Information Sciences 325 (2015) 87–97

Contents lists available at ScienceDirect

Information Sciences

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

The relationship between attribute reducts in rough sets and minimal vertex covers of graphs

Jinkun Chen[a,*], Yaojin Lin[b,**], Guoping Lin[a], Jinjin Li[a], Zhouming Ma[a]

[a] School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, PR China
[b] School of Computer Science, Minnan Normal University, Zhangzhou 363000, PR China

ARTICLE INFO

Article history:
Received 15 October 2013
Revised 20 May 2015
Accepted 4 July 2015
Available online 9 July 2015

Keywords:
Attribute reduction
Vertex covers
Graph theory
Rough sets

ABSTRACT

The problems to find attribute reduction in rough sets and to obtain the minimal vertex cover for graphs are both NP-hard problems. This paper studies the relationship between the two problems. The vertex cover problem for graphs from the perspective of rough sets is first investigated. The attribute reduction of an information system is then studied in the framework of graph theory. The results in this paper show that finding the minimal vertex cover of a graph is equivalent to finding the attribute reduction of an information system induced from the graph. Conversely, the attribute reduction computation can be translated into the calculation of the minimal vertex cover of a derivative graph. Finally, a new algorithm for the vertex cover problem based on rough sets is presented. Furthermore, experiments are conducted to verify the effectiveness of the proposed method.

1. Introduction

Rough set theory, as proposed by Pawlak [35,36], has been proven to be an effective tool for managing uncertainty in information systems [38,39]. Attribute reduction plays an important role in rough set theory. In the framework of rough sets, the main aim of attribute reduction is to find a minimal subset of attributes that preserves the same classification ability as the original attributes [35]. Over the past ten years, attribute reduction has been successfully applied in many fields, such as pattern recognition [19,20,48,63,67], machine learning [8,27,47,65] and data mining [11,42,56,68].

Many approaches have been proposed to find attribute reduction in the literature [1,15,21,24,25,28,30,31,37,44–46,55,57,61,63,64,66]. A beautiful theoretical result is based on the notion of a discernibility matrix [45]. Skowron and Rauszer [45] showed that the set of all reducts is in fact the set of prime implicants of a discernibility function. A representation of the useful information on the set of reducts in simple graphical form was given in [32,33]. To do this, we need to know all the reducts of a given information system. However, it is time consuming in terms of computation time. In fact, as was shown by Wong and Ziarko [59], finding the set of all reducts or an optimal reduct (a reduct with the minimum number of attributes), is NP-hard. Therefore, heuristic methods such as positive-region methods [13,18,41], information entropy methods [26,40,46,62] and discernibility matrix methods [6,7,51–54] have been developed.

The vertex cover problem, which is a classical problem in graph theory, is that of finding a minimal vertex cover with the least number of vertices in a given graph [3]. Except the application in graph theory, the vertex cover problem has also been used in a wide variety of real-world applications such as crew scheduling [43], VLSI design [2], nurse rostering [5], and industrial machine

* Corresponding author. Tel.: +13 615090602.
** Corresponding author.
E-mail addresses: cjk99@163.com (J. Chen), zzlinyaojin@163.com (Y. Lin).

http://dx.doi.org/10.1016/j.ins.2015.07.008
0020-0255/

assignments [60]. Similar to the discernibility function method used for attribute reduction, the minimal vertex cover computation can be also translated into the calculation of prime implicants of a Boolean function [3,10,29]. Although one can generate all the minimal vertex covers of a graph by using the Boolean function method, it is a well known NP-hard optimization problem [10,22]. There are some approximation algorithms that have been proposed for solving this problem with various performance guarantees [9,12,16,17,34]. An overview of these methods was given in [14,49]. However, the best known performance bound on the approximation ratio is two [14]. It is therefore desirable to develop techniques to improve the performance bound on the approximation ratio.

As we have stated above, the attribute reduction and vertex cover problems are both NP-hard and can be obtained via the Boolean logical operation. It seems that there is some kind of natural connection between them. The purpose of this paper is to establish the relationship between the two problems. We have shown that the two problems can be transformed into each other. This study may open new research directions and provide new methods for the two problems. The rest of this paper is organized as follows. In Section 2, some basic notions about rough sets and graph theory are reviewed. In Section 3, a new information system induced from a given graph is introduced, and the relationship between attribute reduction of the derivative information system and the minimal vertex cover of the graph is established. In Section 4, we investigate attribute reduction of a given information system from the viewpoint of the vertex cover problem. In Section 5, a new approximation algorithm for the vertex cover problem based on rough sets is presented. Experiments are also given to show the effectiveness of the proposed method. Finally, some conclusions are drawn in Section 6.

2. Preliminaries

In this section, we recall some basic notions about rough sets and graph theory [3,4,35,48,50].

2.1. Attribute reduction with rough sets

The starting point of rough set theory is an information system. Formally, an information system (IS for short) can be seen as a pair $S = (U, A)$, where U and A, are finite, non-empty sets of objects and attributes, respectively. With each attribute $a \in A$, we define an information function $a: U \longrightarrow V_a$, where V_a is the set of values of a, called the domain of a.

Each non-empty subset $B \subseteq A$ determines an indiscernibility relation:

$$R_B = \{(x, y) \in U \times U | a(x) = a(y), \forall a \in B\}.$$

Obviously, R_B is an equivalence relation on U, it forms a partition $U/R_B = \{[x]_B | x \in U\}$, where $[x]_B$ denotes the equivalence class containing x w.r.t. B, i.e., $[x]_B = \{y \in U | (x, y) \in R_B\}$.

In the theory of rough sets, attribute reduction is one of key processes for knowledge discovery. Given an IS $S = (U, A)$, a reduct of S is a minimal subset of attributes $B \subseteq A$ such that $R_B = R_A$. Many approaches to attribute reduction have been proposed. For our purpose, we introduce the following method based on the discernibility matrix and logical operation [45]. By the discernibility matrix method, one can get all the reducts of an IS.

Let $S = (U, A)$ be an IS with n objects and $(x, y) \in U \times U$. We define

$$M(x, y) = \{a \in A | a(x) \neq a(y)\},$$

$M(x, y)$ is referred to as the discernibility attribute set of x and y in S, and $\mathcal{M} = \{M(x, y) | (x, y) \in U \times U\}$ is called the discernibility set of S. Note that the discernibility set can be also stored in a matrix form, which is a symmetric $n \times n$ matrix with the entry $M(x, y)$.

A discernibility function f_S for an IS is a Boolean function of m Boolean variables $a_1^*, a_2^*, \ldots, a_m^*$ corresponding to the attributes $a_1, a_2, \ldots, a_m$, respectively, and it is defined as follows:

$$f_S(a_1^*, a_2^*, \ldots, a_m^*) = \wedge\{\vee M(x, y) | M(x, y) \in \mathcal{M}, M(x, y) \neq \emptyset\},$$

where $\vee M(x, y)$ is the disjunction of all variables a^* such that $a \in M(x, y)$.

By the operations of disjunction ($\vee$) and conjunction ($\wedge$), Skowron and Rauszer [45] showed that the attribute reduction computation can be translated into the calculation of prime implicants of a Boolean function.

Lemma 1 ([45]). *Let $S = (U, A)$ be an IS. An attribute subset $B \subseteq A$ is a reduct of S iff $\bigwedge_{a_i \in B} a_i^*$ is a prime implicant of the discernibility function f_S.*

From Lemma 1, we can see that if

$$f_S(a_1^*, a_2^*, \ldots, a_m^*) = \wedge\{\vee M(x, y) | M(x, y) \in \mathcal{M}, M(x, y) \neq \emptyset\} = \bigvee_{i=1}^{t} \left(\bigwedge_{j=1}^{s_i} a_j^* \right),$$

where $\bigwedge_{j=1}^{s_i} a_j^*, i \leq t$, are all the prime implicants of the discernibility function f_S, then $B_i = \{a_j | j \leq s_i\}, i \leq t$ are all the reducts of S. Without any confusion, we will write a_i instead of a_i^* in the sequel.