# A New Information Filling Technique Based On Generalized Information Entropy

S. Han, L. Chen, Z. Zhang, J.-X. Li

**Shan Han\*, Lin Chen, Zhi Zhang, Jianxun Li**
Department of Automation, Shanghai Jiao Tong University,
Key Laboratory of System Control and Information Processing,
Ministry of Education of China, Shanghai 200240, China
E-mail: hanshan@sjtu.edu.cn, clcyk2002@aliyun.com,
mark1896bc@gmail.com, lijx@sjtu.edu.cn
\*Corresponding author: hanshan@sjtu.edu.cn

**Abstract:** Multi-sensor decision fusion used for discovering important facts hidden in a mass of data has become a widespread topic in recent years, and has been gradually applied in failure analysis, system evaluation and other fields of big data process. The solution to incompleteness is a key problem of decision fusion during the experiment and has been basically solved by proposed technique in this paper. Firstly, as a generalization of classical rough set, interval similarity relation is employed to classify not only single-valued data but also interval-valued data in the information systems. Then, a new kind of generalized information entropy called "H'-Information Entropy" is suggested based on interval similarity relation to measure the uncertainty and the classification ability in the information systems. Thus, the innovated information filling technique using the properties of H'-Information Entropy can be applied to replace the missing data by some smaller estimation intervals. Finally, the feasibility and advantage of this technique are testified by two actual applications of decision fusion, whose performance is evaluated by the quantification of E-Condition Entropy.
**Keywords:** Multi-Sensor Decision Fusion, Rough Set Theory, Generalized Information Entropy, Information Classification, Information Filling.

## 1 Introduction

Decision fusion is the general name of data-based decision making methods. The purposes of these methods are automatically discovering important facts hidden in a mass of data collected from varieties of sensors or other sources, and expressing them by the natural language of decision rules. Then, people can effectively use the simplified decision rules to aid decision making in future applications. Obviously, this target puts forward higher requirements for both sensors capability and the information processing method. Besides basic statistical method, neural networks and Bayesian networks, rough set theory as a favorable mathematical tool with high performance of information acquisition and classification takes into account consequently [1] [2].

The theory of rough set was firstly proposed by Pawlak in 1982 [3]. It is an extension of classical set theory for the study of information systems characterized by inexact, uncertain and vague, and has been widely used in the fields of knowledge discovery, decision fusion, data mining, pattern recognition, and so on. With the rapid development of rough set, some new frameworks were proposed to extend its application range. The tolerance relation [4], valued tolerance relation [5], similarity relation [6], together with the limited tolerance relation [7], laid a solid foundation for the progress of generalized rough set. Until now, the improvements of binary relations are still attracting the scholars attentions [8]- [11].

As to decision fusion of multi-sensor, data acquired from different sensors is real-valued which characterizes objects of interest. Because of the imprecision of acquisition, the fuzziness of cognition and the limitation of knowledge [12] [13], these attributes with real-valued data usually expressed as Single-valued Information System (SIS), Interval-valued Information System (IIS)

or the combination of SIS and IIS cannot be classified by any binary relation mentioned above absolutely. The discretization of a long interval by dividing the range into a certain number of partitioning small intervals and using symbolic values to replace these small intervals before further calculation is the popular handling method [14] [15], but existing dispute about the cut-off points [16]. Additionally, the classical rough set may generate an unacceptably large number of classifications from the discretized data resulting in too many classification rules to make final decisions. More important, data missing and uncertain called incompleteness may occur unavoidably because of loss of documents, difficulties of measurement or malfunction of sensors, and increases the difficulties of information processing [17]. The direct information filling methods have not been paid enough attention to, and existed methods [12] [13] are lack of precision to make optimal decisions. Thereupon, the researches of solutions to the classifications and the incompleteness in information systems, together with the following calculations of rough set become significant [18].

In this paper, the discussions are paying attention to solve above issues based on rough set theory emphatically. The notion, properties and measurement of generalized rough set are reviewed firstly in Section 2. In Section 3, a brand new "H'-Information Entropy" is defined for the measurement of uncertainty and classification ability in the information systems according to the interval similarity relation. In Section 4, an innovated information filling technique is proposed by the use of H'-Information Entropy to solve the incompleteness in both SIS and IIS, whose performance is evaluated by the quantification of E-Condition Entropy. In Section 5, the advantage and practicability of this new information filling technique are testified by two integrated examples. The missing data in the decision table is filled by smaller estimation intervals and the certain decision rules are extracted by the calculation of rough set finally.

## 2    The Theory of Generalized Rough Set

Information systems with the form of decision tables provide a convenient basis for the representation of objects in terms of their attributes. But the uncertainty and incoordination take place in most of decision tables. Thereupon, rough set as a powerful method is used to deal with these issues.

### 2.1    Basic Concept

Let $S = (U, A)$ be an information system, where $U = \{u_1, u_2, \ldots, u_{|U|}\}$ ($|U|$ means the cardinality of $U$) is a non-empty finite set of objects, denoting the whole research objects of the information system. $A = \{a_1, a_2, \ldots, a_n\}$ is a non-empty finite set of attributes, denoting the whole attributes of the information system. $f : U \to V_a$ is a mapping for $a \in A$ ,where $V_a$ is called the domain of attribute $a$. This constitutes the basic research contents of rough set.

Let $P \subseteq A$ be a subset of attributes. $T$ is a binary relation (including but not limited to equivalence relation). $S_p(u)$ denotes the object set $\{v \in U|(u, v) \in T\}$ , called a class or an information granule. If $(u_i, u_j) \in T$ , then $S_P(u_i) \neq S_P(u_j)$ in general conditions.

The whole classes of $U$ are obtained according to $T$ [19]:

$$K(P) = (S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|}))  \tag{1}$$

At this time, $K(P)$ is not restricted to "partition", but may be the "covering" of $U$, namely:

$$\cup S_P(u_i) = U \qquad i = 1, 2, \ldots, |U|  \tag{2}$$

Then, a rough set can be described by the definitions of lower and upper approximations.

**Definition 1.** Let $S = (U, A)$ be an information system, $X \subseteq U$. With an arbitrary binary equivalence relation on $U$, the lower and upper approximations of $X$ are [3]:

$$\underline{R}(X) = \cup\{S_P(u_i)|S_P(u_i) \subseteq X\} \qquad i = 1, 2, \ldots, |U| \tag{3}$$

$$\overline{R}(X) = \cup\{S_P(u_i)|S_P(u_i) \cap X \neq \emptyset\} \qquad i = 1, 2, \ldots, |U| \tag{4}$$

The covering-based rough set is the great generalization of classical rough set. So, the application range of rough set is extended significantly.

## 2.2    The Measurement of Rough Set

The uncertainty in rough set refers to the size of classes called information granules generally determined by specified attributes in the information system. It can be measured by information entropy [20]. Small information granules imply a precise description of rough set.

**Definition 2.** Let $S = (U, A)$ be an information system, and $P \subseteq A$ be a subset of attributes. $K(P) = (S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|}))$ are the classes of $U$ created by some binary relation according to $P$. Information entropy of $P$ is defined as [19]:

$$H(P) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|S_P(u_i)|}{|U|} \tag{5}$$

This definition of entropy is called "H-Information Entropy". In other words, H-Information Entropy quantifies the classification ability of objects set $U$ according to the attributes subset $P$.

The relationship between different classes can be compared by partial relation " $\preceq$ ". Let $P, Q \subseteq A$ be subsets of attributes. If $K(P) = (S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|}))$ and $K(Q) = (S_Q(u_1), S_Q(u_2), \ldots, S_Q(u_{|U|}))$. The partial relation " $\preceq$ " is defined as:

$$K(P) \preceq K(Q) \Leftrightarrow \forall i = 1, 2, \ldots, |U|, S_P(u_i) \subseteq S_Q(u_i) \tag{6}$$

**Theorem 3.** *If $K(P) \preceq K(Q)$, then $H(Q) \leq H(P)$ [19].*

This theorem explains that H-Information Entropy is monotone increasing with the decrease of elements in each class. The thinner classes are obtained by stricter binary relations. The bigger the H-Information Entropy is, the smaller the information granules will be.

In order to mine the internal relations of subsets $P$ and $Q$, the research of coordination of these two attributes becomes significant. The result embodies the support degree of one subset of attributes to another. It is just the major application of rough set in decision fusion.

The condition entropy is introduced to measure this support degree [21]. It is proved that E-Condition Entropy is suitable for both partition-based and covering-based rough set [19] [22].

**Definition 4.** Let $P, Q \subseteq A$ be two subsets of attributes. If $K(P) = (S_P(u_1), S_P(u_2), \ldots, S_P(u_{|U|}))$ and $K(Q) = (S_Q(u_1), S_Q(u_2), \ldots, S_Q(u_{|U|}))$, then E-Condition Entropy is:

$$E(Q/P) = \sum_{i=1}^{|U|} \frac{|S_P(u_i)| - |S_P(u_i) \cap S_Q(u_i)|}{|U|^2} \tag{7}$$

In particular, if $K(P) \preceq K(Q)$, which means the classes obtained by $P$ are totally included in the classes obtained by $Q$, then $E(Q/P) = 0$ [23]. Thus, some attributes "entirely" support the some others. Rough set degenerates to the classical set. It is considered that if E-Condition Entropy is small, the support degree is high. Then, the coordination in rough set is better.

# 3   Generalized Information Entropy

In the decision tables, the values of attributes may be some interval ranges instead of simple single values. In order to measure the uncertainty in such kind of system, the new information entropy is defined in this section, called H' information entropy.

## 3.1   Interval Similarity Relation

The information system with interval-valued data is now called Interval-valued Information System (IIS).

**Definition 5.** Let $S = (U, A)$ be an IIS. If $x \in U$ and $a \in A$, then the interval range of $f(x, a)$ is denoted as $f(x, a) = [f(x, a)^L, f(x, a)^U]$.

As mentioned in Introduction, it is clear that the classical binary relations are not suitable for the classifications of IIS now. Additionally, it is considered that the length of interval implies the uncertainty of cognition. Because the intersection degrees among objects are highly different, the negligence will enlarge the possibility of intersection factitiously. Thus, the classification ability and the cognition of the system will be reduced. So, the interval similarity relation is employed as the binary relation to classify the objects of IIS in this paper.

**Definition 6.** Let $S = (U, A)$ be an IIS. $x, y \in U$ and $a \in A$. The interval similarity degree of $f(x, a)$ and $f(y, a)$ is defined as:

$$P_{xy}^a = \frac{|f(x, a) \cap f(y, a)|}{|f(x, a) \cup f(y, a)|} \tag{8}$$

where $||$ means the absolute length of interval.

In particular, there may exist some single-valued data in IIS in practical applications. That is $f(x, a) = f(x, a)^L = f(x, a)^U = Constant$. In order to generalize the definition of interval similarity degree, if $f(y, a) = [f(y, a)^L, f(y, a)^U]$ and $Constant \in [f(y, a)^L, f(y, a)^U]$, then we define $P_{xy}^a = 1$, and vice versa. Thus, the definition of similarity degree can be improved.

**Definition 7.** Let $S = (U, A)$ be an IIS. $x, y \in U$ and $a \in A$. the interval similarity relation is defined as:

$$S(A) = \{(x, y) \in U \times U | P_{xy}^a \geq \alpha, \forall a \in A\} \tag{9}$$

where $\alpha$ is the similarity threshold value.

Obviously, if $f(x, a)$ and $f(y, a)$ are both single-valued data, then the interval similarity relation will degenerate to simple equivalence relation by this definition.

*Deduction* 1. The interval similarity relation satisfies reflexivity and symmetry.

It is significant that if an information system is classified by interval similarity relation, the interval length has direct effect on the classification ability. That is interval similarity may affect information entropy of the system.

## 3.2   Definition of Generalized Information Entropy

As the improvement, interval similarity degree defined in Equation (8) will be added to classical information entropy. A new kind of generalized information entropy is defined to measure the uncertainty and classification ability in the information systems.

**Definition 8.** Let $S = (U, A)$ be an information system and $B \subseteq A$ be a subset of attributes. $K(B) = (S_B(u_1), S_B(u_2), \ldots, S_B(u_{|U|}))$ are the classes of $U$ created by interval similarity relation according to $B$. $P_{xy}$ is the minimum interval similarity degree between two objects ($x$ and $y$).

$$P_{xy} = \begin{cases} min\{P_{xy}^k\} & P_{xy}^k > 0, \forall k \in B \\ 0 & else \end{cases} \tag{10}$$

The generalized information entropy called "H'-Information Entropy" is defined as:

$$H\prime(B) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{\sum_{j=1}^{|U|} P_{u_i u_j}}{|U|} \tag{11}$$

*Deduction* 2. H'-Information Entropy satisfies non-negativity, symmetry, continuity, monotonicity and extremum property.

Proof.

1) By the definition of interval similarity degree, it is obvious that $0 \leq P_{u_i u_j} \leq 1$ . Then,

$$0 < \sum_{j=1}^{|U|} P_{u_i u_j} \leq |U| \Rightarrow 0 < \frac{\sum_{j=1}^{|U|} P_{u_i u_j}}{|U|} \leq 1 \Rightarrow -\log_2 \frac{\sum_{j=1}^{|U|} P_{u_i u_j}}{|U|} \geq 0 \tag{12}$$

That is H'-Information Entropy is nonnegative.

2) Moreover, because of the symmetry of interval similarity degree ($P_{u_i u_j} = P_{u_j u_i}$), H'-Information Entropy is also symmetric.

3) The numerical ranges of $P_{u_i u_j}$ are continuous and $P_{u_i u_j}$ constitute the whole variables of H'-Information Entropy. The operations of addition, multiplication, and logarithm are continuous totally. Hence, H'-Information Entropy keeps the continuity.

4) Let $S = (U, A)$ be an information system and $B_1, B_2 \subseteq A$ be subsets of attributes. The condition that $B_1 \subseteq B_2$ is assumed, then there is $K(B_2) \preceq K(B_1)$ [24].

Because $P_{xy}^{B_1} = min\{P_{xy}^k\}$ ($P_{xy}^k > 0, \forall k \in B_1$) and $P_{xy}^{B_2} = min\{P_{xy}^k\}$ ($P_{xy}^k > 0, \forall k \in B_2$), with the assumption $B_1 \subseteq B_2$ , the consequent that $P_{xy}^{B_1} \geq P_{xy}^{B_2}$ is obtained. Hence, $H\prime(B_1) \leq H\prime(B_2)$. As a result,

$$K(B_2) \preceq K(B_1) \Rightarrow H\prime(B_1) \leq H\prime(B_2) \tag{13}$$

It should be noticed that this monotonicity is under the constraint ($B_1 \subseteq B_2$) in this paper. If there is no inclusion relation between two subsets of condition attributes, this monotonicity may be broken down.

5) If $\sum_{j=1}^{|U|} P_{u_i u_j} = 1$, then $H\prime(B) = \log_2 |U|$. If $\sum_{j=1}^{|U|} P_{u_i u_j} = |U|$, then $H\prime(B) = 0$. Hence, there exist two extreme values.

## 4   The Solution to Incompleteness

Data missing and uncertain called incompleteness occur commonly in data acquisition by multi-sensor. In this section, the research focus will turn to this incomplete information system and try to find the solution to incompleteness.

## 4.1   The New Information Filling Technique

Now, a new information filling method is proposed based on the H'-Information Entropy defined above.

**Theorem 9.** *In an IIS with one attribute, it will assuredly increase the H'-Information Entropy when a new object of minimal interval is added into the system.*

Proof. Let $S = (U, A)$ be an IIS. $\forall a \in A$, $S = (U, a)$ is an IIS with one attribute. $u_{|U+1|}$ is the newly added object. There are:

$$f(u_{|U+1|}, a) = [f(u_{|U+1|}, a)^L, f(u_{|U+1|}, a)^U] \tag{14}$$

$$|f(u_{|U+1|}, a)| = f(u_{|U+1|}, a)^U - f(u_{|U+1|}, a)^L = \delta \tag{15}$$

When $u_{|U+1|}$ is added into original IIS, the new H'-Information Entropy is:

$$
\begin{aligned}
H\prime(a)_1 &= -\sum_{i=1}^{|U+1|} \frac{1}{|U+1|} \log_2 \frac{\sum_{j=1}^{|U+1|} P_{u_i u_j}}{|U+1|} \\
&= -\frac{1}{|U+1|}\left(\log_2 \frac{\sum_{j=1}^{|U|} P_{u_1 u_j} + P_{u_1 u_{|U+1|}}}{|U+1|} + \log_2 \frac{\sum_{j=1}^{|U|} P_{u_2 u_j} + P_{u_2 u_{|U+1|}}}{|U+1|} + \cdots \right. \\
&\quad \left. + \log_2 \frac{\sum_{j=1}^{|U|} P_{u_{|U|} u_j} + P_{u_{|U|} u_{|U+1|}}}{|U+1|} + \log_2 \frac{\sum_{j=1}^{|U|} P_{u_{|U+1|} u_j} + P_{u_{|U+1|} u_{|U+1|}}}{|U+1|}\right)
\end{aligned} \tag{16}
$$

Obviously, $P_{u_{|U+1|} u_{|U+1|}} = 1$. $\sum_{j=1}^{|U|} P_{u_i u_j} (i = 1, 2, \ldots, |U|)$ is unchanged since the newly added object does not affect the intersection relations in original system. When $\delta$ is a minimal interval, there is $\delta \ll |f(u_i, a)| (i = 1, 2, \ldots, |U|)$. Therefore, whether $\delta$ has an intersection with other interval or not, it is obtained:

$$P_{u_i u_{|U+1|}} = P_{u_{|U+1|} u_i} = 0 \qquad i = 1, 2, \ldots, |U| \tag{17}$$

As a result:

$$
\begin{aligned}
H\prime(a)_1 &= -\frac{1}{|U+1|}\left(\log_2 \frac{\sum_{j=1}^{|U|} P_{u_1 u_j}}{|U+1|} + \log_2 \frac{\sum_{j=1}^{|U|} P_{u_2 u_j}}{|U+1|} + \cdots + \log_2 \frac{\sum_{j=1}^{|U|} P_{u_{|U|} u_j}}{|U+1|} + \log_2 \frac{1}{|U+1|}\right) \\
&\stackrel{\text{def}}{=} H\prime(a)_{10}
\end{aligned} \tag{18}
$$

Let $\sum_{j=1}^{|U|} P_{u_i u_j} = P_i$ for convenience, then:

$$
\begin{aligned}
H\prime(a)_{10} &= -\frac{1}{|U+1|}\left(\log_2 \frac{P_1}{|U+1|} + \log_2 \frac{P_2}{|U+1|} + \cdots + \log_2 \frac{P_{|U|}}{|U+1|} + \log_2 \frac{1}{|U+1|}\right) \\
&= -\frac{1}{|U+1|} \log_2 \frac{P_1 P_2 \cdots P_{|U|}}{|U+1|^{|U+1|}} \\
&= \frac{1}{|U+1|} \log_2 \frac{|U+1|^{|U+1|}}{P_1 P_2 \cdots P_{|U|}}
\end{aligned} \tag{19}
$$

The H'-Information Entropy of original system is:

$$
\begin{aligned}
H\prime(a)_0 &= -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{\sum\limits_{j=1}^{|U|} P_{u_i u_j}}{|U|} \\
&= \frac{1}{|U|} \log_2 \frac{|U|^{|U|}}{P_1 P_2 \cdots P_{|U|}}
\end{aligned}
\tag{20}
$$

Because $|U| \geq 1$ and $P_i \geq 1 (i = 1, 2, \ldots, |U|)$, Hence,

$$
\begin{aligned}
H\prime(a)_{10} &= \frac{1}{|U+1|} \log_2 \frac{|U+1|^{|U+1|}}{P_1 P_2 \cdots P_{|U|}} \\
&= \log_2 |U+1| - \frac{1}{|U+1|} \log_2(P_1 P_2 \cdots P_{|U|}) \\
&> \log_2 |U| - \frac{1}{|U+1|} \log_2(P_1 P_2 \cdots P_{|U|}) \\
&> \log_2 |U| - \frac{1}{|U|} \log_2(P_1 P_2 \cdots P_{|U|}) \\
&= H\prime(a)_0
\end{aligned}
\tag{21}
$$

It is clear that, Theorem 9 can be expanded to a SIS. That is in a SIS with one attribute, when a new object of minimal interval is added into the system, it will assuredly increase the H'-Information Entropy if the interval range of newly added object doesn't include any existed single-valued data in the attribute. As long as this promise exists, $P_{u_i u_{|U+1|}} = P_{u_{|U+1|} u_i} = 0$ as well according to Equation (8). Consequently, the simplifying process from Equation (16) to (18) stays the same.

In order to make the research more intuitive, the following discussions are based on the assumption that the increase of interval is unidirectional. Hence, the lower limit or upper limit is fixed.

**Theorem 10.** *In an information system with one attribute, the H'-Information Entropy will be monotone decreasing with the interval increase of newly added object within a limit.*

Proof. Followed by previous proof, $\delta$ is increased gradually.

1) If the original system is an IIS, obviously, only the terms of $p_i \overset{\text{def}}{=} P_{u_i u_{|U+1|}} = P_{u_{|U+1|} u_i} = \frac{|f(u_i,a) \cap \delta|}{|f(u_i,a) \cup \delta|}$ are changing during the increase of $\delta$. Firstly, $H\prime(a)_1$ will be monotone decreasing because of the continuous increase of $p_i$. The increase trend of $p_i$ will be sustained until the intersection of $u_i(i = 1, 2, \ldots, |U|)$ and $u_{|U+1|}$ is steady, which means the upper /lower limit of $\delta$ reaches the upper /lower limit of arbitrary $u_i$ (as showed in Fig.1). After that, $|f(u_i, a) \cap \delta|$ is a constant while $|f(u_i, a) \cup \delta|$ is still increased. Then, $p_i$ will turn to decrease. Once this happens to any $u_i \in U$, the trend of $H\prime(a)_1$ may be changed.

2) If the original system is a SIS, there should exist growing amount of $p_i = 1 (i = 1, 2, \ldots, |U|)$ with the increase of $\delta$. Thus, $H\prime(a)_1$ will be monotone decreasing until the interval range of $\delta$ is long enough to cover the largest number of original single-valued data in the domain of the attribute $a$.

Theorem 10 shows that the newly added object will have more opportunity to be classified with other objects with interval increase under the interval similarity relation. This phenomenon will cause the weakness of classification ability and the reduction of information entropy.
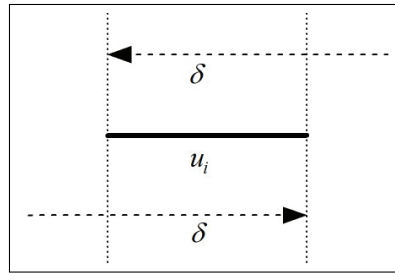
Figure 1: Intersection of Two Intervals.

*Deduction* 3. In an information system with one attribute, there is a minimum value of H'-Information Entropy with the interval increase of newly added object.

$H\prime(a)_1$ is continuous according to Deduction 2 and it has been proved that $H\prime(a)_1$ is monotone decreasing within a limit. Consequently, there is a minimum value in the limit.

*Deduction* 4. In an IIS with one attribute, after the upper /lower limit of newly added object reaches the maximal /minimal value in the domain of the attribute $a$, the H'-Information Entropy will be monotone increasing with the interval increase.

*Deduction* 5. In an SIS with one attribute, after the upper /lower limit of newly added object reaches the maximal /minimal value in the domain of the attribute $a$, the H'-Information Entropy will be a constant with the interval increase.

In an IIS, once the upper /lower limit of newly added object reaches the upper /lower limit of interval-valued data of entire original objects, $p_i$ will turn to decrease with the increase of $\delta$ for all $u_i \in U$. Then, $H\prime(a)_1$ is changed to increase assuredly. Meanwhile, in a SIS, once the upper /lower limit of newly added object reaches the maximal /minimal value of single-valued data of entire original objects, $p_i$ will be the number of covered single-valued data and remains unchanged regardless of the interval extension. Then, $H\prime(a)_1$ is a constant assuredly.

The deductions above explain that when the interval of the newly added object is increased from zero to the maximal /minimal value in the domain of the attribute $a$, there exists at least one global minimum value of H'-Information Entropy.

Now, the missing data in the information systems can be dealt with above-mentioned results. The new information filling technique follows the assumption that the objects with missing data have no effect with the classifications of information. In other words, regardless of the existing of the existence of these incomplete objects, the information entropy of corresponding condition attribute should be unchanged.

The process of the new information filling technique is considered as a new object of minimal interval adds into the complete information system attribute by attribute. Then, the H'-Information Entropy is surely increased implying the rapid improvement of classification ability. According to Deduction 3, 4 and 5, the trend of H'-Information Entropy is decreased firstly and increased or flat finally. The filling interval is expected to make the H'-Information Entropy close to the original one as much as possible before the upper /lower limit reaches the maximal /minimal value in the domain of corresponding attribute. Thus, a smaller filling interval is obtained. Because of the expanded definition of interval similarity relation, this new information filling technique can be applied for both IIS and SIS.

In general, the process is listed below.

1) Calculate the H'-Information Entropy ($H\prime(a)_0$) of original information system for each condition attribute without incomplete objects.

2) Add a new object with minimal interval ($\delta = 10^{-3}$ is recommended) to one attribute. If

the original system is SIS, the lower /upper limit of the interval should be set at the left /right side of minimal /maximal value in the domain of corresponding attribute of the objects with the same decision. If the original system is IIS, the lower /upper limit of the interval should be known and fixed in advance just like the forms of $f(x,a) = [f(x,a)^L, *]$ or $f(x,a) = [*, f(x,a)^U]$. Thus, the initial position marked as $A$ is established. The H'-Information Entropy is $H\prime(a)_1|_{start}$ at the beginning.

3) Forward /backward increase the length of the new interval by a setting step size until its upper /lower limit reaches the maximal /minimal value in the domain of the attribute. Fig.2 is just showing the process of forward increasing for instance. Calculate the new H'-Information Entropy for each step. Record the following data:

a. Mark the position of interval's upper /lower limit as $B$ when the corresponding new H'-Information Entropy reaches global minimum value $H\prime(a)_1|_{min}$ for the first time.

b. The position of maximal /minimal value in the domain of the attribute is marked as $C$, and the corresponding H'-Information Entropy is $H\prime(a)_1|_{end}$.

c. Mark the position of interval's upper (lower) limit as $D$ when the corresponding new H'-Information Entropy is equal to the original H'-Information Entropy $(H\prime(a)_1 = H\prime(a)_0)$ for the first time. $D$ may not exist in some cases.

4) If $H\prime(a)_1|_{min} \leq H\prime(a)_0$ , then the estimation interval is $|A - D|$ (as showed in Fig. 2(Left)).If $H\prime(a)_1|_{min} > H\prime(a)_0$ ,then the estimation interval is $|A - B|$ (as showed in Fig. 4(Right)). Thus, there are $|A - D| \leq |A - C|$ and $|A - B| \leq |A - C|$, so the estimation interval is shortened.
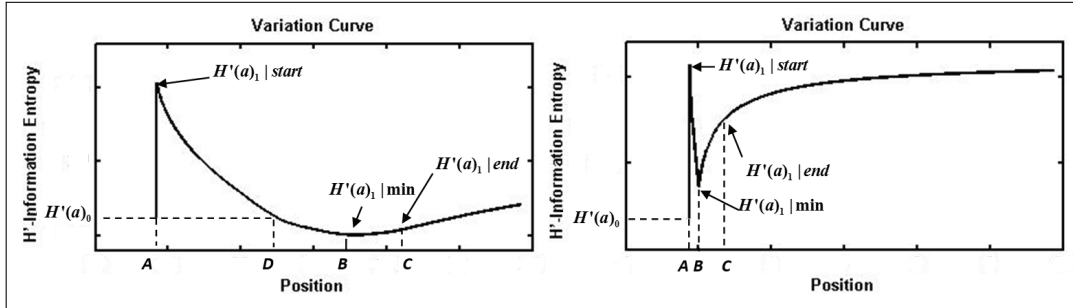


Figure 2: Variation Curve of H'-Information Entropy.

## 4.2 Performance Analysis

In order to check the performance of proposed information filling technique, a deduction is introduced firstly.

*Deduction* 6. In an information system, the shortening of estimation intervals will improve the coordination of the system.

Proof. Let $S = (U, A)$ be an information system. $A = C \cup D$ , where $C$ is the set of condition attributes and $D$ is set of the decision attributes. If $B \subseteq C$ , then the classes obtained according to $B$ and $D$ are:

$$K(B) = (S_B(u_1), S_B(u_2), \ldots, S_B(u_{|U|})) \tag{22}$$

$$K(D) = (S_D(u_1), S_D(u_2), \ldots, S_D(u_{|U|})) \tag{23}$$

With the shortening of estimation intervals in $B$, there exists a new classification:

$$K(B)\prime = (S_B(u_1)\prime, S_B(u_2)\prime, \ldots, S_B(u_{|U|})\prime) \tag{24}$$

Obviously, short estimation interval will decrease the intersection degree and reduce the objects in the classes made by interval similarity relation. That is $S_B(u_i)\prime \subseteq S_B(u_i)$, $|S_B(u_i)\prime| \leq |S_B(u_i)|$ $(i = 1, 2, \ldots, |U|)$. Hence, $S_B(u_i)\prime \cap S_D(u_i) \subseteq S_B(u_i) \cap S_D(u_i)$ and $|S_B(u_i)\prime \cap S_D(u_i)| \leq |S_B(u_i) \cap S_D(u_i)|$. E-Condition Entropy studied in Section 2.2 is used to measure the coordination in rough set. According to its definition, there is:

$$
\begin{aligned}
E(D/B) - E(D/B)\prime &= \sum_{i=1}^{|U|} \frac{|S_B(u_i)| - |S_B(u_i) \cap S_D(u_i)|}{|U|^2} - \sum_{i=1}^{|U|} \frac{|S_B(u_i)\prime| - |S_B(u_i)\prime \cap S_D(u_i)|}{|U|^2} \\
&= \sum_{i=1}^{|U|} \frac{|S_B(u_i)| - |S_B(u_i)\prime| - (|S_B(u_i) \cap S_D(u_i)| - |S_B(u_i)\prime \cap S_D(u_i)|)}{|U|^2} \\
&= \sum_{i=1}^{|U|} \frac{|S_B(u_i)| - |S_B(u_i)\prime| - |(S_B(u_i) - S_B(u_i)\prime) \cap S_D(u_i)|}{|U|^2} \\
&\geq 0
\end{aligned}
\tag{25}
$$

The equality holds if and only if $(S_B(u_i) - S_B(u_i)\prime) \subseteq S_D(u_i)$. So, $E(D/B) \geq E(D/B)\prime$.

It is mentioned that small E-Condition Entropy implies great support degree, which means the shortening of estimation intervals improves the coordination between corresponding condition attributes and decision attributes. Because of the generality of this result to the whole condition attributes, the coordination of the information system is improved consequently.

Now, the benefits of smaller estimation intervals obtained in this paper are reflected in two aspects. On the one hand, the coordination between condition attributes and decision attributes is improved according to Deduction 6. That is the positive region in the rough set is expanded, while the uncertainty is decreased. The reliability of decision will be increased as a result. On the other hand, the shortened estimation intervals are beneficial for decision making. The new decision rules will be fine and precise with the extending of covering range.

## 5    Experimental Studies

The new information filling technique can be applied to process the incompleteness in large amounts of data acquired by different sensors for decision fusion in a variety of fields. Here, two integrated experiments will be studied below to validate the feasibility and superiority of proposed method.

### 5.1    Single-valued Data Estimation and Filling

In the study of single-valued data evaluation and filling, a part of ICU data of one patient is acquired by different sensors and listed in Table 1. In this decision table, the Column "$U$" refers to the times of examinations. Columns "$c1$" to "$c5$" denote the data acquired by sensors of temperature, heart rate, systolic blood pressure, spo2 and respiration rate, and compose five condition attributes. The Column "$d$" denotes four degrees of ICU decisions and composes the decision attribute. Furthermore, the missing data is denoted as "$*$" in the table.

By applying the new information filling technique, a small interval can be estimated to fill the missing data. Figure 3 shows this process of information filling and we can get $f(u17, c1) = [36.3, 36.4]$ (as showed in Figure 3(left)) and $f(u3, c3) = [103.0, 115.9]$ (as showed in Figure 3(right)) as a result.

Then, E-Condition Entropy discussed before can be used to evaluate the performance of above result. We get $E(D/C) = 0$ without the objects with missing data, meaning the condition

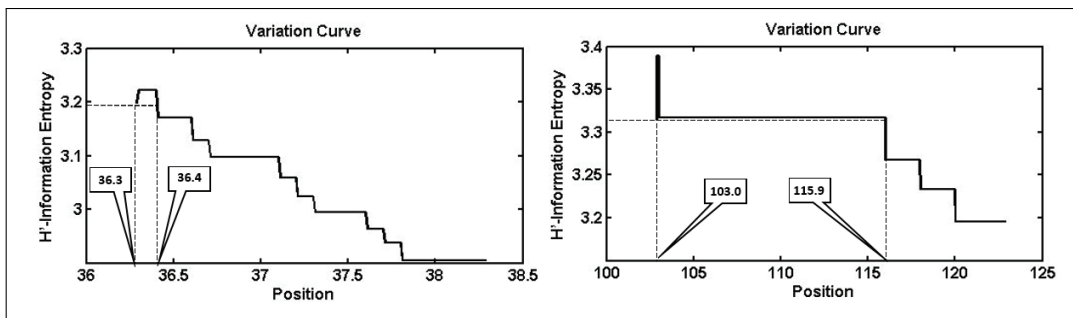| $U$ | $c1$ | $c2$ | $c3$ | $c4$ | $c5$ | $d$ |
|---|---|---|---|---|---|---|
| 1 | 37.8 | 62.0 | 103.0 | 94.0 | 18.0 | 2 |
| 2 | 37.7 | 61.0 | 120.0 | 97.0 | 16.0 | 2 |
| 3 | 37.7 | 61.0 | * | 97.0 | 16.0 | 2 |
| ... | ... | ... | ... | ... | ... | ... |
| 14 | 37.3 | 62.0 | 116.0 | 98.0 | 18.0 | 3 |
| 15 | 37.3 | 62.0 | 116.0 | 98.0 | 18.0 | 3 |
| 16 | 37.3 | 62.0 | 116.0 | 98.0 | 18.0 | 3 |
| 17 | * | 111.0 | 142.0 | 95.0 | 20.0 | 1 |
| 18 | 36.3 | 111.0 | 142.0 | 95.0 | 20.0 | 1 |
| 19 | 36.4 | 111.0 | 142.0 | 95.0 | 20.0 | 1 |
| ... | ... | ... | ... | ... | ... | ... |
| 49 | 36.7 | 98.0 | 133.0 | 92.0 | 20.0 | 4 |
| 50 | 36.7 | 98.0 | 133.0 | 92.0 | 20.0 | 4 |

Table 1: Incomplete ICU Data.



Figure 3: Information Filling of ICU data.

attributes entirely support each decision in the original system. Then, this index is still 0 when u3 and u17 with filling intervals are added to the system. That is the new information filling technique doesn't change the coordination between condition attributes and decision attribute in this example and the results make sense.

## 5.2   Interval-valued Data Estimation and Filling

In the study of interval-valued data estimation and filling, a part of thruster experimental data is acquired by different sensors. In order to simplify the table, only the objects with missing data denoted as "∗" are listed in Table 2. The Column "$U$" refers to the times of experiment. Columns "$c1$" to "$c4$" denote the data acquired by different pressure and temperature sensors, and compose four condition attributes. These acquired data is formed as interval values including the minimum and maximum from the each sensor without data preprocessing. The Column "$d$" denotes three types of failure simplified as symbolic values and composes the decision attribute. We need to fill the missing data effectively and discover the failure rules hidden in the data as accurately as possible.

1) Firstly, the decision table is completed by the method in Ref. [13] and showed in Table 3. Briefly, "∗" is replaced by the minimal /maximal value in the domain of each condition attribute.

2) Then, the decision table is completed by the new information filling technique proposed in this paper and showed in Table 4.

In order to reveal the difference between above two methods, the classifications of condition attributes are made by interval similarity relation. The cardinality of each class obtained by the two methods can be compared in Figure 4.

| U | c1 | c2 | c3 | c4 | d |
|---|-----|-----|-----|-----|---|
| 3 | [4.344, 4.604] | [0.029, 0.037] | [4.600, *] | [23.865, 26.819] | 1 |
| 6 | [4.283, *] | [0.024, 0.033] | [4.613, 4.889] | [23.872, 25.826] | 1 |
| 21 | [4.801, 4.997] | [0.027, 0.035] | [*, 4.516] | [28.700, 30.848] | 2 |
| 25 | [*, 4.993] | [0.023, 0.030] | [4.627, 4.904] | [28.663, 30.806] | 2 |
| 28 | [4.445, 4.711] | [0.025, 0.037] | [4.236, 4.490] | [28.826, *] | 2 |
| 43 | [2.629, *] | [0.024, 0.037] | [4.600, 4.876] | [30.881, 32.116] | 3 |
| 46 | [2.653, 3.812] | [0.025, 0.033] | [4.619, 4.896] | [30.016, *] | 3 |
| 55 | [1.470, 2.551] | [0.027, 0.030] | [4.507, *] | [30.497, 31.717] | 3 |
| ... | ... | ... | ... | ... | ... |

Table 2: Incomplete Thruster Experimental Data.

| U | c1 | c2 | c3 | c4 | d |
|---|-----|-----|-----|-----|---|
| 3 | [4.344, 4.604] | [0.029, 0.037] | [4.600, 4.999] | [23.865, 26.819] | 1 |
| 6 | [4.283, 4.999] | [0.024, 0.033] | [4.613, 4.889] | [23.872, 25.826] | 1 |
| 21 | [4.801, 4.997] | [0.027, 0.035] | [4.062, 4.516] | [28.700, 30.848] | 2 |
| 25 | [1.015, 4.993] | [0.023, 0.030] | [4.627, 4.904] | [28.663, 30.806] | 2 |
| 28 | [4.445, 4.711] | [0.025, 0.037] | [4.236, 4.490] | [28.826, 32.150] | 2 |
| 43 | [2.629, 4.999] | [0.024, 0.037] | [4.600, 4.876] | [30.881, 32.116] | 3 |
| 46 | [2.653, 3.812] | [0.025, 0.033] | [4.619, 4.896] | [30.016, 32.150] | 3 |
| 55 | [1.470, 2.551] | [0.027, 0.030] | [4.507, 4.999] | [30.497, 31.717] | 3 |
| ... | ... | ... | ... | ... | ... |

Table 3: Completed Thruster Experimental Data 1.

The histogram indicates that, with the same binary relation, more than 30% classes obtained from the completed decision table by the new information filling technique are "thinner" than before. That is, the new information filling technique can achieve a classification ability no less than the traditional one. The classes are more precise and the uncertainty is lower now. Additionally, E-Condition Entropy is calculated as $E(D/C) = 0.0033$ in Table 4, while this index is 0.0106 in Table 3. The result also implies a better coordination obtained by the new information filling technique between condition attributes and decision attributes.

Then, the reduct of the system is $c1, c3, c4$ obtained by heuristic algorithm of attribute reduction based on E-Condition Entropy [22] [25]. The certain decision rules (failure rules) can be achieved explicitly (as showed in Table 5) since all the missing data has been filled.

At present, the estimation intervals in each attribute are relatively small in this group of certain failure rules. It has a better significance for the practical decision making. The result is agreed with the performance analysis in Section 4.2. Consequently, the problems of incompleteness in thruster experimental data are solved and the failure rules are achieved favorably.
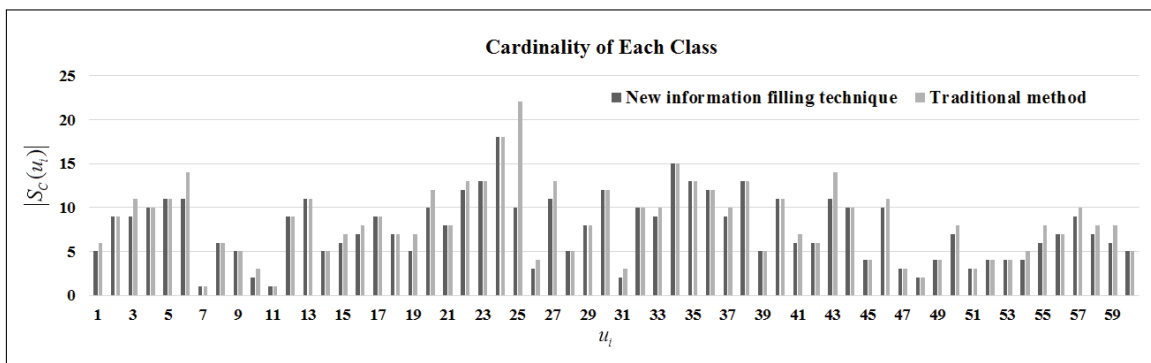


Figure 4: Cardinality of Each Class.

| $U$ | $c1$ | $c2$ | $c3$ | $c4$ | $d$ |
|---|---|---|---|---|---|
| 3 | [4.344, 4.604] | [0.029, 0.037] | [4.600, 4.753] | [23.865, 26.819] | 1 |
| 6 | [4.283, 4.601] | [0.024, 0.033] | [4.613, 4.889] | [23.872, 25.826] | 1 |
| 21 | [4.801, 4.997] | [0.027, 0.035] | [4.235, 4.516] | [28.700, 30.848] | 2 |
| 25 | [4.877, 4.993] | [0.023, 0.030] | [4.627, 4.904] | [28.663, 30.806] | 2 |
| 28 | [4.445, 4.711] | [0.025, 0.037] | [4.236, 4.490] | [28.826, 30.471] | 2 |
| 43 | [2.629, 3.810] | [0.024, 0.037] | [4.600, 4.876] | [30.881, 32.116] | 3 |
| 46 | [2.653, 3.812] | [0.025, 0.033] | [4.619, 4.896] | [30.016, 31.240] | 3 |
| 55 | [1.470, 2.551] | [0.027, 0.030] | [4.507, 4.837] | [30.497, 31.717] | 3 |
| ... | ... | ... | ... | ... | ... |

Table 4: Completed Thruster Experimental Data 2.

| Rules | Descriptions |
|---|---|
| $R1$ | If $c1 \in [4.458, 4.725] \wedge c3 \in [4.867, 4.997] \wedge c4 \in [23.779, 25.730]$, then $d = 1$ |
| $R2$ | If $c1 \in [4.456, 4.732] \wedge c3 \in [4.358, 4.619] \wedge c4 \in [23.717, 25.658]$, then $d = 1$ |
| ... | ... |
| $R19$ | If $c1 \in [4.801, 4.997] \wedge c3 \in [4.235, 4.516] \wedge c4 \in [28.700, 30.848]$, then $d = 2$ |
| $R20$ | If $c1 \in [4.817, 4.989] \wedge c3 \in [4.350, 4.611] \wedge c4 \in [29.200, 30.328]$, then $d = 2$ |
| ... | ... |
| $R35$ | If $c1 \in [2.947, 4.215] \wedge c3 \in [4.868, 4.999] \wedge c4 \in [30.391, 31.607]$, then $d = 3$ |
| $R36$ | If $c1 \in [3.678, 4.047] \wedge c3 \in [4.358, 4.619] \wedge c4 \in [29.950, 31.148]$, then $d = 3$ |
| ... | ... |

Table 5: Certain Decision Rules.

# 6 Conclusions

The solution to incompleteness in decision fusion is presented with the help of a new information filling technique based on generalized information entropy in this paper. Interval similarity relation is employed as the binary relation to classify both single-valued and interval-valued data acquired by different sensors. As the improvement of classical rough set, a new kind of generalized information entropy called H'-Information Entropy is provided to solve the incompleteness occurring unavoidably in information systems. So, the missing data in the decision tables can be filled evidently and favorably. The calculation procedure of this innovated information filling technique is listed in detail. Finally, two integrated examples are given to evaluate the performance of proposed method. The results turn out the superiority and practicality in real applications.

# Acknowledgements

# Bibliography

[1] Pawlak, Z.; Skowron, A. (2007); Rudiments of rough sets. *Information Science*, ISSN 0020-0255, 177(1): 3-27.

[2] Tay, F.E.H.; Shen, L.-X.(2003); Fault diagnosis based on rough set theory. *Engineering Applications of Artificial Intelligence*, ISSN 0952-1976, 16(1): 39-43.

[3] Pawlak, Z.(1982); Rough sets. *International Journal of Computer & Information Sciences* , ISSN 0885-7458, 11(5): 341-356.

[4] Kryszkiewicz, M.(1998); Rough Set Approach to Incomplete Information Systems. *Information Sciences*, ISSN 0020-0255, 112(1-4): 39-49.

[5] Stefanowski, J.; Tsoukiàs, A.(1999); On the Extension of Rough Sets under Incomplete Information. *New Directions in Rough Sets, Data Mining, and Granular-Soft Computing, Lecture Notes in Computer Science*, ISSN 0302-9743, 1711: 73-81.

[6] Stefanowski, J.; Tsoukiàs, A.(2001); Incomplete Information Tables and Rough Classification. *Computational Intelligence*, ISSN 0824-7935, 17(3): 545-566.

[7] Wang, G.-Y.(2002); Extension of Rough Set Under Incomplete Information Systems. *Journal of Computer Research and Development*, ISSN 1000-1239, 39(10): 1238-1243.

[8] Greco, S.; Matarazzo, B.; Slowinski, R.(1999); Rough Approximation of A Preference Relation By Dominance Relations. *European Journal of Operational Research*, ISSN 0377-2217, 117(1): 63-83.

[9] Guan, Y.-Y.; Wang, H.-K.(2006); Set-valued information systems. *Information Sciences*, ISSN 0020-0255, 176(17): 2507-2525.

[10] Leung, Y.; Wu, W.-Z.; Zhang, W.-X.(2006); Knowledge acquisition in incomplete information systems: A rough set approach. *European Journal of Operational Research*, ISSN 0377-2217, 168(1): 164-180.

[11] Yang, X.-B.; Yu, D.-J.; Yang, J.-Y.; Song, X.-N.(2009); Difference Relation-Based Rough Set And Negative Rules In Incomplete Information System. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, ISSN 0218-4885, 17(5): 649-665.

[12] Yang, X.-B.; Yu, D.-J.; Yang, J-Y.; Wei, L.-H.(2009); Dominance-based rough set approach to incomplete interval-valued information system. *Data & Knowledge Engineering*, ISSN 0169-023X, 68(11): 1331-1347.

[13] Zhao, L.; Zhang, X.; Xue, Z.(2011); Security Assessment for Incomplete Interval-valued Information System. *Computer Engineering*, ISSN 1000-3428, 37(11): 146-148.

[14] Grzymala-Busse, J.W.; Stefanowski, J.(2001); Three Discretization Methods for Rule Induction. *International Journal of Intelligent Systems*, ISSN 1098-111X, 16(1): 29-38.

[15] Beynon, M.J.(2004); Stability of continuous value discretisation: An application within rough set theory. *International Journal of Approximate Reasoning*, ISSN 0888-613X, 35(1): 29-53.

[16] Leung, Y.; Fischer, M.-M.; Wu, W.-Z.; Mi, J.-S.(2008) A rough set approach for the discovery of classification rules in interval-valued information systems. *International Journal of Approximate Reasoning*, ISSN 0888-613X, 47(2): 233-246.

[17] Yang, X.-B.; Yang, J.-Y.(2012); Incomplete Information System and Rough Set Theory-Models and Attribute Reductions, Springer- Verlag: Berlin, Heidelberg, ISBN 978-3-642-25934-0.

[18] Zhang, N.; Miao, D.-Q.; Yue, X.-D.(2010); Approaches to Knowledge Reduction in Interval-Valued Information Systems. *Journal of Computer Research and Development*, ISSN 1000-1239, 47(8): 1362-1371.

[19] Liang, J.-Y.; Qian, Y.-H.(2008); Information granules and entropy theory in information systems. *Science in China Series F: Information Sciences*, ISSN 1674-733X, 51(10): 1427-1444.

[20] Düntsch, I.; Gediga, G.(1998); Uncertainty Measures of Rough Set Prediction. *Artificial Intelligence*, ISSN 0004-3702, 106(1): 109-137.

[21] Wang, J.; Miao, D.-Q.(1998); Analysis on Attribute Reduction Strategies of Rough Set. *Journal of Computer Science and Technology*, ISSN 1000-9000, 13(2): 189-193.

[22] Teng, S.-H.; Zhou, S.-L.; Sun, J.-X.; Li, Z.-Y.(2010); Attribute Reduction Algorithm Based on Conditional Entropy under Incomplete Information System. *Journal of National University of Defense Technology*, ISSN 1001-2486, 32(1): 90-94.

[23] Liang, J.-Y.; Chin, K.S.; Dang, C.-Y.; Yam, R.C.M.(2002); A new method for measuring uncertainty and fuzziness in rough set theory. *International Journal of General Systems*, ISSN 0308-1079, 31(4): 331-342.

[24] Xu, J.-C.; Sun, L.(2010); A New Knowledge Reduction Algorithm Based on Decision Power in Rough Set. *Transactions on Rough Sets XII, Lecture Notes in Computer Science*, ISSN 0302-9743, 6109: 76-89.

[25] Li, F.; Yin, Y.-Q.(2009); Approaches to knowledge reduction of covering decision systems based on information theory. *Information Sciences*, ISSN 0020-0255, 179(11): 1694-1704.