Neurocomputing 188 (2016) 326–338

Contents lists available at ScienceDirect

Neurocomputing

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

Construction method of concept lattice based on improved variable precision rough set

Ruiling Zhang [a,b,*], Shengwu Xiong [a], Zhong Chen [a]

[a] School of Computer Science and Technology, Wuhan University of Technology, Wuhan, Hubei 430070, China
[b] School of Information and Technology, Luoyang Normal University, Luoyang, Henan 471022, China

ARTICLE INFO

Article history:
Received 20 March 2015
Received in revised form
25 May 2015
Accepted 28 May 2015
Available online 3 December 2015

Keywords:
Concept lattice
Variable precision rough set
Formal context
Attribute reduction
Rule acquisition

ABSTRACT

This paper mainly focuses on how to construct concept lattice effectively and efficiently based on improved variable precision rough set. On the basis of preprocessing formal concept, one algorithm that can determine the value range of variable precision parameter β according to the approximate classification quality is proposed. An improved β-upper and lower distribution attribute reduction algorithm is also proposed based on the improved variable precision rough set, the algorithm can be used for attribute reduction on the original data of the concept lattice, and to eliminate the redundant knowledge or noises of the formal context. For the reduced formal context, the paper combines the concept construction algorithm with an improved rule acquisition algorithm seamlessly, and proposes a novel approach of concept lattice construction based on improved variable precision rough set. Finally, a concept lattice generation prototype system is developed, this paper also performs comprehensive experiments, and the effectiveness of the improved algorithm is proved through the experimental results.

1. Introduction

Formal concept analysis is an effective approach for data analysis, especially for given information, it originated from the philosophical notion of concepts. The concept hierarchical constructure can be extracted from the formal context of the specific domain represented with binary relation. Concept, as the core data structure of formal concept analysis, describes the essential relation between the objects and features, and also displays the generalization and instantiation relations among the concepts. Its corresponding Hasse diagram can be used for data visualization. Rule acquisition methods based on concept lattices have been widely used in many applications, i.e., knowledge discovery. However, the existing concept lattice construction approaches have many shortcomings, such as redundant rule acquisition, sensitive to noises, high time and space complexity.

The rough set theory proposed by Z. Pawlak provides a new tool for processing fuzzy, imprecise or incomplete data. It can derive the decision or classification rules through knowledge reduction and does not sacrifice the classification ability of information system [1]. The classical rough set theory can reduce the data set effectively, however, it may cause excessive reduction, so that the necessary information may lose and the generalization ability will decrease. The knowledge reduction approach based on variable precision rough set model (VPRS) can effectively prevent the loss of essential information and increase necessary redundancy by setting the parameter β, thus the system can maintain good performance of generalization, error tolerance and noise controlling.

There have been many research works on concept lattice construction using rough set. Lai et al. studied the concept lattice in fuzzy context based on formal concept analysis and rough set theory, their research results demonstrated that the knowledge representation ability of rough set based concept lattice was weaker than that of the concept lattice based on formal concept analysis [2]; Wang et al. demonstrated the general relations between the concept extension and the equivalency classes of rough set theory, and studied the relations between the formal context reduction in concept lattice and attribute reduction with rough set theory [3]; Kang et al. brought formal concept analysis into rough set theory and proposed a rough set model based on formal concept analysis, and thus provided a new interpretation based on formal concept analysis [4]; Wei et al. studied reduction theory from the perspectives of rough set and formal concept analysis, and conducted comparative studies of the both perspectives [5]; Xiao et al. studied rough sets in the context of

* Corresponding author at: School of Computer Science and Technology, Wuhan University of Technology, Wuhan, Hubei 430070, China.
E-mail address: ruilingzhang@163.com (R. Zhang).

http://dx.doi.org/10.1016/j.neucom.2015.05.136
0925-2312/

concept lattice and provided special rough set attributes through the approach of determining the consistency of ideal lattices [6].

The previous research works brought discernibility matrix method of rough sets into concept lattices reduction [7,8], and obtained the minimal attribute sets on the premise of not changing the concept lattices structure. The reduction approaches of the rough set can be used to reduce the formal context directly [9], so that it can eliminate redundant concepts and improve the efficiency of lattice construction. The above methods integrated rough set approaches into the reduction and construction of concept lattices, but they cannot process the incomplete and noisy data efficiently. On the basis of the previous research works, this paper mainly studied the concept lattice construction method based on VPRS in order to improve the efficiency of concept lattice construction and anti-noise ability.

This paper applies VPRS β-upper and lower distribution attribute reduction algorithm to formal context reduction, and proposes a concept lattice construction algorithm based on improved VPRS. The main idea is as the follows: firstly, by taking the effect of the approximate classification quality on β into consideration on the basis of preprocessing the formal context, the paper proposed VPRS $\gamma-\beta$ algorithm for determining value range of through approximate classification quality (γ). Secondly, an improved β-upper and lower distribution attribute reduction algorithm is also proposed based on the improved variable precision rough set, the algorithm can be used for attribute reduction on the original data of the concept lattice, and to eliminate the redundant knowledge or noises of the formal context. Secondly, attributes of original concept lattice data (the formal context) can be reduced with improved VPRS β-upper and lower distribution attribute reduction algorithm to eliminate redundant knowledge and noises in the formal context. Since the improved VBPS β-upper and lower distribution attribute reduction can simplify redundant knowledge reduction in formal context significantly, construction of concept lattice in reduced formal context can effectively reduce redundant nodes in lattices and increase their anti-noise ability.

Finally, a generation system of concept lattice based on VPRS is developed to verify the proposed algorithm. Each step of concept lattice construction is displayed to users intuitively through the visual interface. And the functionality of the system is described detailedly. The structure of this paper is as follows: the related research works are described in Section 2; Section 3 introduces the constructing method of concept lattice reduction based on variable precision rough set; the experiments are performance in Section 4; Section 5 concludes the paper.

2. Related works

2.1. Rough set attribute reduction method

Attribute reduction is a very important research topic in rough set theory. In many large systems, only part of the data tables must be preserved, the rest are redundant. Excessive redundancy not only leads to waste of storage, but also causes difficulties in decision making. It can enhance the clarity of prospective knowledge in the system if the redundant attributes can be removed. Up to now, many researchers have studied the attribute reduction and related problems, and proposed a variety of attribute reduction algorithms, such as: generalized distribution table reduction [10], information entropy based reduction [11], decision logic reasoning, and the integration of discernibility matrix and logic computation [12].

Attribute reduction refers to elimination of redundant attributes effectively under the precondition of maintaining the classification ability of an information system. Finding the optimal reduction set of an information system is a proven NP-hard problem [13], and the research works about the attribute reduction methods are based on heuristic methodology. The current domestic researches in this area mainly focus on bringing intelligent computation methods into heuristic approaches [14].

The methods commonly used to deal with attribute reduction can be categorized as follows:

(1) *Attribute importance based approaches*: This kind of methods treat the attribute importance as the heuristic conditions €then select the attributes with maximum importance step by step until obtaining the attribute reduction.
(2) *Discernibility matrix based approach*: This kind of methods can discriminate the information contained by the matrix, for example, taking the attribute frequency as the heuristic condition€then get the reduced attribute one by one.
(3) *Genetic algorithm based approach*: Genetic algorithm can obtain effective reduction. There are many algorithms of this category, of which the differences are mainly in presentation and fitness function.
(4) *Extension rule reduction approach*: Starzyk, Nelson and Sturtz proposed a new concept called strong equivalence, which has been developed into a generalized law and is used to reduce discernibility function repaidly. Experimental results indicate that this algorithm is much more efficient than the base algorithms, so it can process larger data sets.
(5) *Dynamic reduction approach*: Dynamic reduction is the most stable reduction given decision table to some extent, it is the most common reduction of sub tables generated by random sampling from given decision table, and it can also effectively enhance the ability of anti-noise reduction. The calculation of dynamic reduction is very simple, the main task is to sample the decision tables, then calculate all reductions after sampling. The reductions that remain unchanged or approximately unchanged are the dynamic reductions.

The calculation of reduction generally begins from the core, the computation of the core can be obtained through the definition, or through the discernibility matrix. The completeness of attribute reduction should be deserved much attention: the reduction resulted from the calculation may not be the minimal reduction. After the reduction of the decision system attribute sets are found through attribute reduction technique, the decision rules can be obtained by using the reduction, and the left part of the rules will be a little short. However, if there are still too many attributes in the reduction, the length of the rule's left part is still too long. The shorter reduction can be obtained by calculating approximate reduction, which can be seen as the approximation of conditional attribute set, the attributes number of which is less than that of common attribute reduction algorithm.

At present, variable precision rough set model has been successfully used in many fields, such as knowledge mining [15], decision support system [16] and expert system [17]. Likewise, knowledge reduction is an important research direction of VPRS. However, the related methodologies and theories are still immature. The current reduction methods based on VPRS mainly include β reduction [18], β-upper and lower approximation reduction [19], β distribution reduction [20] and structure-retaining reduction method [21]. Variable precision threshold β is usually decided by the experts. Some researchers have proposed the selection method of β [22,23], which can reduce some difficulties in determining β caused by the lack of prior knowledge. Presently, there are not many research works in knowledge acquisition by combining the VPRS theory with concept lattice, it is worth well studying.