Applied Soft Computing 45 (2016) 129–149

Contents lists available at ScienceDirect

Applied Soft Computing

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

An incremental algorithm for attribute reduction with variable precision rough sets

Degang Chen[a], Yanyan Yang[b,*], Ze Dong[c]

[a] Department of Mathematics and Physics, North China Electric Power University, Beijing, PR China
[b] School of Control and Computer Engineering, North China Electric Power University, Beijing, PR China
[c] Hebei Engineering Research Center of Simulation & Optimized Control for Power Generation, North China Electric Power University, Baoding, PR China

ARTICLE INFO

Article history:
Received 16 October 2015
Received in revised form 23 March 2016
Accepted 6 April 2016
Available online 30 April 2016

Keywords:
Variable precision rough sets
attribute reduction
incremental algorithm
dynamic datasets

ABSTRACT

Attribute reduction with variable precision rough sets (VPRS) attempts to select the most information-rich attributes from a dataset by incorporating a controlled degree of misclassification into approximations of rough sets. However, the existing attribute reduction algorithms with VPRS have no incremental mechanisms of handling dynamic datasets with increasing samples, so that they are computationally time-consuming for such datasets. Therefore, this paper presents an incremental algorithm for attribute reduction with VPRS, in order to address the time complexity of current algorithms. First, two Boolean row vectors are introduced to characterize the discernibility matrix and reduct in VPRS. Then, an incremental manner is employed to update minimal elements in the discernibility matrix at the arrival of an incremental sample. Based on this, a deep insight into the attribute reduction process is gained to reveal which attributes to be added into and/or deleted from a current reduct, and our incremental algorithm is designed by this adoption of the attribute reduction process. Finally, experimental comparisons validate the effectiveness of our proposed incremental algorithm.

1. Introduction

Rough set theory (RS) [16,17], as one generalization of set theory, has emerged as a powerful mathematical tool for modeling and tackling uncertainty, vagueness and indiscernibility in data analysis. It has amply been demonstrated that this theory has its usefulness and versatility in successfully solving a variety of problems [19,32]. The basic assumption of RS is that the rigid inclusion relation without permitting errors is taken account into approximation operators of RS formalism. More precisely, RS is used to perform the complete classification of samples belonging to a specified decision class. As a result, RS is sensitive to misclassification and noise in data. To improve the error-tolerance capability of RS, Ziarko [30,31] proposed a robust model, i.e., variable precision rough set model (VPRS), which allows the partial classification. By introducing an inclusion degree, VPRS relaxes the strict inclusion in approximation operators of RS to the majority inclusion, thereby showing certain robustness to misclassification and noise in data.

One primary application of VPRS is attribute reduction which aims to search for a compact and informative subset of condition attributes from a given dataset. As a pioneering work, Ziarko originally proposed the concept of β-reduct based on VPRS [30,31]. By keeping the dependency function, this type of reduct preserves the sum of samples in β-lower approximations of all decision classes [13]. However, it has been pointed out in Refs. [2,34,35] that the dependency function is not monotonic with the change of attributes so that the algorithm for finding a β-reduct may not be convergent. Furthermore, the derived decision rules from β-reduct may be in conflict with the ones from the original system [13]. Considering the drawback, the concepts of β-lower and upper distribution reducts based on VPRS were presented in Refs. [13] to preserve β-lower and upper approximations of each decision class, and thus the decision rules derived from β-lower and upper distribution reducts are compatible with the ones derived from the original system. In view of the superiority of β-lower and upper distribution reducts, the focus of this paper is on such type of reduct based on VPRS.

Recently, much effort has been made to develop the algorithms for finding β-lower and upper distribution reducts based on VPRS. In Ref. [13], Mi et al. proposed the discernibility matrix based procedure, by which Boolean discernibility functions can be constructed

* Corresponding author.
E-mail address: hrsinbba@163.com (Y. Yang).

http://dx.doi.org/10.1016/j.asoc.2016.04.003
1568-4946/

and all β-lower and upper distribution reducts are also obtained with the minimal disjunctive form. Although this approach provides a solid mathematical foundation for the research on β-lower and upper distribution reducts, it has to compute and examine each element in the discernibility matrix. It has been observed in Ref. [27] that other elements are always absorbed by the minimal elements in the discernibility matrix when simplifying the discernibility function, and only minimal elements in the discernibility matrix are sufficient to find β-lower and upper distribution reducts. On the basis of these facts, Yang et al. [27] developed an algorithm for finding one β-lower and upper distribution reduct based on the minimal elements, and experimentally demonstrated the effectiveness of this algorithm. Moreover, an algorithm based on the set covering was proposed in Ref. [10] to calculate a β-lower distribution reduct by converting the reduct problem to the set-covering problem. However, the previous algorithms which fall into the category of the non-incremental algorithms, are not designed to deal with dynamic datasets where data present themselves in successive samples. At the arrival of new samples, these non-incremental algorithms have to compute the whole dataset from scratch due to lacking the scheme of fully utilizing the previous data information. As a consequence, a huge amount of computational time is needed for some re-computations. Besides, these non-incremental algorithms are incapable to deal with large datasets due to the limitation of the computation capability and the memory size. One possible solution is that such datasets can be cut to bulk and added in succession [33], which can boil down to how to deal with dynamic datasets. In a word, it is always desirable to develop an efficient approach to handle dynamic datasets, so that the time complexity of the non-incremental algorithms can be enhanced.

The incremental approach, as a state-of-the-art technique of handling dynamic datasets, has been introduced into attribute reduction with rough sets [1,3,4,11,25,29]. Some incremental algorithms for attribute reduction have been proposed from the perspective of the following three variations: attribute set [20,24,33], attribute values [21,23] and samples [12,26]. With the variation of the attribute set, Zeng et al. [33] analyzed the updating mechanisms of attribute reduction and proposed incremental algorithms for feature selection with fuzzy rough sets. When new attributes are dynamically added, Wang et al. [24] developed a dimension incremental strategy for attribute reduction based on the incremental computation of three measures of information entropy. Shu et al. [20] proposed an efficient algorithm for updating attribute reduction based on the incremental computation of the positive region in incomplete decision systems when the attribute set is dynamically varied. Besides, when attribute values vary dynamically, Wang et al. [23] developed an incremental algorithm for attribute reduction based on the incremental computation of three measures of entropy. Based on the incremental computation of the positive region, Shu et al. [21] developed two incremental feature selection algorithms for single sample and multiple samples with varying attribute values.

With the variation of samples, some incremental algorithms for attribute reduction have been developed in the framework of rough sets. For example, an incremental attribute reduction algorithm was proposed in Ref. [9] to find the minimal reduct, but it is only applicable for information systems without decision attribute. For a decision table, two incremental algorithms were presented in Refs. [15,18] respectively, but experimental results in Ref. [7] show that both of them have high time complexity. To improve the efficiency of the two methods in [15,18], Hu et al. [7] presented an incremental attribute reduction algorithm based on the positive region, which was shown experimentally to be more efficient than the two algorithms in Refs. [15,18]. Based on the modified discernibility matrix, Hu et al. [8] proposed an incremental algorithm for finding all reducts when adding a new sample into the current dataset. For adding a new sample into a decision table, Yang [26] updated the discernibility matrix and proposed an incremental attribute reduction algorithm. To handle the case when the number of samples increases dynamically, Guan [6] proposed an incremental algorithm for updating all reducts based on the discernibility matrix. When adding a new sample into a current dataset, Feng et al. [5] employed the incremental computation for computing attribute core to improve the efficiency of computing one reduct. Shu et al. [22] presented an incremental attribute reduction algorithm to compute one reduct for a dynamically-increasing incomplete decision system. To allow a group of samples to be added into a current decision table, Liang et al. [12] developed an efficient group incremental algorithm for attribute reduction by introducing incremental mechanisms for three measures of information entropy including Shannon's entropy, complementary entropy and combination entropy.

To the author's best knowledge, the incremental algorithm for attribute reduction with VPRS has not yet been studied so far. The focus of our paper is on developing an incremental algorithm to find one β-upper (lower) distribution reduct based on VPRS. In the framework of VPRS, the formalisms of the discernibility matrix and reduct are characterized in the form of a set, which is inconvenient to carry out the incremental computation of attribute reduction. So this paper first introduces two Boolean row vectors to equivalently characterize β-lower and upper distribution discernibility matrices and reducts, and thus presents a minimal element based algorithm for finding one β-upper (lower) distribution reduct which is viewed as the preprocessing step of our proposed incremental algorithm in this paper. At the arrival of a new sample, we then employ the incremental process of the two Boolean row vectors, in order to incrementally find the minimal elements in the discernibility matrix of the new decision table. Based on the incremental computation of minimal elements, we gain a deep insight into the attribute reduction process that reveals which attributes to be added into and/or deleted from a current reduct. By the adoption of the attribute reduction process, our incremental algorithm is furthermore developed. Finally, experimental comparisons are performed to demonstrate the effectiveness of our proposed incremental algorithm.

The remainder of this paper is organized as follows. In Section 2, the basic concepts of VPRS are briefly reviewed. In Section 3, a minimal element based algorithm for one β-upper (lower) distribution reduct is presented as the preprocessing step of our incremental algorithm. In Section 4, the incremental algorithm for finding one β-upper (lower) distribution reduct is developed. In Section 5, experimental results are presented and analyzed. This paper is ended with conclusions.

2. Background knowledge

In this section, some basic concepts in VPRS are briefly introduced.

2.1. VPRS model

An information system is a pair (U, A) with a mapping $a(x) : U \to V_a$ for $\forall a \in A$, where $U = \{x_1, x_2, \cdots, x_n\}$ is a non-empty finite set of samples, $A = \{a_1, a_2, \cdots, a_m\}$ is a non-empty finite set of attributes, and V_a is the domain of the attribute a. With every non-empty subset $B \subseteq A$, we associate a binary relation $IND(B)$, which is called the B-indiscernibility relation and defined as $IND(B) = \{(x, y) \in U \times U : a(x) =$