International Journal of Approximate Reasoning 70 (2016) 1–12

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

A rough set approach to the characterization of transversal matroids

Guoye Xu *, Zhaohao Wang *

School of Mathematics and Computer Science, Shanxi Normal University, Shanxi, Linfen, 041000, PR China

ARTICLE INFO

Article history:
Received 3 April 2015
Received in revised form 11 November 2015
Accepted 4 December 2015
Available online 10 December 2015

Keywords:
Approximation operator
Covering
Rough set
Transversal matroid

ABSTRACT

Rough sets are efficient for data pre-processing during data mining. However, some important problems such as attribute reduction in rough sets are NP-hard and the algorithms required to solve them are mostly greedy ones. The transversal matroid is an important part of matroid theory, which provides well-established platforms for greedy algorithms. In this study, we investigate transversal matroids using the rough set approach. First, we construct a covering induced by a family of subsets and we propose the approximation operators and upper approximation number based on this covering. We present a sufficient condition under which a subset is a partial transversal, and also a necessary condition. Furthermore, we characterize the transversal matroid with the covering-based approximation operator and construct some types of circuits. Second, we explore the relationships between closure operators in transversal matroids and upper approximation operators based on the covering induced by a family of subsets. Finally, we study two types of axiomatic characterizations of the covering approximation operators based on the set theory and matroid theory, respectively. These results provide more methods for investigating the combination of transversal matroids with rough sets.

1. Introduction

Rough set theory was proposed by Pawlak [1] in 1982 for dealing with the vagueness, granularity, and uncertainty in information systems, where it classifies information through indiscernibility relations, which have been used widely in artificial intelligence, data mining, machine learning, and other fields. Covering-based rough sets [8] are a generalization of Pawlak rough sets and they have attracted much attention because of their advantages for describing and solving various problems [9,10,12,13]. However, many optimization issues related to rough sets, including attribute reduction, are NP-hard, and thus they typically require greedy algorithms [31,32]. To find effective methods for solving these problems, covering-based rough sets have been combined with other theories and methods, such as fuzzy sets [3,20], lattice theory [19], algebra [4,5], topology [6,7], and especially matroid theory [22,26–28,30]. These interesting studies have enriched covering-based rough set theory, as well as extending its applications.

Matroids [2] are a generalization of linear independence in vector spaces, which borrow extensively from linear algebra and graph theory. Due to its abundance of theories and perfect system, matroid theory has been used widely in many fields including combinatorial optimization, network flows [24], algorithm design, and especially greedy algorithm design [25]. Recently, there have been several advances in the matroidal approach to rough sets. Liu et al. [29] defined a parametric

* Corresponding authors.
E-mail addresses: xuguoye00@163.com (G. Xu), nysywzh@163.com (Z. Wang).

http://dx.doi.org/10.1016/j.ijar.2015.12.001

set family to tie the concept of rough sets to matroids, thereby obtaining significant results by combining Pawlak's rough sets and matroids. In particular, they explained why several "greedy" class algorithms work for rough sets. Wang et al. [23] proposed four matroidal structures for covering-based rough sets with matroids and concluded that they coincided with each other, as well as identifying some new properties. These results demonstrate the potential for combining covering-based rough sets with matroids. Zhu et al. [21] established a matroid through the upper approximation number and studied generalized rough sets with matroids, where they obtained some new characteristics of rough sets.

Transversal matroids [2] are based on a family of subsets and they provide a useful bridge for combining matroids and rough sets, and thus they play an important role in matroid theory. Therefore, studying rough sets in conjunction with transversal matroids may help to solve more issues [33,34]. Wang et al. [17] proposed a transversal matroidal structure for covering-based rough sets, where they established a close relationship between the coverings and normal matroids, as well as indicating the change in the relationship between coverings and transversal matroids in the matroidal structure. Lin et al. [18] constructed a matroidal structure of coverings through transversal matroids and established certain equivalent characterizations of coverings using the matroids. Furthermore, they induced other matroidal structures of coverings and studied the relationship between these two matroidal structures. Wang et al. [11] established two matroidal structures for covering-based rough sets, i.e., the transversal matroidal structure and function matroidal structure, as well as studying the relationships between them.

Recently, there have been several advances in the connection between transversal matroids and covering-based rough sets. In particular, the transversal matroidal structure of covering-based rough sets has been established [17,11]. In the present study, we aim to construct a covering-based rough set model from a transversal matroid. Transversal matroids are based on a family of subsets and covering-based rough sets are based on a covering, so we obtain a covering from a family of subsets and we characterize the transversal matroid induced by this family in terms of the rough set model based on this covering.

In this study, we construct the approximation operators and the upper approximation number based on a covering induced by a family of subsets, as well as exploring the transversal matroid using a rough sets approach. In particular, we present the sufficient condition under which a subset is a partial transversal, and also the necessary condition. We characterize the transversal matroid by using the covering approximation operators and we construct some types of circuits in the transversal matroid. These studies research on the relationships among the arbitrary family of subsets, covering-based rough sets, and transversal matroid, which has profound significance for enriching the theoretical foundations and the development of practical applications of rough sets. Moreover, we investigate the relationships between upper approximation operators based on the covering induced by a family of subsets and closure operators in transversal matroids. Our results show that two operators are identical if and only if the covering degrades into a partition. Finally, we build two types of axiomatic characterizations of the covering approximation operator, based on set theory and matroid theory, respectively.

The remainder of this paper is organized as follows. In Section 2, we review some fundamental concepts related to rough sets and matroids. In Section 3, we characterize the transversal matroids using rough sets and we construct some new types of circuits. In Section 4, we study the relationships between the closure operators of transversal matroids and the upper approximation operators. In Section 5, we build two types of axiomatic characterizations for the covering approximation operators, based on set theory and matroid theory, respectively. We give our conclusions in Section 6.

2. Preliminaries

To facilitate our discussion, we review some fundamental concepts related to rough sets and matroids in this section.

2.1. Rough set

Rough sets based on equivalence relations provide a systematic approach for data preprocessing during data mining. The approximation space and approximation operators are two important concepts.

Definition 2.1 *(Pawlak approximation space).* (See [1].) Let U be a finite nonempty set and R is a equivalence relation on U, which means that R is reflexive, symmetric, and transitive. Then, the order pair (U, R) is called the Pawlak approximation space.

Definition 2.2 *(Pawlak approximation operators).* (See [1].) Let (U, R) be the Pawlak approximation space. For any $X \subseteq U$,

$$\overline{R}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\} = \bigcup\{K \in U/R \mid K \cap X \neq \emptyset\},$$

$$\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\} = \bigcup\{K \in U/R \mid K \subseteq X\}.$$

We use U/R to denote a partition generated by the equivalent relation R on U and $[x]_R$ denotes an equivalence class in R that contains an element x in U, i.e., $[x]_R = \{y \in U \mid (x, y) \in R\}$, $U/R = \{[x]_R \mid x \in U\}$. $\overline{R}(X)$, $\underline{R}(X)$ are called the upper approximation and the lower approximation of X with respect to the Pawlak approximation space (U, R). The operators $\overline{R}$, $\underline{R}$ from 2^U to 2^U are called the Pawlak upper approximation operators and the Pawlak lower approximation operators.