Information Fusion 48 (2019) 107–118

Contents lists available at ScienceDirect

Information Fusion

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

Full Length Article

Information fusion in rough set theory : An overview

Wei Wei, Jiye Liang*

Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan Shanxi 030006, China

ARTICLE INFO

Keywords:
Information fusion
Rough set
Multi-granulation
Multi-source
Multi-modality
Multi-scale

ABSTRACT

Rough set theory is an efficient tool for dealing with inexact and uncertain information. Numerous studies have focused on rough set theory and associated methodologies, and in recent decades, various models and algorithms have been proposed. To clarify the application of information fusion in rough set theory, this paper presents an overview of existing information fusion approaches and methods for multi-source, multi-modality, multi-scale, and multi-view information systems from the perspective of objects, attributes, rough approximations, attribute reduction, and decision making. We provide a survey of recent applications of these theories and methods in various fields, and identify some potential challenges that require further research.

1. Introduction

First proposed by Pawlak in 1982, rough set theory has become an efficient tool for dealing with inexact and uncertain information [80–83]. In recent decades, it has attracted the attention of numerous researchers from across the world. Rough set theory is a powerful methodology in the field of artificial intelligence, where it is applied as a component of hybrid solutions in data mining and machine learning. In terms of data analysis, the main advantage of rough set theory is that it does not require any prior or additional information about the data, such as the probability distributions, basic probability assignments, membership grades, or possibility values [5,82].

From a practical point of view, it is necessary to define the basic concepts of rough set theory to represent data. Thus, an information system, as the basic representation of information in rough set theory, was defined by Pawlak [80] as a data table containing rows labeled by objects of interest, columns labeled by attributes, and table entries corresponding to attribute values. The set of all objects is called the universe of discourse. A binary relation on the universe of discourse, known as an indiscernibility relation, can be derived from certain subsets of the set of all attributes in an information system. This relation can be employed to define the basic and most important concepts of a rough set model: the lower and upper approximations. Based on these two rough approximations, essential issues such as the uncertainty measure [52,53], attribute reduction [82,157], and decision making (rule induction) [11,12,14,82] have been introduced and subsequently developed to implement knowledge acquisition.

However, with the development of information technology and the advent of the Big Data era, data have an increasingly complex form and an ever-larger scale. Naturally, this makes it desirable to develop more effective methods of knowledge representation. Researchers have thus proposed several new information systems to handle multi-source [58], multi-scale [120], and multi-modality [27], and multi-view data, and have used rough set theory to represent complex and large-scale data. However, for the challenges introduced by these complex information systems, classical rough sets may no longer be useful or valid. To solve modern problems, information fusion, which originated in the processing of sonar signal systems by the United States in 1973 and is popular in military applications [133], was introduced and combined with rough set theory. A number of information fusion models and algorithms have been developed to deal with these complex information systems [108]. According to the work flow of data processing based on rough set (See Fig. 1), they can be classified into the following five categories:

(1) Information fusion from the perspective of objects. Objects are basic elements in an information system. The fusion of objects from different component information systems in a complex information system can change the complex one into a single one. Consequently, complex information can be directly analyzed or processed by existing models and algorithms from the rough set theory.
(2) Information fusion from the perspective of attributes. Each attribute subset (or attribute) determines a binary relation in an information system or a complex information system. If we extract a family of binary relations from a component information system of a complex information system, then certain methods can be employed to combine the family of binary relations into a single relation. Thus, the rough set problem for a complex information system is transformed into the one in a Pawlak approximation space.

* Corresponding author.
E-mail addresses: weiwei@sxu.edu.cn (W. Wei), ljy@sxu.edu.cn (J. Liang).

https://doi.org/10.1016/j.inffus.2018.08.007
Received 4 January 2018; Received in revised form 17 July 2018; Accepted 24 August 2018
Available online 25 August 2018
1566-2535/

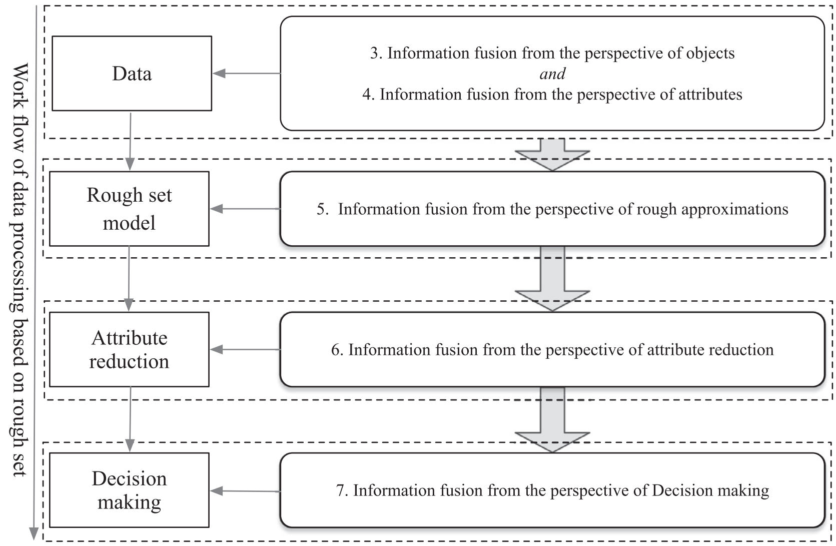

Fig. 1. The roadmap of the survey on information fusion in rough set theory.

Table 1
The types of information systems have being touched for different perspectives of information fusion in rough set theory.

No.	Type of information fusion	Information systems which has been investigated
1	From the perspective of objects	Multi-source, Multi-scale
2	From the perspective of attributes	Multi-modality
3	From the perspective of rough approximations	Multi-source, Multi-scale
4	From the perspective of attribute reduction	Large-scale, Multi-view
5	From the perspective of decision making	Multi-source

(3) Information fusion from the perspective of rough approximations. Rough approximations for a target concept are the most important steps in the process of data analyzing based on the rough set theory. The lower and upper approximations of a target concept are computed for each component information system from a complex information system. These rough approximations are then combined into a single lower approximation and a single upper approximation of the target concept, based on which the classical methods of implementing attribute reduction and extracting decision rules can be used for the complex information system.
(4) Information fusion from the perspective of attribute reduction. Attribute reduction is a necessary pre-processing step in knowledge discovery based on the rough set theory. The different reduces obtained from the component information system of a complex information system, by using the rough set model with different parameters or by the discernibility matrix, can be fused to enhance the efficiency and improve the performance of learners.
(5) Information fusion from the perspective of decision making. Decision making is an important application field of data analysis based on the rough set theory. By merging the decision rules induced from multiple decision information systems and fusing decision results from different decision makers, the performance or stability of decision making could be improved effectively.

These five categories are developed based on the work flow of data processing using rough set, all of which are the necessary aspects of rough set based data processing. They have been successfully used to analyze complex information systems (see Table 1).

In this paper, we provide a comprehensive review of recent and current research on information fusion using rough set theory, and trace the original concepts behind the development of analysis methods for complex information systems using rough set models and algorithms. The remainder of this paper is organized as follows. Section 2 reviews some preliminary details of rough sets and a variety of information systems. In Section 3, information fusion is reviewed from the perspective of objects. Section 4 presents a survey of information fusion from the perspective of attributes. In Section 5, as the richest branch of information fusion using rough set theory, fusion from the perspective of rough approximations is comprehensively analyzed. Section 6 reviews information fusion in view of attribute reducts. Section 7 reviews the existing studies of information fusion from the perspective of decision making. Section 8 concludes the paper with some final remarks.

2. Rough set basis and information systems

2.1. Rough set and classical information system

In rough set theory, a basic knowledge representation method (or information system; to distinguish it from other complex information systems, it is also called a classical information system) is the 4-tuple $IS = (U, A, V, f)$ (for short $S = (U, A)$), where U is a non-empty and finite set of objects called a universe, A is a non-empty and finite set of attributes, V_a is the domain of attribute a, $V = \bigcup_{a \in A} V_a$, and $f : U \times A = V$ is a function, $f(x, a) \in V_a$ $(a \in A)$.

For a given information system $S = (U, A, V, f)$, each attribute subset $B \subseteq A$ determines a binary indiscernibility relation $R_B = \{(x, y) \in U \times U \mid f(x, a) = f(y, a), \forall a \in B\}$, where $f(x, a)$ and $f(y, a)$ are the values of x and y with respect to the attribute a, respectively, and $f(x, B) = \cup_{a \in B}\{f(x, a)\}$. The relation R_B partitions U into some equivalence classes given by $U/R_B = \{[x]_B | x \in U\}$, where $[x]_B$ is the equivalence class determined by x with respect to B, i.e., $[x]_B = \{y \in U \mid (x, y) \in R_B\}$. Furthermore, $(\overline{B}(Y), \underline{B}(Y))$ is called a rough set of Y with respect to B for any $Y \subseteq U$, where $\underline{B}(Y) = \{x \mid [x]_B \subseteq Y\}$ and $\overline{B}(Y) = \{x \mid [x]_B \cap Y \neq \emptyset\}$ are the lower and upper approximations of Y with respect to B, respectively.

To describe a classification problem, a decision information system $DT = (U, C \cup D, V, f)$ is introduced by adding a decision attribute set D into an information system, where C is called a condition attribute set, D is called the decision attribute set, and $C \cap D = \emptyset$. Let $B \subseteq C$, $U/D = \{Y_1, Y_2, \dots, Y_n\}$, and the lower and upper approximations of the decision attribute set D be defined as $\underline{B}D = \{\underline{B}Y_1, \underline{B}Y_2, \dots, \underline{B}Y_n\}$ and $\overline{B}D = \{\overline{B}Y_1, \overline{B}Y_2, \dots, \overline{B}Y_n\}$. Let $POS_B(D) = \bigcup_{i=1}^{n} \underline{B}Y_i$, which is called the positive region of D with respect to B.

In real-world applications, data are often represented by multiple views, as they are usually collected from diverse sources or different feature extractors. These different types of attributes may characterize different specific information [156]. If an information system consists of different types of attributes, one can call it an information system with multiple views, or a multi-view information system for short. In addition, when the number of objects in an information system is very large, it is called a large-scale information system.

2.2. Multi-source information system

To describe information from different sources in the uniform representation of the information systems, Lin et al. [59] employed a group of component information systems with the same domain to represent a multi-source information system. This was defined as $MIS = \{IS_i | IS_i = (U, C_i, \{(V_a)_{a \in AT_i}, f_i\})\}$, where

(1) U is a finite non-empty set of objects;

(2) C_i is a finite non-empty set of the attributes of each component information system;

(3) $\{V_a\}$ is the value of the attribute $a \in C_i$; and

(4) $f_i : U \times C_i \to \{(V_a)_{a \in C_i}\}$ such that, for all $x \in U$ and $a \in C_i$, $f_i(x, a) \in V_a$.

In particular, in [59], a multi-source decision information system was written as $MIS = \{IS_i | IS_i = (U, C_i, \{(V_a)_{a \in C_i}, f_i\}), D, g_d\}$, where D is a finite non-empty set of decision attributes, g_d: $U \to V_d$ for any $d \in D$ is a finite non-empty set of decision attributes, and g_d: $U \to V_d$ for any $d \in D$, with V_d being the domain of a decision attribute d.

2.3. Multi-modality information systems

In some complex pattern recognition tasks, objects may be characterized in terms of multi-modality attributes, i.e., categorical, numerical, text, image, audio. To represent data with multi-modality attributes, Hu et al. [27] defined a multi-modality information system $MMIS = (U, MC)$ in which U is a finite non-empty set of objects and $MC = \{M_1, M_2, \dots, M_P\}$ is a multi-modal condition attribute set; the sets of different modality attributes $M_i \in MC$ may contain different numbers of attributes.

2.4. Multi-scale information systems

To express the hierarchical structure of data measured at different granularities, Wu and Leung [120] introduced the notion of a multi-scale information system in terms of granular computation. A multi-scale information system is a tuple $MSIS = (U, C)$, where $U = \{x_1, x_2, \dots, x_n\}$ is a non-empty, finite set of objects called the universe of discourse, $AT = \{a_l, a_2, \dots, a_t\}$ is a non-empty, finite set of attributes, and each $a_j \in C$ is a multi-scale attribute; i.e., for the same object in U, attribute a_j can take different values at different scales [120–122].

In [121], Wu et al. assumed that all attributes have I levels of granularity. Thus, a multi-scale information system can be represented as $(U, \{a_j^k : k = 1, 2, \dots, I, j = 1, 2, \dots, t\})$, where $a_j^k : U \to V_j^k$ is a surjective function and V_j^k is the domain of the kth-scale attribute a_j^k. For $1 \le k \le I - 1$, there exists a surjective function $g_j^{k,k+1} : V_j^k \to V_j^{k+1}$ such that $a_j^{k+1} = g_j^{k,k+1} \circ a_j^k$, i.e., $a_j^{k+1}(x) = g_i^{k,k+1}(a_j^k(x))$, where $g_j^{k,k+1}$ is called a granular information transformation function [120].

Moreover, $(U, \{a_j^k | k = 1, 2, \dots, m\} \cup \{d\})$ is called a multi-scale decision information system [11,12,14], where $d \notin \{a_j^k \mid k = 1, 2, \dots, I, j = 1, 2, \dots, m\}$ and d: $\{U \to V_d\}$ is a special attribute called the decision.

To represent the hierarchical structure of ordered data, a multi-scale ordered information system and a multi-scale ordered decision information system have been defined [122]. Multi-scale interval information systems [15] have also been introduced to express data using interval information.

3. Information fusion from the perspective of objects

Information fusion from the perspective of objects is an important information fusion approach using the rough set theory. The basic idea is to extract some useful information from some qualified component information systems of a complex information system and integrate this information to construct a classical information system. In what follows, all existing models and algorithms on the rough set theory can be leveraged to acquire knowledge from a complex information system. Thus far, the existing work using such an information fusion approach have mostly focused on multi-source information systems and multi-scale information systems. Some representative related work will be reviewed.

3.1. Fusion of objects in multi-source information systems

In [133], Xu et al. proposed an approach for selecting significant information sources from a multi-source information system, thus transforming a multi-source information system into a classical information system. In this method, two metrics for evaluating information sources (the internal-confidence degree ($IC(IS_i)$) and external-confidence degree ($EC(IS_i)$)) were introduced to assess the quality of each information source, where IS_i indicates ith source in a multi-source information system. To describe the quality of a source more completely, a total score for each information source $IS_i \in MIS$ was defined as: $TotalScore(IS_i) = IC(IS_i) + EC(IS_i)$. By selecting information sources in terms of the metrics proposed above, the more reliable sources (information systems) can be extracted from a multi-source information system. After completing the information source selection procedure, the selected sources were used to construct fuzzy information granules for each object. The fuzzy information granule of each object replaces its description in the multi-source information system. In fact, by using the method, the description of an object from all sources of a multi-source information system was integrated into the description of an object in a classical information system. In other words, the multi-source information system is transformed into a classical information system from the perspective of objects.

To implement information fusion from the perspective of objects for a fuzzy multi-source information system with the same structure (the same universe of discourse and attribute set), Yu et al. [152] proposed two approaches. The first was designed in terms of an operator f that characterizes the degree of consistency between the two component information systems. Using f, the information possessed by each object in all sources of a multi-source information system can be fused to generate the fuzzy description of the corresponding source in a single information system. The second approach was designed based on g, an operator that computes the point with the minimum Euclidean distance to objects in the same row of each source of a multi-source information system. This point can be used to represent all objects in that row for each source in a multi-source information system, thus converting the data to a classical information system. In addition, as real-life information sources often exhibit incomplete fuzzy phenomena, Xu et al. [134] also investigated the fusion of multiple fuzzy incomplete information sources. They leveraged the conditional information entropy [3] to assess the attributes at the same position of different sources, and the attribute with the minimum entropy was selected to describe all objects in a classical information system that consists of objects and their description on the selected attributes.

In all the methods mentioned above, the description of each object in every component information system of a multi-source information is fused to generate a unified description of the object. Then, we can obtain classical information, which consists of these objects with the new description. Based on the classical information system, all existing models and algorithms on the rough set theory can be used to acquire knowledge from a complex information system. Note that the methods can also be used to deal with multi-view information systems.

3.2. Fusion of objects in multi-scale information systems

A multi-scale information system is an attribute-value system in which each object under each attribute is represented by different scales at different levels of granulation, with a granular information transformation from finer to coarser labeled values. Because of its complexity, it is not easy to acquire knowledge from this kind of information system.

To solve this problem, Wu et al. [121] proposed an optimal scale selection method to select the best description for each object. In their method, multi-scale decision information systems are divided into consistent and inconsistent types. For a consistent decision information system with l levels, k $(1 \leq k \leq l)$ is considered to be the optimal scale if and only if k is the maximal number for which the corresponding level (S_k) is a consistent decision information system. It is easy to determine that the optimal scale is precisely the best scale for decision making in the multi-scale decision information system. For an inconsistent decision information system with l levels, the optimal scales are defined using the lower approximation, upper approximation, distribution, maximum distribution, belief distribution, plausibility distribution, and generalized decision. By selecting the optimal scale of a multi-scale decision information system, the multi-scale decision information system is transformed into a classical decision information system. After the fusion process, the existing models and algorithms from rough set theory can be employed to deal with the multi-scale decision information system. Furthermore, for incomplete multi-scale decision information systems, Wu et al. [123] also investigated the issue on selection of scale. In addition, by utilizing sequential three-way decisions [147,148], Hao et al. [21] investigated the optimal scale selection problem for a dynamic multi-scale decision information system in which the number of objects increases over time.

Unfortunately, the approaches described above can only be used to analyze multi-scale decision information systems in which each attribute is granulated across the same number of levels, which presents a restriction in certain applications. To solve this problem, Li et al. [42] introduced a multi-scale decision information system for diverse attributes with different levels of scales, and proposed a complement model and lattice model to select the optimal scale for the type of information system. Although the lattice model can successfully determine the optimal scale from all combinations, it is very time-consuming. Thus, a stepwise optimal scale selection method was designed to efficiently obtain one optimal scale combination [43]. However, in the literature, the optimal scale is defined with respect to all objects in the universe; that is, all objects have a common optimal scale. This kind of optimal scale may affect the qualities of the decision rules. Consequently, She et al. [100] introduced an approach in which the optimal scale is defined with respect to a single object. That approach instinctively creates a classical information system consisting of all objects with different optimal scales.

In all, for a multi-scale information system, by using its optimal scale to represent it, we can implement knowledge discovery and rule extraction from a multi-scale information system by analyzing the optimal scale information system. Essentially, the optimal description of an object from its description in a multi-scale information system is selected to describe the object. Consequently, optimal scale selection methods can be regarded as information fusion from the perspective of objects.

4. Information fusion from the perspective of attributes

In the rough set theory, by employing different methods to extract information from different attributes, one can transform the attributes into a set of binary relations (such as equivalence relation [80,82,83], fuzzy relation [25,75,118,151], neighborhood relation [24,60], dominance relation [16,18,20] and so on), respectively. For a complex information system that implicitly or explicitly possesses a family of binary relations, such methods can extract a family of binary relations from it. Based on the binary relation, it is easy to combine them into a binary relation by which the existing rough set models and approaches can be used to analyze the complex information systems. In this section, the methods for fusing a number of attributes into a binary relation will be systemically reviewed.

4.1. Fusion of equivalence relations derived from attributes

In [150], Yao and She proposed two approaches for fusing many equivalence relations by combining them in terms of set intersections and unions. Let $(U, \mathbb{E})$ be a multi-granulation space, where $\mathbb{E}$ is a family of equivalence relations. Given a subset of equivalence relations $\mathbb{P} \subseteq \mathbb{E}$, one way to construct rough sets in (U, P) is to construct an equivalence relation by combining equivalence relations in $\mathbb{P}$. Based on this idea, two methods of combining a family of equivalence relations were developed: the intersection of relations and the transitive closure of the union of relations. These can be described as follows.

Set intersection was employed to combine equivalence relations in $\mathbb{P} \subseteq \mathbb{E}$ into a single relation P. Note that the intersection of a family of equivalence relations is still an equivalence relation. Thus, the multi-granulation space $(U, \mathbb{P})$ is transformed into a Pawlak approximation space (U, P). The rough set approximations are defined based on the Pawlak approximation space.

The union of equivalence relations was also leveraged to combine equivalence relations. However, as the union of any two equivalence relations is not necessarily an equivalence relation, the union of relations cannot be directly used to define rough approximations. Thus, Yao et al. [150] converted the union of two equivalence relations into an equivalence relation by computing the transitive closure of the union. In this way, a family of equivalence relations in $\mathbb{P}$ can be combined into a single equivalence relation. The corresponding rough set approximations are naturally defined.

Research on the approximations given by combining a set of equivalence relations can be traced back to Pawlak's pioneering work [80] in which a multi-granulation space $(U, \mathbb{E})$ was called a knowledge base, and this was transformed into an approximation space $(U, \cup\mathbb{E})$ using the set intersection.

Although the authors did not explicitly point out that the methods proposed in [150] are suitable for analyzing a complex information system, it is easy to determine if a family of equivalence relations that are derived from a multi-source information system are combined into a single equivalence class by the method presented in [150]. Then the multi-source information system can be dealt with by the existing models and algorithms on the rough set theory.

4.2. Fusion of fuzzy similarity relations derived from attributes

Fuzzy rough sets are powerful models for handling information systems characterized by fuzzy similarity relations [22]. By employing different kernels to extract information from different attributes, one can transform multi-modal attributes into a set of fuzzy similarity relations [26,27]. If these fuzzy similarity relations can be organically merged into a single fuzzy relation, then it is natural to think of using fuzzy rough sets to process a multi-modality information system. In fact, there is an existing tool for handling multi-modality attributes in a multi-modality information system: multi-kernel learning. This tool enables information to be extracted from certain types of attributes in terms of particular kernel functions [97].

Following this idea, Hu et al. [27] utilized a match kernel [77] to extract equivalence relations from categorical data, and leveraged a Gaussian kernel [97], histogram intersection kernel [71], cosine kernel [109], and Cauchy kernel [70] to obtain fuzzy similarity relations from numerical data, image data, text data, and audio data, respectively. They then combined the fuzzy similarity relations in terms of the T-norm. Four combinations based on fuzzy T-norm operations [27] were defined as:

1. Min T-norm: $K_{T_m} = \min(k_i(x, y), k_j(x, y))$;
2. Product T-norm: $K_{T_P}(x, y) = k_i(x, y) \times k_j(x, y)$;
3. Lukasiewicz T-norm: $K_{T_l}(x, y) = \max(k_i(x, y) + k_j(x, y) - 1, 0)$;
4. T_{cos}-norm: $K_{T_{cos}}(x, y) = \max(k_i(x, y) \times k_j(x, y) - \sqrt{1 - k_i(x, y)^2}\sqrt{1 - k_j(x, y)^2}, 0)$, where k_i and k_j represent two fuzzy similarity relations computed with modalities M_i and M_j, respectively.

By fusing these fuzzy similarity relations, Hu et al. further constructed a multi-kernel fuzzy rough sets model and designed an efficient attribute reduction algorithm for a multi-modality information system. The multi-modality information system can be dealt with by the classifier constructed based on the results of attribute reduction.

Note that we can also extract a family of fuzzy relations from a fuzzy multi-source information system and obtain a single fuzzy information system by using T-norm operations to combine the fuzzy relations. Thus, the methods for analyzing a classical fuzzy information system can be used to deal with a fuzzy multi-source information system.

5. Information fusion from the perspective of rough approximations

Indisputably, rough approximation approaches are one of the two cornerstones of rough set theory. The fusion of the rough approximations plays an important role in the analysis of complex information systems [45,47,48,89]. In these approaches, lower and upper approximations of a certain target concept are computed for each component of the system, and these are fused to produce united lower and upper approximations of the target concept. To fuse the information, Qian and Liang [84,85] pioneered the concept of multi-granulation rough sets. Following their work, a series of research results (see Table 2) have been published over the past decade, and multi-granulation rough sets have gradually become a distinct field of rough set theory. In this section, these related studies are comprehensively surveyed in view of information fusion.

5.1. Classic multi-granulation rough sets

To extend Pawlak's rough set model, Qian and Liang [84,85,89,91] proposed multi-granulation rough sets in which the many partitions induced by different attribute sets are used to construct lower and upper approximations of a target concept. These rough approximations are then combined into a single lower approximation and a single upper approximation using certain rules. Qian and Liang used two effective fusion rules, namely disjunctive and conjunctive combination rules. The multi-granulation rough approximations based on disjunction are defined as follows.

Definition 5.1. [84,89] Let $S = (U, C)$ be an information system, $X \subseteq U$, and $\widehat{P_1}, \widehat{P_2}, \dots, \widehat{P_m}$ be the m partitions induced by attributes $P_1, P_2, \dots, P_m$, respectively. The lower approximation and the upper approximation of X with respect to $P_1, P_2, \dots, P_m$ are defined as:

$$\left(\underline{\sum_{i=1}^{m} \widehat{P}_i(X)}\right)^O = \{x : \bigvee \widehat{P}_i(x) \subseteq X, i \leq m\}, \ and$$

$$\left(\overline{\sum_{i=1}^{m} \widehat{P}_i(X)}\right)^O = \sim \left(\underline{\sum_{i=1}^{m} \widehat{P}_i(\sim X)}\right)^O$$

Note that the optimistic lower (or upper) approximation given by Definition 5.1 is the same as the weak lower (or strong upper) approximation presented in [37–39].

The multi-granulation rough approximations based on conjunction are defined as:

Definition 5.2. [85,91] Let $S = (U, C)$ be an information system, $X \subseteq U$, and $\widehat{P_1}, \widehat{P_2}, \dots, \widehat{P_m}$ be the m partitions induced by attributes $P_1, P_2, \dots, P_m$, respectively. The lower approximation and the upper approximation of X related to $P_1, P_2, \dots, P_m$ are defined as:

$$\left(\underline{\sum_{i=1}^{m} \widehat{P}_i(X)}\right)^P = \{x : \bigwedge \widehat{P}_i(x) \subseteq X, i \leq m\}, \ and$$

$$\left(\overline{\sum_{i=1}^{m} \widehat{P}_i(X)}\right)^P = \sim \left(\underline{\sum_{i=1}^{m} \widehat{P}_i(\sim X)}\right)^P$$

Note that multi-granulation rough sets were originally proposed to process an information system with multiple granular structures and an information system with multiple views (a multi-view information system). These structures may be derived from multiple binary relations (e.g., equivalence relation, tolerance relation, and neighborhood relation) on the universe according to user requirements or the specific problem being solved. However, although originally introduced for classical information systems, the multi-granulation rough sets can also be used to analyze multi-source information systems and multi-scale information systems. There were some related works that have been presented below.

For a multi-source information system consisting of q single-source information systems, if a binary relation is employed to granulate each single-source information system, we obtain q granular structures $K_1, K_2, \dots, K_q$. A multi-source information system can then deduce a multi-granulation approximation space, denoted as $F = (U, K_1, K_2, \dots, K_q)$. In terms of granular computing, a partition $\widehat{P}_i$ can be regarded as a granular structure, represented by K_i [87]. It is natural to construct a granular structure that corresponds to a partition using a binary indiscernibility relation R_B [86]. Thus, one can use a multi-granulation rough set model to combine the lower approximations and upper approximations of all sources into one lower approximation and one upper approximation, respectively. The knowledge acquisition approaches for classical rough sets can then be directly employed to process the multi-source information system.

For a multi-scale information system, by considering each scale of a multi-scale information system as a single information system, Gu et al. [13] proposed optimistic and pessimistic multi-granulation rough approximations for a multi-scale information system. However, this approach can only fuse rough approximations derived from each single scale. In other words, the approach fails to fuse information from different scales. Consequently, the fusion of different scale of a multi-scale information is still a question worthy of further discussion.

In addition, it should be noted that the multi-granulation rough set model can also be used for a multi-modality information system for knowledge discovery if one can obtain a partition of universe of discourse from each modality.

5.2. Generalized multi-granulation rough approximations

The classical multi-granulation rough set is established from two qualitative combination rules generated from pessimistic and optimistic viewpoints, respectively. The two rules lack a certain practicality, because one is too restrictive and the other is too relaxed. To overcome this disadvantage, Xu et al. [127,135] proposed a generalized multi-granulation rough set model. Later, Xu et al. [132] presented two generalized multi-granulation double-quantitative decision-theoretic rough sets in terms of the relative and absolute quantitative information between the class and the concept.

Table 2
Information fusion from the perspective of rough approximations.

Types of fusion	Rough approximation	References
Classical	Optimistic multi-granulation approximations	Qian and Liang [84,89]
	Pessimistic multi-granulation approximations	Qian and Liang [85,91]
Generalized	Generalized multi-granulation rough approximations	Xu [127,135]
	Generalized multi-granulation double-quantitative decision-theoretic rough approximations	Xu [132]
Crisp binary relations	Multi-granulation rough approximations based on tolerance relations	Qian [85,88]
	Pessimistic multi-granulation rough approximations based on tolerance relations	Xu [130]
	Multi-granulation rough approximations based on tolerance, similarity, and limited tolerance relations	Yang [139]
	Numerical characterization of multi-granulation rough approximations based on tolerance relations	Tan [105]
Covering	Optimistic multi-granulation covering rough approximations	Liu [65]
	Pessimistic multi-granulation covering rough approximations	Lin [55]
	Multi-granulation covering fuzzy rough approximations	Liu [66]
	Four types of multi-granulation covering rough approximations	Liu [68]
	Three types of multi-granulation covering rough approximations	Lin [57]
Fuzzy	Optimistic and pessimistic multi-granulation fuzzy rough approximations based on fuzzy tolerance relations	Xu [128]
	Multi-granulation fuzzy rough approximations multi-granulation fuzzy rough sets based on *T*-similarity relations	Yang [137]
	Variable precision multi-granulation fuzzy rough approximations	Feng [7,8]
	Intuitionistic fuzzy multi-granulation rough approximations	Huang [29]
	Multi-granulation fuzzy rough set theory over two universes	Sun [104]
	Interval-valued hesitant fuzzy multi-granulation rough approximations over two universes	Zhang [155]
	Multi-granulation rough approximations of a fuzzy set	Xu et al. [131]
Ordered information systems	Optimistic and pessimistic multi-granulation rough set models based on dominance relation	Xu [129]
	Multi-granulation fuzzy rough set approximations based on fuzzy preference relation	Pan [79]
	Optimistic, pessimistic, and mean multi-granulation graded rough approximations based on dominance relations	Yu [153]
Decision-theory rough sets	Optimistic, pessimistic and mean multi-granulation decision-theoretic rough approximations	Qian [92]
	Optimistic and pessimistic fuzzy multi-granulation decision-theoretic rough approximations	Lin [59]
	Generalized multi-granulation double-quantitative decision-theoretic rough approximations	Xu [132]
	Multi-granulation decision-theoretic rough approximations based on dominance relations	Li [50]
	Local multi-granulation decision-theoretic rough approximations	Qian [93]
Variable precision	Optimistic variable precision multi-granulation rough sets	Wei [116]
	Pessimistic variable precision multi-granulation rough sets	Dou [6] and Ju [35]
	Variable precision multi-granulation fuzzy rough approximations	Feng [7,8]
Evidence theory	Rough approximations of fusion by combining evidence theory with multi-granulation rough sets	Lin [58]
	The belief structure of multi-granulation rough sets	Tan [106]
	Numerical characterization of multi-granulation rough approximations based on tolerance relations	Tan [105]
Dynamic data	Updating multi-granulation rough approximations with increasing of granular structures	Yang [141]
	Updating of multi-granulation fuzzy rough approximations	Ju [34]
	Updating multi-granulation rough approximations with refining or coarsening attribute values	Hu [28]
Other types	Neighborhood-based multi-granulation rough approximations	Lin [56]
	Cost-sensitive multi-granulation rough approximations	Yang [140] and Ju [36]
	Non-dual and hybrid multi-granulation rough approximations	Zhang [158]
	Five-valued semantics for multi-granulation rough approximations	She [101]

5.3. Multi-granulation rough approximations based on generalized crisp binary relations

To extend the multi-granulation rough approximation to some information systems with some generalized crisp binary relations, some studies have been conducted in the last decade. Qian et al. [88] proposed a multi-granulation rough set model using a tolerance relation. However, only optimistic multi-granulation rough set models were introduced in that study. To enrich Qian's work, Xu et al. [130] introduced a pessimistic multi-granulation rough set model based on the tolerance relation. Furthermore, Yang et al. [139] leveraged the tolerance relation [41], similarity relation [103], and limited tolerance relations [113] to construct optimistic and pessimistic multi-granulation rough sets. In addition, Tan et al. [105] borrowed the belief and plausibility functions from evidence theory [98] to characterize rough approximations and attribute reductions in multi-granulation rough set theory for incomplete information systems.

5.4. Multi-granulation rough approximations based on covering

Rough sets based on covering are important extensions of rough set models, particularly for handling more complex practical problems [149,161,162]. In terms of the combination of multi-granulation and covering rough sets, Liu [65] introduced the concepts of optimistic multi-granulation covering rough sets. Around the same time, Lin et al. [55] proposed covering-based pessimistic multi-granulation rough sets. Multi-granulation covering fuzzy rough sets in a covering approximation space were also developed [66], which allowed Liu et al. [68] to construct four types of rough set models in which the target concept was approximated by employing the maximal or minimal descriptors of objects in a given universe of discourse. Using the three covering approximation operators in [161,162], Lin et al. [57] constructed three multi-granulation covering rough sets.

5.5. Multi-granulation rough approximations in fuzzy environment

Dubois and Prade proposed the concepts of rough fuzzy sets and fuzzy rough sets. Their definitions were based on approximations of fuzzy sets by crisp approximation spaces and crisp sets by fuzzy approximation spaces, respectively [4]. Other research on fuzzy rough sets and rough fuzzy sets has been reported [1,22,72,78,95,110,111,115,118,143]. Many researchers have extended multi-granulation rough sets to fuzzy environments in recent years. Xu et al. [128] introduced optimistic and pessimistic multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space. Similarly, multi-granulation fuzzy rough sets based on *T*-similarity relations were defined by Yang et al. [137]. Following this work, Feng and Mi [7] proposed two variable-precision multi-granulation fuzzy rough approximations, and uncertainty measures based on these two rough approximations have been investigated [8]. To make multi-granulation fuzzy rough set models suitable for different environments, Huang et al. [29] developed an intuitionistic fuzzy multi-granulation rough set by combining multi-granulation rough sets with intuitionistic fuzzy rough sets; Sun et al. [104] established a multi-granulation fuzzy rough set theory over two universes, while Zhang et al. [155] proposed a new rough set model combining interval-valued hesitant fuzzy sets with multi-granulation rough sets over two universes. In addition, Xu et al. [131] presented rough approximations of a fuzzy set based on multiple classical equivalence relations. The relationships between multi-granulation and classical T-fuzzy rough sets was investigated in [49].

5.6. Multi-granulation rough approximations for ordered information systems

Preference relations are very useful in expressing decision makers' requirements in ordinal decision problems. Based on preference relations, Greco et al. [17–19] proposed a novel rough set model for an ordered information system and constructed a corresponding dominance relation. For an ordered information system, Xu et al. [129] introduced two new multi-granulation rough set models in which a target concept is approximated from different views using the dominant classes induced by multiple granulations. Moreover, Pan et al. [79] introduced an additive consistent multi-granulation fuzzy preference relation rough set model to solve a multi-criteria preference analysis problem. Following the idea of grades in [67], Yu et al. [153] constructed optimistic, pessimistic, and mean multi-granulation graded rough approximations based on dominance relations.

5.7. Multi-granulation decision-theoretic rough approximations

Rough set models based on Bayesian decision theory are an important direction in rough set theory [96,142,147,148]. To combine the advantages of both multi-granulation rough sets and decision-theoretic rough sets, Qian et al. [92] developed a multi-granulation decision-theoretic rough set; this was extended to fuzzy environments by constructing a fuzzy multi-granulation decision-theoretic rough set model [59]. By considering relative and absolute quantitative information between classes and concepts, Xu et al. [132] proposed two generalized multi-granulation double-quantitative decision-theoretic rough sets. A multi-granulation decision-theoretic rough set model was extended to process ordered information systems by Li et al. [50]. Recently, Qian et al. [93] combined local rough sets with multi-granulation decision-theoretic rough sets to construct local multi-granulation decision-theoretic rough sets as a semi-unsupervised learning method, which is helpful when processing data with a large number of unlabeled objects.

5.8. Variable-precision multi-granulation rough approximations

As some information systems contain uncertain information and noisy attributes, Ziarko [154] introduced a variable-precision rough set model. This is an important method for handling uncertain information and eliminating noisy attributes, and reduces the likelihood of incorrect decisions in rough set theory. To enhance the capability of multi-granulation rough set models dealing with noisy data, Wei et al. [116] presented a variable-precision multi-granulation rough set approach by loosening the accuracy requirement on each granulation. Dou et al. [6] and Ju et al. [35] also investigated variable-precision multi-granulation rough sets and introduced optimistic and pessimistic models. Later, Feng et al. [7] proposed a variable-precision multi-granulation fuzzy rough set and analyzed the decision theory underlying this model. Feng et al. [8] then studied the uncertainty and reduction of variable-precision multi-granulation fuzzy rough sets based on three-way decisions.

5.9. Fusion of rough approximations based on evidence theory

Dempster–Shafer theory [98] is an effective reasoning approach in the presence of uncertainty. It provides a simple and effective method for combining evidence from different sources and determining the degree of belief. In recent decades, Dempster–Shafer theory has become an important method in the study of information fusion. Some pioneering research has combined rough set theory with evidence theory [2,117,119,126,144]. On the basis of these studies, Lin et al. [58] further investigated the relationship between Dempster–Shafer theory and multi-granulation rough sets [84,85,89,91], and found that the former and latter can be interpreted as quantitative and qualitative representation rules, respectively. A two-stage fusion approach involving evidence theory and multi-granulation rough sets was designed to utilize their advantages. However, Lin et al. did not construct a belief structure for the multi-granulation rough sets [58], and so their numerical characteristics were not revealed. To solve this problem, Tan et al. [105,106] investigated multi-granulation rough sets by taking belief and plausibility functions from evidence theory. The rough approximations and attribute reductions of the multi-granulation rough sets were numerically characterized by these functions.

5.10. Multi-granulation rough approximations on dynamic data

A key factor in the success of rough set theory is the ability to dynamically update the rough approximations, because many real datasets vary dynamically. Yang et al. [141] explored how to update multi-granulation rough approximations when the multi-granular structure evolves over time, i.e., when new granular structures are merged into the original multi-granular structure. Their work was later extended to the update of multi-granulation fuzzy rough approximations [34]. Motivated by the requirements of dynamic knowledge acquisition in information systems, Hu et al. [28] presented the dynamic mechanisms for updating approximations in multi-granulation rough sets under refined or coarsened attribute values. [33]

5.11. Other types of multi-granulation rough approximations

Neighborhood-based rough sets have been widely investigated [24,60–62,146], and represent a significant strand of rough set theory. To extend the theory to multi-granulation environments, Li et al. [56] proposed two neighborhood multi-granulation rough sets according to different representations of neighborhood information granules.

In classical multi-granulation rough sets, the problem of the test cost [73,74] has not been explicitly considered. To solve this problem, a number of test-cost-sensitive multi-granulation rough set models were constructed by Yang [140] and Ju [36], and these were demonstrated to be a generalization of optimistic, pessimistic, and β-multi-granulation rough sets.

By investigating the basic algebraic operations of rough approximations pairs based on multiple approximation spaces, Zhang et al. [158] derived four constructive methods for rough approximation operators, and proposed non-dual multi-granulation rough sets and hybrid multi-granulation rough sets.

Additionally, She et al. [101] investigated multi-granulation rough set theory from the viewpoint of multiple-valued logic. They extended existing work on rough sets in terms of a three-valued logic framework to multi-granulation environments, and constructed a five-valued semantics for a multi-granulation rough set model.

5.12. Structure analysis of multi-granulation rough sets

A number of hierarchical structures [52,53,87,90,107,145] have been proposed to analyze the finer or coarser relationships between two granulation spaces, but these structures can only be used single granulation spaces. To compare two multi-granulation spaces, Yang et al. [138] presented three different hierarchical structures for two multi-granulation spaces. The properties of these hierarchical structures were analyzed, and the relationships between them and multi-granulation rough sets were investigated.

Moreover, the topological or lattice-theoretic approach to rough set theory has also been extensively investigated over the past three decades [40,69,94]. Nevertheless, these studies mainly focus on the single-granulation rough set model. With the development of multi-granulation rough sets, exploring the mathematical structure of definable sets in the context of multiple granularities is becoming an urgent issue. As a representative work in this direction, She et al. [99] dissected the topological structure and lattice structure of three different kinds of definable sets in the optimistic and pessimistic multi-granulation rough set models, respectively.

6. Information fusion from the perspective of attribute reduction

Attribute reduction is a challenging issue in rough set theory. Improving the effectiveness and efficiency of attribute reduction algorithms has been a persistent goal for researchers. In recent years, a number of algorithms have been developed based on information fusion. Some of these are intended for large-scale information systems, which are first decomposed into a number of small information subsystems for which the reducts are computed and combined to give a reduct of the large-scale information system. Others enhance the effectiveness of attribute reduction for an information system, essentially deriving diverse reducts from the information system by means of certain methods and combining the learners constructed based on the reducts into one single learner or fusing the reducts into a single reduct. These algorithms are reviewed in the following subsections.

6.1. Fusion of attribute reducts

To deal with large-scale decision information systems, Slezak et al. [102] decomposed a decision information system with respect to the universe of objects, and combined rough set theory with Dempster–Shafer and statistical theories to construct a synthesis model of subsystems on the basis of the quality of reasoning. Huang et al. [31] proposed an incremental approach for updating approximations in a distributed information system (a decomposed information system) under attribute generalization. Inspired by the idea of decomposition, Liang et al. [54] developed an efficient attribute reduction algorithm for multi-granulation information to find a reduct for a large-scale information system. The algorithm consists of three main steps:

1. Select smaller information systems (called information subsystems) from the large-scale system;
2. Find a reduct on each information subsystem;
3. Fuse all the reducts of the information subsystems together.

Note that in [54], the ratio of decision attribute values in a subsystem is set to be equal to the ratio of the original large-scale system, and each subsystem has an identical number of objects.

As the total time spent computing the reducts for each information subsystem is much less than that for the original large-scale system, the algorithm yields a reduct much faster than in the previous approach.

Because data sets in real data bases usually take on hybrid forms, i.e., the coexistence of categorical and numerical data. In [112], based on the idea of decomposition and fusion, Wang et al. introduced an efficient feature selection approach for large-scale hybrid data sets. In [112], different attribute rank lists can be obtained from different information subsystems, and a final rank list can be generated by fusing the rank lists derived from different information systems.

The algorithm mentioned above are suitable for a large-scale information system. This is because the large-scale information system is divided into information subsystems in the algorithm, and reducts are computed in the information subsystem simultaneously. Then, the reducts are fused into a single reduct for the large-scale information system; thus, it is an efficient algorithm. However, it must be noted that this algorithm can only obtain an approximate reduct of large-scale information systems, and the quality of the final reduct can be guaranteed by some theoretical results. To solve this problem, by means of the similar decomposition–fusion strategy, Jing et al. [33] developed a method of computing equivalence classes for large-scale information systems. Their technique computes the equivalence classes of information subsystems, and then merges all these equivalence classes into a single one with respect to each object. These are used to compute a reduct of the large-scale information system. On the basis of the integrated equivalence classes, an incremental attribute reduction approach for multi-granulation data has been developed for dynamic information systems.

Note that the fusion of rank lists was used by Lin et al. [63]. Lin et al. [63] presented an attribute reduction algorithm that fuses the attribute rank lists at different granularities. In their algorithm, different neighborhood rough sets are constructed at each granularity with different neighborhood parameters. Thus, different attribute rank lists can be obtained in terms of the neighborhood dependency degree at different granularities. As different attribute rank lists represent different metrics with different approximation abilities, fusing all the rank lists.

6.2. Fusion of the learners constructed based on attribute reducts

To improve the intrusion detection precision, Zhao et al. [159] presented an ensemble algorithm, in which multiple training subsets are first generated using bootstrap technology. Then, attribute reduction in these subsets are obtained using neighborhood rough sets with different radii, and particle swarm optimization is used to optimize the parameters of support vector machine to obtain base classifiers with greater difference and higher precision. Finally, the base classifiers were integrated by a weighted synthesis method. In [160], Zhu et al. developed a fast neighborhood attribute reduction algorithm based on sample pair selection to find all reducts and proposed a randomized attribute reduction algorithm based on neighborhood dependency for large-scale datasets. Because these subspaces generated by these reducts contain complementary information for classification, a classifier based on joint subspace representation was proposed to fully exploit the complementary information in different subspaces.

In the attribute-reduction-based classifier ensemble methods mentioned above, the first step is to generate a set of reducts. The second step is to build base classifiers with all the generated reducts. Finally, different strategies are used to integrate these base classifiers. However, using all reducts generated by the first step to build base classifiers would have a bad influence on the ensemble classifier because the classifiers trained by different reducts may have the same classification results, which will have a bad influence on the diversity of a classifier ensemble. To solve this problem, some studies on attribute-reduction-based selective ensemble were proposed as follows. Hu et al. [23] employed rough-set-based attribute reduction algorithm to obtain a set of reducts of the original information system, and base classifiers were constructed using the reducts. An accuracy-guided forward search and post-pruning strategy was adapted to select a part of the base classifiers, and all selected classifiers were fused to construct an efficient and effective ensemble classifier. Gao et al. [10] proposed an attribute-reduction-based selective classifier ensemble, which consists of four steps: (1) one reduct is randomly chosen from the original reduct pool; (2) the searching space of reducts is reduced by computing the maximum dependency attribute of selected reduct; (3) depending on the accuracy-diversity assessment function, a reduct from the newly searching space is selected (By repeating the second and third steps, a group of proper reducts can be selected); (4) the selected reducts are used to train a set of classifiers, by which the classifier ensemble is formed. For the clustering analysis, Gao et al. [9] proposed a method by which the attributes of categorical data are decomposed into a number of rough subspaces by means of attribute reduction. A method of measuring the performance of rough subspaces was proposed to rank the rough subspaces, and a few high performance rough subspace-based clustering partitions were combined to obtain the final result of clustering.

Because the learners generated by these reducts contain complementary information for the learning task and provide diversity of learner ensemble, the information fusion from the perspective of attribute reduction can improve the performance of learners.

7. Information fusion from the perspective of decision making

The rough set is a promising approach to deal with qualitative information and provide an approach based on an individual object model, which makes the rough set suitable to deal with a few problems of decision making. Many investigations have been conducted from the following two aspects: (1) fusion of decision rules induced from different decision makers; (2) fusion of decision results of different decision makers. In this section, information fusion from the perspective of decision making, which is related to the rough set theory, will be systematically reviewed.

7.1. Fusion of decision rules

In the rough set theory, after attribute reduction, a rule induction algorithm generates decision rules, which may reveal profound knowledge and provide new insights [44,114,125]. These decision rules are more useful for experts to analyze and understand the problem at hand. For the problem of the rough-set-based group decision making, merging the decision rules obtained from multiple decision information systems, which represents decision results from different decision makers respectively, is an important and efficient method to accomplish a consensus decision.

Inuiguchi et al. [32] investigated rule induction from two decision tables as a basis of rough set analysis of more than one decision information systems. They regarded the rule induction process as enumerating minimal conditions satisfied with positive examples but unsatisfied with negative examples and/or with negative decision rules. It is feasible to extend the method to process rule induction from multiple decision information systems. Lin et al. [64] presented a decision rules synthesization model by weighting to mine global decision rules from multi-source information systems. A representation of decision rules based on neighborhood granulation was defined, which is composed of an instance, neighborhood size, and its label. Furthermore, the weight of each information source based on the consistency between different information sources. Based on the weight, global decision rules were generated via a weight-based synthesization model.

Although some existing studies have been conducted on the fusion of decision rules, there are several topics that need to be investigated in this direction. Especially, the fusion of decision rules for multi-view, multi-scale information systems is very meaningful.

7.2. Fusion of decision results

In real life, many important decision problems are not determined by a single decision-maker but by a group of decision makers. The fusion of decision results of a group of decision makers can solve this type of decision problems.

In the group decision-making field, the importance (weights) of individuals in group decision-making has been widely investigated. However, personal judgment biases may influence the determination of initial weights in some subjective approaches. In some objective approaches, the weight is often affected by insufficient knowledge of some group members. Because of these reasons, neither subjective approaches nor objective approaches may be the best method for weight determination. To combine the subjective and objective approaches, Xie et al. [124] combined the variable precision rough set and AHP to obtain the weight of condition attribute sets decided by each decision-maker in group decision-making. Moreover, Yang et al. [136] presented a novel approach for determining the weights of decision makers based on rough group decision in multiple-attribute group decision-making problems. Based on the weight, the results of all different decision makers are weighed to obtain a final decision result.

For the problem of three-way decision, Huang et al.[30] introduced three-way cognitive processes for single-source information systems that are realized by the idea of lower and upper approximations in the rough set theory. Furthermore, for multi-source information systems, three-way decision results obtained in each single-source information system are aggregated by an evaluation function [46], which are used to induce three-way decisions for the multi-source information systems. This similar fusion strategy was used to obtain final concepts by integrating the concepts from all single-source information system in an multi-source information system [76]. Based on decision theory rough sets, Liang et al. [51] gathered evaluation information of each decision maker, aggregated the evaluation to generate interval-valued information granules for each loss function, and then transform the interval-valued information granules with a certain value and determine the precise values of loss functions. Thus, three-way decisions can be deduced easily.

In all the methods mentioned above, each member in a group of decision makers makes their own decision, and fusing their results to generate the final decision results. The ideas of the methods are useful and helpful for a complex decision problem. Although the methods were proposed for group decision-making, the idea can be used to deal with complex information systems when each component information system can be regarded as the description of a decision maker.

8. Conclusion

The exploitation of complex information systems (e.g., multi-source, multi-scale, multi-modality, and multi-view information systems) is a very active area of research. The idea of information fusion in rough set theory is a natural motivation to deal with complex information systems, and has spawned numerous models and powerful algorithms. Existing studies can be divided into five main categories: information fusion from the perspective of objects, information fusion from the perspective of attributes, information fusion from the perspective of rough approximations, information fusion from the perspective of attribute reducts, and information fusion from the perspective of decision making. To date, information fusion has experienced considerable development in all five categories.

However, when considering the many studies on rough set theory, there is still much work to be done to present a complete picture of information fusion by using the rough set theory. Future studies should extend the existing fusion approaches to complex information systems that have not yet been considered. For example, the fusion of objects has just been investigated for multi-source information systems and multi-scale information systems; therefore, it is meaningful to investigate the fusion of other types of complex information systems. Moreover, it would be interesting and promising to explore new fusion approaches that have barely been covered in the literature, such as the fusion of uncertainty measures in the rough set theory, rough-set-based information fusion in the process of constructing learners, rough-set-based information fusion in the context of big data. In addition, all existing approaches focus on complex information systems with the same object sets. Therefore, information fusion for complex information systems with different object sets could be a promising direction in the field of information fusion in rough set.

Acknowledgement

The research was supported by the National Natural Science Foundation of China (Nos. 61402272, U1435212, 61772323 and 61303008) and the National Key Basic Research and Development Program of China (973) (No. 2013CB329404).

References

[1] D.G. Chen, S.Y. Zhao, Local reduction of decision system with fuzzy rough sets, Fuzzy Sets Syst. 161 (2010) 1871–1883.
[2] D.G. Chen, W.L. Li, X. Zhang, S. Kwong, Evidence-theory-based numerical algorithms of attribute reduction with neighborhood-covering rough sets, Int. J. Approx. Reason. 55 (2014) 908–923.
[3] J.H. Dai, W. Wang, Q. Xu, An uncertainty measure for incomplete decision tables and its applications, IEEE Trans. Cybern. 43 (4) (2013) 1277–1289.
[4] D. Dubois, H. Prade, Putting rough sets and fuzzy sets together, in: R. Slowinski (Ed.), Intelligent Decision Support: Hand book of Applications and Advances of the Sets Theory, Kluwer, Dordrecht, 1992, pp. 203–232.
[5] I. Düntsch, G. Gediga, Rough Set Data Analysis: A Road to Non-invasive Knowledge Discovery, Methodos Publishers, Bangor, UK, 2000.
[6] H.L. Dou, X.B. Yang, J.Y. Fan, S.P. Xu, The models of variable precision multigranulation rough sets, in: RSKT 2012, LNCS, volume 7414, 2012, pp. 465–473.
[7] T. Feng, J.S. Mi, Variable precision multigranulation decision-theoretic fuzzy rough sets, Knowl. Based Syst. 91 (2016) 93–101.
[8] T. Feng, H.T. Fan, J.S. Mi, Uncertainty and reduction of variable precision multigranulation fuzzy rough sets based on three-way decisions, Int. J. Approx. Reason. 85 (2017) 36–58.
[9] C. Gao, W. Pedrycz, D. Miao, Rough subspace-based clustering ensemble for categorical data, Soft Comput. 17 (9) (2013) 1643–1658.
[10] Y.W. Guo, L.C. Jiao, S. Wang, S. Wang, F. Liu, K.X. Rong, T. Xiong, A novel dynamic rough subspace based selective ensemble, Pattern Recognit. 48 (2015) 1638–1652.
[11] S.M. Gu, W.Z. Wu, Y. Zheng, Rule acquisition in consistent multi-scale decision systems, in: Proceeding of the FSKD 2011 Conference, Shanghai, 2011, pp. 390–393.
[12] S.M. Gu, W.Z. Wu, Knowledge acquisition in inconsistent multi-scale decision systems, in: Proceeding of the RSKT 2011 Conference, Banff, 2011, pp. 669–678.
[13] S.M. Gu, X. Li, W.Z. Wu, H. Nian, Multi-granularion rough sets in multi-scale information systems, in: Proceedings of the 2013 International Conference on Machine Learning and Cybernetics, Tianjin, 2013, pp. 108–113.
[14] S.M. Gu, W.Z. Wu, On knowledge acquisition in multi-scale decision systems, Int. J. Mach. Learn. Cybern. 4 (2013) 477–486.
[15] S.M. Gu, Y.H. Wan, W.Z. Wu, T.J. Li, Rough set approximations in multi-scale interval information systems, in: Y.Y. Yao, et al. (Eds.), RSFDGrC 2015, LNAI 9437, 2015, pp. 73–81.
[16] S. Greco, B. Matarazzo, R. Slowinski, A new rough set approach to multicriteria and multiattribute classification, Lect. Notes Artif. Intell. 1424 (1998) 60–67.
[17] S. Greco, B. Matarazzo, R. Slowinski, Rough approximation of a preference relation by dominance relations, Eur. J. Oper. Res. 117 (1999) 63–83.
[18] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multi-criteria decision analysis, Eur. J. Oper. Res. 129 (2001) 1–47.
[19] S. Greco, B. Matarazzo, R. Slowinski, Rough approximation by dominance relations, Int. J. Intell. Syst. 17 (2002) 153–171.
[20] S. Greco, B. Matarazzo, R. Slowinski, Rough sets methodology for sorting problems in presence of multiple attributes and criteria, Eur. J. Oper. Res. 138 (2002) 247–259.
[21] C. Hao, J.H. Li, M. Fan, W.Q. Liu, E.C.C. Tsang, Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions, Inf. Sci. (Ny) 415-416 (2017) 213–232.
[22] Q.H. Hu, Z.X. Xie, D.R. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognit. 40 (12) (2007) 3509–3521.
[23] Q.H. Hu, D.R. Yu, Z.X. Xie, X.D. Li, EROS: Ensemble rough subspaces, Pattern Recognit. 40 (12) (2007) 3728–3739.
[24] Q.H. Hu, D.R. Yu, J.F. Liu, C. X. Wu, Neighborhood rough set based heterogeneous feature selection, Inf. Sci. (Ny) 178 (2008) 3577–3594.
[25] Q.H. Hu, S. An, X. Yu, D.R. Yu, Robust fuzzy rough classifiers, Fuzzy Sets Syst. 183 (2011) 26–43.
[26] Q.H. Hu, D.R. Yu, W. Pedrycz, D.G. Chen, Kernelized fuzzy rough sets and their applications, IEEE Trans. Knowl. Data Eng. 23 (11) (2011) 1649–1667.
[27] Q.H. Hu, L.J. Zhang, Y.C. Zhou, W. Pedrycz, Large-scale multi-modality attribute reduction with multi-kernel fuzzy rough sets, IEEE Trans. Fuzzy Syst. (2017). 10.1109/TFUZZ.2017.2647966
[28] C.X. Hu, S.X. Liu, X.L. Huang, Dynamic updating approximations in multigranulation rough sets while refining or coarsening attribute values, Knowl. Based Syst. 130 (2017) 62–73.
[29] B. Huang, C.X. Guo, Y.L. Zhuang, H.X. Li, X.Z. Zhou, Intuitionistic fuzzy multigranulation rough sets, Inf. Sci. (Ny) 277 (2014) 299–320.
[30] C.C. Huang, J.H. Li, C.L. Mei, W.Z. Wu, Three-way concept learning based on cognitive operators: an information fusion viewpoint, Int. J. Approx. Reason. 83 (2017) 218–242.
[31] Y.Y. Huang, T.R. Li, C. Luo, S.J. Horng, Dynamic updating rough approximations in distributed information systems, in: Proceedings of the 2015 International Conference on Intelligent Systems and Knowledge Engineering, Taipei, Taiwan, 2015, pp. 170–175.
[32] M. Inuiguchi, T. Miyajima, Rough set based rule induction from two decision tables, Eur. J. Oper. Res. 181 (2007) 1540–1553.
[33] Y.G. Jing, T.R. Li, H. Fujita, Z. Yu, B. Wang, An incremental attribute reduction approach based on knowledge granularity with a multi-granulation view, Inf. Sci. (Ny) 411 (2017) 23–38.
[34] H.R. Ju, X.B. Yang, X.N. Song, Y.S. Qi, Dynamic updating multi-granulation fuzzy rough set: approximations and reducts, Int. J. Mach. Learn. Cybern. 5 (2014) 981–990.
[35] H.R. Ju, X.B. Yang, H.L. Dou, J.J. Song, Variable precision multi-granulation rough set and attributes reduction, Trans. Rough Set XVIII (2014) 52–68.
[36] H.R. Ju, H.X. Li, X.B. Yang, X.Z. Zhou, B. Huang, Cost-sensitive rough set: a multi--granulation approach, Knowl. Based Syst. 123 (2017) 137–153.
[37] M.A. Khan, M. Banerjee, Formal reasoning with rough sets in multiple-source approximation systems, Int. J. Approx. Reason. 49 (2008) 466–477.
[38] M.A. Khan, M. Banerjee, Multiple-source approximation systems: membership functions and indiscernibility, Lect. Notes Comput. Sci. 5009 (2008) 80–87.
[39] M.A. Khan, Formal reasoning in preference-based multiple-source rough set model, Inf. Sci. (Ny) 334–335 (2016) 122–143.
[40] M. Kondo, On the structure of generalized rough sets, Inf. Sci. (Ny) 176 (2006) 589–600.
[41] M. Kryszkiewicz, Rough set approach to incomplete information systems, Inf. Sci. (Ny) 112 (1998) 39–49.
[42] F. Li, B.Q. Hu, A new approach of optimal scale selection to multi-scale decision tables, Inf. Sci. (Ny) 381 (2017) 193–208.
[43] F. Li, B.Q. Hu, J. Wang, Stepwise optimal scale selection for multi-scale decision tables via attribute significance, Knowl. Based Syst. 129 (2017) 4–16.
[44] J.H. Li, C.L. Mei, Y.J. Lv, Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction, Int. J. Approx. Reason. 54 (2013) 149–165.
[45] J.H. Li, C.L. Mei, W.H. Xu, Y.H. Qian, Concept learning via granular computing: a cognitive viewpoint, Inf. Sci. (Ny) 298 (2015) 447–467.
[46] J.H. Li, C.H. Huang, W.H. Xu, Y.H. Qian, W.Q. Liu, Cognitive concept learning via granular computing for big data, in: Proceedings of the 2015, International Conference on Machine Learning and Cybernetics, 2015, pp. 289–294.
[47] J.H. Li, Y. Ren, C.L. Mei, Y.H. Qian, X.B. Yang, A comparative study of multigranulation rough sets and concept lattices via rule acquisition, Knowl. Based Syst. 91 (2016) 152–164.
[48] J.H. Li, C.C. Huang, J.J. Qi, Y.H. Qian, W.Q. Liu, Three-way cognitive concept learning via multi-granularity, Inf. Sci. (Ny) 378 (2017) 244–263.
[49] W.T. Li, X.Y. Zhang, W.X. Sun, Further study of multigranulation t-fuzzy rough sets, Sci. World J. 14 (2) (2014) 1–18.
[50] W.T. Li, W.H. Xu, Multi-granulation decision-theoretic rough set in ordered information system, Fundam. Inform. 139 (2015) 67–89.
[51] D. Liang, D. Liu, A. Kobina, Three-way group decisions with decision-theoretic rough sets, Inf. Sci. (Ny) 345 (2016) 46–64.
[52] J.Y. Liang, K.S. Chin, C.Y. Dang, A new method for measuring uncertainty and fuzziness in rough set theory., Int. J. Gen. Syst. 31 (4) (2002) 331–342.
[53] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory. int, J. Uncertain. Fuzziness Knowl.-Based Syst. 12 (1) (2004) 37–46.
[54] J.Y. Liang, F. Wang, C.Y. Dang, Y.H. Qian, An efficient rough feature selection algorithm with a multi-granulation view, Int. J. Approx. Reason. 53 (2012) 912–926.
[55] G.P. Lin, Y.H. Qian, J.J. Li, A covering-based pessimistic multigranulation rough set, in: Proceedings of the International Conference on Intelligent Computing, Zhengzhou, China, 2011.
[56] G.P. Lin, Y.H. Qian, J.J. Li, NMGRS: neighborhood-based multigranulation rough sets, Int. J. Approx. Reason. 53 (2012) 1080–1093.
[57] G.P. Lin, J.Y. Liang, Y.H. Qian, Multigranulation roughsets: from partition to covering, Inf. Sci. (Ny) 241 (2013) 101–118.
[58] G.P. Lin, J.Y. Liang, Y.H. Qian, An information fusion approach by combining multigranulation rough sets and evidence theory, Inf. Sci. (Ny) 314 (2015) 184–199.
[59] G.P. Lin, J.Y. Liang, Y.H. Qian, J.J. Li, A fuzzy multigranulation decision-theoretic approach to multi-source fuzzy information systems, Knowl. Based Syst. 91 (2016) 102–113.
[60] T.Y. Lin, Neighborhood systems and approximation in database and knowledge base systems, in: Proceedings of the Fourth International symposium on Methodologies of Intelligent Systems, Poster Session, 1989, pp. 75–86.
[61] T.Y. Lin, Neighborhood systems: mathematical models of information granulations, in: Proceedings of the IEEE International Conference on Systems, Man & Cybernetics, 2003, pp. 5–8.
[62] T.Y. Lin, Granular and nearest neighborhood: rough set approach, in: Granulation Computing: An Emerging Paradigm, Physica-Verlag, 2001, pp. 125–142.
[63] Y.J. Lin, J.J. Li, P.R. Lin, G.P. Lin, J.K. Chen, Feature selection via neighborhood multi-granulation fusion, Knowl. Based Syst. 67 (2014) 162–168.
[64] Y.J. Lin, H.H. Chen, G.P. Lin, J.K. Chen, Z.M. Ma, J.J. Li, Synthesizing decision rules from multiple information sources: a neighborhood granulation viewpoint, Int. J. Mach. Learn. Cybern.org/10.1007/s13042-018-0791-z.
[65] C.H. Liu, D.Q. Miao, Covering rough set model based on multigranulations, in: Proceedings of the RSFDGRc2011, 2011, pp. 87–90.
[66] C.H. Liu, M.Z. Wang, Covering fuzzy rough set based on multi-granulations, in: Proceedings of the International Conference on Uncertainty Reasoning and Knowledge Engineering, 2011, pp. 146–149.
[67] C.H. Liu, D.Q. Miao, N. Zhang, C. Gao, Graded rough set model based on two universes and its properties, Knowl.-Based Syst. 33 (2012) 65–72.
[68] C.H. Liu, D.Q. Miao, J. Qian, On multi-granulation covering rough sets, Int. J. Approx. Reason. 55 (2014) 1404–1418.
[69] G.L. Liu, W. Zhu, The algebraic structures of generalized rough set theory, Inf. Sci. (Ny) 178 (2008) 4105–4113.
[70] L. Lu, B. Huang, Q.Y. Zhang, D.F. Ke, Y.Y. Xu, Research on algorithm of combing lda-based discriminant classifier and MFCC feature extraction for pure acoustic listening similarity, Int. J. Adv. Comput. Technol. 4 (5) (2012) 106–113.
[71] S. Maji, A.C. Berg, J. Malik, Classification using intersection kernel support vector machines is efficient, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.

[72] J.S. Mi, Y. Leung, H.Y. Zhao, T. Feng, Generalized fuzzy rough sets determined by a triangular norm, Inf. Sci. (Ny) 178 (2008) 3203–3213.
[73] F. Min, Q.H. Liu, A hierarchical model for test-cost-sensitive decision systems, Inf. Sci. (Ny) 179 (2009) 2442–2452.
[74] F. Min, H.P. He, Y.H. Qian, W. Zhu, Test-cost-sensitive attribute reduction, Inf. Sci. (Ny) 181 (2011) 4928–4942.
[75] N.N. Morsi, M.M. Yakout, Axiomatics for fuzzy rough sets, Fuzzy Sets Syst. 100 (1998) 327–342.
[76] J.J. Niu, C.C. Huang, J.H. Li, M. Fan, Parallel computing techniques for concept cognitive learning based on granular computing, Int. J. Mach. Learn. Cybern.10.1007/s13042-018-0783-z.
[77] F. Odone, A. Barla, A. Verri, Building kernels from binary strings for image matching, IEEE Trans. Image Process. 14 (2) (2005) 169–180.
[78] Y. Ouyang, Z.D. Wang, H.P. Zhang, On fuzzy rough sets based on tolerance relations, Inf. Sci. (Ny) 180 (2010) 532–542.
[79] W. Pan, K. She, P.Y. Wei, Multi-granulation fuzzy preference relation rough set for ordinal decision system, Fuzzy Sets Syst. 312 (2017) 87–108.
[80] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991.
[81] Z. Pawlak, A. Sknwron, Rough sets and boolean reasoning, Inf. Sci. (Ny) 177 (1) (2007) 41–73.
[82] Z. Pawlak, A. Skowron, Rudiments of rough sets, Inf. Sci. (Ny) 177 (1) (2007) 3–27.
[83] Z. Pawlak, Rough set theory and its applications to data analysis, Cybern. Syst. Int. J. 29 (7) (2010) 661–688.
[84] Y.H. Qian, J.Y. Liang, Rough set method based on multi-granulations, in: Proceedings of the 5th IEEE Conference on Cognitive Informatics, volume 1, 2006, pp. 297–304.
[85] Y.H. Qian, C.Y. Dang, J.Y. Liang, MGRS in incomplete information systems, in: Proceedings of the IEEE International Conference Granular Computation, 2007, pp. 163–168.
[86] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Information granularity in fuzzy binary grc model, IEEE Trans. Fuzzy Syst. 50 (2009) 174–188.
[87] Y.H. Qian, J.Y. Liang, C.Y. Dang, Knowledge structure, knowledge granulation and knowledge distance in a knowledge base, Int. J. Approx. Reason. 50 (1) (2009) 174–188.
[88] Y.H. Qian, C.Y. Dang, J.Y. Liang, Incomplete multi-granulations rough set, IEEE Trans. Syst. Man Cybern.-Part A 40 (2) (2010) 420–431.
[89] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: a multi-granulation rough set, Inf. Sci. (Ny) 180 (2010) 949–970.
[90] Y.H. Qian, J.Y. Liang, W.Z. Wu, C.Y. Dang, Information granularity in fuzzy binary GRC model, IEEE Trans. Fuzzy Syst. 19 (2) (2011) 253–264.
[91] Y.H. Qian, S.Y. Li, J.Y. Liang, Z.Z. Shi, F. Wang, Pessimistic rough set based decisions: a multigranulation fusion strategy, Inf. Sci. (Ny) 264 (2014) 196–210.
[92] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 225–237.
[93] Y.H. Qian, X.Y. Liang, G.P. Lin, Q. Guo, J.Y. Liang, Local multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 82 (2017) 119–137.
[94] K.Y. Qin, Z. Pei, On the topological properties of fuzzy rough sets, Fuzzy Sets Syst. 151 (3) (2005) 601–613.
[95] A.M. Radzikowska, E.E. Kerre, A comparative study of fuzzy rough sets, Fuzzy Sets Syst. 126 (2002) 137–155.
[96] Y.L. Sang, J.Y. Liang, Y.H. Qian, Decision-theoretic rough sets under dynamic granulation, Knowl. Based Syst. 91 (2016) 84–92.
[97] B. Schlkopf, A. Smola, Learning with kernels, in: Proceedings of the 21st International Conference on Machine Learning, 2001, pp. 639–646.
[98] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[99] Y.H. She, X.L. He, On the structure of the multigranulation rough set model, Knowl. Based Syst. 36 (2012) 81–92.
[100] Y.H. She, J.H. Li, H.L. Yang, A local approach to rule induction in multi-scale decision tables, Knowl. Based Syst. 89 (2015) 398–410.
[101] Y.H. She, X.L. He, H.X. Shi, Y.H. Qian, A multiple-valued logic approach for multigranulation rough set model, Int. J. Approx. Reason. 82 (2017) 270–284.
[102] D. Slezak, Decision value oriented decomposition of data tables, in: Proceedings of the 1997 International Conference on International Symposium on Methodologies for Intelligent Systems, Charlotte, North Carolina, United States of America, 1997, pp. 487–496.
[103] J. Stefanowski, A. Tsoukiàs, Incomplete information tables and rough classification, Comput. Intell. 17 (2001) 545–566.
[104] B.Z. Sun, W.M. Ma, Y.H. Qian, Multigranulation fuzzy rough set over two universes and its application to decision making, Knowl. Based Syst. 123 (2017) 61–74.
[105] A.H. Tan, W.Z. Wu, J.J. Li, G.P. Lin, Evidence-theory-based numerical characterization of multigranulation rough sets in incomplete information systems, Fuzzy Sets Syst. 294 (2016) 18–35.
[106] A.H. Tan, W.Z. Wu, Y.Z. Tao, On the belief structures and reductions of multigranulation spaces with decisions, Int. J. Approx. Reason. 88 (2017) 39–52.
[107] X.Q. Tang, P. Zhu, Hierarchical clustering problems and analysis of fuzzy proximity relation on granular space, IEEE Trans. Fuzzy Syst. 21 (5) (2013) 814–824.
[108] Y.Q. Tang, M. Fan, J.H. Li, An information fusion technology for triadic decision contexts, Int. J. Mach. Learn. Cybern. 7 (1) (2016) 13–24.
[109] S. Tata, J.M. Patel, Estimating the selectivity of tf-idf based cosine similarity predicates, Acm Sigmod Record 36 (2) (2007) 7–12.
[110] C.Z. Wang, Y.L. Qi, M.W. Shao, Q.H. Hu, D.G. Chen, Y.H. Qian, Y.J. Lin, A fitting model for feature selection with fuzzy rough sets, IEEE Trans. Fuzzy Syst. 25 (4) (2016) 741–753.
[111] C.Z. Wang, M.W. Shao, Q. He, Y.H. Qian, Y.L. Qi, Feature subset selection based on fuzzy neighborhood rough sets, Knowl. Based Syst. 111 (1) (2016) 173–179.
[112] F. Wang, J.Y. Liang, An efficient feature selection algorithm for hybrid data, Neurocomputing 193 (2016) 33–41.
[113] G.Y. Wang, Extension of rough set under incomplete information systems, in: Proceeding of the 11th IEEE International Conference on Fuzzy Systems, 2002, pp. 1098–1103.
[114] X. Wang, J. Yang, R. Jensen, X. Liu, Rough set feature selection and rule induction for prediction of malignancy degree in brain glioma, Comput. Methods Progr. Biomed. 83 (2006) 147–156.
[115] X.Z. Wang, Y. Ha, D.G. Chen, On the reduction of fuzzy rough sets, in: Proceedings of the 2005 International Conference 1006 on Machine Learning and Cybernetics, volume 5, 2005, pp. 3174–3178.
[116] W. Wei, J.Y. Liang, Y.H. Qian, F. Wang, Variable precision multi-granulation rough set, in: Proceedings of the IEEE International Conference on Granular Computing, 2012, pp. 639–643.
[117] W.Z. Wu, Y. Leung, W.X. Zhang, Connections between rough set theory and Dempster-Shafer theory of evidence, Int. J. Gen. Syst. 31 (4) (2002) 405–430.
[118] W.Z. Wu, J.S. Mi, W.X. Zhang, Generalized fuzzy rough sets, Inf. Sci. (Ny) 151 (2003) 263–282.
[119] W.Z. Wu, Y. Leung, J.S. Mi, On generalized fuzzy belief functions in infinite spaces, IEEE Trans. Fuzzy Syst. 17 (2009) 385–397.
[120] W.Z. Wu, Y. Leung, Theory and applications of granular labeled partitions in multi-scale decision tables, Inf. Sci. (Ny) 181 (2011) 3878–3897.
[121] W.Z. Wu, Y. Leung, Optimal scale selection for multi-scale decision tables, Int. J. Approx. Reason. 54 (2013) 1107–1129.
[122] W.Z. Wu, S.M. Gu, X. Wang, Information granules multi-scale ordered information systems, in: Proceedings of the 2015 International Conference on Machine Learning and Cybernetics, Guangzhou, 2015, pp. 182–187.
[123] W.Z. Wu, Y.H. Qian, T.J. Li, S.M. Gu, On rule acquisition in incomplete multi-scale decision tables, Inf. Sci. (Ny) 378 (2017) 282–302.
[124] G. Xie, J. Zhang, K.K. Lai, L. Yu, Variable precision rough set for group decision–making: an application, Int. J. Approx. Reason. 49 (2008) 331–343.
[125] J.P. Xie, M.H. Yang, J.H. Li, Z. Zheng, Rule acquisition and optimal scale selection in multi-scale formal decision contexts and their applications to smart city, Future Gener. Comput. Syst. 73 (1) (2017) 1–30.
[126] W.H. Xu, X.Y. Zhang, J.M. Zhong, Attribute reduction in ordered information systems based on evidence theory, Knowl. Inf. Syst. 25 (2010) 169–184.
[127] W.H. Xu, X.T. Zhang, Q.R. Wang, A generalized multi-granulation rough set approach, in: Proceedings of the 7th International Conference on Intelligent Computing, ICIC, 2011, pp. 681–689.
[128] W.H. Xu, Q.R. Wang, X.T. Zhang, Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space, International Journal of Fuzzy System 3 (4) (2011) 246–259.
[129] W.H. Xu, W.X. Sun, X.T. Zhang, W.X. Zhang, Multiple granulation rough set approach to ordered information systems, International Journal of Genernal Systems 41 (2012) 475–501.
[130] W.H. Xu, Q.R. Wang, X.Y. Zhang, Multi-granulation rough sets based on tolerance relations, Soft Comput. 17 (2013) 1241–1252.
[131] W.H. Xu, Q.R. Wang, S.Q. Luo, Multi-granulation fuzzy rough sets, J. Intell. Fuzzy Syst. 26 (2014) 1323–1340.
[132] W.H. Xu, Y.T. Guo, Generalized multigranulation double-quantitative decision-theoretic rough set, Knowl. Based Syst. 105 (2016) 190–205.
[133] W.H. Xu, J.H. Yu, A novel approach to information fusion in multi-source datasets: a granular computing viewpoint, Inf. Sci. (Ny) 378 (2017) 410–423.
[134] W.H. Xu, M.M. Li, X.Z. Wang, Information fusion based on information entropy in fuzzy multi-source incomplete information system, Int. J. Fuzzy Syst. 19 (4) (2017) 1200–1216.
[135] W.H. Xu, W.T. Li, X.T. Zhang, Generalized multigranulation rough sets and optimal granularity selection, Granular Comput. (2017) 1–18.
[136] Q. Yang, P.A. Du, Y. Wang, B. Liang, A rough set approach for determining weights of decision makers in group decision making, PLoS ONE. 10.1371/journal.pone.0172679.
[137] X.B. Yang, X.N. Song, H.L. Dou, J.Y. Yang, Multi-granulation rough set: from crisp to fuzzy case, Ann. Fuzzy Math. Inf. 1 (2011) 55–70.
[138] X.B. Yang, Y.H. Qian, J.Y. Yang, Hierarchical structures on multigranulation spaces, J. Comput. Sci. Technol. 27 (6) (2012) 1169–1183.
[139] X.B. Yang, X.N. Song, Z.H. Chen, J.Y. Yang, On multi-granulation rough sets in incomplete information system, Int. J. Mach. Learn. Cybern. 3 (2012) 223–232.
[140] X.B. Yang, Y.S. Qi, X.N. Song, Z.H. Chen, J.Y. Yang, Test cost sensitive multigranulation rough set: model and minimal cost selection, Inf. Sci. (Ny) 250 (2013) 184–199.
[141] X.B. Yang, Y. Qi, H.L. Yu, X.N. Song, J.Y. Yang, Updating multi granulation rough approximations with increasing of granular structures, Knowl.-Based Syst. 64 (2014) 59–69.
[142] J.T. Yao, Y.Y. Yao, W. Ziarko, Probabilistic rough sets: approximations, decision–makings, and applications, Int. J. Approx. Reason. 49 (2008) 253–254.
[143] Y.Y. Yao, Combination of rough and fuzzy sets based on α-level sets, in: T.Y. Lin, N. Cercone (Eds.), Rough Sets and Data Mining: Analysis for Imprecise Data, Kluwer Academic Publishers, Boston, 1997, pp. 301–321.
[144] Y.Y. Yao, P.J. Lingras, Interpretations of belief functions in the theory of rough sets, Inf. Sci. (Ny) 104 (1998) 81–106.
[145] Y.Y. Yao, Information granulation and rough set approximation, Int. J. Intell. Syst. 16 (2001) 87–104.
[146] Y.Y. Yao, Neighborhood systems and approximate retrieval, Inf. Sci. (Ny) 176 (23) (2006) 3431–3452.
[147] Y.Y. Yao, Three-way decisions with probabilistic rough sets, Inf. Sci. (Ny) 180 (2010) 341–353.

[148] Y.Y. Yao, The superiority of three-way decisions in probabilistic rough set models, Inf. Sci. (Ny) 181 (2011) 1080–1096.
[149] Y.Y. Yao, B.X. Yao, Covering based rough set approximations, Inf. Sci. (Ny) 200 (1) (2012) 91–107.
[150] Y.Y. Yao, Y.H. She, Rough set models in multigranulation spaces, Inf Sci (Ny) 327 (2016) 40–56.
[151] D.S. Yeung, D.G. Chen, E.C. Tsang, J.W. Lee, X.Z. Wang, On the generalization of fuzzy rough sets, IEEE Trans. Fuzzy Syst. 13 (3) (2005) 343–361.
[152] J.H. Yu, W.H. Xu, Information fusion in multi-source fuzzy information system with same structure, in: Proceedings of the 2015 International Conference on Machine Learning and Cybernetics, Guangzhou, 2015, pp. 170–175.
[153] J.H. Yu, X.Y. Zhang, Z.H. Zhao, W.H. Xu, Uncertainty measures in multigranulation with different grades rough set based on dominance relation, J. Intell. Fuzzy Syst. 31 (2016) 1133–1144.
[154] W. Ziarko, Variable precision rough set model, J. Comput. Syst. Sci. 46 (1993) 39–59.
[155] C. Zhang, D.Y. Li, Y.M. Mu, D. Song, An interval-valued hesitant fuzzy multigranulation rough set over two universes model for steam turbine fault diagnosis, Appl. Math. Model 42 (2017) 693–704.
[156] C.Q. Zhang, H.Z. Fu, Q.H. Hu, P.F. Zhu, X.C. Cao, Flexible multi-view dimensionality co-reduction, IEEE Trans. Image Process. 26 (2) (2017) 648–659.
[157] X. Zhang, C.L. Mei, D.G. Chen, J.H. Li, Feature selection in mixed data: a method using a novel fuzzy rough set-based information entropy, Pattern Recognit. 56 (2016) 1–15.
[158] X.H. Zhang, D.Q. Miao, C.H. Liu, M.L. Le, Constructive methods of rough approximation operators and multigranulation rough sets, Knowl. Based Syst. 91 (2016) 114–125.
[159] H. Zhao, Intrusion detection ensemble algorithm based on bagging and neighborhood rough set, Int. J. Secur. Appl. 7 (5) (2013) 193–204.
[160] P.F. Zhu, Q.H. Hu, Y.H. Han, C.Q. Zhang, Y. Du, Combining neighborhood separable subspaces for classification via sparsity regularized optimization, Inf. Sci. (Ny) 370–371 (2016) 270–287.
[161] W. Zhu, F.Y. Wang, On three types of covering rough sets, IEEE Trans. Knowl. Data Eng. 19 (8) (2007) 1131–1144.
[162] W. Zhu, Relationship among basic concepts in covering-based rough sets, Inf. Sci. (Ny) 179 (14) (2009) 2478–2486.