Knowledge-Based Systems 105 (2016) 190–205

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Generalized multigranulation double-quantitative decision-theoretic rough set

Weihua Xu*, Yanting Guo

School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, 400054, P.R. China

ARTICLE INFO

Article history:
Received 20 January 2016
Revised 1 April 2016
Accepted 10 May 2016
Available online 11 May 2016

Keywords:
Decision-theoretic theory
Double-quantitative
Generalized multigranulation
Granular computing

ABSTRACT

The principle of the minority subordinate to the majority is the most feasible and credible when people make decisions in real world. So generalized multigranulation rough set theory is a desirable fusion method, in which upper and lower approximations are approximated by granular structures satisfying a certain level of information. However, the relationship between a equivalence class and a concept under each granular structure is very strict. Therefore, more attention are paid to fault tolerance capabilities of double-quantitative rough set theory and the feasibility of majority principle. By considering relative and absolute quantitative information between the class and concept, we propose two kinds of generalized multigranulation double-quantitative decision-theoretic rough sets(GMDq-DTRS). Firstly, we define upper and lower approximations of generalized multigranulation double-quantitative rough sets by introducing upper and lower support characteristic functions. Then, important properties of two kinds of GMDq-DTRS models are explored and corresponding decision rules are given. Moreover, internal relations between the two models under certain constraints and GMDq-DTRS and other models are explored. The definition of the approximation accuracy in GMDq-DTRS is proposed to show the advantage of GMDq-DTRS. Finally, an illustrative case is shown to elaborate the theories advantage of GMDq-DTRS which are valuable to deal with practical problems. Generalized multigranulation double-quantitative decision-theoretic rough set theory will be more feasible when making decisions in real life.

1. Introduction

Rough set theory, proposed by Pawlak in his seminal paper [21], is a new mathematical tool for processing uncertain information. Correlational studies spread across many fields [31,48], such as artificial intelligence, machine learning, neural computing, data mining, cloud computing, information security, knowledge discovery, internet of things, biological information processing and so on.

Compared with classical set theory, Pawlak's rough set theory does not require any transcendental knowledge about data, such as membership functions of fuzzy sets, or probability distribution [7,8,39]. The basic idea of rough sets is to describe a concept by the upper and lower approximate definable sets. The lower approximation consists of elements whose equivalence class is completely contained in the concept and the upper approximation is made up of elements whose equivalence class is partially contained in the concept. Without considering intersection degree, so rough sets have no fault tolerance capability. A large number of generalized models have been put forward, such as the grade rough set model(GRS) [44], the rough set model based on tolerance relation [9,37], the dominance-based rough set model [2], the fuzzy rough set model and the rough fuzzy set model [1] and so on. Particularly, many probabilistic rough set models are presented. Wong et al.[32] put forward the definition of probabilistic rough sets by the introduction of probability approximation spaces into rough sets. Pawlak et al [22] proposed a model of probabilistic approaches versus the deterministic approach. Yao et al.[46] presented the decision-theoretic rough set (DTRS) based on conditional probability and two parameters, which provides reasonable semantic interpretation for decision-making process and gives an effective approach for selecting the threshold parameters. Ziarko [50] constructed the variable precision rough set model when the sum of two parameters is equal to 1. Ślezak studied the Bayesian rough set model [28]. Herbert and Yao [4] explored the game-theoretic rough set model by combining game theory with decision making. Yao et al. [49] constructed a model of web-based medical decision support systems based on DTRS model. Liu et al. [10] proposed a multiple-category classification approach with decision-theoretic rough sets, which can effectively reduce misclassification rate. Yu et al. [45] studied a automatic method of clustering analysis with the decision-theoretic rough set theory. Jia [5,6] raised an optimization problem and attribute reduction about

* Corresponding author.
E-mail address: chxuwh@gmail.com (W. Xu).

http://dx.doi.org/10.1016/j.knosys.2016.05.021
0950-7051/

DTRS model under considering the minimization of the decision cost. Yao et al. [43] constructed a model of web-based medical decision support systems based on DTRS model. Liu et al. [11] proposed a method of policy analysis with three-way decisions. Zhao et al. [53] made an intensive study of email information filtering system by using three-way decisions.

In general, the DTRS model mainly describes approximate spaces in terms of relative quantitative information. The GRS model [15,33,44] mainly describes approximate spaces from absolute quantitative information by introducing absolute rough membership. They are two fundamental expansion models which have strong fault tolerance capabilities due to quantitative descriptions, so none can be neglected. Hence, Zhang et al. [52] made a comparative study of variable precision rough set model and graded rough set model. Greco et al. [3] presented a generalized variable precision rough set model using the absolute and relative rough membership. Combining relative and absolute quantitative information, Li and Xu [18] proposed a framework of double-quantitative decision-theoretic rough sets (Dq-DTRS) based on the Bayesian decision procedure and GRS model.

From the perspective of granular computing, either classical rough sets or double-quantitative rough sets are based on single indiscernibility relations. In many circumstances, however, a target concept needs to be described through multiple binary relations on the basis of a user's requirements or goals of problem solving. Therefore, Qian et al. [23–25] introduced multigranulation rough set theory(MGRS). Multigranulation theoretical framework has been greatly enriched, and a lot of generalized models about multigranulation have also been put forward and deeply studied. Wu and Leung [30] proposed a formal approach to granular computing with multi-scale data decision information systems. Raghavan and Tripathy [26] explored topological properties of multigranulation rough sets for the first time. Xu et al. [33–37] considered variable, fuzzy and ordered multigranulation rough set models, respectively. Liu and Miao [14] presented a multigranulation rough set method in covering contexts. Liang et al. [17] established an efficient feature selection algorithm with a multi-granulation view. She et al. [27] deeply studied explored topological structures and properties of multigranulation rough sets. Considering the principle of the minority subordinate to the majority, Xu [38] proposed the generalized multigranulation rough set model(GMGRS). In the multigranulation rough set theory, each of various binary relation determines a corresponding information granulation, which largely impacts the commonality between each of the granulations and the fusion among all granulations. Qian et al. [30] therefore introduced the idea of multigranulation into DTRS, and further proposed three kinds of the multigranulation DTRS model. And Li and Xu [19,20] studied the multigranulation DTRS in an ordered information system.

In fact, there are so many factors need to be considered in the process of making decisions, and every aspect taken into account is unpractical in terms of time, energy, money and material resources. So the whole decision process is divided into model partition. Each part makes decision according to required granulations and the comprehension evaluation is finally made based on the the principle of the minority subordinate to the majority. For example, singing contest judges come from different industries, which have their own aesthetic standards. A record company may consider from an economic point of view. Music producers pay more attention to the ability of expressing the soul of the music. Then the winner is supported by majority people after the vote. Decisions come from different granular structures, and each decision may have a deviation in terms of actual situation throughout the process. Therefore, double-quantitative decision-theoretic rough sets with strong fault tolerance capabilities are consistent with real world situations, and more attention should be paid to the theory. Meanwhile, it is necessary to introduce the idea of generalized multigranulation into decision-theoretic rough sets. Then we further emphasize comparative advantages of Dq-DTRS and GMGRS, which can be illustrated from the following aspects:

- Compared with classical decision-theoretic rough sets, Dq-DTRS [18] exhibit strong double fault tolerance capabilities in terms of both relative and absolute fault tolerance, and have further advantage of completeness.
- A generalized variable precision rough set model using the absolute and relative rough membership [3] has been used extensively in the study of measures, reasoning, applications of uncertainty and approximate spaces.
- Considering the principle of the minority subordinate to the majority, GMGRS [38] theory is a kind of information fusion strategies through single granulation rough sets.
- For some special information systems, such as multi-source information systems, distributive information systems and groups of intelligent agents, the classical decision-theoretic rough sets can not be used to data mining from these information systems, but GMGRS can.

So the motivation of this paper is to explore double-quantitative decision-theoretic rough sets theory in multiple granular structures. Then we develop a new multigranulation decision model, called generalized multigranulation double-quantitative decision-theoretic rough sets (GMDq-DTRS). In accordance with the type of the double-quantitative decision-theoretic rough sets, two kinds of generalized multigranulation double-quantitative decision-theoretic rough set models are constructed.

The rest of this paper is organized as follows. Section 2 provides a review of basic concepts of Pawlak's rough sets, decision-theoretic rough sets, double-quantitative decision-theoretic rough sets and generalized multigranulation rough sets. In Section 3, we define the lower and upper approximations of generalized multigranulation double-quantitative decision-theoretic rough sets, and discuss the basic relation among two kinds of GMDq-DTRS models under certain constraints. Meanwhile, the comparison between GMDq-DTRS and other models is made. The approximation accuracy in GMDq-DTRS is proposed to show the advantage of GMDq-DTRS. In Section 4, an illustrative case was presented to interpret the theory and advantage of GMDq-DTRS. Finally, Section 5 gets conclusions.

2. Preliminary

In this section, we provide a review of some basic concepts such as rough sets, decision-theoretic rough sets, double-quantitative decision-theoretic rough sets, generalized multigranulation rough sets.

2.1. Pawlak's rough sets

Suppose U be a non-empty finite universe and R be an equivalence relation of $U \times U$. The equivalence relation R induces a partition of U, denoted by $U/R = \{[x]_R | x \in U\}$, where $[x]_R$ represents the equivalence class of x with regard to R. Then (U, R) is the Pawlak approximation space. For an arbitrary subset X of U, the lower and upper approximations of X are defined as follows [21]:

$$\overline{R}(X) = \{x \in U | [x]_R \cap X \neq \emptyset\} = \cup\{[x]_R | [x]_R \cap X \neq \emptyset\},$$
$$\underline{R}(X) = \{x \in U | [x]_R \subseteq X\} = \cup\{[x]_R | [x]_R \subseteq X\}.$$

And $pos(X) = \underline{R}(X), neg(X) = \sim \overline{R}(X), bnd(X) = \overline{R}(X) - \underline{R}(X)$ are called the positive region, negative region, and boundary region of X, respectively. Objects definitely and not definitely contained in the set X form positive region $pos(X)$ and negative region $neg(X)$.