Serdica J. Computing **7** (2013), No 4, 355–374

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

A DISTANCE-BASED METHOD FOR ATTRIBUTE REDUCTION IN INCOMPLETE DECISION SYSTEMS*

Janos Demetrovics, Vu Duc Thi, Nguyen Long Giang

Abstract. There are limitations in recent research undertaken on attribute reduction in incomplete decision systems. In this paper, we propose a distance-based method for attribute reduction in an incomplete decision system. In addition, we prove theoretically that our method is more effective than some other methods.

1. Introduction. Attribute reduction is one of the most important problems in data preprocessing, in knowledge discovery and data mining. Attribute reduction based on rough sets is the process of finding a minimal attribute set, known as *reduct*, which preserves some necessary information of decision systems. There have been many methods to find reducts of complete decision systems [17], such as positive region methods, discernibility matrix methods, information entropy methods, granular computing methods. In reality, decision systems often contain *missing values* in the domain values of attributes and these decision systems are called incomplete decision systems. Derived from the idea of rough set

ACM Computing Classification System (1998): I.5.2, I.2.6.
Key words: Rough set, incomplete decision system, attribute reduction, distance, reduct.
*Supported by RFBR, No.12-07-00755-a, RusMES No.1.345.2011.

theory [11], Marzena Kryszkiewicz [5] defines a tolerance relation based on the equivalent relation and proposes tolerance rough set. Recently, much research has been undertaken on measures and methods to find reducts in incomplete decision systems [1, 3, 4, 7, 8, 9, 12, 13, 20]. Though distance has been a popular measure applied to solve some problems in data mining [16, 18, 19], there is limited reseach on attribute reduction in rough set theory. Yuhua Qian et al. [14, 15] propose distances between coverings in incomplete decision systems. Long Giang Nguyen [10] proposes a distance-based method to find reduct of a complete decision system.

In this paper, we propose a distance-based method for attribute reduction in incomplete decision systems. We first generalize Liang entropy [6] in incomplete decision systems. Based on generalized Liang entropy, we establish a distance between attributes and study some properties of the distance. As a result, we use the proposed distance to formally define a reduct and the importance of attribute, and later construct a heuristic algorithm to find the best reduct.

This paper consists of six sections. The concept of tolerance rough set in incomplete systems is introduced in Section 2. The generalized Liang entropy and its propeties are proposed in Section 3. Section 4 establishes a distance between two attributes based on the generalized Liang entropy and studies some properties of the distance. Section 5 proposes a distance-based method and example to find the best reduct. Section 6 presents our conclusions.

2. Basic concepts.

In this section, we summarize the basic concepts of tolerance rough sets in incomplete decision systems [5].

Let U be a set of objects and $Attr$be a set of attributes. Then $IS = (U, Attr)$ is called an information system. A decision system is an information system $DS = (U, Attr \cup \{d\})$ where $Attr$ is a conditional attribute and d is a decision attribute. An incomplete decision system is a decision system where there exists an attribute $a \in Attr$ so that $a$contains a *missing value*. Further on, a missing value is denoted as '$*$'. Table 1 is an example of an incomplete decision system.

Attributes *Price, Mileage, Size* and *Max-speed* are called conditional attributes and *Decision* is the decision attribute. We denote the decision attribute *Decision* as d, and the conditional attributes *Price, Mileage, Size* and *Max-speed* as $a_1, \ldots, a_4$ in order. Consequently, Table 1 is an incomplete decision system $IDS = (U, Attr \cup \{d\})$ where $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ and $Attr = \{a_1, a_2, a_3, a_4\}$.

Table 1. An example of an incomplete decision system

Car	Price	Mileage	Size	Max-speed	*Decision*
x_1	High	High	Full	Low	Good
x_2	Low	$*$	Full	Low	Good
x_3	$*$	$*$	Compact	High	Poor
x_4	High	$*$	Full	High	Good
x_5	$*$	$*$	Full	High	Excellent
x_6	Low	High	Full	$*$	Good

For any attribute set $A \subseteq Attr$, a *tolerance relation* $TLR(A)$ is defined on $U \times U$ for any $x, y \in U$ as follows:

$$(x, y) \in TLR(A) rightarrow \forall_{a \in Attr}(a(x) = a(y) \vee a(x) = * \vee a(y) = *).$$

It is clear that $TLR(A) = \cap_{a \in A} TLR(\{a\})$. The tolerance relation $TLR(A)$ determines a *covering* of U which is denoted by $K(A)$ or $U/TLR(A)$. Then, $K(A) = U/TLR(A) = \{T_A(x) | x \in U\}$ where $T_A(x) = \{y \in U | (x, y) \in TLR(A)\}$. $T_A(x)$ is called a *tolerance class*. It shows that $T_A(x) \neq \emptyset$ for any $x \in U$ and $\cup_{x \in U} T_A(x) = U$. The set of all $K(A)$ where $A \subseteq Attr$ is denoted as $COV(U)$. For coverings in $COV(U)$, $\omega = \{T_{Attr}(x) = \{x\} | x \in U\}$ is called the *discrete covering* and $\delta = \{T_{Attr}(x) = \{U\} | x \in U\}$ is called the *indiscrete covering*. A partial relation is defined on $COV(U)$ as follows:

Definition 2.1 [9]**.** *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$ *and two attribute sets* A, $B \subseteq Attr$,

1) $U/TLR(A) = U/TLR(B)$ *if and only if* $\forall x \in U$, $T_A(x) = T_B(x)$.

2) $U/TLR(A) \preceq U/TLR(B)$ *if and only if* $\forall x \in U$, $T_A(x) \subseteq T_B(x)$.

Property 2.1 [9]**.** *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$ *and two attribute sets* A, $B \subseteq Attr$, *the following properties hold:*

1) *If* $A \subseteq B \subseteq Attr$ *then* $U/TLR(B) \preceq U/TLR(A)$.

2) *If* A, $B \subseteq Attr$ *then* $T_{A \cup B}(x) = T_A(x) \cap T_B(x)$ *for any* $x \in U$.

Let $IDS = (U, Attr \cup \{d\})$ be an incomplete decision system. For any $A \subseteq Attr$ and $x \in U$, $\partial_A(x) = \{d(y) | y \in T_A(x)\}$ is called the *generalized decision*. If $|\partial_{Attr}(x)| = 1$ for any $x \in U$ then IDS is *consistent*. Otherwise, it is *inconsistent*.

One of the most important concepts in tolerance rough sets is *reduct*. According to Kryszkiewicz [5], a reduct of an incomplete decision system is a minimal subset of a conditional attribute set which keeps the generalized decision unchanged for all objects.

Definition 2.2 [5]. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$*, if an attribute set* $R \subseteq Attr$ *satisfies*

(1) $\partial_R(x) = \partial_{Attr}(x)$ *for any* $x \in U$*;*

(2) $R - \{r\}$ *is not satisfied* (1) *for any* $r \in R$*,*

then R *is called a reduct of IDS based on generalized decision.*

Referring to Table 1, $T_{Attr}(x_1) = \{x_1\}$, $T_{Attr}(x_2) = \{x_2, x_6\}$, $T_{Attr}(x_3) = \{x_3\}$, $T_{Attr}(x_4) = \{x_4, x_5\}$, $T_{Attr}(x_5) = \{x_5, x_4, x_6\}$, $T_{Attr}(x_6) = \{x_6, x_2, x_5\}$, we have the covering $K(Attr) = \{\{x_1\}, \{x_2, x_6\}, \{x_3\}, \{x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_2, x_5, x_6\}\}$.

For $R = \{a_3, a_4\}$, we obtain the covering

$$K(R) = U/TLR(R) = \{T_R(x) | x \in U\}$$
$$= \{\{x_1, x_2, x_6\}, \{x_1, x_2, x_6\}, \{x_3\}, \{x_4, x_5, x_6\}, \{x_4, x_5, x_6\}, \{x_1, x_2, x_4, x_5, x_6\}\}.$$

For the attribute set $Attr$, we have $\partial_{Attr}(x_1) = \partial_{Attr}(x_2) = \{good\}$, $\partial_{Attr}(x_3) = \{poor\}$, $\partial_{Attr}(x_4) = \partial_{Attr}(x_5) = \partial_{Attr}(x_6) = \{good, excellent\}$. For the attribute set R, we have $\partial_R(x_1) = \partial_R(x_2) = \{good\}$, $\partial_R(x_3) = \{poor\}$, $\partial_R(x_4) = \partial_R(x_5) = \partial_R(x_6) = \{good, excellent\}$. As a result, we obtain $\partial_R(x) = \partial_{Attr}(x)$ for any $x \in U$. In addition, $\partial_{\{a_3\}}(x) = \partial_{Attr}(x)$ and $\partial_{\{a_4\}}(x) = \partial_{Attr}(x)$ is incorrect for any $x \in U$. According to Definition 2.2, R is a reduct based on generalized decision.

3. Generalized Liang Entropy and Properties.

3.1. Generalized Liang Entropy.

Definition 3.1. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$ *where* $U = \{x_1, \ldots, x_{|U|}\}$, $A \subseteq Attr$ *and* $U/TLR(A) = \{T_A(x_1), T_A(x_2), \ldots, T_A(x_{|U|})\}$*. We define generalized Liang entropy of* P *as*

$$IE(A) = \sum_{i=1}^{|U|} \frac{1}{|U|} \left(1 - \frac{|T_A(x_i)|}{|U|}\right),$$

where $|T_A(x)|$ *is the cardinality of* $T_A(x)$*. If* $U/TLR(A) = \omega$ *then* $IE(A)$ *has the maximum value* $IE(A) = 1 - \dfrac{1}{|U|}$*. If* $U/TLR(A) = \delta$ *then* $IE(A)$ *has the minimum value* $IE(A) = 0$*. Obviously,* $0 \leq IE(A) \leq 1 - \dfrac{1}{|U|}$.

The following Proposition 3.1 proves that Liang entropy $E(A)$ in [6] is a particular case of our generalized Liang entropy.

Proposition 3.1. *Given a complete decision system* $DS = (U,\ Attr \cup \{d\})$, $A \subseteq Attr$, $U = \{x_1, \ldots, x_{|U|}\}$ *and* $U/A = \{A_1, A_2, \ldots, A_m\}$, *then*

$$IE(A) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|T_A(x_i)|}{|U|}\right) = \sum_{i=1}^{m} \frac{|A_i|}{|U|}\left(1 - \frac{|A_i|}{|U|}\right) = E(A),$$

where $E(A)$ *is Liang entropy in* [6].

Proof. Suppose that $A_i = \{x_{i1}, x_{i2}, \ldots, x_{ip_i}\}$ where $|A_i| = p_i$ and $\sum_{i=1}^{m} p_i = |U|$.

$$A_i = T_A(x_{i1}) = T_A(x_{i2}) = \cdots = T_A(x_{ip_i}),$$
$$|A_i| = |T_A(x_{i1})| = |T_A(x_{i2})| = \cdots = |T_A(x_{ip_i})| = p_i$$

$$\begin{aligned}\frac{|A_i|}{|U|}\left(1 - \frac{|A_i|}{|U|}\right) &= \frac{1}{|U|}\left(|A_i| - \frac{|A_i|\,|A_i|}{|U|}\right)\\ &= \frac{1}{|U|}\left(1 - \frac{|T_A(x_{i1})|}{|U|} + 1 - \frac{|T_A(x_{i2})|}{|U|} + \cdots + 1 - \frac{|T_A(x_{ip_i})|}{|U|}\right)\end{aligned}$$

$$\begin{aligned}E(A) &= \sum_{i=1}^{m} \frac{|A_i|}{|U|}\left(1 - \frac{|A_i|}{|U|}\right) = \sum_{i=1}^{m}\sum_{k=1}^{p_i} \frac{1}{|U|}\left(1 - \frac{|T_A(x_{ik})|}{|U|}\right)\\ &= \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|T_A(x_i)|}{|U|}\right) = IE(A).\end{aligned}$$

Consequently, we have $E(A) = IE(A)$. The proposition is proved. □

Definotion 3.2. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$, *where* $U = \{x_1, \ldots, x_{|U|}\}$ *and* $A, B \subseteq Attr$. *We define generalized Liang entropy of* $A \cup B$ *as*

$$IE(A \cup B) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|T_{A\cup B}(x_i)|}{|U|}\right) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|T_A(x_i) \cap T_B(x_i)|}{|U|}\right).$$

3.2. Conditional Generalized Liang Entropy.

Definition 3.3. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$, *where* $U = \{x_1, \ldots, x_{|U|}\}$, *two attribute sets* A, $B \subseteq Attr$ *and two coverings* $U/TLR(A) = \{T_A(x_1), \ldots, T_A(x_{|U|})\}$ *and* $U/TLR(B) = \{T_B(x_1), \ldots, T_B(x_{|U|})\}$.

We define conditional generalized Liang entropy of B *about* A *as*

$$IE(B\,|A) = \frac{1}{|U|}\sum_{i=1}^{|U|}\left(\frac{|T_A(x_i)| - |T_B(x_i) \cap T_A(x_i)|}{|U|}\right).$$

The following Proposition 3.2 proves that conditional Liang entropy $E(B|A)$ in [6] is a particular case of our conditional generalized Liang entropy $IE(B|A)$.

Proposition 3.2. *Given a complete decision system* $DS = (U,\ Attr \cup \{d\})$, *where* $U = \{x_1, \ldots, x_{|U|}\}$, *two attribute sets* A, $B \subseteq Attr$ *and two partitions* $U/A = \{A_1, A_2, \ldots, A_m\}$ *and* $U/B = \{B_1, B_2, \ldots, B_n\}$, *then*

$$\begin{aligned} IE(B|A) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\left(\frac{|T_A(x_i)| - |T_B(x_i) \cap T_A(x_i)|}{|U|}\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{|B_i \cap A_j|}{|U|}\frac{\left|B_i^c - A_j^c\right|}{|U|} = E(B|A), \end{aligned}$$

where $B_i^c = U - B_i$, $A_j^c = U - A_j$ *and* $E(B|A)$ *is the conditional Liang entropy in* [6].

P r o o f. Suppose that $B_i \cap A_j = \{x_{i1}, x_{i2}, \ldots, x_{is_j}\}$, here $|B_i \cap A_j| = p_j$ and $|B_i| = q_i$. We have $\sum_{j=1}^{m} p_j = q_i$ and $\sum_{i=1}^{n} q_i = |U|$. Then

$$B_i \cap A_j = T_B\left(x_{i1}\right) \cap T_A\left(x_{i1}\right) = T_B\left(x_{i2}\right) \cap T_A\left(x_{i2}\right) = \cdots = T_B\left(x_{ip_j}\right) \cap T_A\left(x_{ip_j}\right),$$

$$\begin{aligned} |B_i \cap A_j| &= |T_B\left(x_{i1}\right) \cap T_A\left(x_{i1}\right)| = |T_B\left(x_{i2}\right) \cap T_A\left(x_{i2}\right)| = \cdots \\ &= \left|T_B\left(x_{ip_j}\right) \cap T_A\left(x_{ip_j}\right)\right| = p_j, \end{aligned}$$

$$\begin{aligned} &|B_i \cap A_j|\left|B_i^c - A_j^c\right| = |B_i \cap A_j|\,|B_i^c \cap A_j| = |B_i \cap A_j|\,|A_j - (B_i \cap A_j)| \\ &= |T_A\left(x_{i1}\right) - (T_B\left(x_{i1}\right) \cap T_A\left(x_{i1}\right))| + \cdots + \left|T_A\left(x_{is_i}\right) - \left(T_B\left(x_{ip_j}\right) \cap T_A\left(x_{ip_j}\right)\right)\right| \\ &= \sum_{k=1}^{p_j}|T_A\left(x_{ik}\right) - (T_B\left(x_{ik}\right) \cap T_A\left(x_{ik}\right))| = \sum_{k=1}^{p_j}|T_A\left(x_{ik}\right)| - |T_B\left(x_{ik}\right) \cap T_A\left(x_{ik}\right)|. \end{aligned}$$

Hence

$$\sum_{j=1}^{m} |B_i \cap A_j| \left|B_i^c - A_j^c\right| = \sum_{j=1}^{m}\sum_{k=1}^{p_j} |T_A(x_{ik})| - |T_B(x_{ik}) \cap T_A(x_{ik})|$$
$$= \sum_{k=1}^{q_i} |T_A(x_{ik})| - |T_B(x_{ik}) \cap T_A(x_{ik})|,$$
$$\sum_{i=1}^{n}\sum_{j=1}^{m} |B_i \cap A_j| \left|B_i^c - A_j^c\right| = \sum_{i=1}^{n}\sum_{k=1}^{q_i} |T_A(x_{ik})| - |T_B(x_{ik}) \cap T_A(x_{ik})|$$
$$= \sum_{i=1}^{|U|} |T_A(x_i)| - |T_B(x_i) \cap T_A(x_i)|,$$
$$IE(B|A) = \frac{1}{|U|}\sum_{i=1}^{|U|}\left(\frac{|T_A(x_i)| - |T_B(x_i) \cap T_A(x_i)|}{|U|}\right)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{|B_i \cap A_j|}{|U|}\frac{\left|B_i^c - A_j^c\right|}{|U|} = E(B|A).$$

Consequently, $IE(B|A) = E(B|A)$. The proposition is proved. □

3.3 Some Properties of Generalized Liang Entropy.

Proposition 3.3. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$, *where* $U = \{x_1, \ldots, x_{|U|}\}$ *and* A, B, $C \subseteq Attr$, *the following properties hold:*

a) *If* $U/TLR(A) \preceq; U/TLR(B)$ *then* $IE(A) \geq IE(B)$.

$IE(A) = IE(B)$ *if and only if* $U/TLR(A) = U/TLR(B)$.

b) *If* $U/TLR(A) \preceq U/TLR(B)$ *then* $IE(A \cup B) = IE(A)$.

c) $IE(A \cup B) \geq IE(A)$, $IE(A \cup B) \geq IE(B)$.

d) $IE(A \cup B) = IE(A) + IE(B|A) = IE(A) + IE(A|B)$.

e) $0 \leq IE(B|A) \leq 1 - 1/|U|$. $IE(B|A) = 0$ *if and only if* $U/TLR(A) \preceq U/TLR(B)$. $IE(B|A) = 1 - 1/|U|$ *if and only if* $U/TLR(A) = \delta$ *and* $U/TLR(B) = \omega$.

f) *If* $U/TLR(A) \preceq U/TLR(B)$ *then* $IE(C|B) \geq IE(C|A)$.

g) *If* $U/TLR(A) \preceq U/TLR(B)$ *then* $IE(A|C) \geq IE(B|C)$.

Proof. *a)* This result obtains directly from Definition 3.1 and Definition 2.1.

b) This result obtains directly from Definition 3.1, Definition 3.2, Definition 2.1 and Property 2.1.

c) This result obtains directly from *a*).

d) From Definition 3.1, Definition 3.2 and Definition 3.3, we have

$$\begin{aligned} IE(B\,|A) \quad &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_A(x_i)| - |T_A(x_i)\cap T_B(x_i)|}{|U|} \\ &= 1 - \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_A(x_i)\cap T_B(x_i)|}{|U|} - 1 + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_A(x_i)|}{|U|} \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}1 - \frac{|T_A(x_i)\cap T_B(x_i)|}{|U|} - \frac{1}{|U|}\sum_{i=1}^{|U|}1 - \frac{|T_A(x_i)|}{|U|} = IE(A\cup B) - IE\,(A)\,. \end{aligned}$$

Consequently, we have $IE(A\cup B) = IE(A) + IE(A|B)$. By symmetric property of $IE(A\cup B)$ we have $IE(A\cup B) = IE(B) + IE(A|B)$.

e) It is clear that $IE(B|A) \geq 0$. From *d*) we have $IE(B|A) = IE(A\cup B) - IE(A)$. $IE(B|A) = 0 \Leftrightarrow IE(A\cup B) = IE(A)$. Property 2.1 shows that $U/TLR(A\cup B) \preceq U/TLR(A)$. From *a*) we obtain $IE(A\cup B) = IE(A) \Leftrightarrow U/TLR(A\cup B) = U/TLR(A) \Leftrightarrow U/TLR(A) \preceq U/TLR(B)$. In addition, it follows from *d*) and Definition 3.1 that $IE(B|A) = IE(A\cup B) - IE(A)$, $IE(A\cup B) \leq 1 - 1/|U|$, $IE(A) \geq 0$. So we obtain $IE(B|A) \leq 1 - 1/|U|$. The conditional equality is $IE(A) = 0 \wedge IE(A\cup B) = 1 - 1/|U|$, that is $U/TLR(A) = \delta$ and $U/TLR(A\cup B) = \omega$. This is equivalent to $U/TLR(A) = \delta$ and $U/TLR(B) = \omega$.

f) Suppose that $U/TLR(C) = \{T_C(x_1), T_C(x_2), \ldots, T_C(x_{|U|})\}$. Since $U/TLR(A) \preceq U/TLR(B)$, we have $T_A(x_i) \subseteq T_B(x_i)$ for $\forall x_i \in U$, $i = 1\ldots|U|$ and

$$\begin{aligned} &(T_B(x_i) - T_A(x_i)) \cap T_C(x_i) \subseteq T_B(x_i) - T_A(x_i) \\ &\Leftrightarrow (T_B(x_i)\cap T_C(x_i)) - (T_A(x_i)\cap T_C(x_i)) \subseteq T_B(x_i) - T_A(x_i) \\ &\Leftrightarrow |(T_B(x_i)\cap T_C(x_i)) - (T_A(x_i)\cap T_C(x_i))| \leq |T_B(x_i) - T_A(x_i)| \end{aligned} \tag{3.1}$$

Since $T_A(x_i) \subseteq T_B(x_i)$ we have $T_A(x_i)\cap T_C(x_i) \subseteq T_B(x_i)\cap T_C(x_i)$ and Equation 3.1 is equivalent to

$$\begin{aligned} &|T_B(x_i)\cap T_C(x_i)| - |T_A(x_i)\cap T_C(x_i)| \leq |T_B(x_i)| - |T_A(x_i)| \\ &\Leftrightarrow |T_B(x_i)| - |T_B(x_i)\cap T_C(x_i)| \geq |T_A(x_i)| - |T_A(x_i)\cap T_C(x_i)| \\ &\Leftrightarrow \frac{1}{|U|}\sum_{i=1}^{n}\frac{|T_B(x_i)| - |T_B(x_i)\cap T_C(x_i)|}{|U|} \geq \frac{1}{|U|}\sum_{i=1}^{n}\frac{|T_A(x_i)| - |T_A(x_i)\cap T_C(x_i)|}{|U|} \\ &\Leftrightarrow IE(C|B) \geq IE(C|A). \end{aligned}$$

g) Since $U/TLR(A) \preceq U/TLR(B)$, we have $T_A(x_i) \subseteq T_B(x_i)$ for $\forall x_i \in U$, $i = 1 \ldots |U|$. Suppose that $U/TLR(C) = \{T_C(x_1), T_C(x_2), \ldots, T_C(x_{|U|})\}$, we obtain

$$
\begin{aligned}
&T_A(x_i) \cap T_C(x_i) \subseteq T_B(x_i) \cap T_C(x_i)\\
&\Leftrightarrow |T_A(x_i) \cap T_C(x_i)| \leq |T_B(x_i) \cap T_C(x_i)|\\
&\Leftrightarrow |T_C(x_i)| - |T_A(x_i) \cap T_C(x_i)| \geq |T_C(x_i)| - |T_B(x_i) \cap T_C(x_i)|\\
&\Leftrightarrow \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_C(x_i)| - |T_A(x_i) \cap T_C(x_i)|}{|U|} \geq \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_C(x_i)| - |T_B(x_i) \cap T_C(x_i)|}{|U|}\\
&\Leftrightarrow IE(A|C) \geq IE(B|C).
\end{aligned}
$$

4. Distance between Coverings and Properties. Let X be the set of objects. A distance between two objects x, $y \in X$, denoted as $d(x, y)$, is a measure which satisfies three conditions [2]:

$$
\begin{aligned}
&\text{(C1)} && d(x,y) &&\geq\ 0, \quad d(x,y) = 0 \Leftrightarrow x = y;\\
&\text{(C2)} && d(x,y) &&=\ d(y,x);\\
&\text{(C3)} && d(x,y) + d(y,z) &&\geq\ d(x,z) \ \text{ for any } \ z \in X.
\end{aligned}
$$

In this section, a distance is established between two coverings generated by two attributes based on the generalized Liang entropy. Some properties of the distance are also investigated.

Lemma 4.1. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$ *where* $U = \{x_1, \ldots, x_{|U|}\}$ *and* A, B, $C \subseteq Attr$, *the following properties hold:*

a) $IE(A|C) + IE(B|A \cup C) = IE(A \cup B|C)$;

b) $IE(B|A) + IE(A|C) \geq IE(B|C)$.

P r o o f. Suppose that

$$
\begin{aligned}
U/TLR(A) &= \{T_A(x_1), T_A(x_2), \ldots, T_A(x_{|U|})\},\\
U/TLR(B) &= \{T_B(x_1), T_B(x_2), \ldots, T_B(x_{|U|})\},\\
U/TLR(C) &= \{T_C(x_1), T_C(x_2), \ldots, T_C(x_{|U|})\}.
\end{aligned}
$$

a) $IE(A|C) + IE(B|A \cup C) =$

$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_C(x_i)| - |T_A(x_i) \cap T_C(x_i)| + |T_{A\cup C}(x_i)| - |T_{A\cup C}(x_i) \cap S_B(x_i)|}{|U|}$$

$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_C(x_i)| - |T_{A\cup C}(x_i)| + |T_{A\cup C}(x_i)| - |T_{A\cup C}(x_i) \cap T_B(x_i)|}{|U|}$$

$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_C(x_i)| - |T_A(x_i) \cap T_B(x_i) \cap T_C(x_i)|}{|U|}$$

$$= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|T_C(x_i)| - |T_C(x_i) \cap T_{A\cup B}(x_i)|}{|U|} = IE(A \cup B|C).$$

Consequently, we have $IE(A|C) + IE(B|A \cup C) = IE(A \cup B|C)$.

b) Using Proposition 3.3, item *a*), it follows from $U/TLR(A \cup C) \preceq U/TLR(A)$, $U/TLR(A \cup B) \preceq U/TLR(B)$ that $IE(B|A) \geq IE(B|A \cup C)$ and $IE(A \cup B|C) \geq IE(B|C)$. Using Lemma 4.1 item *a*) we have

$$IE(B|A) + IE(A|C) \geq IE(B|A \cup C) + IE(A|C) = IE(A \cup B|C) \geq IE(B|C).$$

Consequently, we have $IE(B|A) + IE(A|C) \geq IE(B|C)$. □

Theorem 4.1. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$ *and two attributes* A, $B \subseteq Attr$, *for any* $K(A)$, $K(B) \in COV(U)$, *the mapping* $d_E : COV(U) \times COV(U) \to [0, \infty)$ *determined by*

$$d_E(K(A), K(B)) = IE(A|B) + IE(B|A)$$

is a distance between $K(A)$ *and* $K(B)$.

Proof. (C1) According to Proposition 3.3 item *e*) we have $d_E(K(A), K(B)) \geq 0$ for any $K(A)$, $K(B) \in COV(U)$, $d_E(K(A), K(B)) = 0$

$\Leftrightarrow (IE(B|A.) = 0) \wedge (IE(A|B.) = 0)$

$\Leftrightarrow (U/TLR(A) \preceq U/TLR(B)) \wedge (U/TLR(B) \preceq U/TLR(A)) \Leftrightarrow K(A) = K(B)$.

(C2) According to the definition of the distance d_E, we have $d_E(K(A), K(B)) = d_E(K(B), K(A))$ for any $K(A)$, $K(B) \in COV(U)$.

(C3) For any $K(A)$, $K(B)$, $K(C) \in COV(U)$, from Lemma 4.1 item b) we have

$$IE(B|A) + IE(A|C) \quad \geq \quad IE(B|C) \tag{4.1}$$

$$IE(C|A) + IE(A|B) \quad \geq \quad IE(C|B) \tag{4.2}$$

From Equation (4.1) and Equation (4.2), we obtain

$$d_E(K(B), K(A)) + d_E(K(A), K(C)) \geq d_E(K(B), K(C))$$

From (C1), (C2), (C3) we conclude that $d_E(K(A), K(B))$ is a distance on $COV(U)$. The theorem is proved. □

Proposition 4.1. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$, *where* $U = \{x_1, \ldots, x_{|U|}\}$ *and* $A \subseteq Attr$, *then*

$$d_E(K(A), K(Attr)) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|T_A(x_i)| - |T_{Attr}(x_i)|}{|U|}.$$

P r o o f. Since $A \subseteq Attr$ we have $U/TLR(Attr) \preceq U/TLR(A)$ (Property 2.1). From Proposition 3.3 item e) we obtain $IE(A|Attr) = 0$. In addition, it follows from $A \subseteq Attr$ that $T_{Attr}(x_i) \subseteq T_A(x_i)$ or $T_A(x_i) \cap T_{Attr}(x_i) = T_{Attr}(x_i)$ for $\forall x_i \in U$, $i = 1 \ldots |U|$. Consequently,

$$\begin{aligned} d_E(K(A), K(Attr)) &= IE(A|Attr) + IE(Attr|A) = IE(Attr|A) \\ &= \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|T_A(x_i)| - |T_A(x_i) \cap T_{Attr}(x_i)|}{|U|} = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|T_A(x_i)| - |T_{Attr}(x_i)|}{|U|}. \end{aligned}$$

The proposition is proved. □

Proposition 4.2. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$, *if* $A \subseteq Attr$, *then* $d_E(K(A), K(A \cup \{d\})) \geq d_E(K(Attr), K(Attr \cup \{d\}))$.

P r o o f. Suppose that $U = \{x_1, x_2, \ldots, x_{|U|}\}$ and $A \subseteq Attr$. For $\forall x_i \in U$, $i = 1 \ldots |U|$, it is clear that $T_{Attr}(x_i) \subseteq T_A(x_i)$. So we have

$$\begin{aligned} &(T_A(x_i) - T_{Attr}(x_i)) \cap T_{\{d\}}(x_i) \subseteq T_A(x_i) - T_{Attr}(x_i) \\ &\Leftrightarrow (T_A(x_i) \cap T_{\{d\}}(x_i)) - (T_{Attr}(x_i) \cap T_{\{d\}}(x_i)) \subseteq T_A(x_i) - T_{Attr}(x_i) \\ &\Leftrightarrow |(T_A(x_i) \cap T_{\{d\}}(x_i)) - (T_{Attr}(x_i) \cap T_{\{d\}}(x_i))| \leq |T_A(x_i) - T_{Attr}(x_i)|. \end{aligned} \tag{4.3}$$

It follows from $T_{Attr}(x_i) \subseteq T_A(x_i)$ that $T_{Attr}(x_i) \cap T_{\{d\}}(x_i) \subseteq T_A(x_i) \cap T_{\{d\}}(x_i)$. So Equation 4.3 is equivalent to

$$\begin{aligned} &|T_A(x_i) \cap T_{\{d\}}(x_i)| - |T_{Attr}(x_i) \cap T_{\{d\}}(x_i)| \leq |T_A(x_i)| - |T_{Attr}(x_i)| \\ &\Leftrightarrow |T_A(x_i)| - |T_A(x_i) \cap T_{\{d\}}(x_i)| \geq |T_{Attr}(x_i)| - |T_{Attr}(x_i) \cap T_{\{d\}}(x_i)|. \end{aligned} \tag{4.4}$$

Since $T_A(x_i) \cap T_{\{d\}}(x_i) \subseteq T_A(x_i)$, $T_{Attr}(x_i) \cap T_{\{d\}}(x_i) \subseteq T_{Attr}(x_i)$, Equation 4.4 is equivalent to

$$\begin{aligned} &|T_A(x_i) \cup (T_A(x_i) \cap T_{\{d\}}(x_i))| - |T_A(x_i) \cap (T_A(x_i) \cap T_{\{d\}}(x_i))| \geq \\ &|T_{Attr}(x_i) \cup (T_{Attr}(x_i) \cap T_{\{d\}}(x_i))| - |T_{Attr}(x_i) \cap (T_{Attr}(x_i) \cap T_{\{d\}}(x_i))|. \end{aligned} \tag{4.5}$$

Since $T_{A\cup\{d\}}(x_i) = T_A(x_i) \cap T_{\{d\}}(x_i)$, $T_{Attr\cup\{d\}}(x_i) = T_{Attr}(x_i) \cap T_{\{d\}}(x_i)$, Equation 4.5 is equivalent to

$$\sum_{i=1}^{n} \frac{|T_A(x_i)| - |T_{A\cup\{d\}}(x_i)|}{|U|^2} \geq \sum_{i=1}^{n} \frac{|T_{Attr}(x_i)| - |T_{Attr\cup\{d\}}(x_i)|}{|U|^2}. \tag{4.6}$$

From Proposition 4.1 and $A \subset A \cup \{d\}$, $Attr \subset Attr \cup \{d\}$, Equation 4.6 is equivalent to $d_E(K(A), K(A \cup \{d\})) \geq d_E(K(Attr), K(Attr \cup \{d\}))$. The proposition is proved. □

5. Distance-based Attribute Reduction Method. Deriving from results in Section 3 and 4, we propose a distance-based method for attribute reduction in incomplete decision systems. First, we define a reduct based on the distance. Second, we define the importance of an attribute based on the distance as the classification ability of the attribute. As a result, we propose a heuristic algorithm to find the best reduct by using the importance of an attribute as an attribute selection criterion.

Definition 5.1. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$, *if an attribute set* $R \subseteq Attr$ *satisfies*

(1) $d_E(K(R), K(R \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\}))$;

(2) $\forall r \in R$, $d_E(K(R-\{r\}), K((R-\{r\}) \cup \{d\})) \neq d_E(K(Attr), K(Attr \cup \{d\}))$,

then R *is called a reduct of IDS based on distance.*

The following Proposition 5.1 shows the relationship between *the reduct based on generalized decision* and *the reduct based on distance*.

Proposition 5.1. *Given an incomplete decision system* $IDS = (U, Attr \cup \{d\})$ *and* $R \subseteq Attr$, *if* $d_E(K(R), K(R \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\}))$, *then* $\forall x_i \in U$, $\partial_R(x_i) = \partial_{Attr}(x_i)$.

Proof. Suppose that $U = \{x_1, x_2, \ldots, x_{|U|}\}$.
Since $d_E(K(R), K(R \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\}))$, according to Propo-

sition 4.1 we have:

$$\frac{1}{|U|}\sum_{i=1}^{|U|}\left(\frac{|T_R(x_i)|-|T_{R\cup\{d\}}(x_i)|}{|U|}\right)=\frac{1}{|U|}\sum_{i=1}^{|U|}\left(\frac{|T_{Attr}(x_i)|-|T_{Attr\cup\{d\}}(x_i)|}{|U|}\right)$$

$$\Leftrightarrow |T_R(x_i)|-\left|T_{R\cup\{d\}}(x_i)\right|=|T_{Attr}(x_i)|-\left|T_{Attr\cup\{d\}}(x_i)\right| \text{ for any } x_i\in U. \tag{5.1}$$

It is clear that $T_{R\cup\{d\}}(x_i)\subseteq T_R(x_i)$, $T_{Attr\cup\{d\}}(x_i)\subseteq T_{Attr}(x_i)$, so Equation 5.1 is equivalent to

$$|T_R(x_i)-T_{R\cup\{d\}}(x_i)|=|T_{Attr}(x_i)-T_{Attr\cup\{d\}}(x_i)| \text{ for any } x_i\in U. \tag{5.2}$$

Since $T_{Attr}(x_i)\subseteq T_R(x_i)$ we have $T_{Attr}(x_i)-T_{\{d\}}(x_i)\subseteq T_R(x_i)-T_{\{d\}}(x_i)$

$\Leftrightarrow T_{Attr}(x_i)-T_{Attr}(x_i)\cap T_{\{d\}}(x_i)\subseteq T_R(x_i)-T_R(x_i)\cap T_{\{d\}}(x_i)$

$\Leftrightarrow T_{Attr}(x_i)-T_{Attr\cup\{d\}}(x_i)\subseteq T_R(x_i)-T_{R\cup\{d\}}(x_i)$.

So Equation 5.2 is equivalent to

$$T_R(x_i)-T_{R\cup\{d\}}(x_i)=T_{Attr}(x_i)-T_{Attr\cup\{d\}}(x_i) \text{ for any } x_i\in U. \tag{5.3}$$

In addition, we have

$$T_R(x_i)=(T_R(x_i)\cap T_{\{d\}}(x_i))\cup(T_R(x_i)-(T_R(x_i)\cap T_{\{d\}}(x_i))),$$
$$T_{Attr}(x_i)=(T_{Attr}(x_i)\cap T_{\{d\}}(x_i))\cup(T_{Attr}(x_i)-(T_{Attr}(x_i)\cap T_{\{d\}}(x_i))).$$

Suppose that $d_i=d(x_i)$, $R_i=\{d(y_i)|y_i\in T_R(x_i)-(T_R(x_i)\cap T_{\{d\}}(x_i))\}$, $A_i=\{d(y_i)|y_i\in T_{Attr}(x_i)-(T_{Attr}(x_i)\cap T_{\{d\}}(x_i))\}$. Then we have

$$\begin{aligned}\partial_R(x_i)&=\{d(y_i)|y_i.\in(T_R(x_i)\cap T_{\{d\}}(x_i))\cup(T_R(x_i)-(T_R(x_i)\cap T_{\{d\}}(x_i)))\}\\&=\{d_i\}\cup R_i\\ \partial_{Attr}(x_i)&=\{d(y_i)|y_i\in(T_{Attr}(x_i)\cap T_{\{d\}}(x_i))\\&\qquad\cup(T_{Attr}(x_i)-(T_{Attr}(x_i)\cap T_{\{d\}}(x_i)))\}=\{d_i\}\cup A_i.\end{aligned}$$

According to Equation 5.3, we obtain $R_i=A_i$, thus $\partial_R(x_i)=\partial_{Attr}(x_i)$ for any $x_i\in U$. The proposition is proved. □

Proposition 5.1 shows that *if R_D is a reduct based on metric then there exists a reduct based on generalized decision R_∂ so that $R_\partial\subseteq R_D$.*

If IDS is consistent, it follows from the condition $\forall x_i \in U$, $|\partial_R(x_i)| = |\partial_{Attr}(x_i)| = 1$ that $T_R(x_i) = T_{R\cup\{d\}}(x_i)$ and $T_{Attr}(x_i) = T_{Attr\cup\{d\}}(x_i)$ for any $x_i \in U$. Acoording to Proposition 4.1 we have

$$d_E(K(R), K(R \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\})) = 0.$$

Consequently, $d_E(K(R), K(R \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\}))$ if and only if $\forall x_i \in U$, $\partial_R(x_i) = \partial_{Attr}(x_i)$. This means that *reduct based on metric is equivalent to reduct based on generalized decision.*

Definition 5.2. *Given an incomplete decision system $IDS = (U, Attr \cup \{d\})$ and $A \subset Attr$, the importance of attribute $a \in Attr - A$ is defined as*

$$IMP_A(a) = d_E(K(A), K(A \cup \{d\})) - d_E(K(A \cup \{a\}), K(A \cup \{a\} \cup \{d\})).$$

According to Proposition 4.2 we have $IMP_A(a) \geq 0$. When a is added into A, the distance $d_E(K(A), K(A\cup\{d\}))$ changes, which impacts on the importance of the attribute a in the way that the larger the value of $IMP_A(a)$ is, the more important is the attribute a. Using the importance of an attribute as an attribute selection criterion, we design a heuristic algorithm to find the best reduct.

Algorithm 5.1. The algorithm to find the best reduct of an incomplete decision system.

Input: An incomplete decision system $IDS = (U, Attr \cup \{d\})$.

Output: The best reduct R.

1. $R = \emptyset$;
2. Calculate $d_E(K(R), K(R \cup \{d\}))$, $d_E(K(Attr), K(Attr \cup \{d\}))$;
3. While $d_E(K(R), K(R \cup \{d\})) \neq d_E(K(Attr), K(Attr \cup \{d\}))$ do
4. Begin
5. For each $a \in Attr - R$
6. Begin
7. Calculate $d_E(K(R \cup \{a\}), K(R \cup \{a\} \cup \{d\}))$;
8. Calculate $IMP_R(a) = d_E(K(R), K(R \cup \{d\})) - d_E(K(R \cup \{a\}), K(R \cup \{a\} \cup \{d\}))$;

9. End;
10. Select $a_m \in Attr - R$ so that $IMP_R(a_m) = \underset{a \in Attr-R}{\text{Max}}\{IMP_R(a)\}$;
11. $R = R \cup \{a_m\}$;
12. Calculate $d_E(K(R), K(R \cup \{d\}))$;
13. End;
14. For each $a \in R$
15. Begin
16. Calculate $d_E(K(R - \{a\}), K(R - \{a\} \cup \{d\}))$;
17. if $d_E(K(R - \{a\}), K(R - \{a\} \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\}))$ then $R = R - \{a\}$;
18. End;
19. Return R;

Let us consider the command lines of Algorithm 5.1. From 3 to 13, the obtained attribute set R satisfies $d_E(K(R), K(R \cup \{d\})) = d_E(K(Attr), K(Attr \cup \{d\}))$. From 14 to 18, R is minimal, that is

$$\forall r \in R,\ d_E(K(R - \{r\}), K((R - \{r\}) \cup \{d\})) \neq d_E(K(Attr), K(Attr \cup \{d\})).$$

According to Definition 5.1, R is a reduct. Consequently, Algorithm 5.1 is complete.

Complexity of Algorithm 5.1. First we analyse the complexity of While Loop from 3 to 13. Since $T_R(u_i)$ and $T_{R\cup\{d\}}(u_i)$ are calculated in the previous step, we calculate $T_{R\cup\{a\}}(u_i)$, $T_{R\cup\{a\}\cup\{d\}}(u_i)$ only. The complexity of calculating $T_{R\cup\{a\}}(u_i)$ for $\forall u_i \in U$ when $T_R(u_i)$ calculated is $O(|U|^2)$. So the complexity of calculating all $IMP_R(a)$ is:

$$(|Attr| + (|Attr| - 1) + \cdots + 1) * |U|^2 = (|Attr| * (|Attr| - 1)/2) * |U|^2 = O(|Attr|^2|U|^2).$$

where the cardinality $|Attr|$ is the number of conditional attributes and $|U|$ is the number of objects. The complexity of obtaining the attribute with maximum

importance is $|Attr| + (|Attr| - 1) + \cdots + 1 = |Attr| * (|Attr| - 1)/2 = O(|Attr|^2)$. Hence, the complexity of While Loop is $O(|Attr|^2|U|^2)$. Second, in a similar way, the complexity of For Loop from 14 to 18 is $O(|Attr|^2|U|^2)$. Finally, the complexity of Algorithm 5.1 is $O(|Attr|^2|U|^2)$. Consequently, this complexity is better than the complexity of algorithms in [1, 3, 4, 20].

For example, let us consider the incomplete decision system in Table 1. We have the following coverings:

$$U/TLR(Attr) = \{\{x_1\}, \{x_2, x_6\}, \{x_3\}, \{x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_2, x_5, x_6\}\},$$

$$U/TLR(\{a_1\}) = \{\{x_1, x_3, x_4, x_5\}, \{x_2, x_3, x_5, x_6\}, U, \{x_1, x_3, x_4, x_5\}, U, \{x_2, x_3, x_5, x_6\}\},$$

$$U/TLR(\{a_2\}) = \{U, U, U, U, U, U\},$$

$$U/TLR(\{a_3\}) = \{\{x_1, x_2, x_4, x_5, x_6\}, \{x_1, x_2, x_4, x_5, x_6\}, \{x_3\}, \{x_1, x_2, x_4, x_5, x_6\}, \{x_1, x_2, x_4, x_5, x_6\}, \{x_1, x_2, x_4, x_5, x_6\}\},$$

$$U/TLR(\{a_4\}) = \{\{x_1, x_2, x_6\}, \{x_1, x_2, x_6\}, \{x_3, x_4, x_5, x_6\}, \{x_3, x_4, x_5, x_6\}, \{x_3, x_4, x_5, x_6\}, U\},$$

$$U/TLR(\{d\}) = \{\{x_1, x_2, x_4, x_6\}, \{x_1, x_2, x_4, x_6\}, \{x_3\}, \{x_1, x_2, x_4, x_6\}, \{x_5\}, \{x_1, x_2, x_4, x_6\}\}.$$

We calculate the distance

$$d_E(K(Attr), K(Attr \cup \{d\})) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} (|T_{Attr}(u_i) - (T_{Attr}(u_i) \cap T_{\{d\}}(u_i))|) = \frac{4}{36}.$$

Set $R = \emptyset$ and suppose that $T_\emptyset(x) = U$ for any $x \in U$. We calculate

$$T_\emptyset(x_i) = U \text{ for } \forall x_i \in U, \quad i = 1 \ldots |U|.$$

$$SIG_\emptyset(a_1) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} \left(\left| T_\emptyset(u_i) - T_{\{d\}}(u_i) \right| - \left| T_{\{a_1\}}(u_i) - T_{\{a_1, d\}}(u_i) \right| \right) = 0,$$

$$SIG_\emptyset(a_2) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} \left(\left| T_\emptyset(u_i) - T_{\{d\}}(u_i) \right| - \left| T_{\{a_2\}}(u_i) - T_{\{a_2, d\}}(u_i) \right| \right) = 0,$$

$$SIG_{\emptyset}(a_3) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}\left(\left|T_{\emptyset}(u_i) - T_{\{d\}}(u_i)\right| - \left|T_{\{a_3\}}(u_i) - T_{\{a_3,d\}}(u_i)\right|\right) = \frac{10}{36},$$

$$SIG_{\emptyset}(a_4) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}\left(\left|T_{\emptyset}(u_i) - T_{\{d\}}(u_i)\right| - \left|T_{\{a_4\}}(u_i) - T_{\{a_4,d\}}(u_i)\right|\right) = \frac{8}{36}.$$

We choose a_3 which has the most importance and $R = \{a_3\}$, and calculate the distance

$$d_E(K(\{a_3\}), K(\{a_3, d\})) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}(|T_{\{a_3\}}(u_i) - (T_{\{a_3\}}(u_i) \cap T_{\{d\}}(u_i))|) = \frac{8}{36}.$$

So we have $d_E(K(\{a_3\}), K(\{a_3, d\})) \neq d_E(K(Attr), K(Attr \cup \{d\}))$.

We perform the second loop.

$$SIG_{\{a_3\}}(a_1) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}(|T_{\{a_3\}}(u_i) - T_{\{a_3,d\}}(u_i)| - |T_{\{a_1,a_3\}}(u_i) - T_{\{a_1,a_3,d\}}(u_i)|)$$

$$= \frac{2}{36},$$

$$SIG_{\{a_3\}}(a_2) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}(|T_{\{a_3\}}(u_i) - T_{\{a_3,d\}}(u_i)| - |T_{\{a_2,a_3\}}(u_i) - T_{\{a_2,a_3,d\}}(u_i)|)$$

$$= 0,$$

$$SIG_{\{a_3\}}(a_4) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}(|T_{\{a_3\}}(u_i) - T_{\{a_3,d\}}(u_i)| - |T_{\{a_3,a_4\}}(u_i) - T_{\{a_3,a_4,d\}}(u_i)|)$$

$$= \frac{4}{36}.$$

We choose a_4 which has the most importance and we set $R = \{a_3, a_4\}$, and calculate

$$d_E(K(\{a_3, a_4\}), K(\{a_3, a_4, d\})) = \frac{4}{36} = d_E(K(Attr), K(Attr \cup \{d\})).$$

Hence, we go to For Loop. According to the above calculation, we obtain

$$d_E(K(\{a_3\}), K(\{a_3, d\})) \neq d_E(K(Attr), K(Attr \cup \{d\})).$$

In addition,

$$d_E(K(\{a_4\}), K(\{a_4, d\})) = \frac{10}{36} \neq d_E(K(Attr), K(Attr \cup \{d\})).$$

Consequently, the algorithm finishes and $R = \{a_3, a_4\}$ is the best reduct of $Attr$.

6. Conclusions. Attribute reduction is the most important problem in both classical rough sets and tolerance rough sets. In this paper, a generalized Liang entropy is proposed based on Liang entropy [6] and some properties of the generalized Liang entropy are considered. Based on the generalized Liang entropy, a distance is established between attributes and a distance-based method to find the best reduct is proposed. To construct this method, we define a reduct based on the distance, the importance of an attribute based on the distance. We use the importance of an attribute as heuristic information to design an effective heuristic algorithm to find the best reduct. We prove theoretically that the complexity of our algorithm is less than that of the algorithms in [1, 3, 4, 20].

REFERENCES

[1] Dai X. P, D. H Xiong. Research on Heuristic Knowledge Reduction Algorithm for Incomplete Decision Table. In: Proceedings of the International Conference on Internet Technology and Applications, Wuhan, China, 2010, IEEE, 1–3.

[2] Deza M. M., E. Deza. Encyclopedia of Distances, Springer, 2009.

[3] Huang B., X. He, X. Z. Zhou. Rough Computational methods based on tolerance matrix. *Acta Automatica Sinica*, **30** (2004), 363–370.

[4] Huang B., H. X. Li, X. Z. Zhou. Attribute Reduction Based on Information Quantity under Incomplete Information Systems. *Systems Application Theory & Practice*, **34** (2005), 55–60.

[5] Kryszkiewicz M. Rough set approach to incomplete information systems. *Information Science*, **112** (1998), 39–49.

[6] Liang J. Y., K. S. Chin, C. Y. Dang, R. C. M. Yam. A new method for measuring uncertainty and fuzziness in rough set theory. *International Journal of General Systems*, **31** (2002), No 4, 331–342.

[7] Liang J. Y, Y. H Qian. Axiomatic approach of knowledge granulation in information system. Lecture Notes in Artificial Intelligence, Vol. **4304**, Springer-Verlag, Berlin Heidelberg, 2006, 1074–1078.

[8] Liang J. Y., Y. H. Qian. Information granules and entropy theory in information systems. *Information Sciences*, **51** (2008), 1–18.

[9] Liang J. Y, Z. Z. Shi, D. Y. Li, M. J. Wierman. The information entropy, rough entropy and knowledge granulation in incomplete information system. *International Journal of General Systems*, **35** (2006), No 6, 641–654.

[10] Long G. N. Metric Based Attribute Reduction in Decision Tables. In: Proceedings of the Federated Conference on Computer Science and Information Systems, Wroclaw, Poland, IEEE, 2012, 311–316.

[11] Pawlak Z. Rough sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishers, 1991.

[12] Qian Y. H., J. Y. Liang. Combination Entropy and Combination Granulation in Incomplete Information System. In: Proceedings of the First International Conference on Rough Sets and Knowledge Technology RSKT'06, Springer-Verlag Berlin, Heidelberg, 2006, 184–190.

[13] Qian Y. H., J. Y. LiangQ, F. Wang. New method for measuring uncertainty in incomplete information systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **17** (2009), doi: 10.1142/S0218488509006303.

[14] Qian Y. H., J. Y. Liang, C. Y. Dang. Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. *International Journal of Approximate Reasoning*, **50** (2009), 174–188.

[15] Qian Y. H., J. Y. Liang, C. Y. Dang, F. Wang, W. Xu. Knowledge distance in information systems. *Journal of Systems Science and Systems Engineering*, **16** (2007), 434–449.

[16] De Mántaras R. L. A distance-based attribute selection measure for decision tree induction. *Machine Learning*, **6** (1991), 81–92.

[17] Shifei D., D. Hao. Research and Development of Attribute Reduction Algorithm Based on Rough Set. In: Proceedings of the Control and Decision Conference (CCDC), Chinese, IEEE, 2010, 648–653.

[18] Simovici D. A., S. Jaroszewicz. Generalized conditional entropy and decision trees. In: Proceedings of the EGC, Lyon, France, 2003, 369–380.

[19] Simovici D. A., S. Jaroszewicz. A new metric splitting criterion for decision trees. *International Journal of Parallel Emergent and Distributed Systems*, **21** (2006), No 4, 239–256.

[20] Zhou X. Z., B. Huang. Rough set-based attribute reduction under incomplete Information Systems. *Journal of Nanjing University of Science and Technology*, **27** (2003), 630–636.

Janos Demetrovics
Institute for Computer and Control (SZTAKI)
Hungarian Academy of Sciences
Budapest, Hungary
e-mail: demetrovics@sztaki.mta.hu

Vu Duc Thi
Information Technology Institute
Vietnam National University (VNU)
Hanoi, Vietnam
e-mail: vdthi@vnu.edu.vn

Nguyen Long Giang
Institute of Information Technology
Vietnam Academy of Science
and Technology (VAST)
Viet Nam, Vietnam
e-mail: nlgiang@ioit.ac.vn

Received June 10, 2014
Final Accepted June 26, 2014