Measures of Uncertainty for a Four-Hybrid Information System and Their Applications

Bin Qin (binqin8846@126.com)
Guangxi University of Finance and Economics
Fanping Zeng
Guangxi University of Finance and Economics
Kesong Yan
Guangxi University of Finance and Economics

Research Article

Keywords: 4HIS, UM, Tolerance relation, Information structure,

Posted Date: November 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1100046/v1

Version of Record: A version of this preprint was published at Soft Computing on February 15th, 2022. See the published version at https://doi.org/10.1007/s00500-022-06827-0.

Measures of uncertainty for a four-hybrid information system and their applications

Bin Qin

(*School of Information and Statistics, Guangxi University of Finance and Economics, Nanning, Guangxi 530003, P.R.China*)

Fanping Zeng

(*School of Information and Statistics, Guangxi University of Finance and Economics, Nanning, Guangxi 530003, P.R.China*)

Kesong Yan

(*School of Information and Statistics, Guangxi University of Finance and Economics, Nanning, Guangxi 530003, P.R.China*)

Abstract

A four-hybrid information system (4HIS) is an information system (IS) where the dataset of object descriptions consists of categorical, boolean, real-valued and missing data or attributes. This paper studies measures of uncertainty for a 4HIS and its application in attribute reduction. The distance function for each type of attribute in a 4HIS is first provided. Then, this distance is used to produce the tolerance relation induced by a given subsystem in a 4HIS. Next, information structure of this subsystem is proposed in terms of a set vector and dependence between information structures is introduced. Moreover, granulation and entropy measures in a 4HIS are investigated on the basis of information structures. In order to verify the feasibility of the proposed measures, effectiveness analysis is performed from a statistical perspective. Finally, an application of the proposed measures for attribute reduction in a 4HIS is given.

Keywords: 4HIS; UM; Tolerance relation; Information structure;

Email addresses: binqin8846@126.com (Bin Qin), fanpingzeng100@126.com (Fanping Zeng), kesongyan100@126.com (Kesong Yan)

Effectiveness analysis; Attribute reduction.

1. Introduction

1.1. Research background

Due to the complex diversity of the objective world, uncertainty exists in real life. randomness, vagueness and imprecision are the most important concepts for uncertainty which can be appeared everywhere. Uncertainty plays a vital role in practical problems. Measuring uncertainty (UM) is helpful for understanding the nature of various kinds of information and then offer new visual angle for data analysis. UM a significant issue in many research fields, such as machine learning [33], pattern recognition [11], medical diagnosis [14], data mining [8] and so on.

Granular computing (GrC), proposed by Zadeh [39, 40], is a mode of thinking or method for solving practical problems based on granularity structure. Because GrC reflects the global view and approximate solution ability of human beings when dealing with multilevel and multiperspective problems, GrC has gradually become an important theory for solving uncertain. Information granulation is the basic content of GrC. An object is divided into a series of different information granules under given granulation criteria which is called the process of information granulation. Under dissimilar granulation criteria, the different granularity layers can be obtained, and then multi-granularity grid structure. Granular structure is the collection of information granules. Lin [16] and Yao [35] talked about the importance of GrC, it caught people's attention. GrC is a superset that integrates many theoretical methods in artificial intelligence fields such as rough set theory (RST) [25], fuzzy set theory [41], concept lattice [24, 31] and quotient space theory [46].

RST is a considerable mathematical tool. Not only does it offers new scientific logic and research methods for information science and cognitive science, but also provides a tool dealing with uncertainty. Its essential idea is to construct a partition of the universe by means of indistinguishable relations, obtain equivalence classes, and then establish an approximate space. Information system (IS) based on RST is also called knowledge representation system [25]. An IS can be represented by a data table. Furthermore, the data table contains rows labeled by objects of interest, columns labeled by attributes, and entries of the table indicating attribute values. There are many applications in RST, for instance, uncertainty modeling [7], reasoning with uncertainty [10], rule extraction [4, 23], classification and feature selection [6, 13] are associated with IS.

In order to systematically assess uncertainty, the notion of entropy to communicate theory Shannon [30] was introduced to deal with UM. Beaubouef et al. [3] investigated other methods to rough sets' uncertainty. Miao et al. [22] proposed some more effective and significance measure tools, including information, combination and rough entropy. Liang et al. [19] introduced a rough metric method for knowledge in an incomplete information system (IIS). Mi et al. [23] gave some properties of fuzzy approximation operators and a method of uncertainty measurement for generalized fuzzy rough sets. Li et al. [21] studied uncertainty measurement for a fuzzy relation IS. Li et al. [18] measured uncertainty of a fully fuzzy IS by using Gaussian kernel. Dai et. al [9] studied entropy measures and granularity measures for a set-valued IS. Li et. al [20] investigated UM for a covering IS.

For GrC in an IS, the information structure is a significant research topic. An equivalence relation is a special kind of similarity between objects from a dataset. Given an IS, each attribute subset determines an equivalence relation. The object set of this IS is divided into disjoint classes by this equivalence relation, and these classes are said to be equivalence classes. If two objects belong to the same equivalence class, then we may say that they cannot be distinguished under this equivalence relation. Thus, each equivalence class is seen as an information granule consisting of indistinguishable objects. The family of all these information granules constitutes a vector; this vector is said to be an information structure in the IS induced by this attribute subset. Equally, information structures in an IS are also granular structures in the meaning of GrC. Yu [36] proposed information structures in an IIS. Zhang et al. [45] investigated information structures and uncertainty measures in a fully fuzzy IS. Li et. al [17] investigated information structures in a covering IS.

1.2. Motivation and inspiration

If an IS has many kinds of attributes or data, such as boolean attributes, categorical attributes, real-valued attributes, missing value and so on, then this IS can be called a multiple data IS. Zeng et al. [43] called such an IS as a hybrid information system (HIS). How to process this kind of hybrid data? Zeng et al. [43] investigated the measurement problem of mixed data and the incremental updating method when IS changed. Martti et al. [15] introduced two distance measures in the presence of missing values is very useful to study for medical data of mixed-type variables. Han et al. [12] introduced a useful approach to process hybrid data that a database consisting of six data types. Yu [37] considered information structures and UM of a hybrid information system with images (HISI).

In practical applications, hybrid data exists anywhere. It is very meaningful topic to discuss UM of an IS. The main purpose of this paper is to study UM of a 4HIS.

In recent years, some scholars have discussed topics related to information structures and uncertainty in an IS, such as [1, 5, 32]. However, their research lacks numerical experiments and big data analysis support. To make sure our work is more convincing and complete, this paper gives numerical experiments and data analysis.

1.3. Discussion and contribution

In this part, we discuss several references for hybrid data, so as to see the contribution or innovation of this paper more clearly.

(1) Zeng et al. [43] defined a new distance based on the value difference metric and then constructed a novel fuzzy rough set by combining the distance and Gaussian kernel. Considering an IS often vary with time, they analyzed the updating mechanisms for feature selection with the variation of the attribute set. Moreover, they presented fuzzy rough set approaches for incremental feature selection on HIS and proposed two corresponding algorithms. Finally, extensive experiments on eight datasets show that the incremental approaches significantly outperform non-incremental approaches with feature selection in the computational time.

(2) Zeng et al. [44] analyzed the changing mechanisms of the attribute values and fuzzy equivalence relations in fuzzy rough set and then presented fuzzy rough set approaches for incrementally updating approximations in an HIS. Moreover, they gave two corresponding incremental algorithms. Finally, extensive experiments on eight data sets show that incremental approaches can effectively improve the performance of updating approximations and not only significantly shorten the computational time, but also increase approximation classification accuracies.

(3) Yu [37] considered a hybrid information system with images (HISI). First, he developed new hybrid distance in an HISI. Then, he obtained the fuzzy T_cos-equivalence relation by using Gaussian kernel. Next, he described information structures in an HISI by set vectors, and studied dependence

between them by using inclusion degree. Finally, he investigated UM for an HISI by means of its information structures.

(4) Yuan et al. [38] introduced fuzzy rough sets to deal with the problem of outlier detection in hybrid data (numerical, categorical). First, they defined the granule outlier degree to characterize the outlier degree of fuzzy rough granules by employing the fuzzy approximation accuracy. Then, they constructed the outlier factor based on fuzzy rough granules by integrating the outlier degree and the corresponding weights to characterize the outlier degree of objects. Furthermore, they designed the corresponding outlier detection algorithm. Finally, they evaluated the effectiveness of the algorithm through experiments on 16 real-world datasets. The experimental results show that the algorithm is more flexible for detecting outliers and is suitable for hybrid data.

(5) Zhang et al. [42] proposed a fuzzy rough set based information entropy for feature selection for hybrid data (nominal, real-valued). They first proved that the newly-defined entropy meets the common requirement of monotonicity and can equivalently characterize the existing feature selection in the fuzzy rough set theory. Then, they formulated a feature selection algorithm based on the proposed entropy and a filter-wrapper method is suggested to select the best feature subset in terms of classification accuracy. Finally, they carried out an extensive numerical experiment to assess the performance of the feature selection algorithm.

(6) This paper deal with hybrid data (categorical, boolean, real-valued and missing data). The main details are based on the following considerations: a) a 4HIS itself has uncertainty; b) how to define a tolerance relation in a 4HIS; c) Information structure is very helpful for knowledge discovery from a 4HIS; d) the magnitude of the measured value in a 4HIS can be compared by dependence between information structures; e) which measure is chosen to measure the uncertainty of a 4HIS; f) it is very necessary to analyze the effectiveness of the proposed measurement; h) it is important to give an application of the proposed measures for attribute reduction in a 4HIS.

This paper first provides the distance function for each type of attribute in a 4HIS. This distance is used to produce the tolerance relation induced by a given subsystem. Then, information granules of a 4HIS based on the tolerance relation are constructed. By the way, the information structure formed by information granules composed of toleration classes is presented. Next, the dependence between them is discussed. By means of the dependence, four kinds of measurement to estimate the uncertainty of a 4HIS are

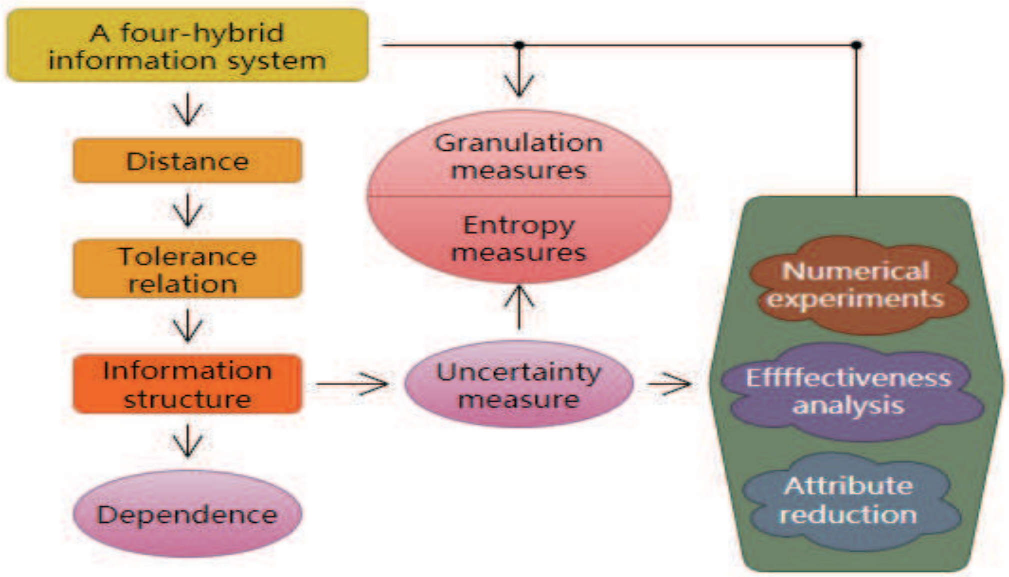


Figure 1: The work process of the paper

put forward. Moreover, the effectiveness analysis about the proposed measures ia carried out from a statistical perspective. We find the influence of θ value on the UM for a 4HIS, which may have potential application value in data mining. Finally, an application of the proposed measures for attribute reduction in a 4HIS is given.

1.4. Structure and organization

The work process of the paper is given in Figure 1.

The remaining part of this paper is organized as follows. Section 2 recalls some notions about a 4HIS. Section 3 constructs the distance between the information values of two objects about each type of attribute in a 4HIS and proposes the tolerance relation induced by a given subsystem of a 4HIS. Section 4 describes information structures in a 4HIS and studies the dependence between them. Section 5 introduces some tools for measuring uncertainty of a 4HIS. Section 6 conducts effectiveness analysis for showing the feasibility of these tools. Section 7 gives an application of the proposed measures for attribute reduction in a 4HIS. Section 8 concludes this paper.

Table 1: A 4HIS

X	Headache(a_1)	Muscle pain(a_2)	Temperature(a_3)	Symptom(a_4)
x_1	Sick	Yes	40	Flu
x_2	Sick	Yes	39.5	Flu
x_3	Middle	*	39	Flu
x_4	Middle	Yes	36.8	Rhinitis
x_5	Middle	No	*	Rhinitis
x_6	No	No	36.6	Health
x_7	No	*	*	Health
x_8	No	Yes	38	Flu
x_9	*	Yes	37	Health

2. Preliminaries

In this section, some basic concepts about a 4HIS are introduced.

Definition 2.1 ([26])**.** *Suppose that X is a finite set of objects. Assume that AT expresses a finite set of attributes. Then the ordered triple (X, AT) is referred to as an information system (IS), if every attribute $a \in AT$ is able to decide a function $a : X \to Y_a$, where $Y_a = \{a(x) : x \in X\}$.*

Let (X, AT) be an IS. If there is $a \in AT$ such that $* \in Y_a$, here $*$ means a null or unknown value, then (X, AT) is called an incomplete information system (IIS).

Definition 2.2. *Suppose that (X, AT) is a IIS. Then (X, AT) is referred to as a four-hybrid information system (4HIS), if $A = A^{cat} \cup A^{boo} \cup A^{rea}$, where A^{cat}, A^{boo} and A^{rea} are the categorical, boolean and real-valued attribute set, respectively.*

Example 2.3. *In Table 1, categorical attribute "Headache", boolean attribute "Muscle pain", real-valued attribute "Temperature" and categorical attribute "Symptom" are denoted as a_1, a_2, a_3 and a_4, respectively, and " * " indicates the missing value. Then, Table 1 depicts a 4HIS (X, AT), where $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$ is an object set, $A = \{a_1, a_2, a_3, a_4\}$ expresses a attribute set.*

Y_a^* is denoted as the set of non-missing information values of the attribute a. Then

$$Y_{a_1}^* = \{Sick, Middle, No\}, \quad Y_{a_2}^* = \{Yes, No\},$$

$$Y^*_{a_3} = \{36.6, 36.8, 37, 38, 39, 39.5, 40\}, \quad Y^*_{a_4} = \{Flu, Rhinitis, Health\}.$$

3. Tolerance relations in a 4HIS

In this section, the distance between the information values of two objects about each type of attribute in a 4HIS is first constructed. Then, the tolerance relation induced by a given subsystem of a 4HIS is proposed.

3.1. The distance function for each type of attribute in a 4HIS

$\forall\ A, B \in 2^X$, denote

$$A \oplus B = A \cup B - A \cap B.$$

$\forall\ a \in A^{rea}$, denote

$$\hat{a} = maxY^*_a - minY^*_a.$$

$\forall\ a \in A^{cat}$ and $x \in X$ with $a(x) \neq *$, denote

$$[x]_a = \{y \in X : a(x) = a(y), a(y) \in Y^*_a\}.$$

For missing data, we have the following thoughts.

1) Consider "$x \neq y, a(x) = *, a(y) \neq *, a \in A$", because "$a(x)$" is treated as "do not care" condition, thus $a(x)$ has the probability of $\frac{1}{|Y^*_a|}$ to equal to one certain value of Y^*_a.

2) Consider "$x \neq y, a(x) \neq *, a(y) = *, a \in A$", because "$a(y)$" is treated as "do not care" condition, thus $a(y)$ has the probability of $\frac{1}{|Y^*_a|}$ to equal to one certain value of Y^*_a.

3) Consider "$x \neq y, a(x) = *, a(y) = *, a \in A$", $a(x)$ and $a(y)$ both have the probability of $\frac{1}{|Y^*_a|}$ to equal to one certain value of Y^*_a, so the joint probability of $a(x)$ and $a(y)$ is $\frac{1}{|Y^*_a|^2}$.

For "$x \neq y, a(x) \neq *, a(y) \neq *, a(x) \neq a(y),\ a \in A^{rea}$", define

$$dis(a(x), a(y)) = \frac{|a(x) - a(y)|}{\hat{a}}.$$

For "$x \neq y, a(x) \neq *, a(y) \neq *, a(x) \neq a(y),\ a \in A^{boo}$", according to the opinion of [43], define

$$dis(a(x), a(y)) = 1.$$

For "$x \neq y, a(x) \neq *, a(y) \neq *, a(x) \neq a(y),\ a \in A^{cat}$", according to the opinion of [37], define

$$dis(a(x), a(y)) = \frac{|[x]_a \oplus [y]_a|}{|[x]_a \cup [y]_a|}.$$

In this way, the following definition is proposed.

Definition 3.1. *Suppose that (X, AT) is a 4HIS with $AT = A^{cat} \cup A^{boo} \cup A^{rea}$. Then $\forall\ u, y \in X$, $\forall\ a \in A$, the distance between $a(x)$ and $a(y)$ is defined as*

$$dis(a(x), a(y)) = \begin{cases} 0, & x = y,\ a \in A; \\ 1 - \frac{1}{|Y_a^*|^2}, & x \neq y, a(x) = *, a(y) = *,\ a \in A; \\ 1 - \frac{1}{|Y_a^*|}, & x \neq y, a(x) = *, a(y) \neq *,\ a \in A; \\ 1 - \frac{1}{|Y_a^*|}, & x \neq y, a(x) \neq *, a(y) = *,\ a \in A; \\ 0, & x \neq y, a(x) \neq *, a(y) \neq *, a(x) = a(y),\ a \in A; \\ 1, & x \neq y, a(x) \neq *, a(y) \neq *, a(x) \neq a(y),\ a \in A^{boo} \quad [43]; \\ \frac{|[x]_a \oplus [y]_a|}{|[x]_a \cup [y]_a|}, & x \neq y, a(x) \neq *, a(y) \neq *, a(x) \neq a(y),\ a \in A^{cat} \quad [37]; \\ \frac{|a(x) - a(y)|}{\hat{a}}, & x \neq y, a(x) \neq *, a(y) \neq *, a(x) \neq a(y),\ a \in A^{rea}. \end{cases}$$

Example 3.2. (*Continued from Examples 2.3*)

(1) *Since a_1 is a categorical attribute, $a_1(x_1) = Sick$ and $(a_1(x_3) = Middle$, we have*

$$[x_1]_{a_1} = \{x_1, x_2\}, \quad [x_3]_{a_1} = \{x_3, x_4, x_5\}.$$

By Definition 3.1,

$$dis((a_1(x_1), a_1(x_3)) = \frac{|[x_1]_{a_1} \oplus [x_3]_{a_1}|}{|[x_1]_{a_1} \cup [x_3]_{a_1}|} = \frac{|[x_1]_{a_1} \cup [x_3]_{a_1} - [x_1]_{a_1} \cap [x_3]_{a_1}|}{|[x_1]_{a_1} \cup [x_3]_{a_1}|} = \frac{5}{5} = 1;$$

(2) *Since $a_2(x_1) = Yes \neq *$, $a_2(x_3) = *$, by Definition 3.1, we have*

$$dis((a_2(x_1), a_2(x_3)) = 1 - \frac{1}{|Y_{a_2}^*|} = 1 - \frac{1}{2} = 0.5;$$

(3) *Since a_3 is a real-valued attribute, by Definition 3.1, we have*

$$dis((a_3(x_1), a_3(x_3)) = \frac{|a_3(x_1) - a_3(x_3)|}{\hat{a_3}} = \frac{|40 - 39|}{40 - 36.6} \approx 0.2941;$$

(4) *Since* $a_4(x_1) = Flu = a_4(x_3)$, *by Definition 3.1, we have*

$$dis((a_4(x_1), a_4(x_3)) = 0.$$

Below, for convenience, the 4HIS $(X, C \cup D)$ with $C = A^{cat} \cup A^{boo} \cup A^{rea}$ is denoted as the 4HIS (X, AT) where every element of D is be viewed as a categorical attribute.

Definition 3.3. *Let* (X, AT) *be a 4HIS.* $\forall\ a \in A$, *put*

$$M_a = (dis(a(x_i), a(a_j)))_{n \times n}.$$

Then M_a *is referred to as the distance matrix of the attribute* a *in* (X, AT).

Example 3.4. (*Continued from Examples 2.3*)

$$M_{a_1} = \begin{pmatrix} 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 \\ 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 \\ 1.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 \\ 1.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 \\ 1.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 0.6667 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.6667 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.6667 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.6667 \\ 0.6667 & 0.6667 & 0.6667 & 0.6667 & 0.6667 & 0.6667 & 0.6667 & 0.6667 & 0.0000 \end{pmatrix},$$

$$M_{a_2} = \begin{pmatrix} 0.0000 & 0.0000 & 0.5000 & 0.0000 & 1.0000 & 1.0000 & 0.5000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.5000 & 0.0000 & 1.0000 & 1.0000 & 0.5000 & 0.0000 & 0.0000 \\ 0.5000 & 0.5000 & 0.0000 & 0.5000 & 0.5000 & 0.5000 & 0.7500 & 0.5000 & 0.5000 \\ 0.0000 & 0.0000 & 0.5000 & 0.0000 & 1.0000 & 1.0000 & 0.5000 & 0.0000 & 0.0000 \\ 1.0000 & 1.0000 & 0.5000 & 1.0000 & 0.0000 & 0.0000 & 0.5000 & 1.0000 & 1.0000 \\ 1.0000 & 1.0000 & 0.5000 & 1.0000 & 0.0000 & 0.0000 & 0.5000 & 1.0000 & 1.0000 \\ 0.5000 & 0.5000 & 0.7500 & 0.5000 & 0.5000 & 0.5000 & 0.0000 & 0.5000 & 0.5000 \\ 0.0000 & 0.0000 & 0.5000 & 0.0000 & 1.0000 & 1.0000 & 0.5000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.5000 & 0.0000 & 1.0000 & 1.0000 & 0.5000 & 0.0000 & 0.0000 \end{pmatrix},$$

$$M_{a_3} = \begin{pmatrix} 0.0000 & 0.1471 & 0.2941 & 0.9412 & 0.8571 & 1.0000 & 0.8571 & 0.5882 & 0.8824 \\ 0.1471 & 0.0000 & 0.1471 & 0.7941 & 0.8571 & 0.8529 & 0.8571 & 0.4412 & 0.7353 \\ 0.2941 & 0.1471 & 0.0000 & 0.6471 & 0.8571 & 0.7059 & 0.8571 & 0.2941 & 0.5882 \\ 0.9412 & 0.7941 & 0.6471 & 0.0000 & 0.8571 & 0.0588 & 0.8571 & 0.3529 & 0.0588 \\ 0.8571 & 0.8571 & 0.8571 & 0.8571 & 0.0000 & 0.8571 & 0.9796 & 0.8571 & 0.8571 \\ 1.0000 & 0.8529 & 0.7059 & 0.0588 & 0.8571 & 0.0000 & 0.8571 & 0.4118 & 0.1176 \\ 0.8571 & 0.8571 & 0.8571 & 0.8571 & 0.9796 & 0.8571 & 0.0000 & 0.8571 & 0.8571 \\ 0.5882 & 0.4412 & 0.2941 & 0.3529 & 0.8571 & 0.4118 & 0.8751 & 0.0000 & 0.2941 \\ 0.8824 & 0.7353 & 0.5882 & 0.0588 & 0.8571 & 0.1176 & 0.8571 & 0.2941 & 0.0000 \end{pmatrix},$$

$$M_{a_4} = \begin{pmatrix} 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \end{pmatrix}.$$

3.2. The tolerance relation induced by a given subsystem of a 4HIS

Below, the tolerance relation induced by a given subsystem of a 4HIS is established.

Definition 3.5. *Suppose that (X, AT) is a 4HIS and $A \subseteq AT$. Pick $\theta \in [0,1]$. Put*

$$R_A^\theta = \{(x,y) \in X \times X : \forall\ a \in A,\ dis(a(x), a(y)) \leq \theta\}.$$

Then R_A^θ is referred as to the relation induced by the subsystem (X, A) with respect to θ.

Clearly, R_A^θ is a tolerance relation on X.

Denote

$$R_A^\theta(x) = \{y \in X : (x,y) \in R_A^\theta\}.$$

Then $R_A^\theta(x)$ is referred as to the tolerance class of the object x under the tolerance relation R_A^θ.

Proposition 3.6. *Let (X, AT) be a 4HIS. Then the following properties hold:*

(1) *If $A \subseteq B \subseteq AT$, then for any $\theta \in [0,1]$ and $x \in X$,*

$$R_B^\theta(x) \subseteq R_A^\theta(x);$$

(2) *If $0 \leq \theta_1 \leq \theta_2 \leq 1$, then for any $A \subseteq AT$ and $x \in X$,*

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

Proof. (1) Suppose $y \in R_B^\theta(x)$. Then $\forall\ a \in B$, $dis(a(x), a(y)) \leq \theta$.

Note that $A \subseteq B$. Then $\forall\ a \in A$, $dis(a(x), a(y)) \leq \theta$. Thus $y \in R_A^\theta(x)$.

Hence $R_B^\theta(x) \subseteq R_A^\theta(x)$.

(2) Suppose $y \in R_A^{\theta_1}(x)$. Then $\forall\ a \in A$, $dis(a(x), a(y)) \leq \theta_1$.

Note that $\theta_1 \leq \theta_2$. Then $\forall\ a \in A$, $dis(a(x), a(y)) \leq \theta_2$. Thus $y \in R_A^{\theta_2}(x)$.

Hence $R_A^{\theta_1}(x) \subseteq R_A^{\theta_2}(x)$. □

Example 3.7. (*Continued from Example 2.3*) *Pick $\theta = 0.5$. By Example 3.4, we have*

$R_{AT}^\theta(x_1) = R_{AT}^\theta(x_2) = \{x_1, x_2\}$, $R_{AT}^\theta(x_3) = \{x_3\}$, $R_{AT}^\theta(x_4) = \{x_4\}$, $R_{AT}^\theta(x_5) = \{x_5\}$, $R_{AT}^\theta(x_6) = \{x_6\}$, $R_{AT}^\theta(x_7) = \{x_7\}$, $R_{AT}^\theta(x_8) = \{x_8\}$, $R_{AT}^\theta(x_9) = \{x_9\}$.

In what follows, an algorithm of computing R_A^θ is designed as follows.

Algorithm 1: Computing R_A^θ.

Input: *A 4HIS* (X, AT), $A \subseteq AT$ *and* $\theta \in [0,1]$.

Output: *The tolerance relation* R_A^θ *on* X.

for $i = 0; i < |X|; i++$ **do**
    for $j = |X| - 1; j > i; j--$ **do**
        for $a(x_i) = *$ *or* $a(x_j) = *$, **do**
            if $a(x_i) = *,\ a(x_j) \neq *$ **then**
                $dis(a(x_i), a(x_j)) = 1 - \frac{1}{|Y_a^*|}$;
            end
            if $a(x_i) \neq *,\ a(x_j) = *$ **then**
                $dis(a(x_i), a(x_j)) = 1 - \frac{1}{|Y_a^*|}$;
            end
            if $a(x_i) = *,\ a(x_j) = *$ **then**
                $dis(a(x_i), a(x_j)) = 1 - \frac{1}{|Y_a^*|^2}$.
            end
        end
        for $a(x_i) \neq *$, $a(x_j) \neq *$, $a(x_i) = a(x_j)$, **do**
            $dis(a(x_i), a(x_j)) = 0$.
        end
        for $a(x_i) \neq *$, $a(x_j) \neq *$, $a(x_i) \neq a(x_j)$, **do**
            if $a \in A^{boo}$ **then**
                $dis(a(x_i), a(x_j)) = 1$;
            end
            if $a \in A^{cat}$ **then**
                $dis(a(x_i), a(x_j)) = \frac{|[x_i]_a \oplus [x_j]_a|}{|[x_i]_a \cup [x_j]_a|}$;
            end
            if $a \in A^{rea}$ **then**
                $dis(a(x_i), a(x_j)) = \frac{|a(x_i) - a(x_j)|}{\hat{a}}$.
            end
        end
        Pick θ. *Put*
        $R_A^\theta = \{(x, y) \in X \times X : \forall\ a \in A,\ dis(a(x), a(y)) \leq \theta\}$.
        Obtain R_A^θ.
    end
end

4. Information structures in a 4HIS

In this section, information structures in a 4HIS are studied.

4.1. The concept of information structures in a 4HIS

Definition 4.1. *Suppose that* (X, AT) *is a 4HIS and* $A \subseteq AT$. *Pick* $\theta \in [0,1]$. *Put*

$$InS^{\theta}(A) = (R_A^{\theta}(x_1), R_A^{\theta}(x_2), \cdots\cdots, R_A^{\theta}(x_n)).$$

Then $InS^{\theta}(A)$ *is referred to as* θ*-information structure of the subsystem* (X, A).

Example 4.2. (*Continued from Example 3.7*)

$$InS^{\theta}(AT) = (\{x_1, x_2\}, \{x_1, x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6, x_8\}, \{x_7\}, \{x_6, x_8\}, \{x_9\}).$$

Definition 4.3. *Suppose that* (X, AT) *is a 4HIS. Given* $A, B \subseteq AT$ *and* $\theta_1, \theta_2 \in [0,1]$. *If for any* i, $R_A^{\theta_1}(x_i) = R_B^{\theta_2}(x_i)$, *then* $InS^{\theta_1}(A)$ *and* $InS^{\theta_2}(B)$ *are called to be the same. It's written as* $InS^{\theta_1}(A) = InS^{\theta_2}(B)$.

4.2. Dependence between information structures in a 4HIS

Definition 4.4. *Let* (X, AT) *be a 4HIS. Given* $A, B \subseteq AT$ *and* $\theta_1, \theta_2 \in [0,1]$.

(1) *If for any* i, $R_A^{\theta_1}(x_i) \subseteq R_B^{\theta_2}(x_i)$, *then* $InS^{\theta_2}(B)$ *is referred to as depend on* $InS^{\theta_1}(A)$. *It is written as* $InS^{\theta_1}(A) \preceq InS^{\theta_2}(B)$.

(2) *If* $InS^{\theta_1}(A) \preceq InS^{\theta_2}(B)$ *and* $InS^{\theta_1}(A) \neq InS^{\theta_2}(B)$, *then* $InS^{\theta_2}(B)$ *is referred to as depend strictly on* $InS^{\theta_1}(A)$. *It is written as* $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$.

Obviously,

$$InS^{\theta_1}(A) = InS^{\theta_2}(B) \Leftrightarrow InS^{\theta_1}(A) \preceq InS^{\theta_2}(B) \text{ and } InS^{\theta_2}(B) \preceq InS^{\theta_1}(A),$$

$$InS^{\theta_1}(A) \prec InS^{\theta_2}(B) \Rightarrow InS^{\theta_1}(A) \preceq InS^{\theta_2}(B).$$

Theorem 4.5. *Suppose that* (X, AT) *is a 4HIS. Given* $A, B \subseteq AT$ *and* $\theta_1, \theta_2 \in [0,1]$. *Then*

$$InS^{\theta_1}(A) = InS^{\theta_2}(B) \Leftrightarrow R_A^{\theta_1} = R_B^{\theta_2}.$$

Proof. Obviously. □

Theorem 4.6. *Suppose that (X, AT) is a 4HIS. Given $A, B \subseteq AT$ and $\theta_1, \theta_2 \in [0,1]$. Then*

$$InS^{\theta_1}(A) \preceq InS^{\theta_2}(B) \Leftrightarrow R_A^{\theta_1} \subseteq R_B^{\theta_2}.$$

Proof. Clearly. □

Corollary 4.7. *Suppose that (X, AT) is a 4HIS. Given $A, B \subseteq AT$ and $\theta_1, \theta_2 \in [0,1]$. Then*

$$InS^{\theta_1}(A) \prec InS^{\theta_2}(B) \Leftrightarrow R_A^{\theta_1} \subset R_B^{\theta_2}.$$

Proof. By Theorems 4.5 and 4.6, it is easy to prove. □

Theorem 4.8. *Suppose that (X, AT) is a 4HIS. Then the following properties hold:*

(1) *If $A \subseteq B \subseteq AT$, then for any $\theta \in [0,1]$, $InS^{\theta}(B) \preceq InS^{\theta}(A)$;*

(2) *If $0 \leq \theta_1 \leq \theta_2 \leq 1$, then for any $A \subseteq AT$, $InS^{\theta_1}(A) \preceq InS^{\theta_2}(A)$.*

Proof. (1) Since $A \subseteq B$, by Proposition 3.6(1), $\forall$ i, we have

$$R_B^{\theta}(x_i) \subseteq R_A^{\theta}(x_i).$$

Thus $InS^{\theta}(B) \preceq InS^{\theta}(A)$.

(2) Since $\theta_1 \leq \theta_2$, by Proposition 3.6(2), $\forall$ $P \subseteq A$, we have

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

Then $InS^{\theta_1}(A) \preceq InS^{\theta_2}(A)$. □

5. UM of a 4HIS

In this section, we studies UM of a 4HIS.

5.1. Granulation measures of a 4HIS

Definition 5.1. *Let (X, AT) be a 4HIS. Given $\theta \in [0,1]$. Suppose that $M^{\theta} : 2^{AT} \to (-\infty, +\infty)$ is a mapping. Then $M^{\theta}(A)$ is referred to as θ-information granulation function of the subsystem (X, A), if M^{θ} satisfies the following conditions:*

(1) $\forall$ $A \subseteq AT$, $M^{\theta}(A) \geq 0$ *(Non-negativity);*

(2) $S^{\theta}(A) = S^{\theta}(B)$ *implies* $M^{\theta}(A) = M^{\theta}(Q)$ *(Invariability);*

(3) $S^{\theta}(A) \prec S^{\theta}(B)$ *implies* $M^{\theta}(A) < M^{\theta}(Q)$ *(Monotonicity).*

Similar to Definition 5 in [29], θ-information granulation in a 4HIS is given as follows.

Definition 5.2. *Let (X, AT) be a 4HIS and $A \subseteq AT$. Pick $\theta \in [0,1]$. Then θ-information granulation of (X, A) is defined by*

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

Proposition 5.3. *Let (X, AT) be a 4HIS and $A \subseteq AT$. Pick $\theta \in [0,1]$. Then*

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

Proof. $\forall\ i$, $1 \leq |R_A^{\theta}(x_i)| \leq n$. Then $n \leq \sum\limits_{i=1}^{n}|R_A^{\theta}(x_i)| \leq n^2$.

By Definition 5.2,

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

$\forall\ i$, $|R_A^{\theta}(x_i)| = 1$. Then $G^{\theta}(A) = \frac{1}{n}$.
$\forall\ i$, $|R_A^{\theta}(x_i)| = n$. Then $G^{\theta}(A) = 1$.

□

Theorem 5.4. *Let (X, AT) be a 4HIS and $A, B \subseteq AT$. Pick $\theta_1, \theta_2 \in [0,1]$. Then the following properties hold:*
(1) If $InS^{\theta_1}(A) \preceq InS^{\theta_2}(B)$, then $G^{\theta_1}(A) \leq G^{\theta_2}(B)$;
(2) If $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$, then $G^{\theta_1}(A) < G^{\theta_2}(B)$.

Proof. (1) This is obvious.
(2) By Definition 5.2,

$$G^{\theta_1}(A) = \frac{1}{n^2}\sum_{i=1}^{n}|R_A^{\theta_1}(x_i)|,\ G^{\theta_2}(B) = \frac{1}{n^2}\sum_{i=1}^{n}|R_B^{\theta_2}(x_i)|.$$

Note that $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$. Then $\forall\ i$, $R_A^{\theta_1}(x_i) \subseteq R_B^{\theta_2}(x_i)$ and $\exists\ j$, $R_A^{\theta_1}(x_j) \subsetneq R_B^{\theta_2}(x_j)$. Thus $\forall\ i$, $|R_A^{\theta_1}(x_i)| \leq |R_B^{\theta_2}(x_i)|$ and $\exists\ j$, $|R_A^{\theta_1}(x_j)| < |R_B^{\theta_2}(x_j)|$.

Hence $G^{\theta_1}(A) < G^{\theta_2}(B)$. □

This theorem shows that when the available information becomes coarse, the θ-information granulation increases, and when the available information becomes finer, 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 draw the conclusion that the θ-information granulation introduced in definition 5.2 can be used to evaluate the degree of a 4HIS.

Proposition 5.5. *Let (X, AT) be a 4HIS. Then the following properties hold:*
(1) *If $A \subseteq B \subseteq AT$, then for any $\theta \in [0, 1]$, $G^\theta(B) \leq G^\theta(A)$;*
(2) *If $0 \leq \theta_1 \leq \theta_2 \leq 1$, then for any $A \subseteq AT$, $G^{\theta_1}(A) \leq G^{\theta_2}(A)$.*

Proof. These follow from Theorems 4.8 and 5.4(1). □

Definition 5.6. *Let (X, AT) is a 4HIS and $A \subseteq AT$. Pick $\theta \in [0, 1]$. Then θ-information amount of (X, A) is defined by*

$$E^\theta(A) = \sum_{i=1}^{n} \frac{1}{n}(1 - \frac{|R_A^\theta(x_i)|}{n}).$$

Theorem 5.7. *Let (X, AT) be a 4HIS and $A, B \subseteq AT$. Pick $\theta_1, \theta_2 \in [0, 1]$. Then the following properties hold:*
(1) *If $InS^{\theta_1}(A) \preceq InS^{\theta_2}(B)$, then $E^{\theta_2}(B) \leq E^{\theta_1}(A)$;*
(2) *If $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$, then $E^{\theta_2}(B) < E^{\theta_1}(A)$.*

Proof. (1) This is clear.
(2) By Definition 5.6,

$$E^{\theta_1}(A) = \sum_{i=1}^{n} \frac{1}{n}(1 - \frac{|R_A^{\theta_1}(x_i)|}{n}),\ E^{\theta_2}(B) = \sum_{i=1}^{n} \frac{1}{n}(1 - \frac{|R_B^{\theta_2}(x_i)|}{n}).$$

Note that $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$. Then $\forall\ i$, $R_A^{\theta_1}(x_i) \subseteq R_B^{\theta_2}(x_i)$ and $\exists\ j$, $R_A^{\theta_1}(x_j) \subsetneqq R_B^{\theta_2}(x_j)$. Thus $\forall\ i$, $|R_A^{\theta_1}(x_i)| \leq |R_B^{\theta_2}(x_i)|$ and $\exists\ j$, $|R_A^{\theta_1}(x_j)| < |R_B^{\theta_2}(x_j)|$.

Hence $E^{\theta_2}(B) < E^{\theta_1}(A)$. □

This theorem shows that when the structure of hybrid information becomes finer, the θ-information amount increases, and when the hybrid information structure becomes rough, the θ-information amount decreases.

Proposition 5.8. *Let (X, AT) be a 4HIS. Then the following properties hold:*

(1) *If $A \subseteq B \subseteq AT$, then for any $\theta \in [0, 1]$, $E^\theta(A) \leq E^\theta(B)$;*

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

Proof. These follow from Theorems 4.8 and 5.7(1). □

Theorem 5.9. *Let (X, AT) be a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then*

$$G^\theta(A) + E^\theta(A) = 1.$$

Proof.

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

□

Corollary 5.10. *Let (X, AT) be a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then*

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

Proof. By Proposition 5.3, $\frac{1}{n} \leq G^\theta(A) \leq 1$. We have

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

By Theorem 5.9, $G^\theta(A) + E^\theta(A) = 1$.

Hence

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

□

5.2. **Entropy measures of a 4HIS**

Definition 5.11. *Let (X, AT) be a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then θ-rough entropy of (X, A) is defined by*

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

Proposition 5.12. *Let (X, AT) be a 4HIS. Given $A \subseteq AT$ and $\theta \in [0,1]$. Then*

$$0 \leq E_r^{\theta}(A) \leq \log_2 n.$$

Proof. $\forall\ i$, $1 \leq |R_A^{\theta}(x_i)| \leq n$. Then

$$0 \leq -\log_2 \frac{1}{|R_A^{\theta}(x_i)|} = \log_2 |R_A^{\theta}(x_i)| \leq \log_2 n.$$

By Definition 5.11,

$$0 \leq E_r^{\theta}(A) \leq \log_2 n.$$

$\forall\ i$, $|R_A^{\theta}(x_i)| = 1$. Then $E_r^{\theta}(A) = 0$.
$\forall\ i$, $|R_A^{\theta}(x_i)| = n$. Then $E_r^{\theta}(A) = \log_2 n$. □

Theorem 5.13. *Let (X, AT) be a 4HIS and $A, B \subseteq AT$. Pick $\theta_1, \theta_2 \in [0,1]$. Then the following properties hold:*
(1) *If $InS^{\theta_1}(A) \preceq InS^{\theta_2}(B)$, then $E_r^{\theta_1}(A) \leq E_r^{\theta_2}(B)$;*
(2) *If $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$, then $E_r^{\theta_1}(A) < E_r^{\theta_2}(B)$.*

Proof. (1) Obviously.
(2) By Definition 5.11,

$$E_r^{\theta_1}(A) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{|R_A^{\theta_1}(x_i)|}, \quad E_r^{\theta_2}(B) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{|R_B^{\theta_2}(x_i)|}.$$

Note that $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$. Then $\forall\ i$, $R_A^{\theta_1}(x_i) \subseteq R_B^{\theta_2}(x_i)$ and $\exists\ j$, $R_A^{\theta_1}(x_j) \subsetneq R_B^{\theta_2}(x_j)$. Thus $\forall\ i$, $|R_A^{\theta_1}(x_i)| \leq |R_B^{\theta_2}(x_i)|$ and $\exists\ j$, $|R_A^{\theta_1}(x_j)| < |R_B^{\theta_2}(x_j)|$.
Hence, $E_r^{\theta_1}(A) < E_r^{\theta_2}(B)$. □

This theorem shows that the greater the uncertainty of the available information, the greater the θ-rough entropy. Therefore, we can draw the conclusion that the θ-rough entropy proposed in Definition 5.11 can be used to evaluate the degree of determination of a 4HIS.

Proposition 5.14. *Let (X, AT) be a 4HIS. Then the following properties hold:*

(1) *If $A \subseteq B \subseteq AT$, then for any $\theta \in [0,1]$, $E_r^\theta(B) \leq E_r^\theta(A)$;*

(2) *If $0 \leq \theta_1 \leq \theta_2 \leq 1$, then for any $A \subseteq AT$, $E_r^{\theta_1}(A) \leq E_r^{\theta_2}(A)$.*

Proof. It can be proved by Theorems 4.8 and 5.13(1). □

Definition 5.15. *Let (X, AT) be a 4HIS. Given $A \subseteq AT$ and $\theta \in [0,1]$. Then θ-information entropy of (X, A) is defined by*

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

Theorem 5.16. *Suppose that (X, AT) is a 4HIS and $A, B \subseteq AT$. Pick $\theta_1, \theta_2 \in [0,1]$. Then the following properties hold:*

(1) *If $InS^{\theta_1}(A) \preceq InS^{\theta_2}(B)$, then $H^{\theta_2}(B) \leq H^{\theta_1}(A)$;*

(2) *If $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$, then $H^{\theta_2}(B) < H^{\theta_1}(A)$.*

Proof. (1) Obviously.

(2) By Definition 5.15,

$$H^{\theta_1}(A) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{|R_A^{\theta_1}(x_i)|}{n}, \quad H^{\theta_2}(B) = -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{|R_B^{\theta_2}(x_i)|}{n}.$$

It should be noted that $InS^{\theta_1}(A) \prec InS^{\theta_2}(B)$. Then $\forall\ i$, $R_A^{\theta_1}(x_i) \subseteq R_B^{\theta_2}(x_i)$ and $\exists\ j$, $R_A^{\theta_1}(x_j) \subsetneqq R_B^{\theta_2}(x_j)$. Thus $\forall\ i$, $|R_A^{\theta_1}(x_i)| \leq |R_B^{\theta_2}(x_i)|$ and $\exists\ j$, $|R_A^{\theta_1}(x_j)| < |R_B^{\theta_2}(x_j)|$.

Hence $H^{\theta_2}(B) < H^{\theta_1}(A)$. □

This theorem shows that when the structure of hybrid information becomes finer, the θ-information amount increases, and when the hybrid information structure becomes rough, the θ-information amount decreases.

Proposition 5.17. *Suppose that (X, AT) is a 4HIS. The the following properties hold:*

(1) *If $A \subseteq B \subseteq AT$, then for any $\theta \in [0,1]$, $H^\theta(A) \leq H^\theta(B)$;*

(2) *If $0 \leq \theta_1 \leq \theta_2 \leq 1$, then for any $A \subseteq AT$, $H^{\theta_2}(A) \leq H^{\theta_1}(A)$.*

Proof. It follows from Theorems 4.8 and 5.16(1). □

Theorem 5.18. *Suppose that* (X, AT) *is a 4HIS. Given* $A \subseteq AT$ *and* $\theta \in [0,1]$. *Then*

$$E_r^\theta(A) + H^\theta(A) = \log_2 n.$$

Proof.

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

□

Corollary 5.19. *Suppose that* (X, AT) *is a 4HIS. Given* $A \subseteq AT$ *and* $\theta \in [0,1]$. *Then*

$$0 \leq H^\theta(A) \leq \log_2 \ n.$$

Proof. By Proposition 5.12, $0 \leq E_r^\theta(A) \leq \log_2 n$. Then

$$-\log_2 n \leq -E_r^\theta(A) \leq 0.$$

By Theorem 5.18, $E_r^\theta(A) + H^\theta(A) = \log_2 n$.
Thus

$$0 \leq H^\theta(A) \leq \log_2 \ n.$$

□

6. Experiments and analysis

In this section, we design a numerical experiment and do effectiveness analysis to evaluate the proposed measures.

6.1. A numerical experiment

In order to show the performance of the proposed measures for the uncertainty in a 4HIS, we select nine data sets that come from UCI (Repository of machine learning databases) which is described in Table 2, where each data set can be expressed as a 4HIS. We carry out a numerical experiment on the nine data sets.

Table 2: Nine data sets from UCI

Date sets	Objects	Features
Scale	Ordinal	Nominal
Annealing	798	39
Automobile	205	26
Contraceptive Method Choice	1473	10
Credit Approval	690	16
Dermatology	366	35
Echocardiogram	132	13
Hepatitis	155	20
Meta-data	528	22
Post-Operative Patient	90	9

Table 3: Description of data sets

No	*Data sets*	*Sample*	*Scale*	*Ordinal*	*Nominal*	*Classes*
1	*Echocardiogram*	132	7	1	3	2
2	*Wine*	178	11	2	0	3
3	*ILPD*	583	5	4	1	2
4	*Credit*	690	3	8	4	2
5	*ThoracicSurgery*	470	2	4	10	2

In Annealing, pick $L_i = \{a_1, \cdots, a_{3\times i}\}$ $(i = 1, \cdots, 13)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Annealing are defined as follows:

$$X_G^{\theta_j}(An) = \{G^{\theta_j}(L_1), \cdots, G^{\theta_j}(L_{13})\}, X_G^{L_i}(An) = \{G^{\theta_1}(L_i), \cdots, G^{\theta_9}(L_i)\};$$

$$X_E^{\theta_j}(An) = \{E^{\theta_j}(L_1), \cdots, E^{\theta_j}(L_{13})\}, X_E^{L_i}(An) = \{E^{\theta_1}(L_i), \cdots, E^{\theta_9}(L_i)\};$$

$$X_{E_r}^{\theta_j}(An) = \{E_r^{\theta_j}(L_1), \cdots, E_r^{\theta_j}(L_{13})\}, X_{E_r}^{L_i}(An) = \{E_r^{\theta_1}(L_i), \cdots, E_r^{\theta_9}(L_i)\};$$

$$X_H^{\theta_j}(An) = \{H^{\theta_j}(L_1), \cdots, H^{\theta_j}(L_{13})\}, X_H^{L_i}(An) = \{H^{\theta_1}(L_i), \cdots, H^{\theta_9}(L_i)\};$$

In Automobile, pick $M_i = \{a_1, \cdots, a_{2\times i}\}$ $(i = 1, \cdots, 13)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Automobile are defined as follows:

$$X_G^{\theta_j}(Au) = \{G^{\theta_j}(M_1), \cdots, G^{\theta_j}(M_{13})\}, \quad X_G^{M_i}(Au) = \{G^{\theta_1}(M_i), \cdots, G^{\theta_9}(M_i)\};$$

$$X_E^{\theta_j}(Au) = \{E^{\theta_j}(M_1), \cdots, E^{\theta_j}(M_{13})\}, \quad X_E^{M_i}(Au) = \{E^{\theta_1}(M_i), \cdots, E^{\theta_9}(M_i)\};$$

$$X_{E_r}^{\theta_j}(Au) = \{E_r^{\theta_j}(M_1), \cdots, E_r^{\theta_j}(M_{13})\}, \quad X_{E_r}^{M_i}(Au) = \{E_r^{\theta_1}(M_i), \cdots, E_r^{\theta_9}(M_i)\};$$

$$X_H^{\theta_j}(Au) = \{H^{\theta_j}(M_1), \cdots, H^{\theta_j}(M_{13})\}, \quad X_H^{M_i}(Au) = \{H^{\theta_1}(M_i), \cdots, H^{\theta_9}(M_i)\};$$

In Contraceptive Method Choice, pick $N_i = \{a_1, \cdots, a_i\}$ $(i = 1, \cdots, 10)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Contraceptive Method Choice are defined as follows:

$$X_G^{\theta_j}(Co) = \{G^{\theta_j}(N_1), \cdots, G^{\theta_j}(N_{10})\}, \quad X_G^{N_i}(Co) = \{G^{\theta_1}(N_i), \cdots, G^{\theta_9}(N_i)\};$$

$$X_E^{\theta_j}(Co) = \{E^{\theta_j}(N_1), \cdots, E^{\theta_j}(N_{10})\}, \quad X_E^{N_i}(Co) = \{E^{\theta_1}(N_i), \cdots, E^{\theta_9}(N_i)\};$$

$$X_{E_r}^{\theta_j}(Co) = \{E_r^{\theta_j}(N_1), \cdots, E_r^{\theta_j}(N_{10})\}, \quad X_{E_r}^{N_i}(Co) = \{N_r^{\theta_1}(M_i), \cdots, E_r^{\theta_9}(N_i)\};$$

$$X_H^{\theta_j}(Co) = \{H^{\theta_j}(N_1), \cdots, H^{\theta_j}(N_{10})\}, \quad X_H^{N_i}(Co) = \{H^{\theta_1}(N_i), \cdots, H^{\theta_9}(N_i)\};$$

In Credit Approval, pick $O_i = \{a_1, \cdots, a_{2\times i}\}$ $(i = 1, \cdots, 8)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Credit Approval are defined as follows:

$$X_G^{\theta_j}(Cr) = \{G^{\theta_j}(O_1), \cdots, G^{\theta_j}(O_8)\}, \quad X_G^{O_i}(Cr) = \{G^{\theta_1}(O_i), \cdots, G^{\theta_9}(O_i)\};$$

$$X_E^{\theta_j}(Cr) = \{E^{\theta_j}(O_1), \cdots, E^{\theta_j}(O_8)\}, \quad X_E^{O_i}(Cr) = \{E^{\theta_1}(O_i), \cdots, E^{\theta_9}(O_i)\};$$

$$X_{E_r}^{\theta_j}(Cr) = \{E_r^{\theta_j}(O_1), \cdots, E_r^{\theta_j}(O_8)\}, \quad X_{E_r}^{O_i}(Cr) = \{O_r^{\theta_1}(M_i), \cdots, E_r^{\theta_9}(O_i)\};$$

$$X_H^{\theta_j}(Cr) = \{H^{\theta_j}(O_1), \cdots, H^{\theta_j}(O_8)\}, \quad X_H^{O_i}(Cr) = \{H^{\theta_1}(O_i), \cdots, H^{\theta_9}(O_i)\};$$

In Dermatology, pick $A_i = \{a_1, \cdots, a_{5\times i}\}$ $(i = 1, \cdots, 7)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Dermatology are defined as follows:

$$X_G^{\theta_j}(De) = \{G^{\theta_j}(A), \cdots, G^{\theta_j}(A_7)\}, \quad X_G^{A_i}(De) = \{G^{\theta_1}(A_i), \cdots, G^{\theta_9}(A_i)\};$$

$$X_E^{\theta_j}(De) = \{E^{\theta_j}(A), \cdots, E^{\theta_j}(A_7)\}, \quad X_E^{A_i}(De) = \{E^{\theta_1}(A_i), \cdots, E^{\theta_9}(A_i)\};$$

$$X_{E_r}^{\theta_j}(De) = \{E_r^{\theta_j}(A), \cdots, E_r^{\theta_j}(A_7)\}, \quad X_{E_r}^{A_i}(De) = \{A_r^{\theta_1}(M_i), \cdots, E_r^{\theta_9}(A_i)\};$$

$$X_H^{\theta_j}(De) = \{H^{\theta_j}(A), \cdots, H^{\theta_j}(A_7)\}, \quad X_H^{A_i}(De) = \{H^{\theta_1}(A_i), \cdots, H^{\theta_9}(A_i)\};$$

In Echocardiogram, pick $Q_i = \{a_1, \cdots, a_i\}$ $(i = 1, \cdots, 13)$ and $\theta_j = 0.1 \times j$

$(j = 1, \cdots, 9)$. Measure sets on Echocardiogram are defined as follows:

$$X_G^{\theta_j}(Ec) = \{G^{\theta_j}(Q_1), \cdots, G^{\theta_j}(Q_{13})\}, \quad X_G^{Q_i}(Ec) = \{G^{\theta_1}(Q_i), \cdots, G^{\theta_9}(Q_i)\};$$

$$X_E^{\theta_j}(Ec) = \{E^{\theta_j}(Q_1), \cdots, E^{\theta_j}(Q_{13})\}, \quad X_E^{Q_i}(Ec) = \{E^{\theta_1}(Q_i), \cdots, E^{\theta_9}(Q_i)\};$$

$$X_{E_r}^{\theta_j}(Ec) = \{E_r^{\theta_j}(Q_1), \cdots, E_r^{\theta_j}(Q_{13})\}, \quad X_{E_r}^{Q_i}(Ec) = \{Q_r^{\theta_1}(M_i), \cdots, E_r^{\theta_9}(Q_i)\};$$

$$X_H^{\theta_j}(Ec) = \{H^{\theta_j}(Q_1), \cdots, H^{\theta_j}(Q_{13})\}, \quad X_H^{Q_i}(Ec) = \{H^{\theta_1}(Q_i), \cdots, H^{\theta_9}(Q_i)\};$$

In Hepatitis, pick $R_i = \{a_1, \cdots, a_{2\times i}\}$ $(i = 1, \cdots, 10)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Hepatitis are defined as follows:

$$X_G^{\theta_j}(He) = \{G^{\theta_j}(R_1), \cdots, G^{\theta_j}(R_{10})\}, \quad X_G^{R_i}(He) = \{G^{\theta_1}(R_i), \cdots, G^{\theta_9}(R_i)\};$$

$$X_E^{\theta_j}(He) = \{E^{\theta_j}(R_1), \cdots, E^{\theta_j}(R_{10})\}, \quad X_E^{R_i}(He) = \{E^{\theta_1}(R_i), \cdots, E^{\theta_9}(R_i)\};$$

$$X_{E_r}^{\theta_j}(He) = \{E_r^{\theta_j}(R_1), \cdots, E_r^{\theta_j}(R_{10})\}, \quad X_{E_r}^{R_i}(He) = \{R_r^{\theta_1}(R_i), \cdots, E_r^{\theta_9}(R_i)\};$$

$$X_H^{\theta_j}(He) = \{H^{\theta_j}(R_1), \cdots, H^{\theta_j}(R_{10})\}, \quad X_H^{R_i}(He) = \{H^{\theta_1}(R_i), \cdots, H^{\theta_9}(R_i)\};$$

In Meta-data, pick $S_i = \{a_1, \cdots, a_{2\times i}\}$ $(i = 1, \cdots, 11)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Meta-data are defined as follows:

$$X_G^{\theta_j}(Me) = \{G^{\theta_j}(S_1), \cdots, G^{\theta_j}(S_{11})\}, \quad X_G^{S_i}(Me) = \{G^{\theta_1}(S_i), \cdots, G^{\theta_9}(S_i)\};$$

$$X_E^{\theta_j}(Me) = \{E^{\theta_j}(S_1), \cdots, E^{\theta_j}(S_{11})\}, \quad X_E^{S_i}(Me) = \{E^{\theta_1}(S_i), \cdots, E^{\theta_9}(S_i)\};$$

$$X_{E_r}^{\theta_j}(Me) = \{E_r^{\theta_j}(S_1), \cdots, E_r^{\theta_j}(S_{11})\}, \quad X_{E_r}^{S_i}(Me) = \{R_r^{\theta_1}(S_i), \cdots, E_r^{\theta_9}(S_i)\};$$

$$X_H^{\theta_j}(Me) = \{H^{\theta_j}(S_1), \cdots, H^{\theta_j}(S_{11})\}, \quad X_H^{S_i}(Me) = \{H^{\theta_1}(S_i), \cdots, H^{\theta_9}(S_i)\};$$

In Post-Operative Patient, pick $T_i = \{a_1, \cdots, a_i\}$ $(i = 1, \cdots, 9)$ and $\theta_j = 0.1 \times j$ $(j = 1, \cdots, 9)$. Measure sets on Post-Operative Patient are defined as follows:

$$X_G^{\theta_j}(Po) = \{G^{\theta_j}(T_1), \cdots, G^{\theta_j}(T_9)\}, \quad X_G^{T_i}(Po) = \{G^{\theta_1}(T_i), \cdots, G^{\theta_9}(T_i)\};$$

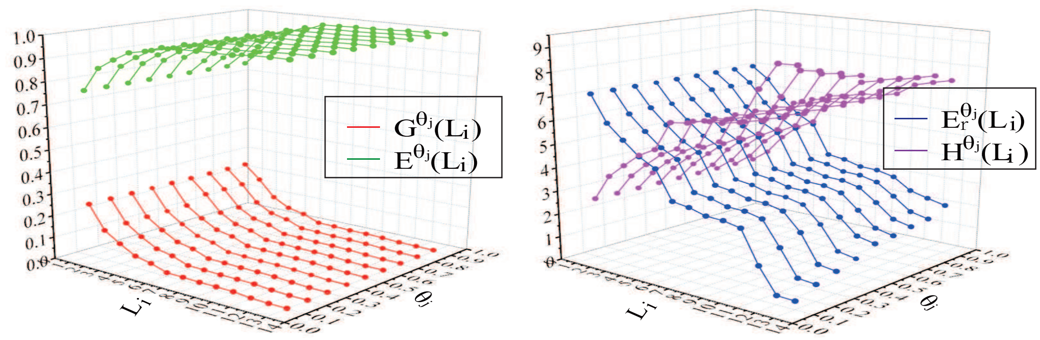
$$X_E^{\theta_j}(Po) = \{E^{\theta_j}(T_1), \cdots, E^{\theta_j}(T_9)\}, \quad X_E^{T_i}(Po) = \{E^{\theta_1}(T_i), \cdots, E^{\theta_9}(T_i)\};$$

$$X_{E_r}^{\theta_j}(Po) = \{E_r^{\theta_j}(T_1), \cdots, E_r^{\theta_j}(T_9)\}, \quad X_{E_r}^{T_i}(Po) = \{R_r^{\theta_1}(T_i), \cdots, E_r^{\theta_9}(T_i)\};$$

$$X_H^{\theta_j}(Po) = \{H^{\theta_j}(T_1), \cdots, H^{\theta_j}(T_9)\}, \quad X_H^{T_i}(Po) = \{H^{\theta_1}(T_i), \cdots, H^{\theta_9}(T_i)\};$$

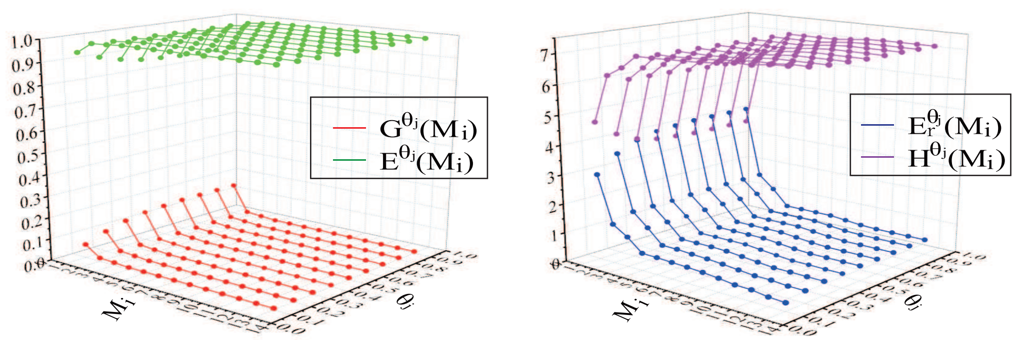
From Figures 1-9, the following conclusions are obtained:

(1) When the threshold θ is a fixed value, G^θ and E_r^θ are both monotonically decreasing as the cardinality of attribute subset increases. Mean-



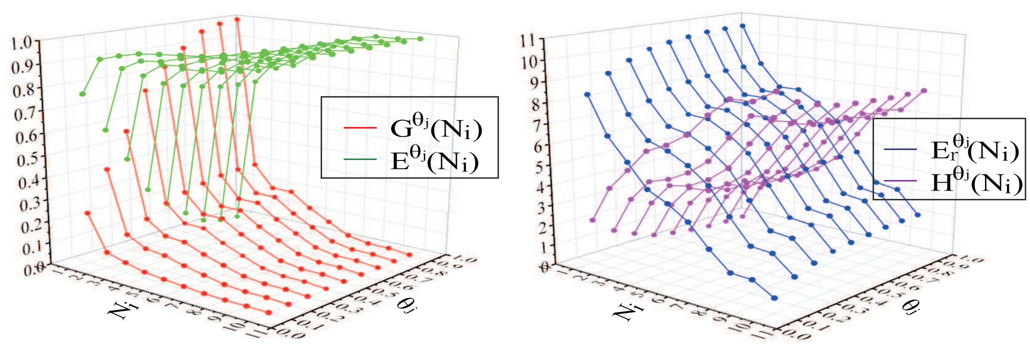
(a) **(b)**

Figure 2: Values of UM on Annealing.



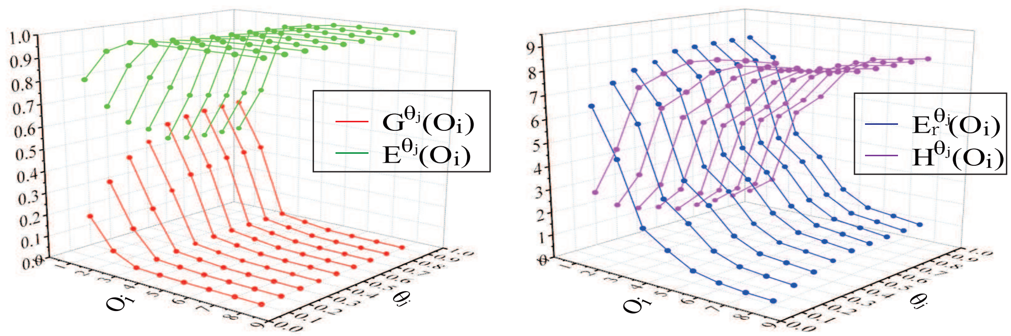
(a) **(b)**

Figure 3: Values of UM on Automobile.



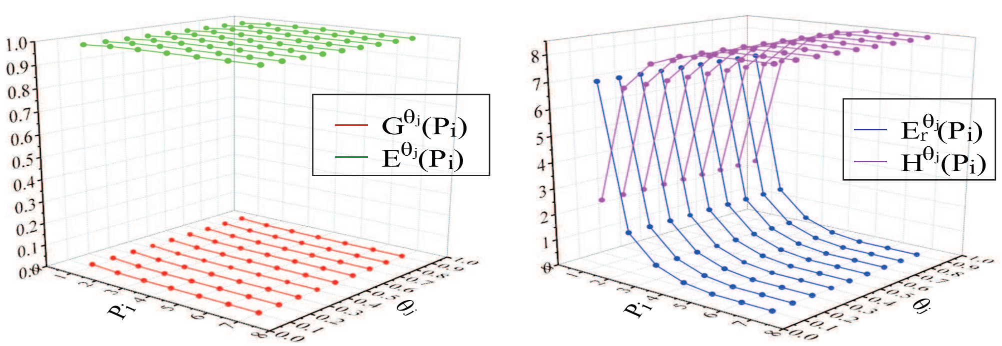
(a) **(b)**

Figure 4: Values of UM on Contraceptive Method Choice.



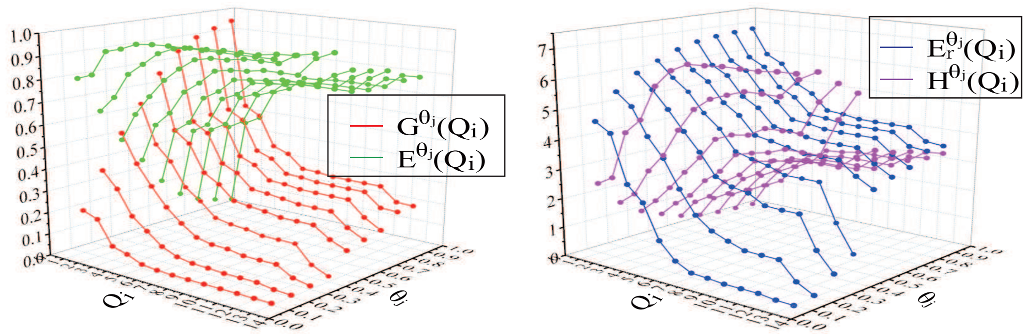
(a) **(b)**

Figure 5: Values of UM on Credit Approval.



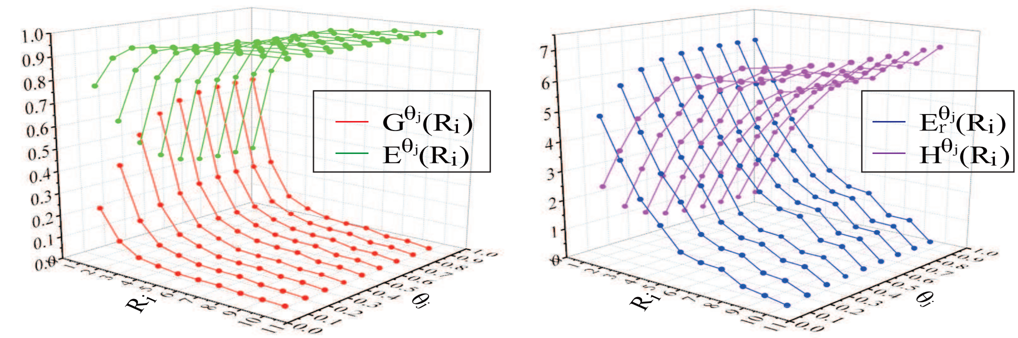
(a) **(b)**

Figure 6: Values of UM on Dermatology.



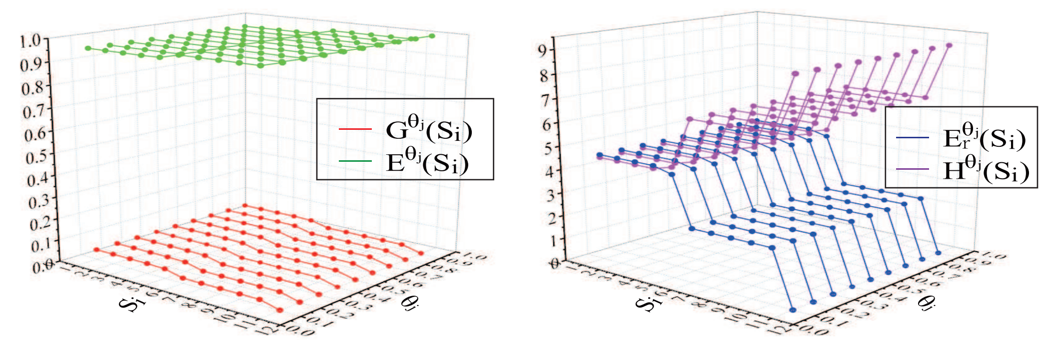
(a) **(b)**

Figure 7: Values of UM on Echocardiogram.



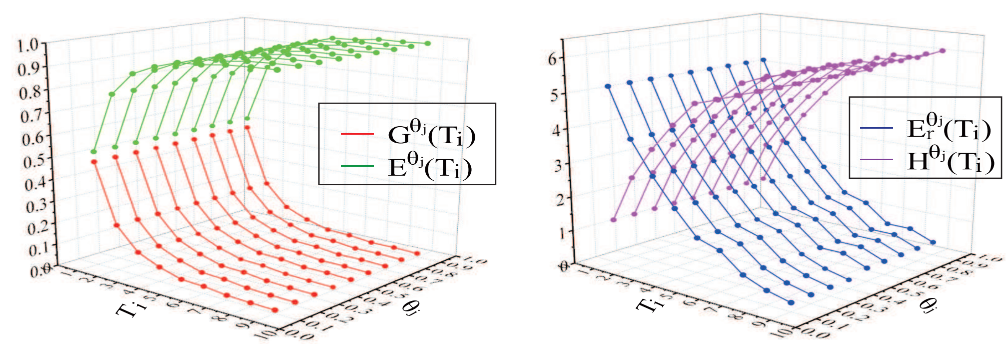
(a) **(b)**

Figure 8: Values of UM on Hepatitis.



(a) **(b)**

Figure 9: Values of UM on Meta-data.



(a) **(b)**

Figure 10: Values of UM on Post-Operative Patient.

while, E^θ and H^θ are both monotonically increasing with the attribute subset growth. It shows that the uncertainty of a 4HIS decreases as the attribute subset growth.

(2) When the attribute subset A is given, G^θ and E_r^θ are both monotonically increasing as the threshold θ increases. Meanwhile, E^θ and H^θ are both monotonically decreasing with the threshold θ growth. It shows that the uncertainty of a 4HIS increases as the threshold θ increases.

Thus, G^θ, E^θ, E_r^θ and H^θ can be applied to measuring uncertainty of a 4HIS.

6.2. Dispersion analysis

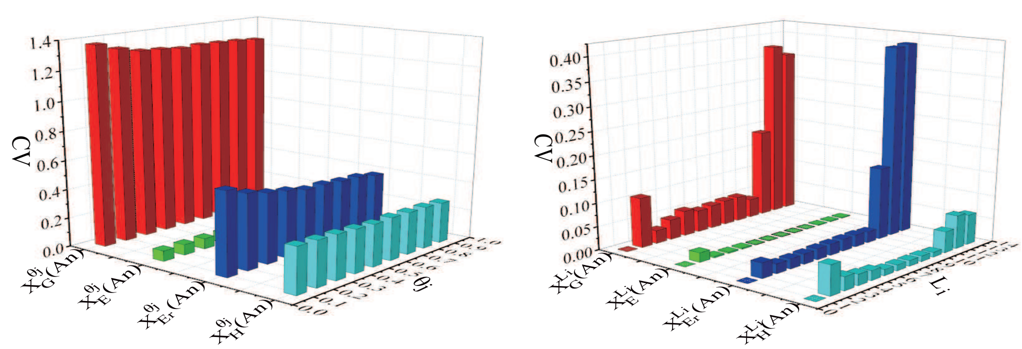
In this part, the standard deviation is applied to do effectiveness analysis of the proposed measures.

Let $X = \{x_1, \cdots, x_n\}$ be a data set. Then arithmetic average value, standard deviation and standard deviation coefficient of X are defined as follows:

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

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

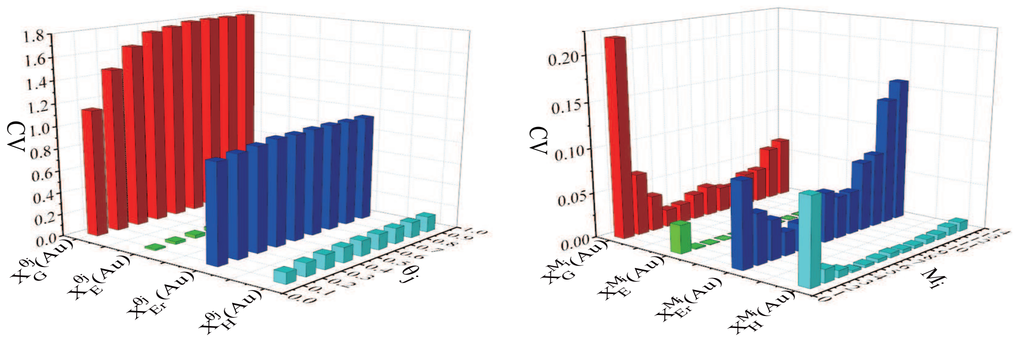
Then, according to the above experiments, the CV values of measure sets on each data set are computed (see Figures 10-18).



(a) (b)

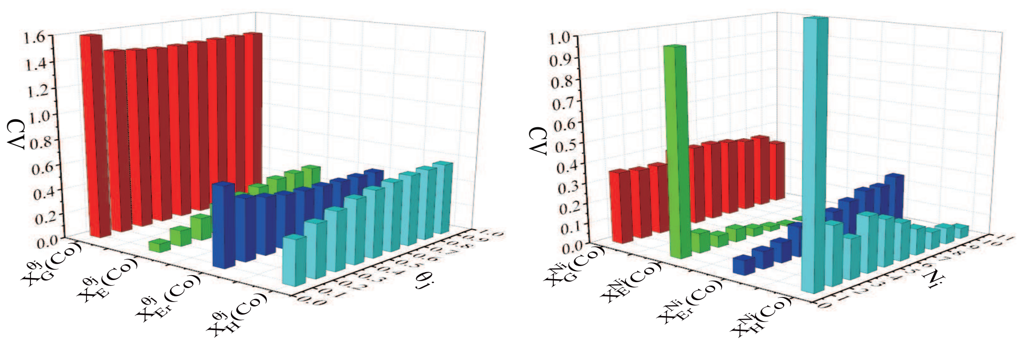
Figure 11: CV-values of measure sets on Annealing.

From Figures 10-18, the following conclusions are obtained:



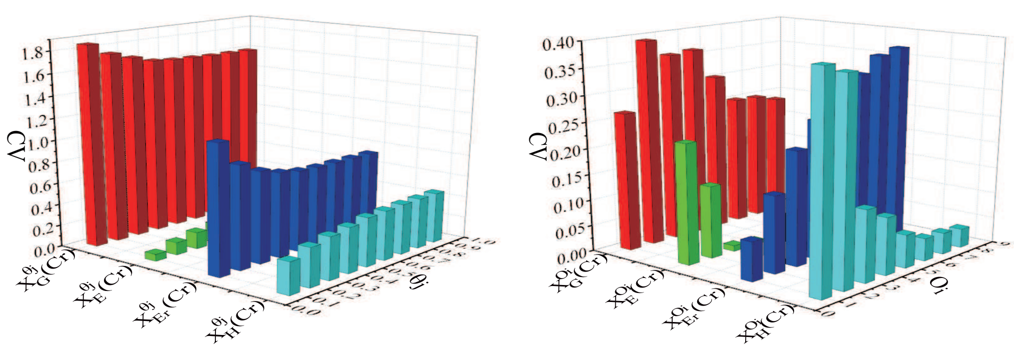
(a) **(b)**

Figure 12: CV-values of measure sets on Automobile.



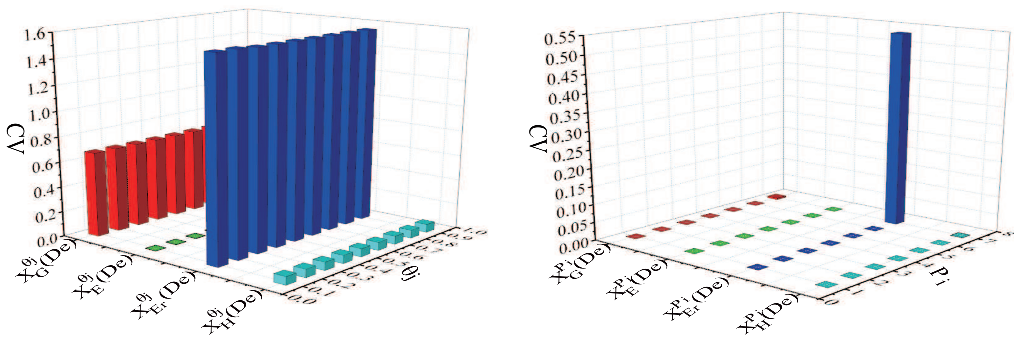
(a) **(b)**

Figure 13: CV-values of measure sets on Contraceptive Method Choice.



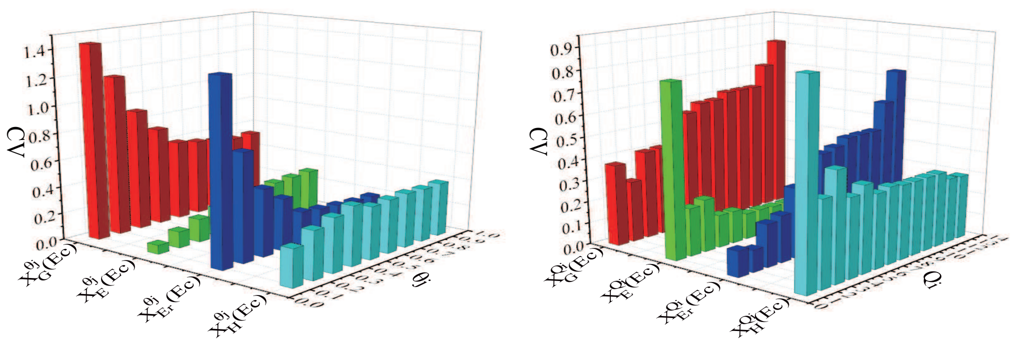
(a) **(b)**

Figure 14: CV-values of measure sets on Credit Approval.



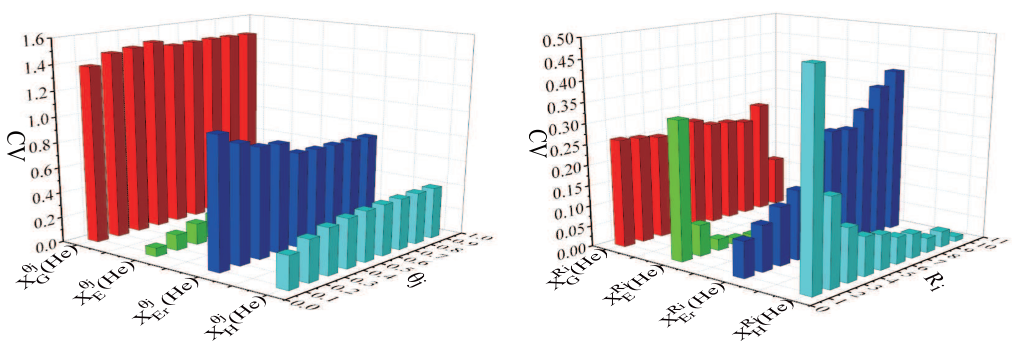
(a) **(b)**

Figure 15: CV-values of measure sets on Dermatology.



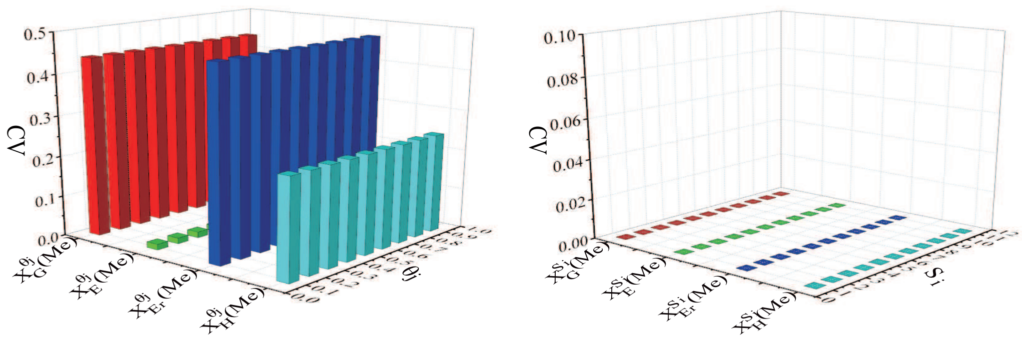
(a) **(b)**

Figure 16: CV-values of measure sets on Echocardiogram.



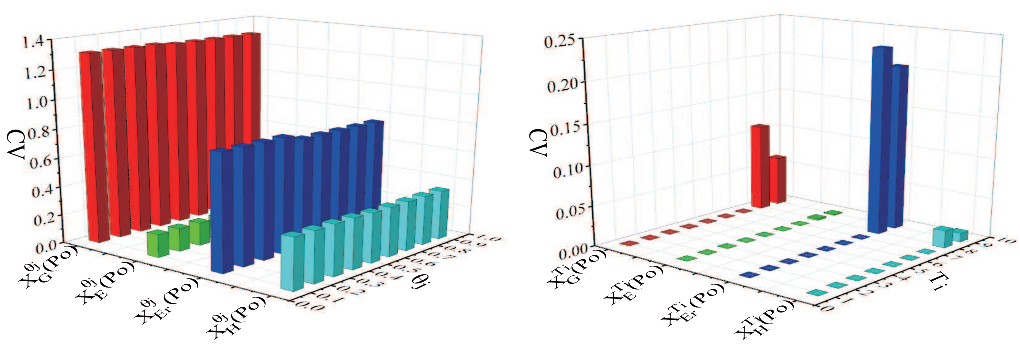
(a) **(b)**

Figure 17: CV-values of measure sets on Hepatitis.



(a) (b)

Figure 18: CV-values of measure sets on Meta-data.



(a) (b)

Figure 19: CV-values of measure sets on Post-Operative Patient.

(1) When the threshold θ is changing, $X_E^{\theta_j}$ is minimum in four measure sets in each data set except for $X_E^{\theta_6}(Ec)$, $X_E^{\theta_7}(Ec)$, $X_E^{\theta_8}(Ec)$ and $X_E^{\theta_9}(Ec)$. It shows that the dispersion degree of E^θ is minimum in most cases.

(2) When the attribute subset A_i is changing, $X_E^{A_i}$ is minimum in four measure sets in each data set except for $X_E^{N_1}(Co)$, $X_E^{N_2}(Co)$, $X_E^{O_1}(Cr)$, $X_E^{Q_1}(Ec)$, $X_E^{Q_2}(Ec)$, $X_E^{Q_3}(Ec)$ and $X_E^{R_1}(He)$. It shows that the dispersion degree of E^θ is minimum in most cases.

Therefore, E^θ has much better performance to measure 4HISs' uncertainty.

By summarizing the above experiments, the following results are obtained:

(1) If people need only monotonicity, then G^θ, E^θ, E_r^θ and H^θ can be used to measure the uncertainty of a 4HIS;

(2) If people investigate only dispersion degree, then E^θ has better performance to measure the uncertainty of a 4HIS;

(3) If people consider both monotonicity and dispersion degree, then E^θ has much better performance to measure the uncertainty of a 4HIS.

7. An application

In this part, an application of the proposed measures for attribute reduction in a 4HIS is given.

Definition 7.1. *Suppose that (X, AT) is a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then A is called θ-consistent, if $R^\theta_A = R^\theta_{AT}$.*

Definition 7.2. *Suppose that (X, AT) is a 4HIS. Given $a \in A \subseteq AT$. Then a is called θ-independent in A, if $R^\theta_A \neq R^\theta_{A-\{a\}}$.*

Definition 7.3. *Suppose that (X, AT) is a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then A is called θ-independent, if for any $a \in A$, a is θ-independent in A.*

Definition 7.4. *Suppose that (X, AT) is a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then A is called a θ-reduct of AT, if A is both θ-consistent and θ-independent.*

In this paper, the family of all θ-coordination subsets (resp., all θ-reducts) of AT is denoted by $co^\theta(AT)$ (resp., $red^\theta(AT)$).

Obviously,

$$A \in red^\theta(AT) \iff A \in co^\theta(AT) \text{ and } \forall\ P' \subset P,\ P' \notin co^\theta(AT).$$

Theorem 7.5. *Suppose that (X, AT) is a 4HIS. Given $A \subseteq AT$ and $\theta \in [0, 1]$. Then the following conditions are equivalent:*

(1) $A \in co^\theta(AT)$;

(2) $G^\theta(A) = G^\theta(AT)$;

(3) $H^\theta(A) = H^\theta(AT)$;

(4) $E^\theta_r(A) = E^\theta_r(AT)$;

(5) $E^\theta(A) = E^\theta(AT)$.

Proof. (1) $\Rightarrow$ (2). This is obvious.

(2) $\Rightarrow$ (1). Suppose $G^\theta(A) = G^\theta(AT)$. Then

$$\frac{1}{n^2}\sum_{i=1}^{n}|R_A^\theta(x_i)| = \frac{1}{n^2}\sum_{i=1}^{n}|R_{AT}^\theta(x_i)|.$$

So

$$\sum_{i=1}^{n}(|R_A^\theta(x_i)| - |R_{AT}^\theta(x_i)|) = 0.$$

Note that $R_{AT}^\theta \subseteq R_A^\theta$. Then $\forall\ i$, $R_{AT}^\theta(x_i) \subseteq R_A^\theta(x_i)$. This implies that

$$\forall\ i,\quad |R_A^\theta(x_i)| - |R_{AT}^\theta(x_i)| \geq 0.$$

So $\forall\ i,\ \ |R_A^\theta(x_i)| - |R_{AT}^\theta(x_i)| = 0$. It follows that $\forall\ i,\ \ R_A^\theta(x_i) = R_{AT}^\theta(x_i)$. Thus $R_A^\theta = R_{AT}^\theta$. Hence

$$A \in co^\theta(AT).$$

(2) $\Leftrightarrow$ (5). It can be obtained by Theorem 5.9.
(1) $\Rightarrow$ (3). This is obvious.
(3) $\Rightarrow$ (1). Suppose $H^\theta(A) = H^\theta(AT)$. Then

$$-\sum_{i=1}^{n}\frac{1}{n}\log_2\frac{|R_A^\theta(x_i)|}{n} = -\sum_{i=1}^{n}\frac{1}{n}\log_2\frac{|R_{AT}^\theta(x_i)|}{n}.$$

So

$$\sum_{i=1}^{n}\log_2\frac{|R_A^\theta(x_i)|}{|R_{AT}^\theta(x_i)|} = 0.$$

Note that $R_{AT}^\theta \subseteq R_A^\theta$. Then $\forall\ i$, $R_{AT}^\theta(x_i) \subseteq R_A^\theta(x_i)$. This implies that

$$\forall\ i,\quad \log_2\frac{|R_A^\theta(x_i)|}{|R_{AT}^\theta(x_i)|} \geq 0.$$

So $\forall\ i,\ \ \log_2\frac{|R_A^\theta(x_i)|}{|R_{AT}^\theta(x_i)|} = 0$. It follows that $\forall\ i,\ \ R_A^\theta(x_i) = R_{AT}^\theta(x_i)$. Thus $R_A^\theta = R_{AT}^\theta$. Hence

$$A \in co^\theta(AT).$$

(3) $\Leftrightarrow$ (4). It follows from Theorem 5.18. $\square$

Corollary 7.6. *Suppose that* (X, AT) *is a 4HIS. Given* $A \subseteq AT$ *and* $\theta \in [0, 1]$*. Then the following conditions are equivalent:*

(1) $A \in red^{\theta}(AT)$;

(2) $G^{\theta}(A) = G^{\theta}(AT)$ *and* $\forall\ a \in A$, $G^{\theta}(A - \{a\}) \neq G^{\theta}(AT)$;

(3) $H^{\theta}(A) = H^{\theta}(AT)$ *and* $\forall\ a \in A$, $H^{\theta}(A - \{a\}) \neq H^{\theta}(AT)$;

(4) $E_r^{\theta}(A) = E_r^{\theta}(AT)$ *and* $\forall\ a \in A$, $E_r^{\theta}(A - \{a\}) \neq E_r^{\theta}(AT)$;

(5) $E^{\theta}(A) = E^{\theta}(AT)$ *and* $\forall\ a \in A$, $E^{\theta}(A - \{a\}) \neq E^{\theta}(AT)$.

Proof. It can be proved by Theorem 7.5. $\square$

8. Conclusions

A HIS contains many types of attributes. It is more difficult to measure an HIS than an IS with a type of attribute. This paper has measured the uncertainty of a 4HIS that contains four types of data or attributes and given its application in attribute reduction. First, a novel distance function for each type of attribute in a 4HIS has been proposed. The proposed distance is more consistent with reality in measuring the difference between two information values on each type of attribute. And then, the tolerance relation has been produced by using the proposed distance. By the way, the information granules composed of the tolerance classes have been constructed, and the information structure formed by the information granules has been presented. Next, four UMs based on the structures have been investigated. Furthermore, the effectiveness of four measures has been verified by statistical analysis. AS an application of the proposed UMs, attribute reduction has been studied. We have found the influence of θ value on the UM for a 4HIS, which may have potential application value in data mining. This paper provides a new idea of UM for hybrid data. The disadvantage is that attribute reduction algorithms are not given. In the future, we will continue to explore attribute reduction algorithms in a 4HIS based on its UM.

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 (11761012, 11861010) and Given

Point on Master of Applied Statistics in Guangxi University of Finance and Economics (2016TJYB03).

Compliance with Ethical Standards

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

[1] M. Aggarwal, Representation of uncertainty with information and probabilistic information granules, International Journal of Fuzzy Systems 19(2017) 1617-1634.

[2] D. Bianucci, G. Cattaneo, Information entropy and granulation co-entropy of partitions and coverings: a summary, Transactions on Rough Sets 10(2009) 15-66.

[3] T. Beaubouef, F.E. Petry, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109(1998) 185-195.

[4] J. Blaszczynski, R. Slowinski, M. Szelag, Sequential covering rule induction algorithm for variable consistency rough set approaches, Information Sciences 181(5)(2011) 987-1002.

[5] Y.M. Chen, Y. Xue, Y. Ma, F.F, Xu, Measures of uncertainty for neighborhood rough sets, Knowledge-Based Systems 120(2017) 226-235.

[6] J.H. Dai, H. Hu, W.Z. Wu, Y.H. Qian, D.B. Huang, Maximal-discernibility-pair-based approach to attribute reduction in fuzzy rough sets, IEEE Transactions on Fuzzy Systems 26(4)(2018) 2175-2187.

[7] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17(2-3)(1990) 191-209.

[8] A. Delgado, I. Romero, Environmental conflict analysis using an integrated grey clustering and entropy-weight method: a case study of a mining project in Peru, Environmental Modelling Software 77(2016) 108-121.

[9] J. Dai, H. Tian, Entropy measures and granularity measures for set-valued information systems, Information Sciences 240(2013) 72-82.

[10] S. Greco, M. Inuiguchi, R. Slowinski, Fuzzy rough sets and multiplepremise gradual decision rules, International Journal of Approximate Reasoning 41(2006) 179-211.

[11] B. Gu, V.S. Sheng, Z.J. Wang, D. Ho, S. Osman, Incremental learning for v-support vector regression, Neural Networks 67(2015) 140-150.

[12] J.W. Han, M. Kamber, Data mining: concepts and techniques, Elsevier, 2011.

[13] Q.H. Hu, W. Pedrycz, D.R. Yu, J. Lang, Selecting discrete and continuous features based on neighborhood decision error minimization, IEEE Transactions on Systems, Man and Cybernetics (Part B) 40(2010) 137-150.

[14] C.F. Hempelmann, U. Sakoglu, V.P. Gurupur, S. Jampana, An entropy-based evaluation method for knowledge bases of medical information systems, Expert Systems with Applications 46(2016) 262-273.

[15] M. Juhola, J. Laurikkala, On metricity of two heterogeneous measures in the presence of missing values, Artificial Intelligence Review 28(2007) 163-178.

[16] T.Y. Lin, Granular computing on binary relations I: data mining and neighborhood systems, In: Rough Sets In Knowledge Discovery, A.Skowron and L.Polkowski (eds), Physica-Verlag (1998) 107-121.

[17] Z.W. Li, D. Huang, X.F. Liu, N.X. Xie, G.Q. Zhang, Information structures in a covering information system, Information Sciences 507(2020) 449-471.

[18] Z.W. Li, X.F. Liu, J.H. Dai, J.L. Chen, H. Fujita, Measures of uncertainty based on Gaussian kernel for a fully fuzzy information system, Knowledge-Based Systems 196(2020) 105791.

[19] J.Y. Liang, K.S. Qu, Information measures of roughness of knowledge and rough sets for information systems, Journal of Systems Science and Systems Engineering 10(2002) 95-103.

[20] Z.W. Li, P.F. Zhang, X. Ge, N.X. Xie, G.Q. Zhang, Uncertainty measurement for a covering information system, Soft Computing 23(2019) 5307-5325.

[21] Z.W. Li, P.F. Zhang, X. Ge, N.X. Xie, G.Q. Zhang, C.F. Wen, Uncertainty measurement for a fuzzy relation information system, IEEE Transactions on Fuzzy Systems 27(2019) 2338-2352.

[22] D.Q. Miao, L.S. Hou, A comparison of rough set methods and representative inductive learning algorithms, Fundamenta Informaticae 59(2)(2004) 203-219.

[23] J.S. Mi, Y. Leung, W.Z. Wu, An uncertainty measure in partition-based fuzzy rough sets, International Journal of General Systems 34(2005) 77-90.

[24] J. Ma, W. Zhang, Y. Leung, X. Song, Granular computing and dual Galois connection, Information Sciences 177(2007) 5365-5377.

[25] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11(1982) 341-356.

[26] Z. Pawlak, Rough sets: Theoretical aspects of reasoning about data, Kluwer Academic Publishers, Dordrecht, 1991.

[27] Y.H. Qian , J.Y. Liang, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16(2008) 179-193 .

[28] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Knowledge structure, knowledge granulation and knowledge distance in a knowledge base, International Journal of Approximate Reasoning 50(2009) 174-188.

[29] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Information granularity in fuzzy binary GrC model, IEEE Transactions on Fuzzy Systems 19(2)(2011) 253-264.

[30] C.E. Shannon, A mathematical theory of communication, The Bell System Technical Journal 27(1948) 379-423.

[31] W.Z. Wu, Y. Leung, J. Mi, Granular computing and knowledge reduction in formal contexts, IEEE Transactions on Knowledge and Data Engineering 21(10)(2009) 1461-1474.

[32] N.X. Xie, M. Liu, Z.W. Li, G.Q. Zhang, New measures of uncertainty for an interval-valued information system, Information Sciences 470(2019) 156C174.

[33] S.D. Xie, Y.X. Wang, Construction of tree network with limited delivery latency in homogeneous wireless sensor networks, Wireless Personal Communications 78(1)(2014) 231-246.

[34] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences 111(1998) 239-259.

[35] Y.Y. Yao, Probabilistic approaches to rough sets, Expert Systems 20(2003) 287-297.

[36] G.J. Yu, Information structures in an incomplete information system: A granular computing viewpoint, International Journal of Computational Intelligence Systems 11(2018) 1179-1191.

[37] G.J. Yu, Information structures and uncertainty measures in a hybrid information system with images, Soft Computing 23(2019) 12961-12979.

[38] Z. Yuan, H.M. Chen, T.R. Li, B.B. Sang, S. Wang, Outlier detection based on fuzzy rough granules in mixed attribute data, IEEE IEEE Transactions on Cybernetics, DOI: 10.1109/TCYB.2021.3058780.

[39] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90(1997) 111-127.

[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 Computing 2(1998) 23-25.

[41] L.A. Zadeh, Fuzzy sets, Information and Control 8(1965) 338-353.

[42] X. Zhang, C.L. Mei, D.G. Chen, J.H. Li, Feature selection in mixed data: A method using a novel fuzzy rough set-based information entropy, Pattern Recognition 56(2016) 1-15.

[43] A.P. Zeng, T.R. Li, D. Liu, J.B. Zhang, H.M. Chen, A fuzzy rough set approach for incremental feature selection on hybrid information systems, Fuzzy Sets and Systems 258(2015) 39-60.

[44] A.P. Zeng, T.R. Li, J. Hu, H.M. Chen, C. Luo, Dynamical updating fuzzy rough approximations for hybrid data under the variation of attribute values, Information Sciences 378(2017) 363-388.

[45] G.Q. Zhang, Z.W. Li, W.Z. Wu, Information structures and uncertainty measures in a fully fuzzy information system, International Journal of Approximate Reasoning 101(2018) 119-149.

[46] L. Zhang, B. Zhang, Theory and application of problem solving-theory and application of granular computing in quotient spaces, Tsinghua University Publishers, Beijing, 2007.