Hindawi
Complexity
Volume 2017, Article ID 8986917, 9 pages
https://doi.org/10.1155/2017/8986917

Research Article

Attribute Reduction Based on Consistent Covering Rough Set and Its Application

Jianchuan Bai,[1,2] Kewen Xia,[1,2] Yongliang Lin,[1,2] and Panpan Wu[1,2]

[1] *School of Electronics and Information Engineering, Hebei University of Technology, Tianjin, China*
[2] *Key Lab of Big Data Computation of Hebei Province, 5340 Xiping Road, Tianjin 300401, China*

Correspondence should be addressed to Kewen Xia; kwxia@hebut.edu.cn

Received 30 March 2017; Revised 25 June 2017; Accepted 24 August 2017; Published 2 October 2017

Academic Editor: Paolo Gastaldo

As an important processing step for rough set theory, attribute reduction aims at eliminating data redundancy and drawing useful information. Covering rough set, as a generalization of classical rough set theory, has attracted wide attention on both theory and application. By using the covering rough set, the process of continuous attribute discretization can be avoided. Firstly, this paper focuses on consistent covering rough set and reviews some basic concepts in consistent covering rough set theory. Then, we establish the model of attribute reduction and elaborate the steps of attribute reduction based on consistent covering rough set. Finally, we apply the studied method to actual lagging data. It can be proved that our method is feasible and the reduction results are recognized by Least Squares Support Vector Machine (LS-SVM) and Relevance Vector Machine (RVM). Furthermore, the recognition results are consistent with the actual test results of a gas well, which verifies the effectiveness and efficiency of the presented method.

1. Introduction

Attribute reduction has become an important step in pattern recognition and machine learning tasks [1, 2]. The main goal of attribute reduction is to remove redundant information in datasets and draw useful information so as to improve classification ability [3]. The theory of classical rough set, as proposed by Pawlak in 1982, has been used as a mathematical tool to deal with various types of insufficient and imperfect data [4]. Rough set theory, which provides a popular mathematical framework for knowledge discovery, feature selection, data mining, and rule extraction, has been concerned by many research scholars since it was first proposed. Generally speaking, the traditional rough set theory can partition the objects of a universe into mutually exclusive equivalence classes, which was based on equivalence relations. The data table that needs to be analyzed by rough set theory is called an information system. Information system, as a mathematical model in artificial intelligence, is deemed as an important application of rough sets [5, 6]. Over the last decades, there has been much work on information systems with rough set, including some successful applications in machine learning, decision analysis, and knowledge discovery. Therefore, rough set theory has been playing a significant role in the unpredictable and uncertain information systems [7, 8].

A drawback of attribute reduction in traditional rough sets is that it can only deal with discrete databases. Therefore, the continuous databases need to be discretized before attribute reduction. Presently, the existing discretization methods can be roughly classified into two categories: supervised discretization method and unsupervised discretization method [9]. Supervised discretization methods generally include discretization based on information entropy and discretization based on ChiMerge algorithm [10], while unsupervised discretization methods arguably include box method for equal frequency or equal width, intuitive division discretization, and discretization based on cluster analysis [11, 12]. There are two limitations in traditional attribute reduction based on rough set theory: (1) databases are numerical in the real world, so that they cannot be handled directly by traditional rough set theory; (2) numerical data have to be discretized before attribute reduction, which inevitably leads to information loss. Therefore, it is desirable to develop an efficient method which can deal with numerical databases directly. The covering rough set theory was proposed to

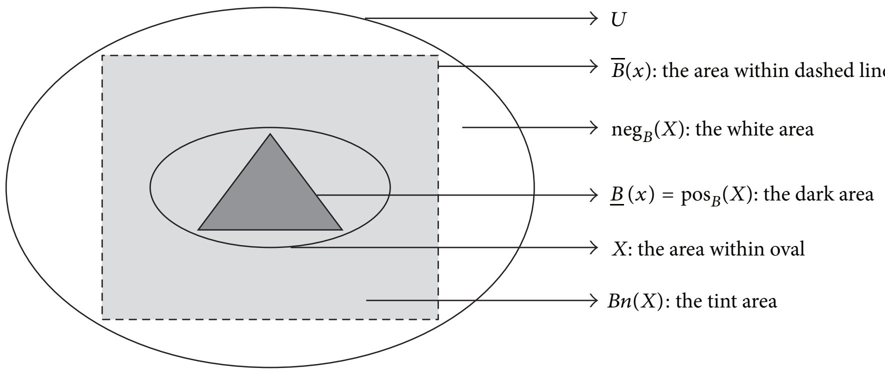

Figure 1: Diagram of approximation relation.

solve this problem efficiently and it avoids the attribute discretization [13].

Covering rough set theory is a generalization of traditional rough set theory, which can deal directly with numerical data. Once launched, covering rough set was of great concern. So far, many researchers conducted studies on the approximation problems based on covering rough set [14–17]. However, to the best of the authors' knowledge, there are relatively few results published on the attribute reduction of covering rough sets and simultaneously its practical application, which motivates the present study.

In this paper, we will first review the theory of traditional rough set and upper-lower approximations, present some basic concepts of consistent covering rough set theory, and establish a model of attribute reduction. Then, the attribute reduction based on consistent covering rough set will be presented and further generalized, which will be compared with attribute reduction based on traditional rough set. Finally, we will apply the studied attribute reduction method to actual logging gas data. LS-SVM and RVM algorithms will be used to recognize the reduction results to confirm its validity.

2. Basic Theory Relate to Rough Set

2.1. Pawlak's Rough Sets. In rough set theory, the quadruplet $S = \langle U, A, V, f \rangle$ is called an information system, where $U = \{x_1, x_2, \ldots, x_n\}$ is a nonempty set of samples, called a universe or a sample space. And $A = \{a_1, a_2, \ldots, a_n\}$ is a nonempty set of attributes or features, which is divided into the set C of conditional attributes and the set D of decision attributes, $A = C \cup D$. Every subset B of attributes can induce binary relation $\text{Ind}(B)$, also called B-indiscernibility relation and defined as $\text{Ind}(B) = \{(x, y) \in U \times U : a(x) = a(y), \forall a \in B\}$. $V = \bigcup_{a \in A} V_n$, V_n is the range of $a \in A$; f is a mapping function of $U \times A \to V$; it gives an attribute value for each object, where $\forall a \in A$, $x \in U$, $f(x, a) \in V$. For $X \subseteq U$, B-upper approximation and B-lower approximation are defined as follows: $\underline{B}(x) = \text{pos}_B(X) = \{x \in U : [x]_B \subseteq X\}$, $\overline{B}(x) = \{x \in U : [x]_B \cap X \neq \varnothing\}$, $Bn(X) = \overline{B}(X) - \underline{B}(X)$, and $\text{neg}_B(X) = U - \overline{B}(x)$. Figure 1 intuitively shows the B-upper approximation, B-lower approximation, and the boundary area.

An attribute $a \in B$ ($B \subseteq A$) is called relatively dispensable in B if $\text{Pos}_B(D) = \text{Pos}_{B-\{a\}}(D)$; otherwise, a is indispensable in B. The set of all indispensable attributes in A is called the core of (U, A), denoted as $\text{Core}(B)$. If B is relatively independent in (U, A, D) and $\text{Pos}_B(D) = \text{Pos}_A(D)$, B is called a reduction of (U, A).

2.2. Basic Nations Related to Covering Rough Set

Definition 1 (see [18]). Let U be a universe of discourse and C be a family of subsets of U. Then, C is called a covering of U if no subset in C is empty and $\bigcup C = U$.

It is obvious that a partition of U is certainly a covering of U and the concept of a covering is an extension of a partition. In [19, 20] the notion of coverings was used to construct lower and upper approximation operators and to study properties of these operators.

Definition 2 (see [21]). Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of U. For every $x \in U$, let $C_x = \bigcap\{K_i : x \in K_i \wedge K_i \in C\}$, $\text{cov}(C) = \{C_x : x \in U\}$ is also a covering of U, and C_x is the minimal set including x in $\text{cov}(C)$; one calls $\text{cov}(C)$ the induced covering of C.

For every $x \in U$, C_x is the minimal set including x in $\text{cov}(C)$. Every element in $\text{cov}(C)$ cannot be written as the union of other elements in $\text{cov}(C)$. $\text{cov}(C) = C$ if and only if C is a partition. For any $x, y \in U$, if $y \in C_x$ then $C_x \supseteq C_y$; so if $y \in C_x$ and $x \in C_y$, then $C_x = C_y$.

Definition 3 (see [21]). Let $\Delta = \{C_i : i = 1, \ldots, m\}$ be a family of coverings of U. For every $x \in U$, let $\Lambda_x = \bigcap\{C_{ix} : C_{ix} \in \text{cov}(C_i), i = 1, 2, \ldots, m\}$, and $\text{cov}(\Delta) = \{\Lambda_x : x \in U\}$ is then also a covering of U; Λ_x is the intersection of all coverings including x in Λ.

Obviously Λ_x is the intersection of all coverings including x in Δ. So, for every $x \in U$, Λ_x is the minimal set including x in $\text{cov}(\Delta)$. $\text{cov}(\Delta)$ can be viewed as the intersection of coverings in Δ. Every element in $\text{cov}(\Delta)$cannot be written as the union of other elements in $\text{cov}(\Delta)$. If every covering in Δ is a partition, then $\text{cov}(\Delta)$ is also a partition and Λ_x is the equivalence class that includes x. For every $x, y \in U$, if $y \in \Lambda_x$, then $\Lambda_x \supseteq \Lambda_y$, so if $y \in \Lambda_x$ and $x \in \Lambda_y$, then $\Lambda_x = \Lambda_y$.

Table 1: Weather information table.

No.	c1	c2	c3	c4	c5	D
1	30000 lux	21°C	35%	5.4 m/s	2 points	Yes
2	3000 lux	12°C	45%	8.9 m/s	23 points	Yes
3	120000 lux	35°C	17%	1.4 m/s	9 point	No
4	450 lux	6°C	76%	17.5 m/s	54 points	No
5	8000 lux	17°C	62%	0.2 m/s	35 points	Yes
6	90000 lux	29.5°C	22%	3.7 m/s	3 points	Yes

Definition 4 (see [21]). Let $\Delta = \{C_i : i = 1, \ldots, m\}$ be a family of coverings of U. For any $X \subseteq U$, the lower and upper approximations of X with respect to cov(Δ) are defined as follows: $\underline{P}X = \{x \in U : \Lambda_x \subseteq X\}$, $\overline{P}X = \{x \in U : \Lambda_x \cap X \neq \varnothing\}$.

The positive domain, negative domain, and boundary domain of X relative to Δ are, respectively, computed by the following formulas: $\underline{P}X = \text{pos}_\Lambda(X)$, $\text{neg}_\Lambda(X) = U - \overline{P}X$, and $Bn(X) = \overline{P}X - \underline{P}X$.

3. Attribute Reduction and Simulation Experiment

3.1. Attribute Reduction Based on Traditional Rough Set. In an information system, attribute reduction is an important application of rough set theory. The key idea is to reduct redundant information while maintaining the indiscernibility relation. Then, traditional reduction methods, such as attribute reduction based on discernibility matrix [22], attribute reduction based on heuristic information [23], and attribute reduction based on evolutionary computation [24], can be used to obtain the attribute reduction results. In the following, we take the particle swarm optimization (PSO) algorithm as an example to study attribute reduction based on an evolutionary algorithm.

3.1.1. Attribute Reduction Based on PSO Algorithm. The basic concepts of attribute reduction in rough set theory and the ideas of particle swarm optimization (PSO) are briefly combined to construct attribute reduction algorithm based on PSO. It reduces algorithm complexity effectively. The steps of its algorithm are as follows.

Step 1. Discretize data in original information table (the discretization method is attribute discretization based on curve inflection points) [25].

Step 2. Initialize the particle swarm randomly.

Step 3. Construct the fitness function: $f(x) = \alpha(1 - \text{card}(x)/m) + (1 - \alpha)(\gamma_a(D)/\gamma_b(D))$; calculate the fitness value of each particle swarm.

Step 4. For each particle swarm, set current fitness value as the new *pbest*, if the current fitness value is better than the past one. Select the best *pbest* as *gbest*, and continue to update the position.

Table 2: Discrete results based on curve inflection points.

No.	c1	c2	c3	c4	c5	D
1	1	1	0	1	0	1
2	0	0	1	1	1	1
3	1	1	0	0	0	0
4	0	0	1	1	1	0
5	0	0	1	0	1	1
6	1	1	0	0	0	1

Step 5. Determine whether the termination condition is satisfied; if yes, go to Step 6. Otherwise, return to Step 2 (or take the iteration times as termination condition).

Step 6. Test each particle by using the reduction definition, get all the candidate reduction sets, remove the redundant attributes, and then get the final reduction sets.

3.1.2. Classic Example Simulation and Comparative Analysis. Table 1 is weather information of a city during daytime. And "No." shows the number of tested days. There are 5 condition attributes which are "c1" (luminosity), "c2" (temperature), "c3" (relative humidity), "c4" (wind speed), and "c5" (precipitation), respectively. "D" is decision attribute which represents travel condition.

After discretization based on curve inflection points [25], see Table 2.

After attribute reduction based on PSO algorithm, the reduction result is $U = \{c2, c4, c5\}$, which means that the condition attributes c1 and c3 are redundant. And 3 key attributes determine the travel condition, which are c2 (temperature), c4 (wind speed), and c5 (precipitation), respectively.

However, if we apply discretization based on information entropy, see Table 3.

After the same reduction method based on PSO algorithm but not same discretization method, the reduction result is $U = \{c1, c3, c4, c5\}$. Apparently, 4 key attributes determine travel condition. They are c1 (luminosity), c3 (relative humidity), c4 (wind speed), and c5 (precipitation), respectively.

We examine that the numerical data have to be discretized through traditional rough set theory in real life. However, it should be pointed out that attribute discretization destroys indiscernibility relations between condition attributes and decision attributes to some extent, and it also leads to lack of

Table 3: Discrete results based on information entropy.

No.	c1	c2	c3	c4	c5	D
1	2	2	2	2	1	1
2	1	1	2	2	1	1
3	2	2	1	1	1	0
4	1	2	2	2	2	0
5	1	1	2	1	2	1
6	2	2	1	1	1	1

information and different reduction results. As a result, the accuracy of attribute reduction is affected. In order to solve the complexity of continuous attribute discretization, we will present a method of attribute reduction based on consistent covering rough set. And the present method can be used to greatly improve the accuracy and efficiency of attribute reduction.

3.2. Attribute Reduction Based on Consistent Covering Rough Set

3.2.1. Basic Definitions and Principles. In practical applications, a large number of databases cannot be directly handled by classical rough sets. For this reason, neighborhood rough sets and similarity relation rough sets were developed. These models induce coverings of a universe instead of partitions and can thus be categorized into covering rough sets. In the following, we review some definitions of consistent covering rough sets.

Definition 5 (see [26]). Let $\Delta = \{C_i : i = 1, 2, \ldots, m\}$ be a family of coverings of U. D is a decision attribute set. U/D is a decision division on U.

If, $\forall x \in U$, $\exists B_j \in U/D$ such that $\Delta_x \subseteq D_j$, then decision system (U, Δ, D) is called a consistent covering decision system and donated as $\mathrm{cov}(\Delta) \leq U/D$. The positive region of D relative to Δ is defined as $\mathrm{Pos}_\Delta(D) = \bigcup_{X \in U/D} \underline{\Delta}(X)$. Otherwise, (U, Δ, D) is called an inconsistent covering decision system.

Definition 6. Let $(U, \Delta, D = \{d\})$ be a consistent covering decision system. Supposing $U = \{x_1, x_2, \ldots, x_n\}$, by $M(U, \Delta, D)$ we donate $n \times n$ matrix (c_{ij}), called the discernibility matrix of (U, Δ, D) and defined as follows.

(1) when $d(\Delta_{x_i}) \neq d(\Delta_{x_j})$,

$$c_{ij} = \left\{C \in \Delta : \left(C_{x_i} \not\subset C_{x_j}\right) \wedge \left(C_{x_j} \not\subset C_{x_i}\right)\right\} \cup \left\{C_s \wedge C_t : \left(\left(C_s\right)_{x_i} \subset \left(C_s\right)_{x_j}\right) \wedge \left(\left(C_t\right)_{x_j} \subset \left(C_t\right)_{x_i}\right)\right\}. \quad (1)$$

(2) When $d(\Delta_{x_i}) = d(\Delta_{x_j})$,

$$c_{ij} = \Delta. \quad (2)$$

If $C \in c_{ij}$ for $d(\Delta_{x_i}) \neq d(\Delta_{x_j})$, C is one of the covers to maintain the relation between x_i and x_j with respect to Δ. Here we should point out that if $c_{ij} = \{C_s : s = 1, \ldots, l\}$, the relations between elements in c_{ij} are a disjunction; if $c_{ij} = \{C, C_s \wedge C_t : s, t \leq n\}$, we mean it is conjunction between C_s and C_t, $s_0, t_0, s_1, t_1 \leq n$. Since $M(U, \Delta, D)$ is symmetric and $c_{ii} = \Delta$, for $c_{ii} = \Delta i = 1, \ldots, n$, we represent $M(U, \Delta, D)$ only by elements in the lower triangle of $M(U, \Delta, D)$.

Theorem 7. *Let $(U, \Delta, D = \{d\})$ be a consistent covering decision system, $M(U, \Delta, D) = (c_{ij} : i, j \leq n)$ is the discernibility matrix of (U, Δ, D), and the discernibility function is as follows:*

$$f(U, \Delta, D) = \bigwedge_{i,j=1}^{n} \left(\bigvee c_{ij}\right), \quad \left(c_{ij} \neq \varnothing\right) \quad (3)$$

where $\bigvee c_{ij}$ means, for every c_{ij}, C or $C_s \wedge C_t$ conducts a disjunctive operation. Make $f(U, \Delta, D) = \bigvee_{k=1}^{l} (\bigwedge \Delta_k)$ $(\Delta_k \subseteq \Delta)$; then the set $\{\Delta_k : k \leq l\}$ is the collection of all the reductions of the system.

3.2.2. Classic Example Simulation. In order to further validate the feasibility of the algorithm, illustrative example is applied for simulation analysis.

Table 4 is logging dataset of a typical well, where "No." represents the number of wells. "c1" represents acoustic time, "c2" represents caliper, "c3" represents natural gamma, "c4" represents plate radius, "c5" represents induction resistivity, and "c6" represents flushed zone resistivity. "D" is the type of a well, "0" represents dry well, and "1" represents oil well [27].

Obviously, all the condition attributes are numerical data in Table 4. So data have to be discretized and they cannot be directly handled by traditional rough sets. Therefore, consistent covering rough set can be applied to deal with the data in Table 4 so that the lack of information by traditional rough set theory is avoided.

According to the definition of covering rough set, $C_{i \leq 7} = \{(A_i)_b(j) : j = 1, 2, \ldots, 7\}$ is a covering of sample set $U = \{1, 2, \ldots, 7\}$. Therefore, $\Delta = \{C_1, C_2, C_3, C_4\}$ constitutes 4 coverings of sample U. In decision attribute table, for every condition attribute, the "descending order" is used to establish the equivalence relation. So we can get $\Delta_1 = \{1\}$, $\Delta_2 = \{2\}$, $\Delta_3 = \{3\}$, $\Delta_4 = \{4\}$, $\Delta_5 = \{5\}$, $\Delta_6 = \{6\}$, $\Delta_7 = \{7\}$. The sample U can be divided into two categories which are D_1, D_2 according to decision attribute D, and $D_1 = \{1, 2, 3\}$, $D_2 = \{4, 5, 6, 7\}$. Obviously, $\mathrm{Pos}_\Delta(D) = \bigcup_{X \in U/D} \underline{\Delta}(X) = \underline{\Delta}(D_1) \cup \underline{\Delta}(D_2)$ due to the definition of consistent covering rough set. And $\underline{\Delta}(D_1) = R_s(1) \cup R_s(3) \cup R_s(6) = \{1, 2, 3\}$, $\underline{\Delta}(D_2) = R_s(2) \cup R_s(4) \cup R_s(5) = \{4, 5, 6, 7\}$. Therefore, the discernibility matrix of (U, Δ, D) is as follows:

$$\begin{bmatrix} \Delta & \Delta & \Delta & \Delta_{14} & \Delta_{15} & \Delta_{16} & \Delta_{17} \\ \Delta & \Delta & \Delta & \Delta_{24} & \Delta_{25} & \Delta_{26} & \Delta_{27} \\ \Delta & \Delta & \Delta & \Delta_{34} & \Delta_{35} & \Delta_{36} & \Delta_{37} \\ \Delta & \Delta & \Delta & \Delta & \Delta & \Delta & \Delta \\ \Delta & \Delta & \Delta & \Delta & \Delta & \Delta & \Delta \\ \Delta & \Delta & \Delta & \Delta & \Delta & \Delta & \Delta \end{bmatrix}. \quad (4)$$

Table 4: Logging data set.

No.	c1	c2	c3	c4	D
1	341.919	30.6871	28.7472	7.67827	0
2	330.67	22.2717	27.4297	6.68225	0
3	384.480	22.1349	32.9302	8.72574	0
4	331.969	30.7547	25.1362	5.4875	1
5	299.244	31.5446	30.8018	5.32932	1
6	267.413	32.3739	31.9981	3.45624	1
7	277.413	22.2649	29.9981	4.45624	1

For every Δ, calculation results are as follows:

$$\begin{aligned}
&\Delta_{14} = \Delta_{34} = \Delta_{35} = \Delta_{36} = \Delta_{37} \\
&\quad = \{C_1 \wedge C_2, C_2 \wedge C_3, C_2 \wedge C_4\} \\
&\Delta_{15} = \Delta_{16} = \Delta_{25} = \Delta_{26} \\
&\quad = \{C_1 \wedge C_2, C_1 \wedge C_3, C_2 \wedge C_4, C_3 \wedge C_4\} \\
&\Delta_{17} = \Delta_{27} = \{C_1 \wedge C_3, C_2 \wedge C_3, C_3 \wedge C_4\} \\
&\Delta_{24} = \{C_1 \wedge C_3, C_1 \wedge C_4, C_1 \wedge C_6, C_2 \wedge C_3, C_2 \wedge C_4\} \\
&\Delta = C_1 \vee C_2 \vee C_3 \vee C_4.
\end{aligned} \tag{5}$$

Reduction results based on discernibility function are as follows:

$$\begin{aligned}
f(U, \Delta, D) &= \bigwedge_{i,j=1}^{n} \left(\bigvee c_{ij}\right) = \Delta \wedge ((C_1 \wedge C_2) \\
&\vee (C_2 \wedge C_3) \vee (C_2 \wedge C_4) \vee (C_1 \wedge C_3) \vee (C_1 \wedge C_4) \\
&\vee (C_2 \wedge C_4)) \wedge ((C_1 \wedge C_3) \vee (C_1 \wedge C_4) \\
&\vee (C_2 \wedge C_3) \vee (C_2 \wedge C_4)) \wedge ((C_1 \wedge C_2) \\
&\vee (C_1 \wedge C_3) \vee (C_2 \wedge C_4) \vee (C_3 \wedge C_4)) \\
&\wedge ((C_1 \wedge C_3) \vee (C_2 \wedge C_3) \vee (C_3 \wedge C_4)) = (C_1 \wedge C_2 \\
&\wedge C_3) \vee (C_2 \wedge C_3 \wedge C_4).
\end{aligned} \tag{6}$$

According to the results, there are two reduction results in this decision system, which are $(C_1 \wedge C_2 \wedge C_3)$, $(C_2 \wedge C_3 \wedge C_4)$, respectively. Apparently, the condition attributes C_1, C_2, are the key information to distinguish between "dry well" and "oil well," so Core$(\Delta) = \{C_1, C_2\}$.

4. Algorithm Description Based on Consistent Covering Rough Set

Attribute reduction is a core application of rough set. In this paper, the main emphasis is laid on the attribute reduction based on consistent covering rough set. For consistent covering decision system, the essence of attribute reduction is to ensure the minimum subset of conditional attribute so as to achieve the purpose of attribute reduction [28]. According to the above classic example (Section 3.2.2), the flow chart of attribute reduction model based on consistent covering rough set is provided in Figure 2.

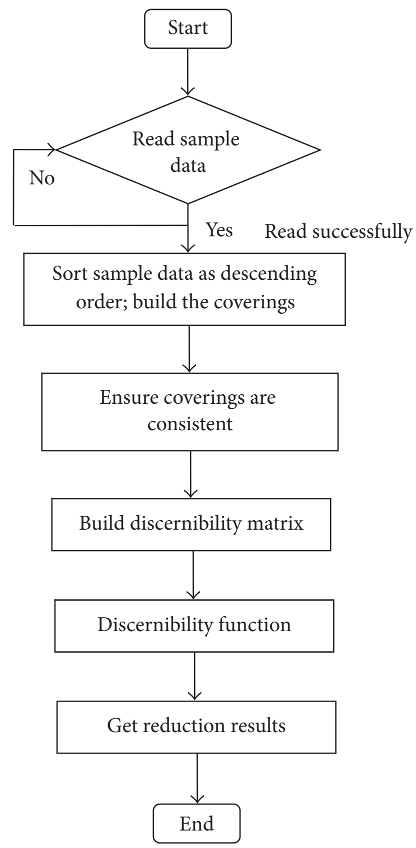

Figure 2: Flow chart of reduction based on consistent covering rough set.

According to Figure 2, algorithm steps are designed as follows. Furthermore, the algorithm of attribute reduction based on consistent rough set is programmed in this paper.

Step 1. Read sample data in decision information table.

Step 2. Sort sample data as descending order, and build coverings of the sample.

Step 3. Ensure the decision system is consistent; then run Step 4 (we only consider consistent covering decision system in this paper).

Step 4. Build discernibility matrix $M(U,\Delta,D)$, that is, (c_{ij}), $1 \le j \le i \le n$.

Step 5. Write discernibility function: $f(U,\Delta,D) = \bigwedge_{i,j=1}^{n} (\bigvee c_{ij})$, $(c_{ij} \ne \varphi)$ according to discernibility matrix.

Step 6. Get the reduction set $\{\Delta_k : k \le l\}$ through conjunctive and disjunctive forms; that is, $\mathrm{Red}_D(\Delta) = \{\Delta_k : k \le l\}$.

5. Practical Application and Experimental Analysis

In order to validate the effectiveness of the studied method for attribute reduction based on consistent covering rough set, we adopt the logging data of a gas well named "Su6" in Xinjiang (China) as showed in Table 5 and conduct a comparative analysis. All condition attributes are numerical. Moreover, 200 experimental sample data types (well depth 3000 m–3400 m) are selected instead of all logging data in order to maintain confidentiality. Among them, they are 80 gas layer points and 120 nongas layer points according to the actual test results. There are 13 condition attributes in Table 5, which are GR (natural gamma), DT (acoustic time), SP (spontaneous potential), WQ (flush zone resistivity), LLD (deep investigated double lateral resistivity), LLS (shallow investigated double lateral resistivity), DEN (density), NPHI (compensated neutron), PE (photoelectric absorption index), U (uranium), TH (thorium), K (potassium), and CALI (borehole diameter). The decision attributes of sample information are the nongas layer and the gas layer, the decision attributes are denoted by D_1, D_2, respectively. And "0" is for nongas layer; "1" is for gas layer. (Note: gas field is abbreviated as natural gas field that is rich in natural gas. Typically, organic matter is buried between 1 and 6 km depth, and oil will be produced with temperatures between 65 and 150 degrees Celsius. Natural gas will be produced while deeper.)

According to the definition of consistent covering rough set, $\mathrm{Pos}_{\Delta}(D) = \bigcup_{X \in U/D} \underline{\Delta}(X) = \underline{\Delta}(D_1) \cup \underline{\Delta}(D_2)$. Obviously, logging data in Table 5 is consistent decision system. The data in Table 5 are input to the program of attribute reduction based on consistent covering rough set. Then, reduction results are $\{\mathrm{GR}, \mathrm{DT}, \mathrm{SP}, \mathrm{LLD}, \mathrm{LLS}, \mathrm{DEN}, \mathrm{K}\}$. The two different traditional reduction methods of the rough set, which are attribute reduction based on identification matrix and particle swarm optimization- (PSO-) based attribute reduction of rough set, are used to deal with the data in Table 5 for comparison and analysis. Reduction results are shown in Table 6.

According to reduction results and running time in Table 6, we know that reduction method based on consistent covering rough set has advantages of fewer reduction attributes and shorter running time.

In order to further validate the effectiveness of attribute reduction based on consistent covering rough set, the reduction results are recognized by Least Squares Support Vector Machine (LS-SVM) [29] and Relevance Vector Machine (RVM) [30]. Recognition results are shown in Table 7. The recognition results show that recognition accuracies of the studied algorithm are 94.2% (LS-SVM) and 91.5% (RVM), respectively, which are higher than the other two reduction algorithms.

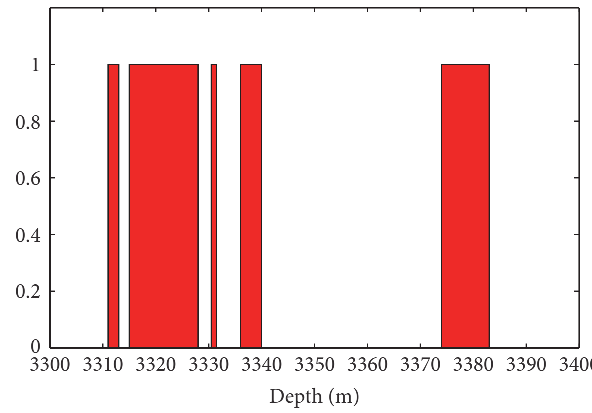

Figure 3: The actual gas distribution.

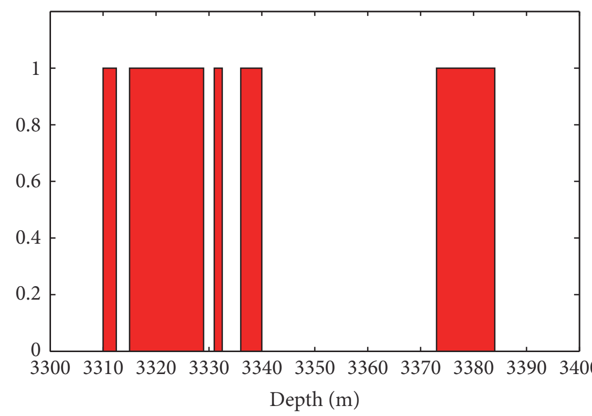

Figure 4: LS-SVM recognition result.

Figure 3 shows the actual gas distribution, Figure 4 shows that recognition results of the studied algorithm by LS-SVM, and the recognition accuracy is 94.2%. Figure 5 shows recognition results of the studied algorithm by RVM and the recognition accuracy is 91.5%.

According to a comparison of recognition results in Figures 3, 4, and 5, we know both recognition accuracies of the studied algorithm by LS-SVM and RVM go up to 90%. It can effectively reduce the tedious work in gas recognition and improve the recognition accuracy. Figures 6 and 7 show classification of the studied algorithm by LS-SVM and RVM, respectively. Among them, the red line indicates classification line, the green points indicate gas layer points, and the black asterisks indicate nongas layer points.

The proposition of attribute reduction based on consistent covering rough set is of great significance. On one hand, it avoids the tedious steps of continuous attribute discretization and reduces the lack of important information in decision information table. For these reasons, the accuracy and efficiency of attribute reduction based on traditional rough set can be improved largely. On the other hand, the proposed algorithm can directly handle the numerical data in the real world and significantly reduce the workload compared with traditional attribute reduction. The presented

Table 5: Logging data.

No.	GR	DT	SP	WQ	LLD	LLS	DEN	NPHI	PE	U	TH	K	CALI	D
1	63.362	209.389	−98.9	60.92	62.405	57.044	2.691	16.237	3.37	1.233	6.859	1.349	24.358	0
2	73.283	206.856	−98.535	106.986	48.157	45.815	2.663	12.105	3.562	2.678	4.607	1.894	24.322	0
3	77.512	210.22	−98.437	39.98	39.923	38.771	2.63	10.003	2.667	0.691	3.933	1.942	24.257	0
4	63.234	213.583	−98.65	43.676	38.349	37.276	3.641	11.285	2.732	1.275	5.906	1.717	24.312	0
5	66.788	211.721	−98.417	46.151	42.694	40.441	2.629	11.715	2.781	1.138	5.319	1.708	24.355	0
					⋮				⋮					
61	54.793	222.773	−109.652	31.408	43.163	36.693	2.533	7.007	1.893	0.902	3.43	1.238	24.254	0
62	47.722	222.53	−107.823	57.011	47.901	41.626	2.532	6.466	1.763	0.987	1.742	1.435	24.291	1
63	31.311	221.165	−105.828	70.879	57.312	47.148	2.521	5.447	2.081	0.842	1.227	1.004	24.254	1
64	39.088	212.365	−103.517	196.409	75.751	65.579	2.562	6.602	2.16	0.508	2.204	0.867	24.26	1
65	74.959	221.004	−102.212	45.069	33.361	31.123	2.618	11.147	2.528	1.55	7.019	1.728	24.24	0
					⋮				⋮					
151	69.009	221.921	−103.021	33.201	34.429	30.911	2.586	12.751	2.523	1.628	8.349	1.904	24.352	1
152	46.494	226.873	−107.029	32.213	36.114	31.939	2.546	10.368	2.345	2.215	6.114	1.046	24.399	1
153	50.546	228.546	−109.857	45.837	39.122	33.695	2.533	9.242	2.285	1.842	4.077	0.832	24.32	1
154	52.725	232.334	−113.119	40.916	44.262	36.261	2.509	7.909	2.218	1.624	4.238	1.156	24.298	1
155	45.29	238.317	−115.629	60.731	50.187	39.995	2.466	7.242	2.113	1.857	1.848	0.864	24.389	1
					⋮				⋮					
196	146.387	238.736	−96.099	36.032	32.811	30.017	2.613	40.85	3.745	5.371	17.927	2.399	24.293	0
197	141.499	243.434	−95.169	15.669	31.371	27.584	2.573	43.18	2.511	4.725	22.255	2.371	24.291	0
198	131.245	232.916	−95.325	43.736	41.085	35.179	2.666	40.857	2.711	2.769	17.662	3.142	24.192	0
199	149.161	238.786	−94.79	46.739	42.172	35.286	2.635	41.694	3.121	3.884	16.046	3.272	24.367	0
200	135.228	235.772	−95.217	37.182	52.036	42.043	2.652	40.582	3.286	2.619	18.602	2.9	24.552	0

Table 6: Comparison of reduction results.

Reduction algorithm	Reduction results	Running time
Attribute reduction based on consistent covering rough set	{GR, DT, SP, LLD, LLS, DEN, K}	2.67 s
Attribute reduction based on identification matrix	{GR, SP, LLD, DEN, NPHI, PE, U, TH, CALI}	3.61 s
PSO-based attribute reduction of rough set	{GR, SP, WQ, LLD, DEN, NPHI, PE, TH, K}	2.98 s

Table 7: Comparison of recognition accuracy.

Reduction algorithm	LS-SVM	RVM
Attribute reduction based on consistent covering rough set	94.2%	91.5%
Attribute reduction based on identification matrix	87.3%	83.6%
PSO-based attribute reduction of rough set	89.1%	86.7%

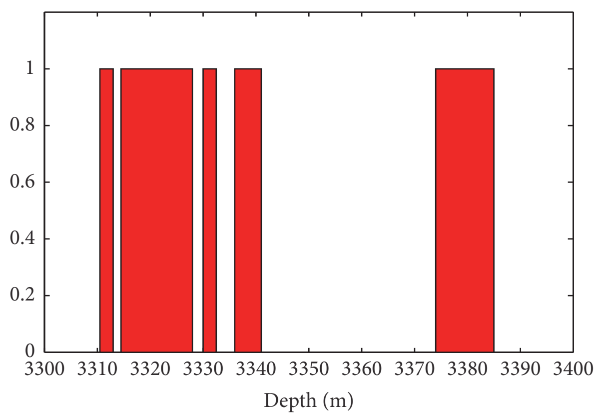

Figure 5: RVM recognition result.

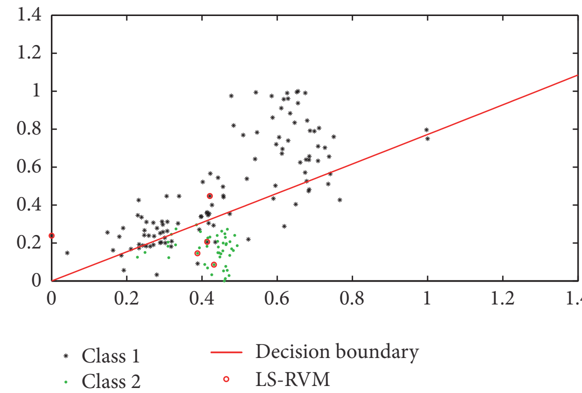

Figure 6: Classification of sample data by LS-SVM.

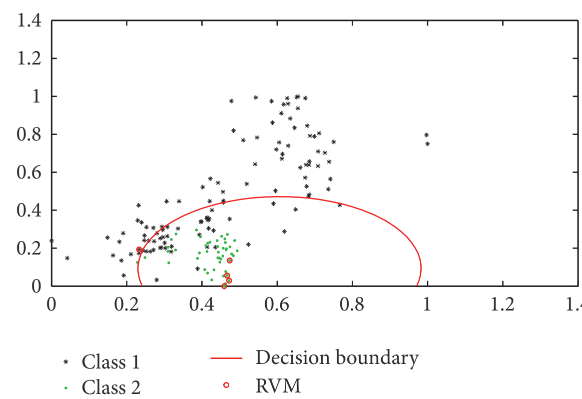

Figure 7: Classification of sample data by RVM.

method was applied to actual lagging data; it proved that gas exploration is effective and the recognition accuracy is high. The presented method is feasible and reasonable and it has important theoretical significance and practical value for artificial intelligence and data mining.

6. Conclusion

An efficient attribute reduction algorithm on the basis of consistent covering rough set has been presented. The knowledge of traditional rough set and covering rough set has been analyzed. The drawbacks of attribute reduction based on traditional rough set and the advantages of covering rough set have been also discussed. The actual logging data have been applied to test the feasibility and efficiency of the presented algorithm. The experimental results have shown that the studied reduction method can effectively handle numerical data and is much more efficient than traditional rough set theory. The reduction results have been compared with actual recognition results by LS-SVM and RVM algorithm so as to validate the algorithm's effectiveness. It has been demonstrated that the proposed recognition results are consistent with the actual gas distribution and the recognition accuracy is high.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Hebei Province Natural Science Foundation (no. E2016202341) and Hebei Province Foundation for Returned Scholars (no. C2012003038).

References

[1] W. Ding, J. Wang, and Z. Guan, "A novel minimum attribute reduction algorithm based on hierarchical elitist role model combining competitive and cooperative co-evolution," *Chinese Journal of Electronics*, vol. 22, no. 4, pp. 677–682, 2013.

[2] D. Chen, S. Zhao, L. Zhang, Y. Yang, and X. Zhang, "Sample pair selection for attribute reduction with rough set," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 11, pp. 2080–2093, 2012.

[3] Z. a. Pawlak, "Rough sets," *International Journal of Computer and Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.

[4] Z. Pawlak, "Rough sets and their applications," *Computational Intelligence in Theory And Practice*, vol. 8, pp. 73–91, 2001.

[5] C. Wang, Q. Hu, X. Wang, D. Chen, and Y. Qian, "Feature selection based on neighborhood discrimination index," *IEEE Transactions on Neural Networks Learning Systems*, vol. 99, pp. 1–14, 2017.

[6] L. Guan and G. Wang, "Generalized approximations defined by non-equivalence relations," *Information Sciences. An International Journal*, vol. 193, pp. 163–179, 2012.

[7] D. Miao and D. Li, *Rough Set Theory, Algorithm And Application*, Tsinghua University Press, Beijing, China, 2008.

[8] Y. Qian, J. Liang, and C. Dang, "Incomplete multigranulation rough set," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, vol. 40, no. 2, pp. 420–431, 2010.

[9] C.-J. Tsai, C.-I. Lee, and W.-P. Yang, "A discretization algorithm based on Class-Attribute Contingency Coefficient," *Information Sciences*, vol. 178, no. 3, pp. 714–731, 2008.

[10] R. Kerber, "ChiMerge: discretization of numeric attributes," in *Proceedings of the Tenth Conference on Artificial Intelligence*, pp. 123–128, 1992.

[11] D.-Q. Miao, "A New method of discretization of continuous attributes in rough sets," *Acta Automatica Sinica*, vol. 27, no. 3, pp. 296–302, 2001.

[12] G. Zhang, W. Jin, and L. Hu, "Generalized discretization of continuous attributes in rough set theory," *Control and Decision*, vol. 20, no. 4, pp. 372–376, 2005.

[13] L. Hou, G. Wang, N. Niu, and Y. Wu, "Discretizaton in rough set theory," *Computer Science*, vol. 27, no. 12, pp. 89–94, 2000.

[14] R. Slowinski and D. Vanderpooten, "A generalized definition of rough approximations based on similarity," *IEEE Transactions on Knowledge and Data Engineering*, vol. 12, no. 2, pp. 331–336, 1996.

[15] N. D. Thuan, "Covering rough sets from a topological pointof view," *International Journal of Computer Theory and Engineering*, vol. 1, no. 5, pp. 601–609, 2012.

[16] W. Zhu, "Topological approaches to covering rough sets," *Information Sciences. An International Journal*, vol. 177, no. 6, pp. 1499–1508, 2007.

[17] G. Resconi and C. Hinde, "Agents and rough sets," *International Journal of Computational Intelligence Systems*, vol. 7, no. 1, pp. 137–157, 2014.

[18] Z. Bonikowski, E. Bryniarski, and U. Wybraniec-Skardowska, "Extensions and intentions in the rough set theory," *Information Sciences. An International Journal*, vol. 107, no. 1-4, pp. 149–167, 1998.

[19] J. A. Pomykała, "Approximation operations in approximation space," *Bulletin of the Polish Academy of Sciences*, vol. 35, no. 1, pp. 653–662, 1987.

[20] J. A. Pomykala, "On definability in the nondeterministic information system," *Bulletin of the Polish Academy of Sciences: Mathematics*, vol. 36, no. 3, pp. 193–210, 1988.

[21] C. Wang, Q. He, D. Chen, and Q. Hu, "A novel method for attribute reduction of covering decision systems," *Information Sciences. An International Journal*, vol. 254, pp. 181–196, 2014.

[22] Y. Lv, H. Liu, and J. Jiang, "The method of attribute reduction based on discernibility matrix," in *Proceedings of the International Conference on Computer Application System Modelling*, vol. 1, pp. 370-370, 2010.

[23] R.-S. Cao, S.-H. Ni, and P. Zhang, "New heuristic attribute reduction algorithm based on sample selection," *Computer Science*, vol. 43, no. 8, pp. 40–43, 2016.

[24] M. Alweshah, O. A. Alzubi, and J. A. Alzubi, "Solving attribute reduction problem using wrapper genetic programming," *International Journal of Computer Science and Network Security*, vol. 16, no. 5, pp. 77–84, 2016.

[25] J. Bai, K. Xia, Y. Chi, and L. Liu, "Continuous attribute discretization based on inflection point," *Journal of Information and Computational Science*, vol. 11, no. 4, pp. 1327–1333, 2014.

[26] C. Degang, W. Changzhong, and H. Qinghua, "A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets," *Information Sciences. An International Journal*, vol. 177, no. 17, pp. 3500–3518, 2007.

[27] C. Wang and D. Chen, *Theory and method of knowledge acquisition based on Rough Sets*, Harbin Institute of Technology Press, China, 2010.

[28] Y. Yao and B. Yao, "Covering based rough set approximations," *Information Sciences. An International Journal*, vol. 200, pp. 91–107, 2012.

[29] M. M. Adankon and M. Cheriet, "Model selection for the LS-SVM. Application to handwriting recognition," *Pattern Recognition*, vol. 42, no. 12, pp. 3264–3270, 2009.

[30] G. M. Foody, "RVM-based multi-class classification of remotely sensed data," *International Journal of Remote Sensing*, vol. 29, no. 6, pp. 1817–1823, 2008.

Advances in
Operations Research
Advances in
Decision Sciences
Journal of
Applied Mathematics
Algebra
Journal of
Probability and Statistics
The Scientific
World Journal
International Journal of
Differential Equations
Hindawi
Submit your manuscripts at
https://www.hindawi.com
International Journal of
Combinatorics
Advances in
Mathematical Physics
Journal of
Complex Analysis
Journal of
Mathematics
Mathematical Problems
in Engineering
Abstract and
Applied Analysis
Discrete Dynamics in
Nature and Society
International
Journal of
Mathematics and
Mathematical
Sciences
Journal of
Discrete Mathematics
Journal of
Function Spaces
International Journal of
Stochastic Analysis
Journal of
Optimization