Information Sciences 539 (2020) 312–326

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Novel matrix-based approaches to computing minimal and maximal descriptions in covering-based rough sets

Caihui Liu [a,*], Kecan Cai [b], Duoqian Miao [b], Jin Qian [c,d]

[a] School of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000 Jiangxi, China
[b] Department of Computer Science and Technology, Tongji University, 201804 Shanghai, China
[c] School of Software, East China Jiaotong University, Nanchang 330013, Jiangxi, China
[d] School of Computer Engineering, Jiangsu University of Technology, Changzhou 213015, Jiangsu, China

ARTICLE INFO

Article history:
Received 5 November 2019
Received in revised form 1 June 2020
Accepted 7 June 2020
Available online 18 June 2020

Keywords:
Covering
Matrix
Maximal description
Minimal description
Rough sets

ABSTRACT

Minimal and maximal descriptions of concepts are two important notions in covering-based rough sets. Many issues in covering-based rough sets (e.g., reducts, approximations, etc.) are related to them. It is well known that, it is time-consuming and error-prone when set representations are used to compute minimal and maximal descriptions in a large scale covering approximation space. To address this problem, matrix-based methods have been proposed in which calculations can be conveniently implemented by computers. In this paper, motivated by the need for knowledge discovery from large scale covering information systems and inspired by the previous research work, we present two novel matrix-based approaches to compute minimal and maximal descriptions in covering-based rough sets, which can reduce the computational complexity of traditional methods. First, by introducing the operation "sum" into the calculation of matrix instead of the operation "$\oplus$", we propose a new matrix-based approach, called approach-1, to compute minimal and maximal descriptions, which does not need to compare the elements in two matrices. Second, by using the binary relation of inclusion between elements in a covering, we propose another approach to compute minimal and maximal descriptions. Finally, we present experimental comparisons showing the computational efficiency of the proposed approaches on six UCI datasets. Experimental results show that the proposed approaches are promising and comparable with other tested methods.

1. Introduction

Covering, as a widely used form of data representation, has been applied to many practical applications. Covering-based rough sets, as an efficient means for dealing with covering data, was first proposed by Zakowski [1]. As one of the meaningful extension models of rough sets [2], it has attracted much attention and induced lots of interesting results [3–14,34]. Pomykala [3] introduced several pairs of dual approximation operators of covering-based rough sets. Bonikowski et al. [4] studied the covering-based rough approximation operators based on the mutual correspondence of the concepts of extension and intension. Mordeson [5] applied covering-based rough sets to ideal theory and discussed basic properties of the upper approximation operator and showed how it can be used to give algebraic structural properties of certain subsets. Tsang

* Corresponding author.
E-mail addresses: liucaihui@gnnu.edu.cn (C. Liu), caikecan@tongji.edu.cn (K. Cai), dqmiao@tongji.edu.cn (D. Miao), 3116@ecjtu.edu.cn (J. Qian).

https://doi.org/10.1016/j.ins.2020.06.022

et al. [6] proposed a new upper approximation operator of covering-based rough sets, which may help to extract more efficient rules. Zhu and Wang [7,8] proposed three types of covering-based rough sets and introduced the notions of reduct and exclusion of a covering. Grzymala-Busse [15–18] has done a series of studies on the missing attribute values, which also an important aspect of covering-based rough approximation operators. Liu and Sai gave a comparison of two types of rough sets induced by coverings [19]. Qin et al. discussed some properties of five pairs of dual covering approximation operators, and presented conditions with which these covering approximation operators are identical [20]. Zhang and Luo changed five pairs of covering approximation operators into the same pair of relation approximation operators [21]. Restrepo et al. investigated partial order relation for approximation operators in covering based rough sets [22]. Restrepo et al. used the concepts of duality, conjugacy and adjointness to establish relationships between the most commonly used covering approximation operators [23]. Zhu et al. [9,10] discussed some measurements of uncertainty in covering-based rough sets. Ma [11] introduced some types of neighborhood-related covering rough sets. Zhang et al. [24] have shown the closeness of union and intersection operations of rough approximation pairs. As a guideline and directional research of covering-based rough sets, Yao et al. [12] proposed a framework for the study of covering based rough set approximations. They summarized and classified the existing approximation operators into element-based, granule-based and subsystem-based definitions, which enables us to reproduce many existing approximation operators and introduce some new approximation operators.

Minimal and maximal descriptions of concepts are two important notions in covering-based rough sets. Many issues in covering-based rough sets, e.g., reducts, approximations, etc., are related to them. In most cases, it is time-consuming and error-prone when set representations are used to compute minimal and maximal descriptions in a large scale covering approximation space. Especially, with the increase of the volume of data, it is becoming more and more difficult to process them. To address this problem, matrix-based methods have been proposed in which calculations can be conveniently implemented by computers. Recently, research on the matrix-based representation for covering-based rough sets has been a hot topic in the area of rough set theory. For example, Wang et al. [25] defined two characteristic matrices of a covering and represented equivalently three types of existing covering approximation operators by using those two boolean characteristic matrices. Tan et al. [26] introduced matrix-based methods to compute set approximations and reducts in a covering decision information system. Lang et al. [27] presented incremental approaches for computing the second and sixth lower and upper approximations of sets in the dynamic covering approximation spaces, where the characteristic matrices of dynamic coverings are updated by the way of immigration and emigration of objects. Huang et al. [28] presented a matrix-based representation of rough fuzzy approximations by a Boolean matrix associated with a matrix operator in dynamic fuzzy decision systems. Wang and Zhu [29] first constructed a bipartite graph by using a covering, and then presented two equivalent representations of a pair of covering upper and lower approximation operators according to the constructed bipartite graph. Yang and Hu [30] studied the matrix representations and interdependency of three pairs of L-fuzzy covering-based approximation operators and proposed some necessary and sufficient conditions under which two L-fuzzy coverings to generate the same L-fuzzy covering-based approximation operators. In Ref. [32], Wang and Zhang defined several matrices and matrix operations for computing the minimal and maximal descriptions, together with the corresponding approximation operators, and presented the method to calculate reductions of coverings. Although the results in Ref. [32] showed an interesting view to investigate the combination between matrices and covering-based rough sets, there still exist some drawbacks that should be addressed. For example, noting that the result of $A(C) \cdot A^T(C)$ is symmetric, we only need to calculate the upper triangle elements in the matrix, which can reduce the number of calculations. In this paper, motivated by the need for knowledge discovery from large scale covering information systems and inspired by the work of Wang et al. [32], we present two novel matrix-based approaches to compute minimal and maximal descriptions in covering-based rough sets, which can reduce the computational complexity of traditional methods.

The motivation of this paper is outlined as follows.

- New calculation strategies are developed, which provide new methods to compute minimal and maximal descriptions in covering-based rough sets.
- The proposed calculation methods are promising and comparable with the existing methods in the literature.

The paper is organized as follows. The next section introduces some preliminary concepts and properties regarding covering-based rough sets. In Section 3, we present two kinds of calculating methods for minimal and maximal descriptions in covering-based rough sets. The experimental results are given in Section 4. Finally, some conclusions are included in the last section.

2. Preliminaries

In this section, we give some basic definitions about covering approximation space, minimal description and maximal description in covering-based rough sets. Throughout the paper, we suppose that U is a nonempty and finite set, named universe.

Definition 1. [1] Let U be a universe of discourse and C a family of nonempty subsets of U. If $\cup C = U$, then C is called a covering of U. The ordered pair $\langle U, C\rangle$ is called a covering approximation space.

Coverings are a type of common and important data organization mode, and covering-based rough sets which serve as a generalization of Pawlak's rough sets [2] are an effective tool to deal with this kind of data. One example is that an attribute or an attribute subset in incomplete information systems induces a covering instead of a partition. Another example is that a tolerance relation on a universe also generates a covering. It has been attracted increasing research interest about covering and covering-based rough sets [35–39]. Since a partition of a universe is a family of disjoint subsets of the universe, then it is a special case of a covering. It is part of the reason why covering-based rough sets attract increasing research interest.

Definition 2 [7]. Let $\langle U, C\rangle$ be a covering approximation space, $x \in U$, then $md_C(x) = \{K \in C : x \in K \wedge \ (\forall S \in C)\ (x \in S \wedge S \subseteq K \Rightarrow K = S)\}$ is called the minimal description of x.

Given an approximation space $\langle U, C\rangle$, for any object $x \in U$, the minimal descriptor of x contains the core objects in the approximation space that are related to x, and the minimal descriptor may provide a simple and key description for x when we discuss the issue of set approximations in $\langle U, C\rangle$.

Definition 3 [7]. Let $\langle U, C\rangle$ be a covering approximation space, $x \in U$, then $MD_C(x) = \{K \in C : x \in K \wedge (\forall S \in C)(x \in S \wedge K \subseteq S \Rightarrow K = S)\}$ is called the maximal description of x.

The maximal descriptor of x contains all objects in the approximation space that are related to x, and the maximal descriptor may provide a detailed and comprehensive description for x when we discuss the issue of set approximations in $\langle U, C\rangle$.

Yao et al. [12] have pointed out that the utilization of the maximal descriptors of objects is equally reasonable as the utilization of the minimal ones in a covering approximation space.

3. Two matrix-based calculation approaches for minimal and maximal descriptions in covering-based rough sets

In this section, we first introduce the matrix-based approaches proposed by Wang and Zhang [32], and give some examples to show how to calculate the minimal and maximal descriptions by using Wang's approaches. Second, we propose two novel matrix-based approaches, called matrix-based approach-1 and matrix-based approach-2, respectively, to compute the minimal and maximal descriptions in covering-based rough sets. In other words, in this paper we study the minimal and maximal descriptions in covering-based rough sets from the view of modified matrix, where matrix-based approaches proposed in Ref. [32] are served as a basis of matrix-based approach-1 and matrix-based approach-2. We also give some examples to show how to calculate the minimal and maximal descriptions by using matrix-based approach-1 and matrix-based approach-2. In addition, we analyze the time complexities of different approaches.

3.1. Wang's approach for calculating minimal and maximal descriptions

Definition 4 [25]. Let $U = \{x_1, x_2, \ldots, x_m\}$ be a universe of discourse, $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of U, then matrix $A(C) = (a_{ij})_{n\times m}$ is called a matrix representation of C, where for any $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$, $a_{ij} = \begin{cases} 1, & x_j \in K_i, \\ 0, & x_j \notin K_i, \end{cases}$.

We can also express $A(C)$ in the form of vectors, that is, $A(C) = (a_{ij})_{n\times m} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$, where for any $1 \leqslant i \leqslant n, \alpha_i = (a_{i1}, a_{i2}, \ldots, a_{im})$.

The following example is employed to show how we obtain a matrix representation of a covering.

Example 1 [32]. Let $U = \{x_1, x_2, x_3, x_4, x_5\}$ and $C = \{K_1, K_2, K_3, K_4\}$, where $K_1 = \{x_1, x_2\}$, $K_2 = \{x_1, x_3\}, K_3 = \{x_1, x_2, x_3\}, K_4 = \{x_4, x_5\}$. From Definition 4, the matrix representation of C can be figured out as follows.

$$A(C) = \begin{array}{c} \\ K_1 \\ K_2 \\ K_3 \\ K_4 \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

Lemma 1 ([31,32]). *Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, then $A(C) \cdot A^T(C) = (b_{ij})_{n\times n}$, where $b_{ij} = |K_i \cap K_j|$.*

In fact, Lemma 1 can be used to calculate the number of conjoint elements of K_i and K_j, where $1 \leqslant i,j \leqslant n$. It can be easily proved that the result of $A(C) \cdot A^T(C)$ is a symmetric matrix.

Definition 5 [32]. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, then for any $1 \leqslant j \leqslant m$,

$$A(C)_{x_j} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, \text{ where for any } 1 \leqslant i \leqslant n, \beta_i = \begin{cases} \alpha_i, & a_{ij} \neq 0 \\ \mathbf{0}, & a_{ij} = 0 \end{cases}.$$

The function of Definition 5 is to find those K_i which contains x_j in $A(C)$, and if the K_i does not contain x_j, then we set β_i as zero vector.

Definition 6 [32]. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, then $A(\hat{C}) = A(C) \cdot \mathbf{1}_{m \times n}$.

In fact, the function of Definition 6 is to calculate how many elements in each K_i. Example 2 is employed to show how Definitions 6 and 7 work.

Example 2 [32]. (Example 1; continuation)

$$A(C)_{x_1} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A(\hat{C}) = \begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 \end{pmatrix}.$$

From Example 2, one can find that the last row of $A(C)_{x_1}$ is zero vector because the element x_1 does not belong to K_4. Moreover, one can also see that the elements in each row of $A(\hat{C})$ are the same because each row just shows the number of elements in each K_i. For instance, the third row of $A(\hat{C})$ is $(3,3,3,3)$, which indicates that there are three elements in K_3.

Definition 7 [32]. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$ and $A(C)_{x_j} \cdot A^T(C)_{x_j} = (a_{st})_{n \times n}$, then $a_{st} = |K_s \cap K_t|$ and $a_{st} > 0$ if and only if $x_j \in K_s \cap K_t$, where "$\cdot$" denotes the multiplication of any two matrices and A^T the transposition of A.

Definition 8 [32]. Let $A_1 = (a_{ij})_{n \times n}$ and $A_2 = (b_{ij})_{n \times n}$ be two matrices. We define an operation $\oplus$ as follows: $A_3 = A_1 \oplus A_2 = (d_i)_{n \times 1}$, where $d_i = \begin{cases} 1, & a_{ii} = b_{ii} \wedge (i \neq j \Rightarrow a_{ij} \neq b_{ij}) \\ 0, & otherwise \end{cases}$.

Definition 9 [32]. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, for any $C_1 \subseteq C$, we define:

(1) $f(C_1) = (y_i)_{n \times l}$, where $y_i = 1$, if $K_i \in C_1$, else $y_i = 0$.
(2) $f(md(x_i)) = A(C)_{x_i} A^T(C)_{x_i} \oplus A^T(\hat{C})$.
(3) $f(MD(x_i)) = A(C)_{x_i} \cdot A^T(C)_{x_i} \oplus A(\hat{C})$.

In fact, Definitions 7–9 characterize a method for computing minimal and maximal descriptions.

Example 3 [32]. (Example 1; continuation)

$$f(md(x_1)) = A(C)_{x_1} \cdot A^T(C)_{x_1} \oplus A^T(\hat{C})$$

$$= \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \oplus \begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 \end{pmatrix}^T = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Hence, $md(x_1) = \{K_1, K_2\}$.

$$f(MD(x_1)) = A(C)_{x_1} \cdot A^T(C)_{x_1} \oplus A(\hat{C}) = \begin{pmatrix} 2 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Hence, we have that $MD(x_1) = \{K_3\}$.

As a summary, the approach mentioned in Ref. [32] can be summed up as Algorithm 1. The total time complexity of Algorithm 1 is $O(|C||U| + (|C| + |C|^2)|U|)$, and steps 9–20 are the main part of Algorithm 1 for calculating the minimal descriptions. The time complexity of steps 9–20 is $O((|C| + |C|^2)|U|)$.

Algorithm 1. Wang's Matrix approach for computing minimal descriptions

Input: $U = \{x_1, \cdots, x_m\}$, $C = \{K_1, \cdots, K_n\}$, matrix representation $A(C) = (a_{ij})_{mn}$

Output: $f(md(x_j))$, $md(x_j)$

1 $m \leftarrow |U|$, $n \leftarrow |C|$;
2 **for** $i = 1 \rightarrow n$ **do**
3   **for** $j = 1 \rightarrow m$ **do**
4   **if** $x_j \in K$ **then**
5       $a_{ij} = 1$;
6   **else**
7       $a_{ij} = 0$;
8 compute $A^T(\hat{C})$;
9 **for** $j = 1 \rightarrow m$ **do**
10   **for** $i = 1 \rightarrow n$ **do**
11   compute $A(C)_{x_j}$;
12   compute $A^T(C)_{x_j}$;
13   compute $A(C)_{x_j} \cdot A^T(C)_{x_j}$;
14   **for** $i = 1 \rightarrow n$ **do**
15   **for** $k = 1 \rightarrow n$ **do**
16   **if** $((A(C)_{x_j} \cdot A^T(C)_{x_j})(i,i) = A^T(\hat{C})(i,i)) \wedge (i \neq k \Rightarrow (A(C)_{x_j} \cdot A^T(C)_{x_j})(i,k) \neq A^T(\hat{C})(i,k)))$ **then**
17         $f(md(x_j)(i)) \leftarrow 1$;
18         $md(x_j) \leftarrow K_i$;
19   **else**
20         $f(md(x_j)(i)) \leftarrow 0$;
21 Return $f(md(x_j))$, $md(x_j)$.

3.2. Modified approaches for calculating minimal and maximal descriptions

In this subsection, derived from Algorithm 1, we shall propose two new matrix-based approaches, called approach-1 and approach-2 respectively, to compute the minimal and maximal descriptions. Two detailed examples are employed to help better understand the proposed approaches. Moreover, we also analyze the time complexity of the proposed algorithms.

3.2.1. Approach-1 for calculating minimal and maximal descriptions

In approach-1, we replace the operation "$\oplus$" in Algorithm 1 with a "sum" operation. The sum operation does not need to compare the elements in two matrices, which can reduce the calculation complexity and improve the efficiency of Algorithm 1.

Definition 10. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of universe $U = \{x_1, x_2, \ldots, x_m\}$, and $A(C)_{x_j} \cdot A^T(C)_{x_j} = (c_{st})_{n \times n}$, then

(1) We define $(c_{st})_{n \times n} = (\chi_1, \chi_2, \ldots, \chi_n)$, where $\chi_i = \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix}$, $i = 1, 2, \ldots, n$.

(2) We define $S(C)_{x_j} = (d_i)_{n\times 1} = \chi_1 + \chi_2 + \cdots + \chi_n = \begin{pmatrix} \sum_{i=1}^{n} c_{1i} \\ \sum_{i=1}^{n} c_{2i} \\ \vdots \\ \sum_{i=1}^{n} c_{ni} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$.

(3) We define $y_i = \begin{cases} 1, & d_i = \min\{d_j | d_j > 0, j \in \{1,2,\ldots n\}\} \\ 0, & otherwise \end{cases}$. where "min" denotes the minimum value, and $\min\{d_j | d_j > 0, j \in \{1,2,\ldots n\}\}$ means that we should first find all the values which satisfy $d_j > 0$, and then select the minimum one from them.

Definition 11. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of universe U, for any $C_1 \subseteq C$, the characteristic function of C_1 is defined as $g(C_1) = (q_i)_{n\times 1}$, where $q_i = \begin{cases} 1, & K_i \in C_1 \\ 0, & otherwise \end{cases}$.

Theorem 1. *Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of universe $U = \{x_1, x_2, \ldots, x_m\}$*, and $A(C)_{x_j} \cdot A^T(C)_{x_j} = (a_{st})_{n\times n}$, *the characteristic function $g(md(x_j)) = (y_i)_{n\times 1}$*. If $y_i = 1$ *then $K_i \in md(x_j)$*, else $K_i \notin md(x_j)$.

Proof. Suppose that $x_j \in K_s$ and $x_j \in K_t$, $A(C)_{x_j} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$ then

$$A(C)_{x_j} \cdot A^T(C)_{x_j} = (a_{st})_{n\times n} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \cdot \begin{pmatrix} \beta_1^T & \beta_2^T & \cdots & \beta_n^T \end{pmatrix} = \begin{pmatrix} \beta_1 \cdot \beta_1^T & \beta_1 \cdot \beta_2^T & \cdots & \beta_1 \cdot \beta_n^T \\ \beta_2 \cdot \beta_1^T & \beta_2 \cdot \beta_2^T & \cdots & \beta_2 \cdot \beta_n^T \\ \cdots & \cdots & \cdots & \cdots \\ \beta_n \cdot \beta_1^T & \beta_n \cdot \beta_2^T & \cdots & \beta_n \cdot \beta_n^T \end{pmatrix}.$$

Therefore, $a_{st} = \beta_s \cdot \beta_t^T$, where $s, t \in \{1, 2, \ldots, n\}$. According to Proposition 2 in [20], we have that $a_{st} = \beta_s \cdot \beta_t^T > 0$. Then, on the one hand, according to Definition 10, we have that.

$d_i = a_{i1} + a_{i2} + \cdots + a_{in} = \beta_i \cdot \beta_1^T + \beta_i \cdot \beta_2^T + \cdots + \beta_i \cdot \beta_n^T$, noting that $\beta_s \cdot \beta_t^T > 0$, then $d_i > 0$, where $i \in \{1, 2, \ldots, n\}$.

On the other hand, if $d_i = \min\{d_s, d_i\}$, then $d_s - d_i > 0$, where $s, i \in \{1, 2, \cdots, n\}$.

At same time, we know that

$$\begin{aligned} d_s - d_i &= (\beta_s \cdot \beta_1^T + \beta_s \cdot \beta_2^T + \cdots + \beta_s \cdot \beta_n^T) - (\beta_i \cdot \beta_1^T + \beta_i \cdot \beta_2^T + \cdots + \beta_i \cdot \beta_n^T) \\ &= (\beta_s \cdot \beta_1^T - \beta_i \cdot \beta_1^T) + (\beta_s \cdot \beta_2^T - \beta_i \cdot \beta_2^T) + \cdots + (\beta_s \cdot \beta_n^T - \beta_i \cdot \beta_n^T). \end{aligned}$$

As we know $A(C)_{x_j} \cdot A^T(C)_{x_j}$ is a special case of $A(C) \cdot A^T(C)$, where $A(C) \cdot A^T(C) = (b_{ij})_{n\times n}$, and $A(C) = (a_{tj})_{n\times m} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, a_{tj} \neq 0$.

Therefore,

$$\begin{aligned} d_s - d_i &= (\beta_s \cdot \beta_1^T - \beta_i \cdot \beta_1^T) + (\beta_s \cdot \beta_2^T - \beta_i \cdot \beta_2^T) + \cdots + (\beta_s \cdot \beta_n^T - \beta_i \cdot \beta_n^T) \\ &= (\alpha_s \cdot \alpha_1^T - \alpha_i \cdot \alpha_1^T) + (\alpha_s \cdot \alpha_2^T - \alpha_i \cdot \alpha_2^T) + \cdots + (\alpha_s \cdot \alpha_n^T - \alpha_i \cdot \alpha_n^T) = (b_{s1} - b_{i1}) + (b_{s2} - b_{i2}) + \cdots + (b_{sn} - b_{in}). \end{aligned}$$

Then, according to Lemma 1,

$$d_s - d_i = (b_{s1} - b_{i1}) + (b_{s2} - b_{i2}) + \cdots + (b_{sn} - b_{in}) = (|K_s \cap K_1| - |K_i \cap K_1|) + \cdots + (|K_s \cap K_n| - |K_i \cap K_n|) > 0.$$

Then, we have that $K_s \supseteq K_i$, that is to say $K_i \in md(x_j)$. This completes the proof. □.

Definition 12. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of universe $U = \{x_1, x_2, \ldots, x_m\}$, and $A(C)_{x_j} \cdot A^T(C)_{x_j} = (a_{st})_{n\times n}, S(C)_{x_j} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$, We define

$$h_i = \begin{cases} 1, & d_i = \max\{d_j | d_j > 0, j \in \{1, 2, \ldots n\}\} \\ 0, & otherwise, \end{cases}$$

where "max" denotes the maximum value, and $\max\{d_j | d_j > 0,\ j \in \{1, 2, \ldots n\}\}$ means that we should first find all the values which satisfy $d_j > 0$, and then select the maximum one from them.

Theorem 2. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of universe $U = \{x_1, x_2, \ldots, x_m\}$, and $A(C)_{x_j} \cdot A^T(C)_{x_j} = (a_{st})_{n\times n}$, the characteristic function $g(MD(x_j)) = (h_i)_{n\times 1}$. If $h_i = 1$ then $K_i \in MD(x_j)$, else $K_i \notin MD(x_j)$.

Proof. We omit the proof of Theorem 2, since it can be proved in a similar way as Theorem 1.

Example 4 Example 1; continuation.

$$A(C)_{x_1} \cdot A^T(C)_{x_1} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2 & 0 \\ 1 & 2 & 2 & 0 \\ 2 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow S(C)_{x_1} = \begin{pmatrix} 2+1+2+0 \\ 1+2+2+0 \\ 2+2+3+0 \\ 0+0+0+0 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 7 \\ 0 \end{pmatrix} \Rightarrow g(md(x_1)) = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, g(MD(x_1)) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Hence, $md(x_1) = \{K_1, K_2\}, MD(x_1) = \{K_3\}$.

$$A(C)_{x_2} \cdot A^T(C)_{x_2} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow S(C)_{x_2} = \begin{pmatrix} 2+0+2+0 \\ 0+0+0+0 \\ 2+0+3+0 \\ 0+0+0+0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 5 \\ 0 \end{pmatrix} \Rightarrow g(md(x_2)) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad g(MD(x_2)) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Hence, $md(x_2) = \{K_1\}, MD(x_2) = \{K_3\}$.

$$A(C)_{x_3} \cdot A^T(C)_{x_3} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow S(C)_{x_3} = \begin{pmatrix} 0+0+0+0 \\ 0+2+2+0 \\ 0+2+3+0 \\ 0+0+0+0 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 5 \\ 0 \end{pmatrix} \Rightarrow g(md(x_3)) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad g(MD(x_3)) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Hence, $md(x_3) = \{K_2\}, MD(x_3) = \{K_3\}$.

$$A(C)_{x_4} \cdot A^T(C)_{x_4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \Rightarrow S(C)_{x_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \Rightarrow g(md(x_4)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad g(MD(x_4)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Hence, $md(x_4) = \{K_4\}, MD(x_4) = \{K_4\}$.

$$A(C)_{x_5} \cdot A^T(C)_{x_5} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \Rightarrow S(C)_{x_5} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \Rightarrow g(md(x_5)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad g(MD(x_5)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Hence, $md(x_5) = \{K_4\}, MD(x_5) = \{K_4\}$.

As a summary, the above mentioned approach can be summed up as Algorithm 2. Algorithm 2 is an improved matrix-based algorithm for computing minimal and maximal descriptions in covering-based rough sets. The total time complexity of Algorithm 2 is $O(|C||U| + (|C| + |C|)|U|)$, and steps 9–26 are the main part of Algorithm 2 for calculating the minimal descriptions. The time complexity of steps 9–26 is $O((|C| + |C|)|U|)$, which shows that Algorithm 2 is more efficient than Algorithm 1.

Algorithm 2. Approach-1 for computing minimal descriptions

Input: $U = \{x_1, \cdots, x_m\}$, $C = \{K_1, \cdots, K_n\}$, matrix representation $A(C) = (a_{ij})_{mn}$
Output: $md(x_j)$, $MD(x_j)$
1 $m \leftarrow |U|$, $n \leftarrow |C|$;
2 **for** $i = 1 \rightarrow n$ **do**
3   **for** $j = 1 \rightarrow m$ **do**
4   **if** $x_j \in K$ **then**
5             $a_{ij} = 1$;
6   **else**
7             $a_{ij} = 0$;
8 compute $A^T(\hat{C})$;
9 **for** $j = 1 \rightarrow m$ **do**
10   **for** $i = 1 \rightarrow n$ **do**
11   compute $A(C)_{x_j}$;
12   compute $A^T(C)_{x_j}$;
13   compute $A(C)_{x_j} \cdot A^T(C)_{x_j}$;
14   compute $S(C)_{x_j}$;
15   **for** $i = 1 \rightarrow n$ **do**
16   **if** $S(C)_{x_j}(i) = \min(([S(C)_{x_j})(i)]_{n\times 1} > 0)$ **then**
17             $g(md(x_j)(i)) \leftarrow 1$;
18             $md(x_j) \leftarrow K_i$;
19   **else**
20             $g(md(x_j)(i)) \leftarrow 0$;
21   **for** $i = 1 \rightarrow n$ **do**
22   **if** $S(C)_{x_j}(i) = \max(([S(C)_{x_j})(i)]_{n\times 1} > 0)$ **then**
23             $g(MD(x_j)(i)) \leftarrow 1$;
24             $MD(x_j) \leftarrow K_i$;
25   **else**
26             $g(MD(x_j)(i)) \leftarrow 0$;
27 Return $md(x_j)$, $MD(x_j)$.

In this subsection, derived from Algorithm 1, we shall propose two new matrix-based approaches, called approach-1 and approach-2 respectively, to compute the minimal and maximal descriptions. Two detailed examples are employed to help better understand the proposed approaches. Moreover, we also analyze the time complexity of the proposed algorithms.

3.2.2. Approach-2 for calculating minimal and maximal descriptions

In this subsection, we shall propose another kind of matrix-based approach to compute the minimal and maximal descriptions in covering-based rough sets.

Let $U = \{x_1, x_2, \ldots, x_m\}$ be a universe of discourse, $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of U, then matrix $A(K_s) = (a_{si})_{1\times m}$ is called a matrix representation of K_s and matrix $A(\sim K_s) = (\sim a_{si})_{1\times m}$ is called a matrix representation of $\sim K_s$, where for any $1 \leqslant s \leqslant n$ and $1 \leqslant i \leqslant m$, $a_{si} = \begin{cases} 1, & x_j \in K_s, \\ 0, & x_j \notin K_s \end{cases}$, $\sim a_{si} = \begin{cases} 0, & x_j \in K_s, \\ 1, & x_j \notin K_s \end{cases}$

Theorem 3. *Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}, A(C) = (a_{ij})_{n\times m}$ is a matrix representation of C, for any $K_s, K_t \in C(s \neq t)$, $A(K_s) \cdot A^T(\sim K_t) = 0$ if and only if $K_s \subseteq K_t$.*

Proof. For any $K_s, K_t \in C(s \neq t)$, if $A(K_s) = (a_{sj})_{1\times m} = (a_{s1}, a_{s2}, \cdots, a_{sm})$, $A(K_t) = (a_{tj})_{1\times m} = (a_{t1}, a_{t2}, \cdots, a_{tm})$, then $A(K_s) \cdot A^T(\sim K_t) = 0 \Longleftrightarrow K_s \cap \sim K_t = \varnothing \Longleftrightarrow K_s \subseteq K_t$. This completes the proof. □

Definition 13. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, $K_s \in C$, we define $\varphi_t = \begin{cases} 1, x_j \in K_t \wedge (\forall K_s \in C)(x_j \in K_s \wedge K_s \subseteq K_t \Rightarrow K_t = K_s) \\ 0, otherwise \end{cases}$.

Theorem 4. *Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, $K_s \in C$, the characteristic function $\psi(md(x_j)) = (\varphi_t)_{n\times 1}$. If $\varphi_t = 1$, then $K_t \in md(x_j)$, else $K_t \notin md(x_j)$.*

Proof. Theorem 4 can be easily proved by using Definition 2.

Definition 14. Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, $K_s \in C$, we define $\eta_s = \begin{cases} 1, & x_j \in K_t \wedge (\forall K_s \in C)(x_j \in K_s \wedge K_t \subseteq K_s \Rightarrow K_t = K_s \\ 0, & otherwise \end{cases}$.

Theorem 5. *Let $C = \{K_1, K_2, \ldots, K_n\}$ be a covering of $U = \{x_1, x_2, \ldots, x_m\}$, $K_s \in C$, the characteristic function $\zeta(MD(x_j)) = (\eta_s)_{n\times 1}$. If $\eta_s = 1$, then $K_s \in MD(x_j)$, else $K_s \notin MD(x_j)$.*

Example 5 Example 1; continuation. According to Example 1, we can obtain the following results:

$A(K_1) \cdot A^T(\sim K_2) = (1,1,0,0,0) \cdot (0,1,0,1,1)^T = 1,$
$A(K_1) \cdot A^T(\sim K_3) = (1,1,0,0,0) \cdot (0,0,0,1,1)^T = 0,$
$A(K_1) \cdot A^T(\sim K_4) = (1,1,0,0,0) \cdot (1,1,1,0,0)^T = 2,$
$A(K_2) \cdot A^T(\sim K_1) = (1,0,1,0,0) \cdot (0,0,1,1,1)^T = 1,$
$A(K_2) \cdot A^T(\sim K_3) = (1,0,1,0,0) \cdot (0,0,0,1,1)^T = 0,$
$A(K_2) \cdot A^T(\sim K_4) = (1,0,1,0,0) \cdot (1,1,1,0,0)^T = 2,$
$A(K_3) \cdot A^T(\sim K_1) = (1,1,1,0,0) \cdot (0,0,1,1,1)^T = 1,$
$A(K_3) \cdot A^T(\sim K_2) = (1,1,1,0,0) \cdot (0,1,0,1,1)^T = 1,$
$A(K_3) \cdot A^T(\sim K_4) = (1,1,1,0,0) \cdot (1,1,1,0,0)^T = 3,$
$A(K_4) \cdot A^T(\sim K_1) = (0,0,0,1,1) \cdot (0,0,1,1,1)^T = 2,$
$A(K_4) \cdot A^T(\sim K_2) = (0,0,0,1,1) \cdot (0,1,0,1,1)^T = 2,$

$A(K_4) \cdot A^T(\sim K_3) = (0,0,0,1,1) \cdot (0,0,0,1,1)^T = 2.$
From Theorem 3, we can conclude that $K_1 \subseteq K_3$ and $K_2 \subseteq K_3$.

For x_1: As x_1 is included in K_1, K_2 and K_3, we can obtain the vector $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$.

According to Theorems 4 and 5, and considering that $K_1 \subseteq K_3$ and $K_2 \subseteq K_3$, we can obtain the following results:

$$\psi(md(x_1)) = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow md(x_1) = \{K_1, K_2\}, \quad \zeta(MD(x_1)) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow MD(x_1) = \{K_3\}.$$

In a similar way, we can obtain the following results:

$$\psi(md(x_2)) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow md(x_2) = \{K_1\}, \quad \zeta(MD(x_2)) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow MD(x_2) = \{K_3\}.$$

$$\psi(md(x_3)) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow md(x_3) = \{K_2\}, \quad \zeta(MD(x_3)) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow MD(x_3) = \{K_3\}.$$

$$\psi(md(x_4)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow md(x_4) = \{K_4\}, \quad \zeta(MD(x_4)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow MD(x_4) = \{K_4\}.$$

$$\psi(md(x_5)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow md(x_5) = \{K_4\}, \quad \zeta(MD(x_5)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow MD(x_5) = \{K_4\}.$$

The method discussed above can be summarized as Algorithm 3, which is also a matrix-based method for calculating minimal and maximal descriptions. The total time complexity of Algorithm 3 is $O(|C||U| + |C| + |C|^2|U|)$. Steps 10–20 are the main part of Algorithm 3 for calculating the minimal and maximal descriptions, whose time complexity is $O(|C| + |C|^2|U|)$.

Algorithm 3. Approach-2 for computing minimal and maximal descriptions

Input: $U = \{x_1, x_2, \cdots, x_m\}$, $C = \{K_1, K_2, \cdots, K_n\}$, matrix representation $A(C) = (a_{ij})_{m\times n}$
Output: $md(x_j)$, $MD(x_j)$

$m \leftarrow |U|$, $n \leftarrow |C|$;
for $i = 1 \rightarrow n$ **do**
    for $j = 1 \rightarrow m$ **do**
        if $x_j \in K$ **then**
            $a_{ij} = 1$;
        else
            $a_{ij} = 0$;
for $i = 1 \rightarrow n$ **do**
    compute $A(K_i)$ and $A^T(\sim K_i)$;
for $i = 1 \rightarrow n$ **do**
    for $j = 1 \rightarrow m$ **do**
        for $k = 1 \rightarrow n$ **do**
            if $i \neq k \wedge A(K_i) \cdot A^T(\sim K_k) = 0$ **then**
                $\psi(md(x_j)(k)) \leftarrow 0$; $\zeta(MD(x_j)(i)) \leftarrow 0$;
            else
                $\psi(md(x_j)(k)) \leftarrow 1$; $\zeta(MD(x_j)(i)) \leftarrow 1$;
        if $\psi(md(x_j)(i)) \neq 0$ **then**
            $md(x_j) \leftarrow K_i$;
        if $\zeta(MD(x_j)(i)) \neq 0$ **then**
            $MD(x_j) \leftarrow K_i$;
Return $md(x_j)$, $MD(x_j)$.

4. Experimental analysis

In this section, in order to evaluate our algorithms, we conduct some experiments on a personal computer with 64-bit Windows10, Intel(R) Core(TM) i5-6500 CPU@3.2 GHz, and 8 GB memory. The software is MATLAB R2016b. The objective of the following experimental results is to compare the efficiency of the proposed methods and the existing methods. For the sake of clarification, the methods described in Ref. [32], approach-1 and approach-2 proposed in this paper are denoted as Matrix approach-0, Matrix approach-1 and Matrix approach-2, respectively. Since our approaches only deal with discrete attributes, we employ Rosetta software (http://www.lcb.uu.se/tools/rosetta/) to fill in some missing values and transform the numerical and continuous attributes into the discrete ones. We perform the experiments on six datasets available from the UCI machine leaning repository [33]. The characteristics of the six datasets are summarized in Table 1.

In the experiments, we gradually increase the size of data sets, and compare the time of various approaches (i.e., Matrix approach-0, Matrix approach-1 and Matrix approach-2) for calculating the minimal and maximal descriptions in covering-based rough sets. More concretely, we divide each date set T into ten subsets, which is denoted by $\{U_1, U_2, \ldots, U_{10}\}$, where U_1 contains the top 10% elements in T, U_2 contains the top 20% elements in T,..., and U_{10} is the whole data set, so the size of U_1

Table 1
Description of the datasets.

No.	Data sets	Number of objects	Number of attributes
1	Iris	150	4
2	German	1000	21
3	Statlog (Image Segmentation)	2310	19
4	Chess	3196	36
5	Bach Choral Harmony	5665	17
6	Anuran Calls (MFCCs)	7195	22

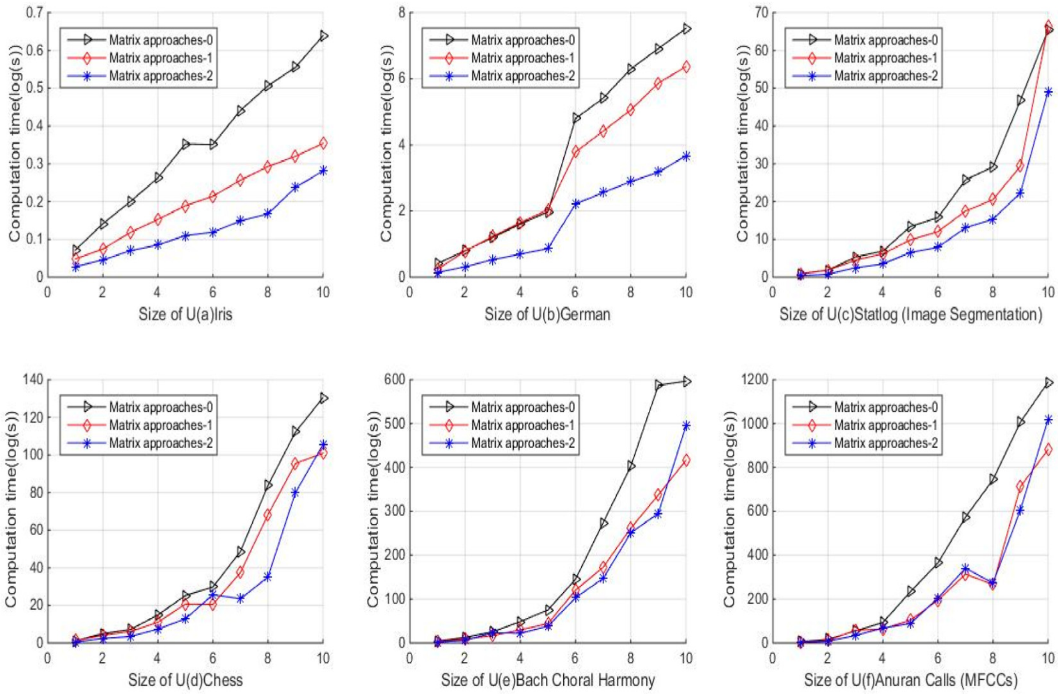

Fig. 1. Comparisons of computational time of minimal descriptions with the size increasing gradually.

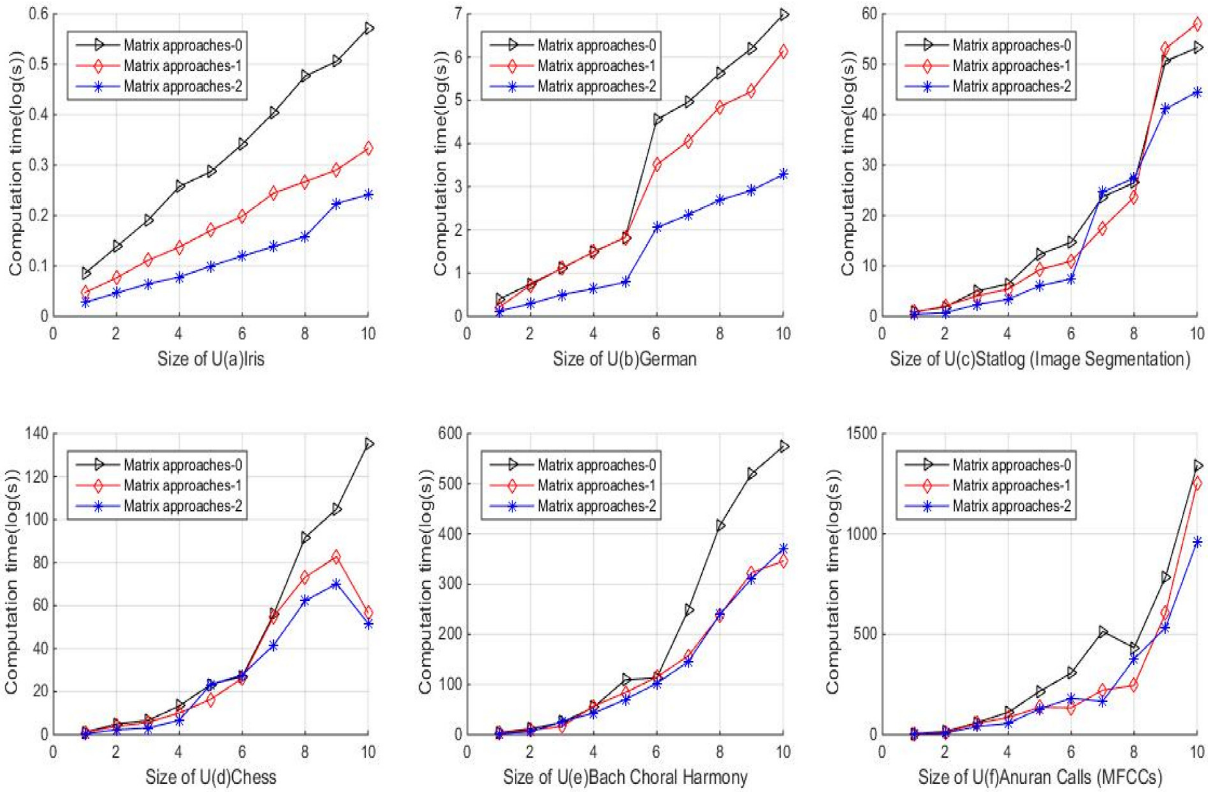

Fig. 2. Comparisons of computational time of maximal descriptions with the size increasing gradually.

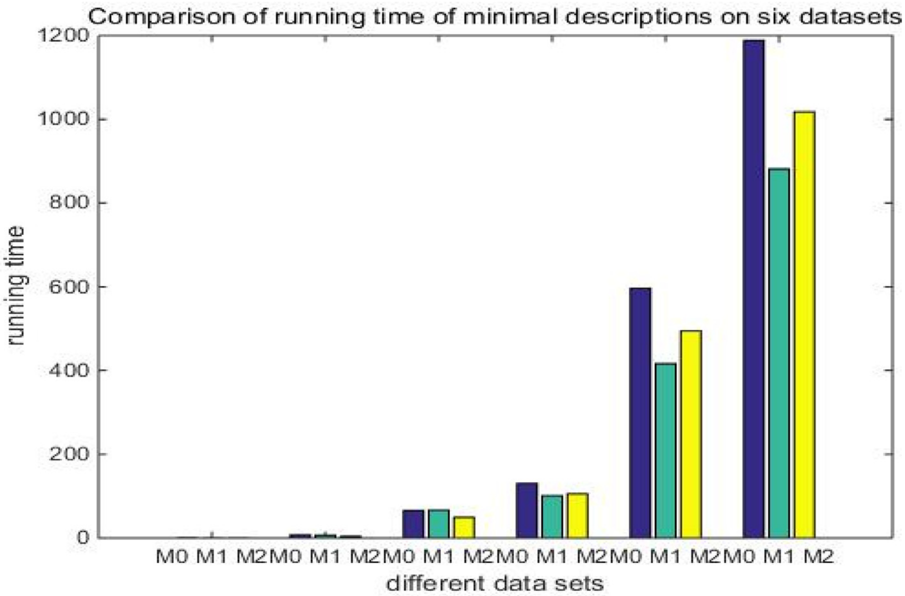

Fig. 3. Comparisons of running time of minimal descriptions on six datasets.

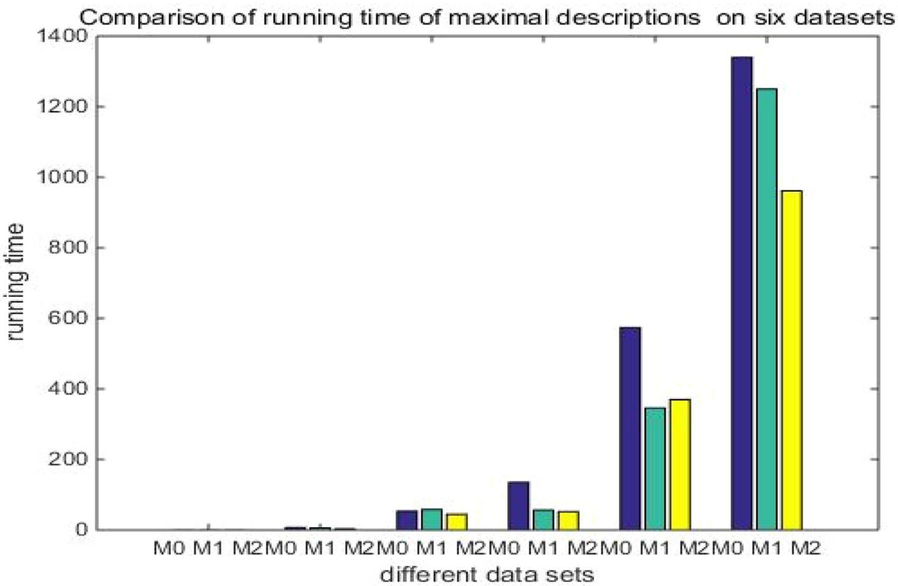

Fig. 4. Comparisons of running time of maximal descriptions on six datasets.

to U_{10} is gradually increasing. For each $1 \leqslant i \leqslant 10$, U_i is chosen as a temporary data set to compute the minimal or maximal descriptions.

Experimental results of three kinds of matrix-based approaches on various datasets are given in Figs. 1–4. According to the above experimental results, we can see that the computation time of each approach is growing when the size of data set is gradually increasing, but the computation times of Matrix approach-1 and Matrix approach-2 are less than or equal to that of Matrix approach-0. From the above results, we can also see that when dealing with a small data set which contains less than 1000 samples, Matrix approach-2 performs better than Matrix approach-0 and Matrix approach-1. However, when dealing with a larger data set which contains more than 1000 samples, both Matrix approach-1 and Matrix approach-2 are more efficient than Matrix approach-0. Therefore, we can conclude that Matrix approach-1 and Matrix approach-2 proposed in the paper are more efficient than Matrix approach-0 for computing the minimal and maximal descriptions in covering-based rough sets.

5. Conclusions

In this paper, based on some work of the minimal description and the maximal description, we mainly discuss matrix approaches to computing minimal and maximal descriptions in covering-based rough sets. Within this framework, we have proposed two new methods for computing minimal and maximal descriptions in covering-based rough set. Approach-1 is based on the operation "sum" which does not need to compare the elements in two matrix and Approach-2 is based on the binary relation of inclusion between elements in a covering. Finally, we present experimental comparisons showing the computational efficiency of the proposed methods and the experimental results show that the proposed methods are effective and efficient.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant Nos. 61663002, 61976158. Double thousand plan of Jiangxi Province of China, Qinglan Project of Jiangsu Province of China.

References

[1] W. Zakowski, Approximations in the space $(U, \P)$, Demonstratio Mathematica 16 (1983) 761–769.
[2] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341–356.
[3] J.A. Pomykala, Approximation operations in approximation space, Bulletin of Polish Academy of Sciences 35 (9–10) (1987) 653–662.
[4] Z. Bonikowski, E. Bryniarski, U.W. Skardowska, Extension and intentions in the rough set theory, Information Sciences 107 (1998) 149–167.
[5] J. Mordeson, Rough set theory applied to (fuzzy) ideal theory, Fuzzy Set Systems 121 (2001) 315–324.
[6] E. Tsang, D. Cheng, J. Lee, D. Yeung, On the upper approximations of covering generalized rough sets, in: Proceedings of the 3rd International Conference on Machine Learning and Cybernetics, 2004, pp. 4200–4203.
[7] W. Zhu, F.Y. Wang, Reduction and axiomatization of covering generalized rough sets, Information Sciences 152 (2003) 217–230.
[8] W. Zhu, Generalized rough sets based on relations, Information Sciences 177 (22) (2007) 4997–5011.
[9] P. Zhu, Covering rough sets based on neighborhoods: an approach without using neighborhoods, International Journal of Approximate Reasoning 52 (3) (2011) 461–472.
[10] P. Zhu, Q.Y. Wen, Entropy and co-entropy of a covering approximation space, International Journal of Approximate Reasoning 53 (4) (2012) 528–540.
[11] L.W. Ma, On some types of neighborhood-related covering rough sets, International Journal of Approximate Reasoning 53 (6) (2012) 901–911.
[12] Y.Y. Yao, B. Yao, Covering based rough set approximations, Information Sciences 200 (2012) 91–107.
[13] Y.Y. Yao, Information granulation and rough set approximation, International Journal of Intelligent Systems 16 (1) (2001) 87–104.
[14] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences 111 (1–4) (1998) 239–259.
[15] Jerzy W. Grzymala-Busse, Data with missing attribute values: generalization of indiscernibility relation and rule induction, in: Transactions on Rough Sets I, Springer, Berlin, Heidelberg, 2004, pp. 78–95.
[16] Jerzy W. Grzymala-Busse, Rough set strategies to data with missing attribute values, in: T.Y. Lin, S. Ohsuga, C.-J. Liau, X. Hu (Eds.), Foundations and Novel Approaches in Data Mining, Studies in Computational Intelligence, vol. 9, Springer-Verlag, 2006, pp. 197–212. .
[17] Jerzy W. Grzymala-Busse, Witold J. Grzymala-Busse, An experimental comparison of three rough set approaches to missing attribute values, in: Transactions on Rough Sets VI, Springer, Berlin, Heidelberg, 2007, pp. 31–50.
[18] Jerzy W. Grzymala-Busse, Three approaches to missing attribute values: a rough set perspective, in: Data Mining: Foundations and Practice, Springer, Berlin, Heidelberg, 2008, pp. 139–152.
[19] G.L. Liu, Y. Sai, A comparison of two types of rough sets induced by coverings, International Journal of Approximate Reasoning 50 (2009) 521–528.
[20] K. Qin, Y. Gao, Z. Pei, On covering rough sets, in: The Second International Conference on Rough Sets and Knowledge Technology (RSKT 2007), Lecture Notes in Computer Science, vol. 4481, 2007, pp. 34–41. .
[21] Y.L. Zhang, M.K. Luo, Relationships between covering-based rough sets and relation-based rough sets, Information Science 225 (2013) 55–71.
[22] M. Restrepo, C. Cornelis, J. Gomez, Partial order relation for approximation operators in covering based rough sets, Information Science 284 (2014) 44–59.
[23] M. Restrepo, C. Cornelis, J. Gomez, Duality, conjugacy and adjointness of approximation operators in covering-based rough sets, International Journal of Approximate Reasoning 55 (2014) 469–485.
[24] X.H. Zhang, J.H. Dai, Y.C. Yu, On the union and intersection operations of rough sets based on various approximation spaces, Information Sciences 292 (2015) 214–229.
[25] S.P. Wang, W. Zhu, Q.X. Zhu, F. Min, Characteristic matrix of covering and its application to boolean matrix decomposition, Information Sciences 263 (2014) 186–197.
[26] A.H. Tan, J.J. Li, Y.J. Lin, G.P. Lin, Matrix-based set approximations and reductions in covering decision information systems, International Journal of Approximate Reasoning 59 (2015) 68–80.
[27] G.M. Lang, Q.G. Li, M.J. Cai, T. Yang, Characteristic matrixes-based knowledge reduction in dynamic covering decision information systems, Knowledge-Based Systems 85 (2015) 1–26.
[28] Y.Y. Huang, T.R. Li, C. Luo, H. Fujita, S.J. Horing, Matrix-based dynamic updating rough fuzzy approximations for data mining, Knowledge-Based Systems 119 (2017) 273–283.
[29] J.Q. Wang, W. Zhu, Applications of bipartite graphs and their adjacency matrices to covering-based rough sets, Fundamenta Informaticae 156 (2) (2017) 237–254.

[30] B. Yang, B.Q. Hu, Matrix representations and interdependency on L-fuzzy covering-based approximation operators, International Journal of Approximate Reasoning 96 (2018) 57–77.
[31] Z. Lin, J. Wang, W. Zhu, Computing minimal description and maximal description in covering-based rough sets through matrices, Journal of Shandong University (Natural Science) 49 (2014) 97–101 (in Chinese).
[32] J.Q. Wang, X.H. Zhang, Matrix approaches for some issues about minimal and maximal descriptions in covering-based rough sets, International Journal of Approximate Reasoning 104 (2019) 126–143.
[33] A. Asuncion, D.J. Newman, UCI Machine Learning Repository, University of California, Department of Information and Computer Science, Irvine, CA, 2007, URL:http://www.ics.uci.edu/learn/MLRepository.html. .
[34] L. D'eer, C. Cornelis, Y.Y. Yao, A semantically sound approach to Pawlak rough sets and covering-based rough sets, International Journal of Approximate Reasoning 78 (2016) 62–72..
[35] I. Couso, D. Dubois, Rough sets, coverings and incomplete information, Fundamenta Informaticae XXI (2001) 1001–1025.
[36] W. Bartol, J. Miro, K. Pioro, F. Rossello, On the coverings by tolerance classes, Information Sciences 166 (1–4) (2004) 193–211.
[37] D. Bianucci, G. Cattaneo, D. Ciucci, Entropies and co-entropies of coverings with application to incomplete information systems, Fundamenta Informaticae 75 (1–4) (2007) 77–105.
[38] W. Zhu, Relationship between generalized rough sets based on binary relation and covering, Information Sciences 179 (3) (2009) 210–225.
[39] W. Zhu, Relationship among basic concepts in covering-based rough sets, Information Sciences 179 (14) (2009) 2478–2486.