i)|} - \sum_{i=1}^{n} \frac{|S_P^\theta(x_i)|}{n} \log_2 \frac{|S_P^\theta(x_i)|}{n} \\ &= -\sum_{i=1}^{n} \frac{|S_P^\theta(x_i)|}{n} (\log_2 \frac{1}{|S_P^\theta(x_i)|} + \log_2 \frac{|S_P^\theta(x_i)|}{n}) \\ &= -\sum_{i=1}^{n} \frac{|S_P^\theta(x_i)|}{n} \log_2 \frac{1}{n} \\ &= \frac{\log_2 n}{n} \sum_{i=1}^{n} |S_P^\theta(x_i)|. \end{aligned}$$

Since $1 \le |S_P^\theta(x_i)| \le n$,

$$\log_2 n \le E_r^\theta(P) + H^\theta(P) \le n \log_2 n.$$

□

Corollary 4.23. *Assume that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta \in (0, 1]$, then*

$$0 \le H^\theta(P) \le \log_2 \ n.$$

Proof. Clearly, $0 \le H^\theta(P)$.
By Definition 4.21, $0 \le E_r^\theta(P) \le n \log_2 n$. Then,

$$-n \log_2 n \le -E_r^\theta(P) \le 0.$$

By Theorem 4.22, $\log_2 n \le E_r^\theta(P) + H^\theta(P) \le n \log_2 n$.
Thus,
$H^\theta(P) \le n \log_2 \ n.$ □

5. An application

As an application of the proposed measures, this section presents the rough entropy of a rough set in an interval-valued information system and shows that it is more accurate than the roughness in the same system.

5.1. Accuracy and roughness in an information system

Accuracy and roughness represent the complete and incomplete extent of the knowledge of a given object subset, respectively. They represent the number of elements in each approximation region and can be used to evaluate the uncertainty of the boundary region.

If (U, A) is an information system and $P \subseteq A$, then an equivalence relation $IND(P)$ can be defined by

$$IND(P) = \{(x, y) \in U \times U \mid \forall\ a \in P,\ a(x) = a(y)\}.$$

Definition 5.1. [17] Suppose that (U, A) is an information system. Given $P \subseteq A$, for any $X \in 2^U$,

$$\underline{P}X = \{x \in U \mid [x]_P \subseteq X\}, \quad \overline{P}X = \{x \in U \mid [x]_P \cap X \neq \emptyset\},$$

where $[x]_P = \{y \in U \mid (x, y) \in IND(P)\}$. Then, $\underline{P}(X)$ and $\overline{P}(X)$ are called the lower and upper approximations with respect to P, respectively.

If $\underline{P}X = \overline{P}X$, then X is said to be definable with respect to P; otherwise, X is said to be rough with respect to P.

Pawlak [18] proposed two numerical measures for evaluating the uncertainty of a rough set X: accuracy and roughness. The two measures are defined as

$$\alpha_P(X) = \frac{|\underline{P}X|}{|\overline{P}X|}, \quad \rho_P(X) = 1 - \alpha_P(X).$$

5.2. Limitation of classical measures in interval-valued information systems

Definition 5.2 [5]**.** Let (U, A) be an interval-valued information system. Given $P \subseteq A$ and $\theta \in (0, 1]$, for any $X \in 2^U$, define

$$\underline{P}^\theta(X) = \{x \in U \mid S_P^\theta(x) \subseteq X\}, \quad \overline{P}^\theta X = \{x \in U \mid S_P^\theta(x) \cap X \neq \emptyset\}.$$

Then, $\underline{P}^\theta(X)$ and $\overline{P}^\theta(X)$ are called the θ-lower and θ-upper approximations of X with respect to P, respectively. X is said to be θ-definable with respect to P if $\underline{P}^\theta(X) = \overline{P}^\theta(X)$; otherwise, X is said to be θ-rough with respect to P.

Definition 5.3 [5]**.** Let (U, A) be an interval-valued information system. Given $P \subseteq A$ and $\theta \in (0, 1]$, suppose that X is a θ-rough set with respect to P. Then, the θ-accuracy and θ-roughness of X with respect to P are, respectively, defined as

$$\alpha_P^\theta(X) = \frac{|\underline{P}^\theta X|}{|\overline{P}^\theta X|}, \quad \rho_P^\theta(X) = 1 - \frac{|\underline{P}^\theta X|}{|\overline{P}^\theta X|}.$$

If X is θ-definable with respect to P, then $\rho_P^\theta(X) = 0$; if X is θ-rough with respect to P, then $\rho_P^\theta(X) \neq 0$.

Example 5.4. (Continued from Example 3.1) Select $X = \{x_2, x_5\}$, $P = \{a_1, a_2\}$, $Q = \{a_1, a_2, a_3\}$ and $\theta = 0.4$. Then,

$$\underline{P}^{0.4}(X) = \underline{Q}^{0.4}(X) = \{x_5\}, \ \overline{P}^{0.4}(X) = \overline{Q}^{0.4}(X) = \{x_2, x_5, x_8\};$$

$S^\theta(P) = (\{x_1, x_3, x_6, x_7\}, \{x_2, x_5, x_8\}, \{x_1, x_3, x_7\}, \{x_4, x_9, x_{10}\}, \{x_2, x_5\}, \{x_1, x_6, x_7, x_8\}, \{x_1, x_3, x_6, x_7, x_8\}, \{x_2, x_6, x_7, x_8\}, \{x_4, x_9, x_{10}\}, \{x_4, x_9, x_{10}\})$,

$S^\theta(Q) = (\{x_1, x_7\}, \{x_2, x_5, x_8\}, \{x_3\}, \{x_4\}, \{x_2, x_5\}, \{x_6, x_8\}, \{x_1, x_7\}, \{x_2, x_6, x_8\}, \{x_9, x_{10}\}, \{x_9, x_{10}\})$.

Clearly,

$\underline{P}^\theta(X) \neq \overline{P}^\theta(X)$, $\underline{Q}^\theta(X) \neq \overline{Q}^\theta(X)$.

Then, X is a θ-rough set with respect to both P and Q.

Thus,

$$\rho_P^\theta(X) = 0.67 = \rho_Q^\theta(X).$$

Note that $S^\theta(Q) \prec S^\theta(P)$. Then, $S^\theta(P)$ depends strictly on $S^\theta(Q)$. By Theorem 4.4(2), $G^\theta(Q) < G^\theta(P)$. Then, the uncertainty of subsystem (U, P) with respect to θ is larger than that of subsystem (U, Q) with respect to θ. However, the θ-roughness of rough set X with respect to P is equal to the θ-roughness of X with respect to Q, i.e., the roughness of X in subsystem (U, P) with respect to θ is equal to the roughness of X in subsystem (U, Q) with respect to θ. Thus, X has the same roughness in subsystems (U, P) and (U, Q) with respect to θ. Therefore, a new and more accurate uncertainty measure is required for rough sets in interval-valued information systems.

5.3. Rough entropy of a rough set in interval-valued information systems

Definition 5.5. Let (U, A) be an interval-valued information system, and let $P \subseteq A$. Suppose $X \in 2^U$ is θ-rough with respect to P. Then, the θ-rough entropy of rough set X with respect to P is defined as

$$RE_P^\theta(X) = \rho_P^\theta(X) G^\theta(P).$$

Proposition 5.6. *Suppose that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta \in (0, 1]$, then*

$$0 \le RE_P^\theta(X) \le n.$$

Moreover, if $X = U$, then for any $P \subseteq A$ and $\theta \in (0, 1]$, RE_P^θ achieves the minimum value 0; if $X = \emptyset$ and S_P^θ is a universal relation on U, then RE_P^θ achieves the maximum value n.

Proof. Since for each i, $1 \le |S_P^\theta(x_i)| \le n$, we have $0 \le \rho_P^\theta(X) \le 1$ and $\frac{1}{n} \le G^\theta(P) \le n$.

By Definition 5.1,

$$0 \le RE_P^\theta(X) \le n.$$

If $X = U$, then $\rho_P^\theta(U) = 0$. Therefore, $RE_P^\theta(U) = 0$.

If $X = \emptyset$ and S_P^θ is a universal relation on U, then $\rho_P^\theta(\emptyset) = 1$ and $G^\theta(P) = n$. Therefore, $RE_P^\theta(\emptyset) = n$. □

Theorem 5.7. *Assume that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta_1, \theta_2 \in (0, 1]$:*

(1) If $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$, then $RE_P^{\theta_1}(X) \le RE_Q^{\theta_2}(X)$.

(2) If $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$, then $RE_P^{\theta_1}(X) < RE_Q^{\theta_2}(X)$.

Proof. (1) The proof is trivial.

(2) Since $S^{\theta_1}(P) \prec S^{\theta_2}(Q)$, we have $S^{\theta_1}(P) \preceq S^{\theta_2}(Q)$. Then, for $\forall$ i, $S_P^{\theta_1}(x_i) \subseteq S_Q^{\theta_2}(x_i)$, such that $\underline{P}^{\theta_1}(X) \ge \underline{Q}^{\theta_2}(X)$ and $\overline{P}^{\theta_1}(X) \le \overline{Q}^{\theta_2}(X)$. Then, $\rho_P^{\theta_1}(X) \le \rho_Q^{\theta_2}(X)$. By Theorem 4.4, $G^\theta(P) < G^\theta(Q)$. Then, $RE_P^\theta(X) < RE_Q^\theta(X)$. □

Corollary 5.8. *Assume that (U, A) is an interval-valued information system. Given $P, Q \subseteq A$ and $\theta_1, \theta_2 \in (0, 1]$, if $S^{\theta_1}(P) = S^{\theta_2}(Q)$, then $RE_P^{\theta_1}(X) = RE_Q^{\theta_2}(X)$.*

Proof. This holds by Theorem 4.5. □

Proposition 5.9. *Assume that (U, A) is an interval-valued information system.*

(1) If $P \subseteq Q$ with $P, Q \subseteq A$, then for any $\theta \in (0, 1]$, $RE_Q^\theta(X) \le RE_P^\theta(X)$.

(2) If $0 < \theta_1 \le \theta_2 \le 1$, then for any $P \subseteq A$, $RE_P^{\theta_2}(X) \le RE_P^{\theta_1}(X)$.

Proof. This holds by Theorem 3.8 and Proposition 5.7 (1). □

Corollary 5.10. *Let (U, A) be an interval-valued information system. Given $\theta_1, \theta_2 \in (0, 1]$ and $P, Q \subseteq A$, if $P \subseteq Q$, $\theta_1 \le \theta_2$, then*

$$RE_Q^{\theta_2}(X) \le RE_Q^{\theta_1}(X) \le RE_P^{\theta_1}(X), \quad RE_Q^{\theta_2}(X) \le RE_P^{\theta_2}(X) \le RE_P^{\theta_1}(X).$$

Proof. This follows from Proposition 5.9. □

Example 5.11 (Continued from Example 5.4)**.** We have

$$\rho_P^\theta(X) = 0.67 = \rho_Q^\theta(X).$$

By Definition 4.2, $G_P^\theta = 1.22$ *and* $G_Q^\theta = 0.44$.

Then, $RE_P^\theta(X) = 0.817$, *and* $RE_Q^\theta(X) = 0.295$.

Thus,

$$G_P^\theta > G_Q^\theta,$$

and

$$RE_P^\theta(X) > RE_Q^\theta(X).$$

This example illustrates that the rough entropy of a rough set in a subsystem with respect to a given parameter is more accurate than is its roughness in the same subsystem with respect to the same parameter in an interval-valued information system. Clearly, this subsystem can take into account the whole system.

6. Numerical experiments and effectiveness analysis

6.1. Numerical experiments

Example 6.1. To demonstrate the performance of the uncertainty measurement, we construct a numerical experiment on the Face recognition dataset (see Table 1 in [2] or Table 2 in [5]). This dataset consists of the 27 observations shown in Table 2. There are 6 attributes, including the length spanned by the eyes, the length between the eyes, and the length from the outer right eye to the upper middle lip at the point between the nose and mouth. There are two groups of experiments.

In the first experiment, we compare four tools for measuring the uncertainty of subsystems.

Table 2
Face recognition dataset.

Subject	a_1	a_2	a_3	a_4	a_5	a_6
FRA1	[155.00,157.00]	[58.00,61.01]	[100.45,103.28]	[105.00,107.30]	[61.40,65.73]	[64.20,67.80]
FRA2	[154.00,160.01]	[57.00,64.00]	[101.98,105.55]	[104.35,107.30]	[60.88,63.03]	[62.94,66.47]
FRA3	[154.01,161.00]	[57.00,63.00]	[99.36,105.65]	[101.04,109.04]	[60.95,65.60]	[60.42,66.40]
HUS1	[168.86,172.84]	[58.55,63.39]	[102.83,106.53]	[122.38,124.52]	[56.73,61.07]	[60.44,64.54]
HUS2	[169.85,175.03]	[60.21,64.38]	[102.94,108.71]	[120.24,124.52]	[56.73,62.37]	[60.44,66.84]
HUS3	[168.76,175.15]	[61.40,63.51]	[104.35,107.45]	[120.93,125.18]	[57.20,61.72]	[58.14,67.08]
INC1	[155.26,160.45]	[53.15,60.21]	[95.88,98.49]	[91.68,94.37]	[62.48,66.22]	[58.90,63.13]
INC2	[156.26,161.31]	[51.09,60.07]	[95.77,99.36]	[91.21,96.83]	[54.92,64.20]	[54.41,61.55]
INC3	[154.47,160.31]	[55.08,59.03]	[93.54,98.98]	[90.43,96.43]	[59.03,65.86]	[55.97,65.80]
ISA1	[164.00,168.00]	[55.01,60.03]	[120.28,123.04]	[117.52,121.02]	[54.38,57.45]	[50.80,53.25]
ISA2	[163.00,170.00]	[54.04,59.00]	[118.80,123.04]	[116.67,120.24]	[55.47,58.67]	[52.43,55.23]
ISA3	[164.01,169.01]	[55.00,59.01]	[117.38,123.11]	[116.67,122.43]	[52.80,58.31]	[52.20,55.47]
JPL1	[167.11,171.19]	[61.03,65.01]	[118.23,121.82]	[108.30,111.20]	[63.89,67.88]	[57.28,60.83]
JPL2	[169.14,173.18]	[60.07,65.07]	[118.85,120.88]	[108.98,113.17]	[62.63,69.07]	[57.38,61.62]
JPL3	[169.03,170.11]	[59.01,65.01]	[115.88,121.38]	[110.34,112.49]	[61.72,68.25]	[59.46,62.94]
KHA1	[149.34,155.54]	[54.15,59.14]	[111.95,115.75]	[105.36,111.07]	[54.20,58.14]	[48.27,50.61]
KHA2	[149.34,155.32]	[52.04,58.22]	[111.20,113.22]	[105.36,111.07]	[53.71,58.14]	[49.41,52.80]
KHA3	[150.33,157.26]	[52.09,60.21]	[109.04,112.70]	[104.74,111.07]	[55.47,60.03]	[49.20,53.41]
LOT1	[152.64,157.62]	[51.35,56.22]	[116.73,119.67]	[114.62,117.41]	[55.44,59.55]	[53.01,56.60]
LOT2	[154.64,157.62]	[52.24,56.32]	[117.52,119.67]	[114.28,117.41]	[57.63,60.61]	[54.41,57.98]
LOT3	[154.83,157.81]	[50.36,55.23]	[117.59,119.75]	[114.04,116.83]	[56.64,61.07]	[55.23,57.80]
PHI1	[163.08,167.07]	[66.03,68.07]	[115.26,119.60]	[116.10,121.02]	[60.96,65.30]	[57.01,59.82]
PHI2	[164.00,168.03]	[65.03,68.12]	[114.55,119.60]	[115.26,120.97]	[60.96,67.27]	[55.32,61.52]
PHI3	[161.01,167.00]	[64.07,69.01]	[116.67,118.79]	[114.59,118.83]	[61.52,68.68]	[56.57,60.11]
ROM1	[167.15,171.24]	[64.07,68.07]	[123.75,126.59]	[122.92,126.37]	[51.22,54.64]	[49.65,53.71]
ROM2	[168.15,172.14]	[63.13,68.07]	[122.33,127.29]	[124.08,127.14]	[50.22,57.14]	[49.93,56.94]
ROM3	[167.11,171.19]	[63.13,68.03]	[121.62,126.57]	[122.58,127.78]	[49.41,57.28]	[50.99,60.46]

Table 3
Four θ-measure values of the uncertainty for (U, P_i) on the Face recognition dataset.

i	$G^\theta(P_i)$	$H^\theta(P_i)$	$E_r^\theta(P_i)$	$E^\theta(P_i)$
1	2.550	13.638	24.578	5.487
2	0.896	11.381	10.633	3.734
3	0.306	9.214	4.347	2.546
4	0.295	9.037	4.171	2.483
5	0.281	8.889	3.967	2.423
6	0.257	8.565	3.586	2.299

(1) We compare four tools for measuring the uncertainty of systems with different P. Select $P_1 = \{a_1\}$, $P_2 = \{a_1, a_2\}$, $P_3 = \{a_1, a_2, a_3\}$, $P_4 = \{a_1, a_2, a_3, a_4\}$, $P_5 = \{a_1, a_2, a_3, a_4, a_5\}$, $P_6 = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ and $\theta = 0.4$.

By Definitions 4.1, 4.8, 4.15 and 4.21, we obtain Table 3.

The experimental results are shown in Fig. 1. The uncertainty measures of the interval-valued information system (U, A) show certain regularity with growth in attribute subset $P \subseteq A$, which is reflected in the following facts:

1) G^θ, E_r^θ and E^θ are monotonically decreasing with attributes increases;

2) E_r^θ is more sensitive than G^θ, H^θ and E^θ;

3) When i in $\{a_1, a_2, \cdots, a_i\}$ reaches 3, the values of the four uncertainty measurements of the interval-valued information system (U, A) tend to constant values;

4) The differences among G^θ, H^θ and E^θ are almost identical.

Thus, θ-rough entropy is more suitable than other measures for the uncertainty measurement of interval-valued information systems.

(2) We compare four tools for measuring the uncertainty of systems with different θ. Select $\theta_1 = 0.1, \theta_2 = 0.2, \theta_3 = 0.3, \theta_4 = 0.4, \theta_5 = 0.5, \theta_6 = 0.6, \theta_7 = 0.7, \theta_8 = 0.8, \theta_9 = 0.9$ and $P = A$.

By Definitions 4.1, 4.8, 4.15 and 4.21, we obtain Table 4.

The experimental results are shown in Fig. 2. As the threshold θ increases, the uncertainty measures of interval-valued information system (U, A) show a certain regularity, which is reflected in the following facts:

1) G^θ, E_r^θ and E^θ monotonically decreases as the threshold θ increases;

2) H^θ and E_r^θ are almost identical;

3) H^θ and E_r^θ are more sensitive than G^θ and E^θ.

Thus, the changes in the θ-information entropy and θ-rough entropy are closely related to θ, and the change in θ-information granulation is not closely related to θ.

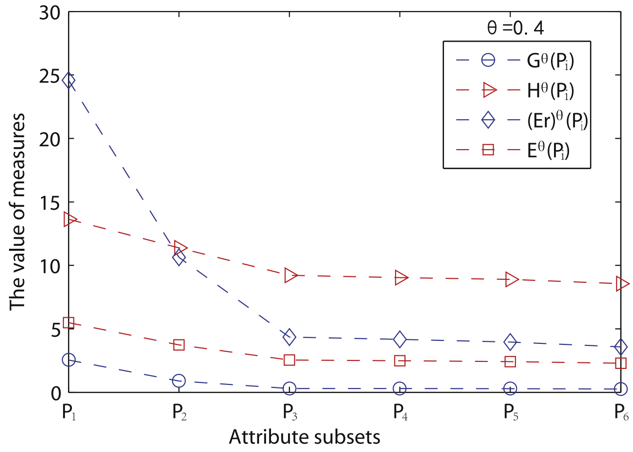

Fig. 1. The change of four tools for measuring uncertainty with different subsystems.

Table 4
Four θ_i-measure values of the uncertainty for (U, P) on the Face recognition dataset.

i	$G^{\theta_i}(P)$	$H^{\theta_i}(P)$	$E_r^{\theta_i}(P)$	$E^{\theta_i}(P)$
1	0.333	9.510	4.755	2.667
2	0.333	9.510	4.755	2.667
3	0.306	9.214	4.347	2.546
4	0.257	8.565	3.586	2.299
5	0.193	7.768	2.622	1.992
6	0.128	6.712	1.565	1.613
7	0.062	5.367	0.444	1.161
8	0.045	4.959	0.148	1.029
9	0.037	4.755	0	0.963

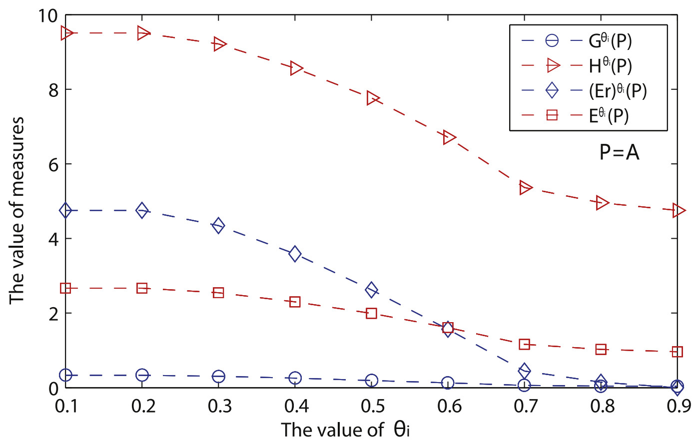

Fig. 2. The change of four tools for measuring uncertainty with different θ.

In the second experiment, we compare $\rho_P^\theta(X)$ and $RE_P^\theta(X)$.
(1) We compare $\rho_P^\theta(X_i)$ and $RE_P^\theta(X_i)$.
Let $P = A$ and $\theta = 0.4$. Select
$X_1 = \{x_1, x_2\}$, $X_2 = \{x_1, x_2, \cdots, x_4\}$, $X_3 = \{x_1, x_2, \cdots, x_6\}$, $X_4 = \{x_1, x_2, \cdots, x_8\}$, $X_5 = \{x_1, x_2, \cdots, x_{10}\}$, $X_6 = \{x_1, x_2, \cdots, x_{12}\}$, $X_7 = \{x_1, x_2, \cdots, x_{14}\}$, $X_8 = \{x_1, x_2, \cdots, x_{16}\}$, $X_9 = \{x_1, x_2, \cdots, x_{18}\}$, $X_{10} = \{x_1, x_2, \cdots, x_{20}\}$, $X_{11} = \{x_1, x_2, \cdots, x_{22}\}$, $X_{12} = \{x_1, x_2, \cdots, x_{24}\}$, $X_{13} = \{x_1, x_2, \cdots, x_{26}\}$, $X_{14} = \{x_1, x_2, \cdots, x_{27}\}$.
Fig. 3 shows that these two measures can be used to evaluate the uncertainty of interval-valued information systems. Furthermore, when the object subset X is not fixed, the curve of RE_P^θ is smoother than the curve of ρ_P^θ, which indicates that RE_P^θ is better than ρ_P^θ.
(2) We compare $\rho_{P_i}^\theta(X^*)$ and $RE_{P_i}^\theta(X^*)$.

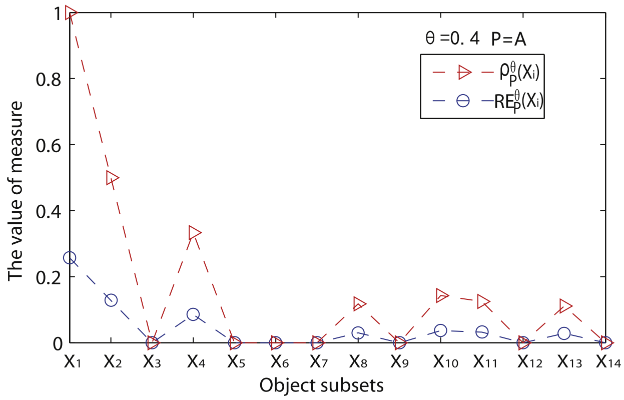

Fig. 3. The results of $\rho_P^\theta(Xi)$ and $RE_P^\theta(Xi)$ with different i.

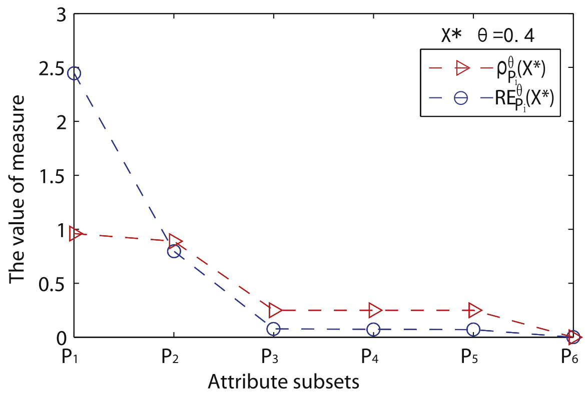

Fig. 4. The results of $\rho_{P_i}^\theta(X^*)$ and $RE_R^\theta(Xi^*)$ with different i.

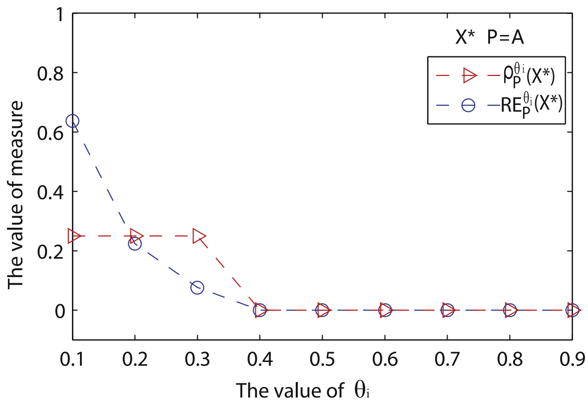

Fig. 5. The results of $\rho_P^\theta(X^*)$ and $RE_P^\theta(X^*)$ with different i.

Let $X^* = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ and $\theta = 0.4$. Select
$P_1 = \{a_1\}$, $P_2 = \{a_1, a_2\}$, $P_3 = \{a_1, a_2, a_3\}$, $P_4 = \{a_1, a_2, a_3, a_4\}$, $P_5 = \{a_1, a_2, a_3, a_4, a_5\}$, $P_6 = \{a_1, a_2, a_3, a_4, a_5, a_6\}$.

Fig. 4 shows that $\rho_P^\theta(X^*)$ and $RE_P^\theta(X^*)$ both decrease as the attribute subset grows. Furthermore, RE_P^θ is more sensitive than ρ_P^θ to growth in the attribute subset, which indicates that RE_P^θ is better than ρ_P^θ.

(3) We compare $\rho_P^{\theta_i}(X^*)$ and $RE_P^{\theta_i}(X^*)$.

Let $X^* = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ and $P = A$. Select
$\theta_1 = 0.1, \theta_2 = 0.2, \theta_3 = 0.3, \theta_4 = 0.4, \theta_5 = 0.5, \theta_6 = 0.6, \theta_7 = 0.7, \theta_8 = 0.8, \theta_9 = 0.9,$

Fig. 5 shows that $\rho_P^\theta(X^*)$ and $RE_P^\theta(X^*)$ decrease as the threshold θ increases. Additionally, when $\theta > 0.4$, $\rho_P^\theta(X^*) = RE_P^\theta(X^*) = 0$; when $\theta < 0.4$, RE_P^θ is more sensitive than ρ_P^θ, which indicates that RE_P^θ is better than ρ_P^θ.

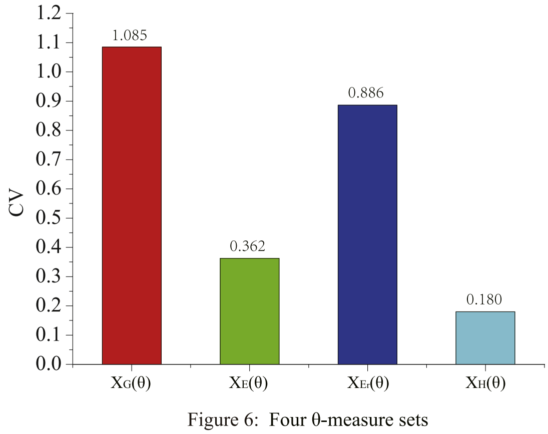

Fig. 6. Four θ-measure sets.

6.2. Effectiveness analysis

In this section, we conduct a numerical experiment and perform a statistical effectiveness analysis from three aspects: dispersion, correlation and variance.

6.2.1. Analysis of dispersion

In statistical analysis, we often research the degree of dispersion of a dataset. A metric used to measure the dispersion degree of a dataset is called a difference measure. Common difference measures include the range, four point difference, average difference, standard deviation, and standard deviation coefficient. In this paper, we apply the standard deviation coefficient to perform the effectiveness analysis of the proposed measures.

Given dataset $X = \{x_1, \cdots, x_n\}$, its arithmetic average $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, its standard deviation $\sigma(X) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$, and its standard deviation coefficient

$$CV(X) = \frac{\sigma(X)}{\bar{x}}.$$

Four θ-measure sets on the Face recognition dataset are defined as follows:

$$X_G(\theta) = \{G^\theta(P_1), \cdots, G^\theta(P_6)\},\ X_E(\theta) = \{E^\theta(P_1), \cdots, E^\theta(P_6)\},$$

$$X_{E_r}(\theta) = \{E_r^\theta(P_1), \cdots, E_r^\theta(P_6)\},\ X_H(\theta) = \{H^\theta(P_1), \cdots, H^\theta(P_6)\}.$$

We compare the CV of the four θ-measure sets. The experimental results are shown in Fig. 6.

$CV(X_G(\theta)) > CV(X_{E_r}(\theta)) > CV(X_E(\theta)) > CV(X_H(\theta))$ means the dispersion degree of H is the minimum.

Four P-measure sets on the Face recognition dataset are defined as follows:

$$X_G(P) = \{G^{\theta_1}(P), \cdots, G^{\theta_9}(P)\},\ X_E(P) = \{E^{\theta_1}(P), \cdots, E^{\theta_9}(P)\},$$

$$X_{E_r}(P) = \{E_r^{\theta_1}(P), \cdots, E_r^{\theta_9}(P)\},\ X_H(P) = \{H^{\theta_1}(P), \cdots, H^{\theta_9}(P)\}.$$

We compare the CV of four P-measure sets. The experimental results are shown in Fig. 7.

$CV(X_{E_r}(P)) > CV(X_G(P)) > CV(X_E(P)) > CV(X_H(P))$ means the dispersion degree of H is the minimum.

Thus, we obtain the following results:

(1) if we require only monotonicity, then G^θ, $(E_r)^\theta$ and E^θ can be applied to measure the uncertainty of an interval-valued information system;

(2) if we consider only the dispersion degree, then H^θ has better performance for measuring the uncertainty of an interval-valued information system;

(3) if we pay are concerned with both monotonicity and the dispersion degree, then E^θ has much the best performance for measuring the uncertainty of an interval-valued information system.

6.2.2. Analysis of correlation

In statistics, the Pearson correlation coefficient is used to measure the strength of a linear correlation between two variables or two datasets.

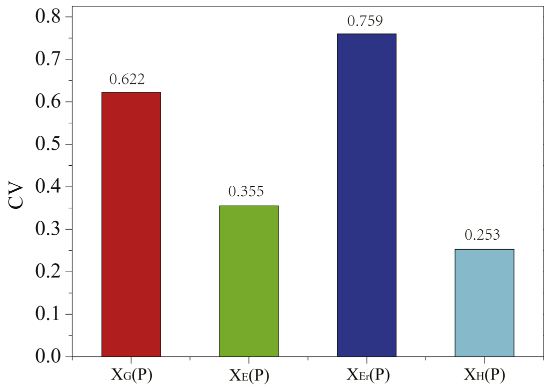

Fig. 7. Four P-measure sets.

Table 5
r-values of sixteen pairs of four θ-measure sets on the Face recognition dataset.

r	$X_G(\theta)$	$X_E(\theta)$	$X_{E_r}(\theta)$	$X_H(\theta)$
$X_G(\theta)$	1			
$X_E(\theta)$	0.997	1		
$X_{E_r}(\theta)$	0.9985	0.9942	1	
$X_H(\theta)$	0.9662	0.995	0.9786	1

Table 6
The correlation between two θ-measures.

	G^θ	E^θ	${E_r}^\theta$	H^θ
G^θ	CPC			
E^θ	HPC	CPC		
${E_r}^\theta$	HPC	HPC	CPC	
H^θ	HPC	HPC	HPC	CPC

Given two datasets $X = \{x_1, \cdots, x_n\}$ and $Y = \{y_1, \cdots, y_n\}$, the Pearson correlation coefficient between X and Y, denoted by $r(X, Y)$, is defined as

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

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

Clearly,

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

If $r(X, Y) = 0$, then there is no correlation between X and Y; if $r(X, Y) > 0$, then the correlation between X and Y is positive; if $r(X, Y) < 0$, then the correlation between X and Y is negative. Specifically, $r(X, Y) = 1$ indicates complete positive correlation between X and Y, and $r(X, Y) = -1$ represents complete negative correlation between X and Y.

The closer $|r|$ is to 0, the smaller the degree of correlation between variables; conversely, the closer the absolute value of the Pearson correlation coefficient r is to 1, the greater the degree of correlation between variables. Generally, the degree of correlation can be classified as follows: $|r| = 1$ is called complete correlation; $0.7 \le |r| < 1$ is called high correlation; $0.4 \le |r| < 0.7$ is called moderate correlation; $0 < |r| < 0.4$ is called low correlation; and $r = 0$ is called no correlation.

The Pearson correlation coefficients on θ-measure sets are shown in Table 5.

$r(X_G(\theta), X_E(\theta)) > 0.7$ indicate high positive correlation between $X_G(\theta)$ and $X_E(\theta)$. Therefore, G^θ and E^θ are highly positively correlated. Similarly, G^θ and E_r^θ, G^θ and H^θ, E^θ and E_r^θ, E^θ and H^θ, and E_r^θ and H^θ are highly positively correlated. The conclusions are shown in Table 6, where "HPC" and "CPC" indicate "high positive correlation" and "complete positive correlation", respectively.

The Pearson correlation coefficients on P-measure sets are shown in Table 7.

$r(X_G(P), X_E(P)) > 0.7$ indicates high positive correlation between $X_G(P)$ and $X_E(P)$. Therefore, G^θ and E^θ are highly positively correlated. Similarly, G^θ and E_r^θ, G^θ and H^θ, E^θ and E_r^θ, E^θ and H^θ, and E_r^θ and H^θ are highly positively correlated. The conclusions are shown in Table 8, where "HPC" and "CPC" indicate "high positive correlation" and "complete positive correlation", respectively.

Table 7
r-values of sixteen pairs of four P-measure sets on the Face recognition dataset.

r	$X_G(P)$	$X_E(P)$	$X_{E_r}(P)$	$X_H(P)$
$X_G(P)$	1			
$X_E(P)$	0.9974	1		
$X_{E_r}(P)$	0.9998	0.9987	1	
$X_H(P)$	0.9949	0.9996	0.9968	1

Table 8
The correlation between two θ-measures.

	G^θ	E^θ	$E_r{}^\theta$	H^θ
G^θ	CPC			
E^θ	HPC	CPC		
$E_r{}^\theta$	HPC	HPC	CPC	
H^θ	HPC	HPC	HPC	CPC

Table 9
Calculation table for the sum of squares of the four sets of data.

x/R	R_G	R_E	R_{Er}	R_H	ΣR	$(\Sigma R)^2$
x_{P_1}	2.55	5.487	24.578	13.638	46.253	2139.340
x_{P_2}	0.896	3.734	10.633	11.381	26.644	709.903
x_{P_3}	0.306	2.546	4.347	9.214	16.413	269.387
x_{P_4}	0.295	2.483	4.171	9.037	15.986	255.552
x_{P_5}	0.281	2.423	3.967	8.889	15.56	242.114
x_{P_6}	0.257	2.299	3.586	8.565	14.707	216.296
Σx	4.585	18.972	51.282	60.724	135.563	3832.591
Σx^2	7.631	67.854	782.029	634.461	1491.974	
$(\Sigma x)^2/n$	3.504	59.989	438.307	614.567	1116.368	
n	6	6	6	6	$N = 24$	
$\bar{x}$	0.764	3.162	8.547	10.121		

6.2.3. Analysis of variance

Analysis of variance (for short, ANOVA) is a collection of statistical models and their associated procedures (such as "variation" among and between groups) used to analyse the differences among group means. ANOVA was developed by statistician and evolutionary biologist Ronald Fisher. In its simplest form, ANOVA provides a statistical test of whether the means of several groups are equal and therefore generalizes the t-test to more than two groups. ANOVA is useful for comparing (testing) three or more means (groups or variables) for statistical significance. ANOVA is conceptually similar to multiple two-sample t-tests and is therefore suited to a wide range of practical problems.

We conduct a comprehensive F test for the ANOVA to verify whether there are significant differences in the effects of the proposed measures. The test steps are as follows:

Step 1 Propose the hypothesis.

$H_0: \ \mu_1 = \mu_2 = \mu_3 = \mu_4,$

$H_1: \ \textit{At least two population averages are not equal}.$

Step 2 Calculate the value of the F test statistic.

(1) Calculate the sums of square according to the data in Table 9.

Total sum of squares

$$SS_t = \sum\sum x^2 - \frac{(\sum\sum x)^2}{N} = 1491.974 - \frac{135.563^2}{24} = 726.253,$$

Inter-group sum of squares

$$SS_b = \sum\left[\frac{(\sum x)^2}{n}\right] - \frac{(\sum\sum x)^2}{N} = 1116.368 - \frac{135.563^2}{24} = 350.646,$$

Block sum of squares

$$SS_r = \sum\left[\frac{(\sum R)^2}{K}\right] - \frac{(\sum\sum x)^2}{N} = \frac{3832.591}{4} - \frac{135.563^2}{24} = 192.426,$$

where $K = 4$ is the number of groups.

Sum of squares of errors

$$SS_e = SS_t - SS_b - SS_r = 183.181.$$

(2) Calculate the degrees of freedom
Total degrees of freedom

$$df_t = N - 1 = 24 - 1 = 23,$$

Inter-group degrees of freedom

$$df_b = K - 1 = 4 - 1 = 3,$$

Block degrees of freedom

$$df_r = n - 1 = 6 - 1 = 5,$$

Error degree of freedom

$$df_e = 23 - 3 - 5 = 15.$$

(3) Calculate the variance
Inter-group variance

$$MS_b = \frac{SS_b}{df_b} = \frac{350.646}{3} = 116.882,$$

Block variance

$$MS_r = \frac{SS_r}{df_r} = \frac{192.426}{5} = 38.485,$$

Error variance

$$MS_e = \frac{SS_e}{df_e} = \frac{183.181}{15} = 12.212.$$

(4) Calculate the F value

$$F = \frac{MS_b}{MS_e} = \frac{116.882}{12.212}.$$

Step 3 Statistical decision.

$df_b = 3,\ df_e = 15,\ \alpha = 0.01$. Check the F-value table, $F_{(3,15)0.01} = 5.42$. The actual value of the calculated F test statistic $F = 9.571$. Since $F = 9.571 > F_{(3,15)0.01} = 5.42$, we have $P(F > 5.42) < 0.01$.

Note that the value of the sample statistic falls within the rejection region. Thus, the null hypothesis H_0 is rejected, and the alternative hypothesis H_1 is accepted. Therefore, at least two of the overall averages in the four θ-measure sets are not equal. Thus, there are significant differences in the effects of the proposed measures.

7. Conclusions

In this paper, we have proposed four measuring tools (i.e., θ-information granulation, θ-information amount, θ-rough entropy and θ-information entropy) to evaluate the uncertainty of a given interval-valued information system by means of its information structure. As an application of information granulation, we have proposed the rough entropy of a rough set in interval-valued information systems. We have presented a numerical experiment on the Face recognition dataset and conducted a statistical effectiveness analysis from three aspects, namely, dispersion analysis, correlation analysis and variance analysis, to demonstrate the feasibility of the proposed measures. The measures proposed in this paper can be applied to data mining from interval data. In the future, we will consider additional applications of the proposed measures.

Acknowledgements

The authors would like to thank the editors and the anonymous reviewers for their valuable comments and suggestions which have helped immensely in improving the quality of the paper. This work is supported by National Natural Science Foundation of China (11461005), Natural Science Foundation of Guangxi Province (2016GXNSFAA380045, 2016GXNSFAA380282, 2016GXNSFAA380286), Undergraduate Teaching Reform Project of Guangxi (2017JGA179), Key Laboratory of Optimization Control and Engineering Calculation in Department of Guangxi Education and Special Funds of Guangxi Distinguished Experts Construction Engineering.

References

[1] R. Al-Hmouz, W. Pedrycz, A. Balamash, D. Computing, Description and prediction of time series: a general framework of Granular Computing, Expert Syst. Appl. 42 (10) (2015) 4830–4839.
[2] L. Billard, A. Douzal-Chouakria, E. Diday, Symbolic Principal Component for Interval-Valued Observations, University of Georgia, USA, 2007 Technical report.
[3] J.H. Dai, Rough 3-valued algebras, Inf. Sci. 178 (2008) 1986–1996.
[4] J.H. Dai, H.W. Tian, Entropy measures and granularity measures for set-valued information systems, Inf. Sci. 240 (2013) 72–82.

[5] J.H. Dai, W.T. Wang, J.S. Mi, Uncertainty measurement for interval-valued information systems, Inf. Sci. 251 (2013) 63–78.
[6] J.H. Dai, W.T. Wang, Q. Xu, H.W. Tian, Uncertainty measurement for interval-valued decision systems based on extended conditional entropy, Knowl. Based Syst. 27 (2012) 443–450.
[7] J.H. Dai, B.J. Wei, X.H. Zhang, Q.L. Zhang, Uncertainty measurement for incomplete interval-valued information systems based on α-weak similarity, Knowl. Based Syst. 136 (2017) 159–171.
[8] G. Facchinetti, R. Ricci, S. Muzzioli, Note on ranking fuzzy triangular numbers, Int. J. Intell. Syst. 13 (1998) 613–622.
[9] X.C. Hu, W. Pedrycz, X.M. Wang, Fuzzy classifiers with information granules in feature space and logic-based computing, Pattern Recognit. 80 (2018) 156–167.
[10] Y. Leung, M.M. Fischer, W.Z. Wu, J.S. Mi, A rough set approach for the discovery of classification rules in interval-valued information systems, Int. J. Approx. Reason. 47 (2008) 233–246.
[11] V. Loia, F. Orciuoli, W. Pedrycz, Towards a granular computing approach based on formal concept analysis for discovering periodicities in data, Knowl. Based Syst. 146 (2018) 1–11.
[12] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Sci. China (Series F) 51 (2008) 1427–1444.
[13] J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wierman, The information entropy, rough entropy and knowledge granulation in incomplete information systems, Int. J. Gen. Syst. 35 (6) (2006) 641–654.
[14] J.M. Ma, W.X. Zhang, Y. Leung, X.X. Song, Granular computing and dual Galois connection, Inf. Sci. 177 (2007) 5365–5377.
[15] Y. Nakahara, User oriented ranking criteria and its application to fuzzy mathematical programming problems, Fuzzy Sets Syst. 94 (1998) 275–286.
[16] Y. Nakahara, M. Sasaki, M. Gen, On the linear programming problems with interval coefficients, Comput. Ind. Eng. 23 (1992) 301–304.
[17] Z. Pawlak, Rough sets, Int. J. Comput. Inf.Sci. 11 (1982) 341–356.
[18] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[19] W. Pedrycz, From fuzzy data analysis and fuzzy regression to granular fuzzy data analysis, Fuzzy Sets Syst. 274 (2015) 12–17.
[20] W. Pedrycz, W. Homenda, Building the fundamentals of granular computing: a principle of justifiable granularity, Appl. Soft Comput. 13 (10) (2013) 4209–4218.
[21] A. Pedrycz, K. Hirota, W. Pedrycz, F.Y. Dong, Granular representation and granular computing with fuzzy sets, Fuzzy Sets Syst. 203 (2012) 17–32.
[22] Z. Pawlak, A. Skowron, Rudiments of rough sets, Inf. Sci. 177 (2007) 3–27.
[23] Z. Pawlak, A. Skowron, Rough sets: some extensions, Inf. Sci. 177 (2007) 28–40.
[24] Z. Pawlak, A. Skowron, Rough sets and boolean reasoning, Inf. Sci. 177 (2007) 41–73.
[25] Y.H. Qian, J.Y. Liang, C.Y. Dang, Interval ordered information systems, Comput. Math. Appl. 56 (2008) 1994–2009.
[26] Y.H. Qian, Y.B. Li, J.Y. Liang, G.P. Lin, C.Y. Dang, Fuzzy granular structure distance, IEEE Trans. Fuzzy Syst. 23 (6) (2015) 2245–2259.
[27] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Information granularity in fuzzy binary GRC model, IEEE Trans. Fuzzy Syst. 19 (2) (2011) 253–264.
[28] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423.
[29] H. Sakai, M. Nakata, D. Slezak, A prototype system for rule generation in lipskis incomplete information databases, in: Proceedings of 13th Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, 2011, pp. 175–182.
[30] Q. Wu, Z.T. Liu, Real formal concept analysis based on grey-rough set theory, Knowl. Based Syst. 22 (2009) 38–45.
[31] W.Z. Wu, Y. Leung, J.S. Mi, Granular computing and knowledge reduction in formal contexts, IEEE Trans. Knowl. Data Eng. 21 (10) (2009) 1461–1474.
[32] H. Wang, H.B. Yue, Entropy measures and granularity measures for interval and set-valued information systems, Soft Comput. 20 (2016) 3489–3495.
[33] W.H. Xu, X.Y. Zhang, W.X. Zhang, Knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems, Appl. Soft Comput. 9 (2009) 1244–1251.
[34] Y.Y. Yao, X.N. Li, Comparison of rough-set and interval-set models for uncertain reasoning, Fundam. Inform. 27 (1996) 289–298.
[35] X.B. Yang, D.J. Yu, J.Y. Yang, L.H. Wei, Dominance-based rough set approach to incomplete interval-valued information system, Data Knowl. Eng. 68 (2009) 1331–1347.
[36] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[37] L.A. Zadeh, Fuzzy logic equals computing with words, IEEE Trans. Fuzzy Syst. 4 (1996) 103–111.
[38] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets Syst. 90 (1997) 111–127.
[39] L.A. Zadeh, A new direction in AI-toward a computational theory of perceptions, AI Mag. 22 (1) (2001) 73–84.
[40] 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. 2 (1998) 23–25.
[41] G.Q. Zhang, Z.W. Li, W.Z. Wu, X.F. Liu, N.X. Xie, Information structures and uncertainty measures in a fully fuzzy information system, Int. J. Approx. Reason. 101 (2018) 119–149.
[42] C.F. Zhong, W. Pedrycz, D. Wang, L.N. Li, Z.W. Li, Granular data imputation: a framework of granular computing, Appl. Soft Comput. 46 (2016) 307–316.
[43] L. Zhang, B. Zhang, Theory and Application of Problem Solving-Theory and Application of Granular Computing in Quotient Spaces, Tsinghua University Publishers, Beijing, 2007.