

Chapter 14

The Improved Discernibility Matrix Based on Decision Vector

Jingfu Li and Denghui Bu

Abstract Skowron's discernibility matrix is the evolvement of the attribute reduction models under algebra view. Aiming at the generating process of discernibility matrix, this paper simplifies decision system by introducing decision vector, and designs an improved discernibility matrix, which is capable of suiting inconsistency of decision system, as well as largely reducing comparison times and storage space for constructing discernibility matrix. Finally, an illustrative example is used to show the feasibility and efficiency of this improved discernibility matrix.

Keywords Rough set · Discernibility matrix · Attribute reduction · Decision vector

14.1 Introduction

The theory of rough set, as proposed by Pawlak, is an efficient mathematical theory to deal with uncertainty and vagueness of decision system, and it has many successful applications in the fields of artificial intelligence such as experts system, pattern recognition, machine learning and knowledge discovery [1, 2]. Attribute reduction plays important role in analyzing enormous data, for it provides the equivalent information descriptive ability as the entire set of attributes [3].

Skowron and Rauszer introduced two basic notions in 1991, namely discernibility matrix and discernibility function [4], which provides a canonical and precise mathematical model for finding core and reduction of decision system [5, 6]. However, this method has some disadvantages: generating discernibility

J. Li (✉) · D. Bu
International College, Huanghuai University, 6# Kaiyuan Road,
Zhumadian, Henan, China
e-mail: sysgl007@gmail.com

matrix would consume enormous time and space; and it is difficult to suit inconsistent circumstance of decision system [7, 8].

In this paper, these disadvantages mentioned above are analyzed in detail, and an efficient method of simplifying decision system is proposed to find minimum discernibility set. This method not only considers inconsistent circumstance of decision system sufficiently, but also saves executive time and space greatly. The theoretic analysis and simulation instance shows that this algorithm is feasible and effective in practice.

14.2 Preliminaries

Formally, an decision information system DS is denoted as $DS = \langle U, A, V, f \rangle$, where $U = \{x_1, x_2, \dots, x_n\}$ is a nonempty finite set of objects, U is also called as the universe; $A = C \cup D$ is a nonempty finite set of attributes, C is condition attribute set and D is decision attribute set; $V = \cup_{a \in A} V_a$, V_a is the value domain of the attribute a ; $f : U \times A \rightarrow V$ is a information mapping function such that $f(x_i, a) \in V_a (x_i \in U, a \in A)$.

Definition 1 Given a subset of attributes $R \subseteq A$, an indiscernible relation $IND(R) \subseteq U \times U$ is denoted as: $IND(R) = \{(x_i, x_j) \in U \times U : \forall a \in R, f(x_i, a) = f(x_j, a)\}$. For any object pair $\langle x_i, x_j \rangle$, if $x_i IND(R) x_j$ then x_i and x_j have the same value on the attribute set R , i.e. they are indiscernible on the attribute set R . The partition of the universe induced by C and D is denoted by $U/IND(C)$ and $U/IND(D)$, which are called as condition partition and decision partition respectively.

Definition 2 Given an set of object $X \subseteq U$, the lower and upper approximation of X are defined as below:

$$\begin{aligned} \underline{R}(X) &= \cup \{X_i : X_i \in U/IND(R) \wedge X_i \subseteq X\}, \\ \bar{R}(X) &= \cup \{X_i : X_i \in U/IND(R) \wedge X_i \cap X \neq \emptyset\}. \end{aligned}$$

According to indiscernible relation $IND(R)$, the lower approximation $\underline{R}(X)$ is the greatest definable set contained in X , and the upper approximation $\bar{R}(X)$ is the least definable set containing X .

Definition 3 Given decision system $DS = \langle U, A = C \cup D, V, f \rangle$, the positive region is defined by:

$$POS_c(D) = \cup_{X \in U/D} C(X).$$

The positive region is the union of such condition partitions that can be fully contained in the decision partitions.

Definition 4 Given decision system $DS = \langle U, A, V, f \rangle$, considering the characteristic Ψ of DS , an attribute set $B \subseteq A$ is called as a reduction of DS iff:

1. $\Psi(B) = \Psi(A)$;
2. $\forall B' \subset B, \Psi(B') \neq \Psi(B)$.

The intersection of all reductions is called as core. Different characteristic Ψ is corresponding to different function standard of reduction goal, for example keeping condition partition unchanged, keeping relative positive region unchanged, and keeping information entropy unchanged.

As important development of attribute reduction in rough set theory, Skowron and Rauszer introduced the notion of discernibility matrix based on which discernibility function was designed, and proved that the set of prime implicants in minimal disjunctive form of discernibility function is corresponding to all reduction of given information system.

Definition 5 Given decision system $DS = \langle U, A = C \cup D, V, f \rangle$, Discernibility matrix is a $|U| \times |U|$ matrix denoted as M , its element $M(i, j)$ is defined by:

$$M(i, j) = \begin{cases} \{a: a(x_i) \neq a(x_j), (a \in C) \wedge (x_i, x_j \in U)\} & (D(x_i) \neq D(x_j)) \wedge \left(\min \left\{ \left| V_D([x_i]_{IND(C)}) \right|, \left| V_D([x_j]_{IND(C)}) \right| \right\} = 1 \right) \\ \emptyset & \text{otherwise} \end{cases}$$

$M(i, j)$ is an element located on the i th row and the j th column of discernibility matrix M ; $[x_i]_{IND(C)}$ is the equivalence class of x_i with respect to condition attribute set C , i.e. $[x_i]_{IND(C)} \in U/IND(C)$; $V_D([x_i]_{IND(C)})$ denotes the value domain of $[x_i]_{IND(C)}$, $|Y|$ denotes the cardinality of the set Y .

“ $\left(\min \left\{ \left| V_D([x_i]_{IND(C)}) \right|, \left| V_D([x_j]_{IND(C)}) \right| \right\} = 1 \right)$ ” indicates that at least one of the two objects x_i and x_j is contained in the positive region.

It consists of all attributes on which two objects x_i and x_j possess different value, simultaneously at least one of x_i and x_j has only one decision value. Definition 5 assures that discernibility matrix must be capable of discerning the object of positive region from other objects, therefore this discernibility matrix may be considered as discernibility matrix based on changeless positive region.

14.3 Discernibility Matrix Based on Decision Vector

Discernibility matrix is a kind of efficient method to find the reduction. However, generating discernibility matrix would consume enormous time and space, and it is also difficult to suit inconsistent circumstance of decision system. This section would analysis these disadvantages in detail, introduce decision vector to simplify decision system for improving discernibility matrix. This improved matrix is

capable of describe inconsistent objects, and decreasing the comparison times and storage spaces.

Definition 6 Given an decision system $DS = \langle U, A = C \cup D, V, f \rangle$, $U = \{x_1, x_2, \dots, x_n\}$, $U/IND(D) = \{Y_1, Y_2, \dots, Y_{|U/IND(D)|}\}$, then the value domain of system decision can be achieved: $F_D = \{f(Y_1, D), f(Y_2, D), \dots, f(Y_{|U/IND(D)|}, D)\}$. For $\forall x_i \in U$, $B \subseteq C$, the decision of $[x_i]_B$ with respect to D can be denoted as a decision vector: $DV([x_i]_B) = \left(\mu_{DV([x_i]_B)}(f(Y_1, D)), f(Y_1, D), \dots, \left(\mu_{DV([x_i]_B)}(f(Y_{U/IND(D)}, D)), f(Y_{U/IND(D)}, D) \right) \right)$ where $f(Y_j, D)$ is the value domain of decision class $Y_j (1 \leq j \leq |U/IND(D)|)$ over decision attribute D , $\mu_{DV([x_i]_B)}(f(Y_j, D)) = |Y_j \cap [x_i]_B| / |[x_i]_B|$. Obviously, $DV([x_i]_B) = \mu_{DV([x_i]_B)}(f(Y_1, D)), \dots, \mu_{DV([x_i]_B)}(f(Y_{|U/IND(D)|}, D))$ is the probability distribution of the equivalence class $[x_i]_B$ on each decision equivalence class, so $\mu_{DV([x_i]_B)}(f(Y_1, D)) + \dots + \mu_{DV([x_i]_B)}(f(Y_{|U/IND(D)|}, D)) = 1$.

Definition 7 For a decision vector DV , its λ -cut defined for $\lambda \in (0, 1]$ by: $DV_\lambda = \{x : DV(x) \geq \lambda\}$

All elements whose membership are higher than λ constitute the λ -cut of decision vector DV .

Definition 8 For an decision system $DS = \langle U, A = C \cup D, V, f \rangle$, its simplification decision system is defined by: $DS' = \langle U/IND(C), C \cup D', V', f' \rangle$, $V' = \cup_{a \in C \cup D'} V_a$, V_a is the value domain of the attribute a ; $f' : U \times (C \cup D') \rightarrow V'$ is an information mapping function such that

$$f'(X_i, a) \in V'_a (X_i \in U/IND(C), a \in C \cup D'), \text{ here } f'(X_i, D') = \tilde{F}_D(X_i).$$

Based on simplification decision system, we can define discernibility matrix which keeps positive region and information entropy unchanged respectively.

Definition 9 Given decision system $DS = \langle U, A = C \cup D, V, f \rangle$ and its simplification decision system $DS' = \langle U/IND(C), C \cup D', V', f' \rangle$, Discernibility matrix is a $|U/IND(C)| \times |U/IND(C)|$ matrix denoted as M' , its element $M'(i, j)$ is defined by:

$$M_{(i,j)} = \begin{cases} \{a: a(x_i) \neq a(x_j), (a \in C) \wedge (x_i, x_j \in U)\} & |(DV(X_i))_1| = 1 \vee |(DV(X_j))_1| = 1 \\ \emptyset & \text{otherwise} \end{cases}$$

where $(DV(X_i))_1$ is 1-level set (1-cut) of $DV(X_i)$, $|(DV(X_i))_1| = 1$ indicates that X_i is contained in positive region. This discernibility matrix assures that at least one of two compared objects is contained in positive region, so it is capable of keeping positive region unchanged.

Definition 10 Given an decision system $DS = \langle U, A = C \cup D, V, f \rangle$ and its simplification decision system $DS' = \langle U/IND(C), C \cup D', V', f' \rangle$, Discernibility matrix is a $|U/IND(C)| \times |U/IND(C)|$ matrix denoted as M' , its element $M'(i, j)$ is defined by:

$$M'(i, j) = \begin{cases} \{a : a(x_i) \neq a(x_j), (a \in C) \wedge (x_i, x_j \in U)\} & DV(X_i) \neq DV(X_j) \\ \emptyset & \text{otherwise} \end{cases}$$

where $DV(X_i) \neq DV(X_j)$ indicates that the probability distribution of X_i and X_j on decision equivalence class are not equal. Therefore, this discernibility matrix is capable of keeping information entropy unchanged.

The improved discernibility matrix defined as definition 9 and definition 10, considering the equivalence class as elementary comparison unit. There is not more than one time of comparison between two condition equivalence classes, so the improved discernibility matrix is able to save the time and spaces consumed in generating discernibility matrix than before.

Definition 11 For an simplification decision system $DS' = \langle U/IND(C), C \cup D', V', f' \rangle$ and its improved discernibility matrix M' , the discernibility function is a Boolean expression defined as follows:

$$f(M') = \wedge \{ \vee M'(i, j) : 1 < j < i \leq |U/IND(C)|, M'(i, j) \neq \emptyset \}$$

The expression $\vee M'(i, j)$ denotes making disjunction operation among all attributes which belong to nonempty element $M'(i, j)$. The expression $\wedge \{ \vee M'(i, j) \}$ means making conjunction operation among all $\vee M'(i, j)$. The physical meaning of discernibility function consists in the fact that the pair (X_i, Y_j) can be discerned by using any attribute in $M'(i, j)$, so disjunction operations are made among all attributes which belong to $M'(i, j)$; and all discernible pairs of condition equivalence class can be assured to be still discernible, therefore conjunction operations are made among all $\vee M'(i, j)$. Thus, the value of $\wedge \{ \vee M'(i, j) \}$ becomes true by selecting some attributes (if $a \in M'(i, j)$ is selected, then the value of $\vee M'(i, j)$ is true; if the value of all $\vee M'(i, j)$ in $f(M')$ is true, then $f(M')$ is true.)

Based on the discernibility matrix and discernibility function, attribute reduction and core can be considered in association with knowledge reduction. The core attributes are common part of all reduction, and they are indispensable to preserve original knowledge in information system.

Theorem 1 Given an decision system $DS = \langle U, A = C \cup D, V, f \rangle$ and its improved discernibility matrix M' , $R \subseteq C$ is a reduction if and only if R satisfies:

1. $\forall M'(i, j) \neq \emptyset (1 \leq j < i \leq n) \Rightarrow R \cap M'(i, j) \neq \emptyset$;
2. $\forall R' \subset R, \exists M'(i, j) \neq \emptyset$ such that $R' \cap M'(i, j) = \emptyset$.

The condition (1) shows R is jointly sufficient to distinguish all object pairs which are discernible in original information system, and the condition (2) indicates R is minimal subset of attributes satisfying condition (1). These two

conditions ensure that $R \subseteq C$ is a minimal subset of attributes which preserve particular property of given information system, so R is a reduction.

Theorem 2 *Suppose $f'(M)$ is the minimal disjunctive form which is equivalent to $f(M)$, $R \subseteq C$ is a reduction if and only if $\wedge R$ is a prime implicant of $f'(M)$.*

Every disjunctive item of the minimal disjunctive form is called a prime implicant. $\wedge R$ denotes the conjunctive form of all attributes in R . Theorem 2 illuminates that the problem of finding all reduction is equivalent to transforming the conjunctive form $f(M)$ to the minimal disjunctive form $f'(M)$ by applying many times of the absorption or multiplication laws on Boolean expression.

Theorem 3 *For improved discernibility matrix M' , the attribute core is the collection of singleton attribute element in M' , i.e. $Core(M') = \{M'(i, j) : |M'(i, j)| = 1\}$.*

Where $M'(i, j)$ is an element of discernibility matrix M' , and $|M'(i, j)|$ is the cardinality of $M'(i, j)$, namely the number of attributes contained in $M'(i, j)$. $|M'(i, j)| = 1$ denotes $M'(i, j)$ is an element of singleton attribute, then it is necessary for $M'(i, j)$ contained in each prime implicant of the minimal disjunctive form $f'(M')$. So $Core(M')$ is indispensable to all reductions, and be contained in every reduction.

14.4 Analysis of an Instance

In order to illustrate the method mentioned above, we give an example of application. This example is simple, but it illustrates essential characteristics. Given an decision system in Table 14.1, where $U = \{x1, \dots, x8\}$ $C = \{a, b, c, d, e, f\}$.

Its simplification decision system DS' can be computed as Table 14.2.

The improved discernibility matrix based on changeless positive region and information entropy are shown as Tables 14.3 and 14.4 respectively.

Based on Table 14.3, we can get the discernibility function $f'(M) = (a \vee b \vee c) \wedge (c \vee d) \wedge (b \vee c \vee f) = c \vee (b \wedge d) \vee (a \wedge d \wedge f)$, so the reduction

Table 14.1 Decision system

U	a	b	c	d	d	f	D
x1	0	0	0	0	0	0	0
x2	0	0	0	0	0	0	1
x3	1	1	1	0	0	0	1
x4	1	1	0	1	0	0	0
x5	1	1	0	1	0	0	1
x6	1	1	0	1	0	0	1
x7	1	0	0	0	0	1	0
x8	1	0	0	0	0	1	1

Table 14.2 Simplification decision system DS'

U	a	b	c	d	d	f	D'
X1	0	0	0	0	0	0	(1/2, 1/2)
X2	1	1	1	0	0	0	(0, 1)
X3	1	1	0	1	0	0	(1/3, 2/3)
X4	1	0	0	0	0	1	(1/2, 1/2)

Table 14.3 The improved discernibility matrix based on positive region

	X1	X2	X3	X4
X1				
X2	{a, b, c}			
X3	Φ	{c, d}		
X4	Φ	{b, c, f}	Φ	

Table 14.4 The improved discernibility matrix based on information entropy

	X1	X2	X3	X4
X1				
X2	{a, b, c}			
X3	{a, b, d}	{c, d}		
X4	Φ	{b, c, f}	{b, d, f}	

set based positive region is $\{\{c\}, \{b, d\}, \{a, d, f\}\}$. Based on Table 14.4, we can get the discernibility function $f'(M) = (a \vee b \vee c) \wedge (a \vee b \vee d) \wedge (c \vee d) \wedge (b \vee c \vee f) \wedge (b \vee d \vee f)$, so we can compute that the reduction set based information entropy is $\{\{a, d, f\}, \{a, c, f\}, \{b, c\}, \{b, d\}, \{c, d\}\}$.

14.5 Conclusions

Attribute reduction based on discernibility matrix is one of important parts researched in rough set theory. This paper analyses some disadvantages of discernibility matrix before in detail, and proposes an efficient method of simplifying decision system to find minimum discernibility set. By introducing decision vector, this paper designs an improved discernibility matrix that can suit inconsistency of decision system, as well as be capable of reducing comparison times and storage space for building discernibility matrix by the greatest extent. The theoretic analysis and simulation instance shows that this algorithm is feasible and effective in practice.

References

1. Skowron A, Pawlak Z (2007) Rough sets and Boolean reasoning. *Inf Sci* 177:41
2. Thangavel K, Pethalakshmi A (2009) Dimensionality reduction based on rough set theory: a review. *Appl Soft Comput* 9:1
3. Skowron A, Pawlak Z (2007) Rudiments of rough sets. *Inf Sci* 177:3
4. Skowron RC (1992) The discernibility matrices and functions in information systems. *Intelligent decision support, handbook of applications and advances of rough sets theory*. Kluwer, The Netherlands, p 331
5. Zhao Y, Yao Y, Luo F (2007) Data analysis based on discernibility and indiscernibility. *Inf Sci* 177:4959
6. Yao Y, Zhao Y (2009) Discernibility matrix simplification for constructing attribute reducts. *Inf Sci* 179:867
7. Huang Z, Sun W, Wu H (2008) Reduction tree algorithm based on discernibility matrix. *Comput Appl* 28:1457
8. Huang Z, Wang D (2009) Study on data reduction algorithm based on rough set. *Comput Eng Des* 30:4284