



An information fusion approach by combining multigranulation rough sets and evidence theory



Guoping Lin ^{a,b}, Jiye Liang ^{a,*}, Yuhua Qian ^a

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

^b School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, Fujian, China

ARTICLE INFO

Article history:

Received 1 April 2014

Received in revised form 20 February 2015

Accepted 22 March 2015

Available online 3 April 2015

Keywords:

Granular computing

Multigranulation rough sets

Evidence theory

Dempster's combination rule

Information fusion

ABSTRACT

Multigranulation rough set (MGRS) theory provides two kinds of qualitative combination rules that are generated by optimistic and pessimistic multigranulation fusion functions. They are used to aggregate multiple granular structures from a set theoretic standpoint. However, the two combination rules seem to lack robustness because one is too relaxed and the other too restrictive to solve some practical problems. Dempster's combination rule in the evidence theory has been employed to aggregate information coming from multiple sources. However, it fails to deal with conflict evidence. To overcome these limitations, we focus on the combination of granular structures with both reliability and conflict from multiple sources, which has been a challenging task in the field of granular computing. We first address the connection between multigranulation rough set theory and the evidence theory. Then, a two-grade fusion approach involved in the evidence theory and multigranulation rough set theory is proposed, which is based on a well-defined distance function among granulation structures. Finally, an illustrative example is given to show the effectiveness of the proposed fusion method. The results of this study will be useful for pooling the uncertain data from different sources and significant for establishing a new direction of granular computing.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

In the information age, complex data is often represented by a *multi-source information system* [7] in which data come from different sources. How to fuse such data has become a challenging task in the community of *granular computing* (GrC) [60]. *Information granulation* is one of three basic issues: *information granulation*, *organization*, and *causation* in granular computing. *Information granulation* involves decomposition of whole data into parts called granules. Then, these granules are organized into a granular structure (or a granular space). In granular computing, the granules induced by an *equivalence relation* (or a *tolerance relation*) form a set of equivalence classes (or tolerance classes), in which each equivalence class (or tolerance class) can be regarded as a Pawlak *information granule* (or a *tolerance information granule*).

A multi-source information system is used to represent information coming from multiple sources. Single-source information system is a special multi-source information system. According to the granulation approach, the objects in a multi-source information system can be granulated into multiple granular structures induced by a family of binary relations, or a

* Corresponding author.

E-mail address: ljiy@sxu.edu.cn (J. Liang).

family of attribute sets. In each information subsystem, the objects are organized into a granular structure by an attribute set. It is natural to put a fundamental issue on how to combine multiple granular structures from a multi-source information system. In this paper, we call this kind of information fusion as *granulation fusion*.

Rough set theory, proposed by Pawlak [17,18] in 1982, has been proved to be an efficient tool for uncertainty management and uncertainty reasoning. This theory is emerging as a powerful methodology in the field of artificial intelligence such as pattern recognition, machine learning and automated knowledge acquisition. The basic structure of the rough set theory is a *known knowledge base* (or an approximate space) consisting of a universe of discourse and an indiscernible relation imposed on it. Based on the known knowledge base, the primitive notion of the *lower and upper approximate operators* can be induced. The lower and upper approximations of a target concept characterize the non-numeric aspect expressed by the known knowledge base. In the view of granular computing (proposed by Zadeh [60]), Pawlak rough set model and its extensions are based on a single relation (such as an equivalence relation, a tolerance relation or a reflexive relation) on the universe, which are called single granulation rough sets [19–21,23]. However, if data come from different sources, the data analysis mechanism of the classical rough set theory is not desirable even not efficient. In this circumstance, one often needs to describe concurrently a target concept through multiple binary relations according to a user's requirements or targets of problem solving, which motivates us to consider how to fuse such data from different sources.

Information fusion is a typical problem that involves the integration of multi-source information in signal processing, image processing, knowledge representation and inference, which has been the objective of many researches over the last few years. Up to now, a variety of *qualitative (non-numeric)* and *quantitative (numeric)* information fusion methods [23,38,51,41] have been developed over the years [1,30,50,55,9]. It is worth pointing out that Qian et al. [22–25] introduced the multigranulation rough set theory (MGRS) which employed conjunctive/disjunctive operators of multiple binary relations to integrate multiple granular structures induced by a family of binary relations. Furthermore, Khan and Banerjee [7,8] proposed a weak lower (or strong upper) approximation of a target concept in the framework of a multi-source approximation space. In fact, the essence of the so-called optimistic lower (or upper) approximation defined in [22] is the same as the weak lower (or strong upper) approximation of a target concept proposed in [7]. Similarly, the essence of the so-called pessimistic lower (or upper) approximation defined in [25] is the same as the strong lower (or weak upper) approximation of a target concept proposed in [7]. The former focuses on multiple granulations and the latter focuses on multiple approximation spaces. Since multigranulation rough set model inception, its theoretical framework has been largely enriched, and many extended multigranulation rough set models, as well as relative properties and applications have been also studied extensively [13,45,10,11,14–16,26,29,32–34,49,56–59,39,40,43,44].

In the view of information fusion, MGRS theory can be regarded as a qualitative fusion strategy through the optimistic and pessimistic fusing paradigms. The optimistic fusion paradigm expresses the idea that in multiple independent granular structures, one needs only at least one granular structure to satisfy with the inclusion condition between an equivalence class and a target concept. Whereas the pessimistic version needs all granular structures to satisfy with the inclusion condition. However, the former seems too relaxed for data analysis and leads to generating a loose uncertain interval and the latter seems too restrictive and generates a tight uncertain interval. Therefore, both of them seem not enough precise to measure the uncertainty in multi-source environment. That is, these two fusing methods are two extreme cases which limit the application scope of MGRS. In addition, previous related work addressing qualitative combination rule deserves to be mentioned here. Yao and Wong [52,42] proposed qualitative combination rule which requires the definition of a binary relation expressing the preference of one proposition or source, over another. However, the qualitative methods are still not desirable in engineering analysis, so one expects to propose the quantitative approach to deal with quantitative data.

Dempster–Shafer (DS) theory (also known as *evidence theory* or *Dempster–Shafer theory of evidence*) [2,31], a general extension of Bayesian theory, provides a simple method for combining the evidence carried by a number of different sources. This method is called *orthogonal sum* or *Dempster's combination rule*. In Dempster–Shafer theory, inference is made by aggregating independent evidence from different sources via the Dempster's combination rule. Unfortunately, the unexpected and rather counterintuitive results of Dempster's combination rule under some situations, as highlighted by Zadeh [61], limit the application of Dempster's combination rule in intelligent fusion process. To overcome the disadvantage, its alternatives have been developed. Of all the alternatives, there are three distinguished improved combination rules, such as Smets's unnormalized combination rule (known as the conjunctive combination rule) [36,37], Yager's combination rule [54] and Dubois and Prade's disjunctive combination rule [3]. However, it is pointed out that all combination rules based on DS theory have a common characteristic. That is, they all require prior information to define a *basic probability assignment (bpa)*.

From the above discussions, we find that both multigranulation fusion rules and Dempster's combination rule have their respective limitations. For this reason, in this paper we focus on a fundamental issue on how to combine multiple granulations from a multi-source information system. To address this issue, we first examine the connection between multigranulation rough set theory and Dempster–Shafer's theory. One may capture the prior information as the mass function according to the relationship between the single granulation Pawlak rough set and the evidence theory [35,53,45–48]. Then, we propose a two-grade quantitative fusion approach integrating Dempster–Shafer theory and multigranulation rough set theory to deal with uncertainty of a multi-source information system. This new approach is based on a new distance between two granular structures. Finally, an illustrative example is given to show the effectiveness of the proposed fusion method.

This paper is organized as follows. Section 2 reviews some basic concepts of a multi-source information system, Dempster–Shafer theory and MGRS. Section 3 discusses the connection between MGRS and the evidence theory. In

Section 4 we propose a new fusion function based on the evidence distance to merge multi-source uncertain information from a multi-source information system. In Section 5, an example is subsequently employed to demonstrate the validity and effectiveness of the integrated information fusion approach. Section 6 concludes the paper with a summary and direction for future.

2. Preliminaries

In this section, we review some basic concepts of a multi-source information system, the Dempster–Shafer theory and MGRS.

2.1. Multi-source information systems

In Pawlak rough set theory, a *single-source information system* [17] is defined as a triple $IS = (U, AT, f)$, where $U = \{x_1, x_2, \dots, x_n\}$ is a finite non-empty set of objects (the universe of discourse) and AT is a finite non-empty set of attributes, and $f_a : U \rightarrow V_a$ for any $a \in AT$ with V_a being the domain of an attribute a .

For any $B \subseteq AT$, there is an associated *indiscernibility relation* R_B :

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

Obviously, the relation R_B is an equivalence relation and it can generate a partition of U , denoted by $U/R_B = \{[x]_B \mid x \in U\}$. $[x]_B$ denotes the equivalence class (called Pawlak information granule). As defined by Pawlak, (U, R_B) is called a Pawlak approximation space, briefly written as (U, R) . Based on (U, R) , a pair of lower and upper approximations of $X \subseteq U$ are defined as: $\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\}$ and $\overline{R}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$, respectively.

In the view of granular computing, one calls U/R_B a *granular structure* which can be represented as $K(R_B) = \{G_{R_B}(x_1), G_{R_B}(x_2), \dots, G_{R_B}(x_n)\}$ [27]. Accordingly, a binary indiscernibility relation R_B is regarded as one of granulation methods for partitioning objects [28].

In particular, the finest granular structure on U is denoted as $K(\delta) = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$ and the coarsest one on U is denoted as $K(\omega) = \{\{x_1, x_2, \dots, x_n\}\}$.

Let us consider the scenario when we obtain information regarding a set of objects from different sources. Information from each source is collected in the form of the above information system [17], and thus a family of the single information systems with the same domain is obtained and called a *multi-source information system* which is formulated as follows.

A *multi-source information system* is defined as $MS = \{IS_i \mid IS_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i)\}$, where,

- (1) U is a finite non-empty set of objects;
- (2) AT_i is a finite non-empty set of attributes of each subsystem;
- (3) $\{V_a\}$ is the value of the attribute $a \in AT_i$; and
- (4) $f_i : U \times AT_i \rightarrow \{(V_a)_{a \in AT_i}\}$ such that for all $x \in U$ and $a \in AT_i, f_i(x, a) \in V_a$.

In particular, a multi-source decision information system is given by $MS = \{IS_i \mid IS_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i, D, g)\}$, where D is a finite non-empty set of decision attributes and $g_d : U \rightarrow V_d$ for any $d \in D$ with V_d being the domain of a decision attribute d .

In this paper, we suppose $MS = \{IS_i \mid IS_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i)\}$ is composed of q single-source information systems. If one employs a binary relation to granulate every single-source information system from a multi-source information system, one gets q granular structures: K_1, K_2, \dots, K_q . Then, we obtain a multigranulation approximation space, denoted as $F = (U, K_1, K_2, \dots, K_q)$ from a multi-source information system.

Throughout this paper, we suppose that a universe of discourse U is a non-empty finite set and a granular structure K_i is represented as $K_i = \{G_i(x_1), G_i(x_2), \dots, G_i(x_{|U|})\}$ where $|\cdot|$ represents the cardinality of K_i [27].

2.2. Basics of Dempster–Shafer theory

Dempster–Shafer theory [31] is an approach to reason under uncertainty. It enables us to combine evidence from different sources and arrives at a degree of belief. It has become an important method for the study of information fusion. To facilitate our discussion, we first briefly review some basic concepts of the theory in this section. Let U be a finite non-empty set of mutually exclusive and exhaustive hypotheses, called *the frame of discernment* (or *the universe of discourse*), and let 2^U be the power set of U .

Definition 2.1. A basic probability assignment (bpa) is a mapping $m: 2^U \rightarrow [0, 1]$ that satisfies

$$m(\emptyset) = 0, \sum_{X \subseteq U} m(X) = 1.$$

A set $X \subseteq U$ with $m(X) \neq 0$ is called a focal element of m . Let \mathcal{M} be a family of all focal elements of m . A pair of (\mathcal{M}, m) is called a belief structure on U .

Definition 2.2. The belief function is a mapping $Bel: 2^U \rightarrow [0, 1]$ that satisfies

$$Bel(X) = \sum_{X' \subseteq X} m(X'), \forall X \in 2^U.$$

Definition 2.3. The plausibility function is a mapping $Pl: 2^U \rightarrow [0, 1]$ that satisfies

$$Pl(X) = \sum_{X \cap X' \neq \emptyset} m(X'), \forall X \in 2^U.$$

According to the belief structure, a pair of belief and plausibility functions can be derived. Based on the same basic probability assignment, the belief and plausibility functions are dual, i.e., $Bel(X) = 1 - Pl(\sim X)$, where $\sim X$ is the complement of X .

$[Bel(X), Pl(X)]$ is the confidence interval which describes the uncertainty about X , and $Pl(X) - Bel(X)$ represents the level of ignorance about X . Furthermore, a belief function satisfies the following axioms:

- (1) $Bel(\emptyset) = 0$,
- (2) $Bel(U) = 1$,
- (3) $Bel(\cup_{i=1}^m X_i) \geq \sum_{\emptyset \neq J \subseteq \{1, 2, \dots, m\}} (-1)^{|J|+1} Bel(\cap_{i \in J} X_i)$ for any $X_1, X_2, \dots, X_m \subseteq U$.

When uncertain information comes from different sources, it is important to desire a consensus by combining such information. Dempster–Shafer theory uses Dempster’s combination rule for combining belief functions defined by independent bodies of evidence.

Definition 2.4. Let m_1 and m_2 be two *bpas* on U which are derived from two distinct sources. Then, the combination (or the joint) m_{12} is calculated from the aggregation of m_1 and m_2 in the following manner:

$$m_{12}(X) = \frac{\sum_{B \cap C = X} m_1(B)m_2(C)}{1 - \lambda}, \forall X \in 2^U, X \neq \emptyset,$$

where $\lambda = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$ and represents the basic probability mass associated with conflict. This is determined by the summing the products of the *bpa*’s of all sets where the intersection is null. In general, we denote $m_{12}(X)$ by $m_1 \oplus m_2(X)$ which is the so-called orthogonal sum operator of combination.

2.3. Qualitative fusion functions based on multigranulation rough sets

According to two different approximating strategies, i.e., seeking common reserving difference and seeking common rejecting difference, Qian and Liang [23] proposed two kinds of multigranulation rough sets which provide two kinds of different lower or upper fusion functions. They are optimistic lower and upper multigranulation fusion functions and pessimistic versions, which are *disjunctive and conjunctive combination rules*, respectively.

Definition 2.5. Let $MS = \{IS_i | IS_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i)\}$ be a multi-source information system. K_1, K_2, \dots, K_q are q granular structures induced by AT_1, AT_2, \dots, AT_q and $X \subseteq U$. The optimistic lower and upper approximations of X in a multigranulation rough set can be formally represented as two qualitative fusion functions, respectively,

$$\underline{\sum_{i=1}^q K_i^o(X)} = f_l^o(K_1, K_2, \dots, K_q) = \{x \in U | [x]_{K_1} \subseteq X \vee \dots \vee [x]_{K_q} \subseteq X\},$$

$$\overline{\sum_{i=1}^q K_i^o(X)} = f_u^o(K_1, K_2, \dots, K_q) = \sim \sum_{i=1}^q K_i^o(\sim X),$$

where f_l^o is called an optimistic lower qualitative fusion function, and f_u^o is called an optimistic upper qualitative fusion function. They are a kind of *disjunctive combination rules*. These two functions are used to compute the lower and upper approximations of a multigranulation rough set through fusing q granular structures. In practical applications of multigranulation rough sets, the fusion function has many different forms according to various semantics and requirements. It is noted that the essence of a so-called optimistic lower (or upper) approximation defined in Definition 2.5 is the same as weak lower (or strong upper) approximation proposed in [7,8].

Similar to the form of representation of the orthogonal sums of evidence, we denote optimistic lower fusion operator as

$$f_l^o(K_1, K_2, \dots, K_q)(X) = \underline{K}_1 \oplus \underline{K}_2 \oplus \dots \oplus \underline{K}_q(X), \quad X \subseteq U,$$

and the optimistic upper fusion operator as

$$f_u^o(K_1, K_2, \dots, K_q)(X) = \overline{K}_1 \oplus \overline{K}_2 \oplus \dots \oplus \overline{K}_q(X), \quad X \subseteq U,$$

where $\underline{K}_i(X)$ is a lower approximation of a target concept X , denoted as $\underline{K}_i(X) = \{G_i(x) \in K_i | G_i(x) \subseteq X, x \in U\}$, $\overline{K}_i(X)$ is an upper approximation of a target concept X , denoted as $\overline{K}_i(X) = \{G_i(x) \in K_i | G_i(x) \cap X \neq \emptyset, x \in U\}$ in each single granulation structure as defined in Pawlak rough sets [17], and \oplus represents the disjunctive operator of combination.

Proposition 2.1. f_l^o is disjunctive, i.e., f_l^o has a behavior of indulgent.

Proof. Suppose $X \subseteq U$, according to the definition of f_l^o , we have

$$|f_l^o(K_1, K_2, \dots, K_q)(X)| \geq |\underline{K}_i(X)|, i = \{1, 2, \dots, q\}.$$

Hence,

$$|f_l^o(K_1, K_2, \dots, K_q)(X)| \geq \max\{|\underline{K}_1(X)|, \dots, |\underline{K}_i(X)|, \dots, |\underline{K}_q(X)|\}.$$

Then, according to the definition given in [4], the optimistic operator is disjunctive. \square

Obviously, the following propositions also hold.

Proposition 2.2. f_l^o satisfies associativity.

Proposition 2.3. f_l^o satisfies commutativity.

From the above properties, we can find that the order and grouping of granular structures do not affect the result by the optimistic lower fusion operator. Correspondingly, due to the duality of the optimistic lower and upper fusion operators, the latter has the same propositions to that the former has.

Definition 2.6. Let $MS = \{S_i | S_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i)\}$ be a multi-source information system. K_1, K_2, \dots, K_q are q granular structures induced by AT_1, AT_2, \dots, AT_q and $X \subseteq U$. The pessimistic lower and upper approximations of X in a multigranulation rough set can be formally represented as two qualitative fusion functions, respectively,

$$\begin{aligned} \sum_{i=1}^q K_i^p(X) &= f_l^p(K_1, K_2, \dots, K_q)(X) = \{x \in U | [x]_{K_1} \subseteq X \wedge \dots \wedge [x]_{K_q} \subseteq X\}, \\ \sum_{i=1}^q K_i^u(X) &= f_u^p(K_1, K_2, \dots, K_q)(X) \approx \sum_{i=1}^q K_i^p(\sim X), \end{aligned}$$

where f_l^p is called a pessimistic lower qualitative fusion function, and f_u^p is called a pessimistic upper qualitative fusion function. f_l^p and f_u^p are a kind of conjunctive combination rules.

Similar to the form of representation of the orthogonal sums of evidence, we denote the pessimistic lower fusion operator as

$$f_l^p(K_1, K_2, \dots, K_q)(X) = \underline{K}_1 \otimes \underline{K}_2 \otimes \dots \otimes \underline{K}_q(X), \quad X \subseteq U,$$

and the pessimistic upper fusion operator as

$$f_u^p(K_1, K_2, \dots, K_q)(X) = \overline{K}_1 \otimes \overline{K}_2 \otimes \dots \otimes \overline{K}_q(X), \quad X \subseteq U,$$

where $\underline{K}_i(X)$ and $\overline{K}_i(X)$ are defined in each single granulation structure as defined in the Pawlak rough sets [17], and \otimes represents the conjunctive operator of combination. Similarly, it is noted that the essence of the so-called pessimistic lower (or upper) approximation defined in Definition 2.6 is the same as the strong lower (or weak upper) approximation proposed in [7,8].

Proposition 2.4. f_l^p is conjunctive.

Proof. Suppose $X \subseteq U$, according to the definition of f_l^p , we have

$$|f_l^p(K_1, K_2, \dots, K_q)(X)| \leq |\underline{K}_i(X)|, i = \{1, 2, \dots, q\}.$$

Hence,

$$|f_1^p(K_1, K_2, \dots, K_q)(X)| \leq \min\{|K_1(X)|, \dots, |K_i(X)|, \dots, |K_q(X)|\}.$$

Then, according to the definition given in [3], the pessimistic operator is conjunctive. □

Similarly, according to Definition 2.6, the following propositions hold.

Proposition 2.5. f_1^p satisfies associativity.

Proposition 2.6. f_1^p satisfies commutativity.

From the above discussions, we can see that the order and grouping of granular structures do not affect the result by the pessimistic lower operator. Due to the duality of the pessimistic lower fusion operator and the upper pessimistic fusion operator, the latter has the same properties to that the former has.

Remark 1. These two kinds of fusion operators are different from Dempster’s combination rule but they have the same purpose to combine possibly conflictive beliefs or reliable beliefs. The semantic interpretation of optimistic fusion function is that when different granular structures coming from different sources are conflict, one employs optimistic rule to fuse information, whereas the semantic interpretation of pessimistic strategy is that when evidence is not conflict, one employs pessimistic strategy to do that.

For the sake of completeness and subsequent discussions, we summarize the main result from the relationship between the rough set theory and Dempster–Shafer theory in the following theorem.

In Pawlak’s rough set models, sets are approximated by a Pawlak approximation space (U, R) . The qualities of lower and upper approximations of a set $X \subseteq U$ are defined, respectively, by

$$\underline{Q}(X) = \frac{|R(X)|}{|U|}, \quad \overline{Q}(X) = \frac{|\overline{R}(X)|}{|U|}.$$

Theorem 2.1. The qualities of lower and upper approximations defined by the above. $\underline{Q}(X)$ and $\overline{Q}(X)$ are a dual pair of belief and plausibility functions and the corresponding basic probability assignment is $m(A) = \frac{|A|}{|U|}$ for all $A \in U/R$; and 0 otherwise. Conversely, if Bel and Pl are a dual pair of belief and plausibility functions on U satisfying two conditions: (i) the set of focal elements of m is a partition of U ; (ii) $m(A) = \frac{|A|}{|U|}$ for every focal element A of m , where m is the basic probability assignment of Bel, then there exists a Pawlak approximation space (U, R) , i.e. there exists an equivalence relation R on U , such that the induced qualities of lower and upper approximations satisfy

$$\underline{Q}(X) = Bel(X), \quad \overline{Q}(X) = Pl(X), \quad \text{for all } X \subseteq U.$$

The first part of this theorem is shown by Skowron [35] and the second part is given by Yao [51].

In this paper, we denote (K, m) as a belief structure, where

$$K = \{G(x_1), G(x_2), \dots, G(x_{|U|})\} \text{ and } m(G(x_i)) = \frac{|G(x_i)|}{|U|}.$$

Obviously, each $G(x_i)$ is a focal set for $m(G(x_i)) > 0$ and K is corresponding focal sets of the belief structure (G, m) . Particularly, if focal elements $G(x_1), G(x_2), \dots, G(x_{|U|})$ satisfy $G(x_1) \subseteq G(x_2) \subseteq \dots \subseteq G(x_{|U|})$, then they are called consonant focal elements [38]. Similarly, if sets of focal elements K_1, K_2 satisfy $K_1 \preceq K_2$, i.e., if for any $S \in K_1$, there exists $L \in K_2$ such that $S \subseteq L$, then we call that K_1 is finer than K_2 (or K_2 is coarser than K_1), denoted by $K_1 \preceq K_2$. If $K_1 \preceq K_2$ and $K_1 \neq K_2$, we say that K_1 is strictly finer than K_2 (or K_2 is strictly coarser than K_1), written as $K_1 \prec K_2$, then they are called the generalized consonant focal elements.

3. The connection between Dempster–Shafer theory and MGRS theory

In the view of Dempster–Shafer theory, we suppose each granulation structure induced by an attribute set is regarded as a body of evidence in multigranulation rough sets. Therefore, the aim of fusing multiple uncertain information is equivalent to combining granulation structures from different sources. According to the relationship between Dempster–Shafer theory and Pawlak rough set theory, Yao [51] proposed a basic function assignment, i.e., $m(G(x_i)) = \frac{|G(x_i)|}{|U|}$. In what follows, we denote the optimistic multigranulation belief function and the pessimistic one as Bel_{\oplus}^o and Bel_{\otimes}^p , denote the optimistic multigranulation plausibility function and the pessimistic one as Pl_{\oplus}^o and Pl_{\otimes}^p , and denote the belief and plausibility functions by the way of D–S combination rule as Bel_{\oplus}^{DS} and Pl_{\oplus}^{DS} , respectively.

3.1. The relationship between Dempster–Shafer theory and the optimistic MGRS theory

Dempster–Shafer theory and MGRS are two different fusion methods by their combination rule. The former captures the numeric aspect and can be interpreted as the quantitative representation rule, whereas, the latter captures the non-numeric aspect of fusing uncertainty of a target concept and can be interpreted as the qualitative representation rule. Though they capture different aspects of fusion rule, they are complementary each other in data fusion. In this section, we first investigate their connection on the foundation of the relationship between rough set theory and the evidence theory which has been addressed by Wu [43,44] and Yao [51].

In the multigranulation rough set model, a target concept is approximated by a multigranulation approximation space $(U, K_1, K_2, \dots, K_q)$. The qualities of lower and upper approximations of a set $X \subseteq U$ are defined, respectively, by

$$\underline{Q}^o(X) = \frac{|\sum_{i=1}^q K_i^o(X)|}{|U|}, \tag{1}$$

$$\overline{Q}^o(X) = \frac{|\sum_{i=1}^q \overline{K_i^o(X)}|}{|U|}. \tag{2}$$

However, a pair of $\overline{Q}^o(X)$ and $\underline{Q}^o(X)$ may not always be equal to that computed by Dempster’s combination approach. It shows that the optimistic MGRS fusion operator is not equal to Dempster’s combination operator and the conjunction operator of multiple single Pawlak approximation operators. In what follows, we employ an example to illustrate these conclusions.

Example 3.1. Given a universe $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ and two granular structures K_1 and K_2 on U , where

$$K_1 = \{\{x_1, x_7\}, \{x_2, x_6\}, \{x_3, x_4, x_5\}, \{x_8\}\}$$

and

$$K_2 = \{\{x_1, x_2\}, \{x_3, x_4, x_5\}, \{x_6, x_7, x_8\}\}.$$

For a target concept $X = \{x_1, x_2, x_6, x_8\} \subseteq U$, then,

- (1) according to Eqs. (1) and (2), one obtains $\underline{Q}^o(X) = \frac{|\sum_{i=1}^2 K_i^o(X)|}{|U|} = \frac{1}{2}$ and $\overline{Q}^o(X) = \frac{|\sum_{i=1}^2 \overline{K_i^o(X)}|}{|U|} = \frac{5}{8}$,
- (2) according to Dempster’s combination rule, one obtains $Bel_{\oplus}^{DS}(X) = \frac{17}{32}$ and $Pl_{\oplus}^{DS}(X) = \frac{17}{32}$,
- (3) according to the conjunction of the Pawlak approximation operators, one gets a new approximation space $K_1 \cap K_2 = \{\{x_1\}, \{x_2\}, \{x_3, x_4, x_5\}, \{x_6\}, \{x_7\}, \{x_8\}\}$ and obtains $\underline{Q}(X) = \frac{1}{2}$ and $\overline{Q}(X) = \frac{1}{2}$.

Based on the above example, we find that the results in Theorem 2.1 do not hold between optimistic MGRS combination method and Dempster’s combination rule. It can be seen that the optimistic multigranulation fusion method is an approximation of Dempster’s combination rule presented in [38]. In what follows, a criterion will be considered to evaluate approximation with respect to the error they make.

Definition 3.1. Let (K_1, m_1) and (K_2, m_2) be two belief structures. Bel_1 and Bel_2 are two belief functions induced by two granular structures (or two distinct bodies of evidence) in a multigranulation approximation space (U, K_1, K_2) . Then the lower approximate error measure $\underline{\Delta}$ and the upper approximate error measure $\overline{\Delta}$ of $X \subseteq U$ are defined as:

$$\underline{\Delta}(X) = Bel_{\oplus}^o(X) - \underline{Q}^o(X),$$

$$\overline{\Delta}(X) = Pl_{\oplus}^o(X) - \overline{Q}^o(X).$$

The main error measure is based on the distance between the correct and approximated combination belief functions computed by the optimistic fusion rule. From the *bpa*, one can calculate the optimistic low approximate quality as follows.

Theorem 3.1. Suppose Bel_1 and Bel_2 are two belief functions induced by two granular structures (or two distinct bodies of evidence) in a multigranulation approximation space (U, K_1, K_2) and $X \subseteq U$. Then an optimistic combination belief function satisfies:

$$\underline{Q}^o(X) = \sum_{B \subseteq X} m_1(B) + \sum_{C \subseteq X} m_2(C) - \sum_{B, C \subseteq X} m_{\oplus}^o(B \cap C),$$

where

$$m_{\oplus}^o(B \cap C) = \begin{cases} \frac{1}{1 - \sum_{B, C \subseteq X} m_1(B)m_2(C)}, & K_1 \cap K_2 \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases} \text{ and } B \in K_1, C \in K_2.$$

Proof. It can be easily proved by Definition 2.5. □

Theorem 3.1 shows the relationship between the lower approximation quality in the multigranulation rough set and the belief function induced from each granular structure.

Example 3.2. (Continued from Example 3.1.) According to Theorem 3.1, we have $\underline{Q}^o(X) = \sum_{B \subseteq X} m_1(B) + \sum_{C \subseteq X} m_2(C) - \sum_{B, C \subseteq X} m_{\oplus}^o(B \cap C) = \frac{2}{8} + \frac{1}{8} + \frac{2}{8} - \frac{1}{8} = \frac{4}{8}$, which is equal to that computed by Eq. (1).

Corollary 3.1. Suppose Bel_1 and Bel_2 are two belief functions induced by two granular structures (or two distinct bodies of evidence) in a multigranulation approximation space (U, K_1, K_2) and $X \subseteq U$. Then an optimistic combination plausibility function satisfies:

$$\bar{Q}^p(X) = 1 - \left(\sum_{B \subseteq \sim X} m_1(B) + \sum_{C \subseteq \sim X} m_2(C) - \sum_{B, C \subseteq \sim X} m_{\oplus}^o(B \cap C) \right).$$

Without loss of generality, we here extend the combination rule with $q, q \geq 2$ belief functions.

Theorem 3.2. Suppose $Bel_1, Bel_2, \dots, Bel_q$ are q belief functions induced by granular structures (bodies of evidence) in a multigranulation approximation space $(U, K_1, K_2, \dots, K_q)$ and $X \subseteq U$. Then,

$$\underline{Q}^o(X) = \sum_{i=1}^q Bel_i(X) - \sum_{1 \leq i < j \leq q} m(K_i(X) \cap K_j(X)) + \dots + (-1)^{q-1} m(\cap_{i=1}^q K_i(X)).$$

Proof. It can easily be proved by the set-theoretic theory. □

Theorem 3.3. Suppose $Bel_1, Bel_2, \dots, Bel_q$ are q belief functions induced by granular structures (bodies of evidence) in a multigranulation approximation space $(U, K_1, K_2, \dots, K_q), K_1 \preceq K_2 \preceq \dots \preceq K_q$, and $X \subseteq U$. Then

$$\underline{Q}^o(X) = \frac{|K_1(X)|}{|U|}.$$

From Theorem 3.3, we can find that when belief structures satisfy the generalized consonant property, the optimistic lower approximation quality is equal to the optimistic multigranulation combination belief function, i.e., $Bel_{\oplus}^o(X) = \underline{Q}^o(X)$ and $Pl_{\oplus}^o(X) = \bar{Q}^o(X)$. Therefore, the error measures $\bar{\Delta}(X) = 0$ and $\underline{\Delta}(X) = 0$.

Theorem 3.4. Suppose $Bel_1, Bel_2, \dots, Bel_q$ are q belief functions induced by granular structures (bodies of evidence) in a multigranulation approximation space $(U, K_1, K_2, \dots, K_q), K_1 \preceq K_2 \preceq \dots \preceq K_q$, and $X_1 \subseteq X_2 \subseteq \dots \subseteq X_l \subseteq U$. Then

$$Bel_{\oplus}^o(X_1) \leq Bel_{\oplus}^o(X_2) \leq \dots \leq Bel_{\oplus}^o(X_l).$$

From the above discussions, we have only addressed the relationship between the lower approximation quality of multigranulation rough set and belief function induced from each granular structure. In what follows, we will give a sufficient and necessary condition under which the optimistic lower and upper approximate qualities can be characterized by belief and plausibility functions, respectively. For convenient discussion, we consider a special case with $|AT_i| = 1$. In other words, we will use the single-source information system to complete our discussion.

Let $IS = (U, AT, f)$ be an information system. For any $x \in U$, we define a minimal granule with respect to x as:

$$Md(x) = \{[x]_{\{a\}} | a \in AT, ([x]_{\{a'\}} \subseteq [x]_{\{a\}}, a' \neq a \Rightarrow [x]_{\{a'\}} = [x]_{\{a\}})\},$$

and denote $j(X) = \{x \in U | X \in Md(x)\}$ for $X \subseteq U$. In general, $|Md(x)| \neq 1$. In the following, we employ K_a to denote a granular structure induced by an attribute a .

It is obvious that $x \in \sum_{a \in AT} K_a^o(X)$ if and only if there exists $X' \in Md(x)$ such that $X' \subseteq X$, i.e., $\sum_{a \in AT} K_a^o(X) = \{x \in U | \exists X' \in Md(x), X' \subseteq X\}$.

Theorem 3.5. Let $IS = (U, AT, f)$ be an information system. A set function $m : 2^U \rightarrow [0, 1]$. Then,

$$m(X) = \frac{1}{|U|} \sum_{x \in j(X)} \frac{1}{|Md(x)|}$$

is a basic probability assignment.

Proof.

- (1) Obviously, $m(\emptyset) = 0$.
- (2) Since $x \in j(X) \iff X \in Md(x)$ for $x \in U$ and $X \subseteq U$, we have

$$m(X) = \frac{1}{|U|} \sum_{x \in j(X)} \frac{1}{|Md(x)|} = \frac{1}{|U|} \sum_{X \in Md(x)} \frac{1}{|Md(x)|}.$$

Hence,

$$\sum_{X \subseteq U} m(X) = \sum_{X \subseteq U} \left(\frac{1}{|U|} \sum_{x \in j(X)} \frac{1}{|Md(x)|} \right) = \frac{1}{|U|} \sum_{X \subseteq U} \left(\sum_{X \in Md(x)} \frac{1}{|Md(x)|} \right) = \frac{1}{|U|} \sum_{x \in U} \left(\sum_{X \in Md(x)} \frac{1}{|Md(x)|} \right).$$

In fact, $\sum_{X \in Md(x)} \frac{1}{|Md(x)|} = 1$, so we have

$$\sum_{X \subseteq U} m(X) = \frac{1}{|U|} \sum_{x \in U} \left(\sum_{X \in Md(x)} \frac{1}{|Md(x)|} \right) = \frac{1}{|U|} \sum_{x \in U} 1 = \frac{|U|}{|U|} = 1.$$

Therefore, we can conclude that m is a basic probability assignment. \square

Particularly, if $|Md(x)| = 1$, then $m(X) = \frac{|j(X)|}{|U|}$ for $x \in U, X \subseteq U$.

The following theorem presents a formula to measure the optimistic lower and upper approximations of MGRS by belief and plausibility functions under a special condition.

Theorem 3.6. Suppose $IS = (U, AT, f)$ is an information system. If $|Md(x)| = 1$ for all $x \in U$ and $X \subseteq U$, then we have

$$Bel_{\oplus}^o(X) = \frac{1}{|U|} \left| \sum_{a \in AT} K_a^o(X) \right|$$

and

$$Pl_{\oplus}^o(X) = \frac{1}{|U|} \left| \sum_{a \in AT} \overline{K_a^o(X)} \right|.$$

Proof. Since $|Md(x)| = 1$ for all $x \in U$, we have $m(X) = \frac{1}{|U|} \sum_{x \in j(X)} \frac{1}{|Md(x)|} = \frac{1}{|U|} \sum_{x \in j(X)} 1 = \frac{|j(X)|}{|U|}$. Hence, $Bel_{\oplus}^o(X) = \sum_{X' \subseteq X} m(X') = \sum_{X' \subseteq X} \frac{|j(X')|}{|U|} = \frac{\sum_{X' \subseteq X} |j(X')|}{|U|}$. For the further proof, we first verify that $j(X) \cap j(X') = \emptyset$ for any $X \neq X', X, X' \in U$. Assume $x \in j(X) \cap j(X')$. Then $X \in Md(x)$ and $X' \in Md(x)$. With the fact of $|Md(x)| = 1$, we have $X = X'$, which contradicts to $X \neq X'$. Thus $j(X) \cap j(X') = \emptyset$ for any $X \neq X', X, X' \in U$. Therefore, $Bel_{\oplus}^o(X) = \frac{\sum_{X' \subseteq X} |j(X')|}{|U|} = \frac{|\cup_{X' \subseteq X} j(X')|}{|U|}$. Since $|Md(x)| = 1$ for all $x \in U$, we can write $Md(x) = \{n(x)\}$ for all $x \in U$. Thus $n(x) \subseteq U$ is the minimal element in $\{[x]_a | a \in AT\}$. On the other hand, we have $j(X') = \{x \in U | X' \in Md(x)\} = \{x \in U | n(x) = X'\}$, thus $\cup_{X' \subseteq X} j(X') = \{x \in U | n(x) \subseteq X\}$. Therefore, $Bel_{\oplus}^o(X) = \frac{|\cup_{X' \subseteq X} j(X')|}{|U|} = \frac{|\{x \in U | n(x) \subseteq X\}|}{|U|}$. In addition, we can easily conclude that $n(x) \subseteq X \iff x \in \sum_{a \in AT} \overline{K_a^o(X)}$ for any $x \in U, X \subseteq U$. Therefore, $Bel_{\oplus}^o(X) = \frac{|\{x \in U | n(x) \subseteq X\}|}{|U|} = \frac{1}{|U|} |\sum_{a \in AT} \overline{K_a^o(X)}|$.

Similarly, we can prove $Pl_{\oplus}^o(X) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X)|$. \square

From the above conclusions, under the condition that $|Md(x)| = 1$ for all $x \in U$, the optimistic multigranulation approximations can be measured by belief and plausibility functions. However, we wonder whether the condition $|Md(x)| = 1$ is necessary or not. We first give a lemma.

Lemma 3.1. Let $IS = (U, AT, f)$ be an information system and $X \subseteq U$. Then $Bel_{\oplus}^o(X) = \frac{1}{|U|} \sum_{x \in \sum_{a \in AT} \overline{K_a^o(X)}} \frac{|[X' \in Md(x) | X' \subseteq X]|}{|Md(x)|}$.

Proof. According to the definition of belief function, we have

$$Bel_{\oplus}^o(X) = \sum_{X' \subseteq X} m(X') = \sum_{X' \subseteq X} \left(\frac{1}{|U|} \sum_{x \in j(X')} \frac{1}{|Md(x)|} \right) = \frac{1}{|U|} \sum_{X' \subseteq X} \left(\sum_{x \in j(X')} \frac{1}{|Md(x)|} \right) = \frac{1}{|U|} \sum_{X' \subseteq X} \left(\sum_{X' \in Md(x)} \frac{1}{|Md(x)|} \right) = \frac{1}{|U|} \sum_{x \in U} \left(\sum_{X' \in Md(x), X' \subseteq X} \frac{1}{|Md(x)|} \right).$$

On the other hand, it follows $\sum_{X' \in Md(x), X' \subseteq X} \frac{1}{|Md(x)|} = \frac{|[X' \in Md(x) | X' \subseteq X]|}{|Md(x)|}$. Thus,

$Bel_{\oplus}^o(X) = \frac{1}{|U|} \sum_{x \in U} (\sum_{X' \in Md(x), X' \subseteq X} \frac{1}{|Md(x)|}) = \frac{1}{|U|} \sum_{x \in U} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|}$. Moreover, with the fact of $x \in \sum_{a \in AT} K_a^o(X) \iff \exists X' \in Md(x), X' \subseteq X$, it is easy to verify that $|\{X' \in Md(x) | X' \subseteq X\}| \neq 0 \iff x \in \sum_{a \in AT} K_a^o(X)$. Therefore,

$$\begin{aligned} Bel_{\oplus}^o(X) &= \frac{1}{|U|} \sum_{x \in U} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|} = \frac{1}{|U|} \sum_{x \in \sum_{a \in AT} K_a^o(X)} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|} + \frac{1}{|U|} \sum_{x \notin \sum_{a \in AT} K_a^o(X)} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|} \\ &= \frac{1}{|U|} \sum_{x \in \sum_{a \in AT} K_a^o(X)} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|}. \quad \square \end{aligned}$$

The following result tells us $|Md(x)| = 1$ for all $x \in U$. It is a basic condition to guarantee that the qualities of the optimistic lower and upper approximations of MGRS can be measured by the belief and plausibility functions.

Theorem 3.7. Let $IS = (U, AT, f)$ be an information system. If $Bel_{\oplus}^o(X) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X)|$ and $Pl_{\oplus}^o(X) = \frac{1}{|U|} |\sum_{a \in AT} \overline{K_a^o(X)}|$ for $X \subseteq U$, then $|Md(x)| = 1$ for all $x \in U$.

Proof. Suppose, by contradiction, that $|Md(x_0)| > 1$ for some $x_0 \in U$, and $X \in Md(x_0)$. By Lemma 3.1, $Bel_{\oplus}^o(X) = \frac{1}{|U|} \sum_{x \in \sum_{a \in AT} K_a^o(X)} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|}$. We can see that $\frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|} \leq 1$ for all $x \in U$. Since $X \in Md(x_0)$, we can verify that $\{X' \in Md(x_0) | X' \subseteq X\} = \{X\}$. That means $|\{X' \in Md(x_0) | X' \subseteq X\}| = 1$. Besides, with the fact of $|Md(x_0)| \geq 1$ we have that $\frac{|\{X' \in Md(x_0) | X' \subseteq X\}|}{|Md(x_0)|} < 1$. Consequently,

$$Bel_{\oplus}^o(X) = \frac{1}{|U|} \sum_{x \in \sum_{a \in AT} K_a^o(X)} \frac{|\{X' \in Md(x) | X' \subseteq X\}|}{|Md(x)|} < \frac{1}{|U|} \sum_{x \in \sum_{a \in AT} K_a^o(X)} 1 = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X)|.$$

It is equal to $Bel_{\oplus}^o(X) < \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X)|$, which contradicts to $Bel_{\oplus}^o(X) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X)|$.

Therefore, we finish the proof. \square

By the above analysis, we have the properties as follows.

Corollary 3.2. Let $IS = (U, AT, f)$ be an information system. Then the following conclusions are equal to each other.

- (1) $|Md(x)| = 1$ for all $x \in U$;
- (2) $Bel_{\oplus}^o(X) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X)|$ and $Pl_{\oplus}^o(X) = \frac{1}{|U|} |\sum_{a \in AT} \overline{K_a^o(X)}|$ for all $X \subseteq U$.

Proof. By Theorems 3.6 and 3.7, it holds. \square

Let us employ an example to characterize the optimistic multigranulation approximations by belief and plausibility functions.

Example 3.3. Let $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ be a universe of discourse. K_1, K_2, K_3 and K_4 are four granular structures on U , where $K_1 = \{\{x_1, x_3, x_4\}, \{x_2, x_6\}, \{x_5\}\}$, $K_2 = \{\{x_1, x_3\}, \{x_2, x_5, x_6\}, \{x_4\}\}$, $K_3 = \{\{x_1, x_3, x_4\}, \{x_2, x_5, x_6\}\}$, and $K_4 = \{\{x_1, x_3\}, \{x_2, x_6\}, \{x_4, x_5\}\}$. Then one gets the minimal granule of x_i , that is, $Md(x_1) = \{\{x_1, x_3\}\}$, $Md(x_2) = \{\{x_2, x_3\}\}$, $Md(x_3) = \{\{x_1, x_3\}\}$, $Md(x_4) = \{\{x_4\}\}$, $Md(x_5) = \{\{x_5\}\}$, and $Md(x_6) = \{\{x_2, x_6\}\}$. We can find that $|Md(x)| = 1$ for all $x \in U$. So the optimistic multigranulation approximations can be measured by belief and plausibility functions.

We randomly choose three target concepts $X_1 = \{x_2, x_5, x_6\}$, $X_2 = \{x_1, x_2, x_4\}$, and $X_3 = \{x_2, x_3, x_6\}$. Based on the definition of optimistic multigranulation approximations, we can obtain that

$$\sum_{a \in AT} K_a^o(X_1) = \{x_5, x_6\}, \sum_{a \in AT} K_a^o(X_2) = \{x_4\} \text{ and } \sum_{a \in AT} K_a^o(X_3) = \{x_2, x_6\}. \text{ Then, } \\ Bel_{\oplus}^o(X_1) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X_1)| = \frac{1}{3}, Bel_{\oplus}^o(X_2) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X_2)| = \frac{1}{6}, \text{ and } Bel_{\oplus}^o(X_3) = \frac{1}{|U|} |\sum_{a \in AT} K_a^o(X_3)| = \frac{1}{3}.$$

3.2. The relationship between Dempster–Shafer theory and the pessimistic MGRS theory

Similar to the discussion of the relationship between Dempster–Shafer theory and the optimistic MGRS, we also examine it between Dempster–Shafer theory and the pessimistic MGRS.

In pessimistic multigranulation rough set model, a target concept is approximated by a multigranulation approximation space $(U, K_1, K_2, \dots, K_q)$. The qualities of lower and upper approximations of a set $X \subseteq U$ are defined, respectively, by

$$\underline{Q}^p(X) = \frac{|\sum_{i=1}^q K_i^p(X)|}{|U|}, \tag{3}$$

$$\bar{Q}^p(X) = \frac{|\sum_{i=1}^q \overline{K_i^p(X)}|}{|U|}. \tag{4}$$

Similar to the results of Yao' study [51], they are a dual pair of belief and plausibility functions. The main results are summarized in the following theorem.

Theorem 3.8. *The qualities of lower and upper approximations defined by Eqs. (3) and (4), $\underline{Q}^p(X)$ and $\bar{Q}^p(X)$ are a dual pair of belief and plausibility functions and the corresponding basic probability assignment is $m(j(A))$, where $j(A)$ is any focal element which is the possible intersections between pairs of focal element of $K_i, i = \{1, 2, \dots, q\}$, respectively, for all $j(A) \subseteq U$; and 0 otherwise. Conversely, if Bel_{\otimes}^p and Pl_{\otimes}^p are a dual pair of belief and plausibility functions on U satisfying two conditions: (i) the set of focal elements of m_{\otimes}^p is a partition of U ; (ii) $m_{\otimes}^p(j(A))$ for every focal element A of m_{\otimes}^p , where m is the basic probability assignment of Bel_{\otimes} , then there exists a Pawlak approximation space (U, R) , i.e. there exists an equivalence relation R on U , such that the induced qualities of lower and upper approximations of $X \subseteq U$ satisfy*

$$\underline{Q}^p(X) = Bel_{\otimes}^p(X), \tag{5}$$

$$\bar{Q}^p(X) = Pl_{\otimes}^p(X). \tag{6}$$

Proof. According to (2) of Theorem 8 from [25], i.e., $\underline{Q}^p(\emptyset) = \emptyset$ and $\underline{Q}^p(U) = U$, hence, $Bel_{\otimes}^p(\emptyset) = 0$ and $Bel_{\otimes}^p(U) = 1$, respectively. Considering a collection $X_1, X_2, \dots, X_n \subseteq U$, we have

$$\begin{aligned} \underline{Q}^p(X_1 \cup \dots \cup X_n) &= \frac{|\sum_{i=1}^q \overline{K_i^p(X_1 \cup \dots \cup X_n)}|}{|U|} \geq \frac{|\sum_{i=1}^q \overline{K_i^p(X_1)} \cup \dots \cup \sum_{i=1}^q \overline{K_i^p(X_n)}|}{|U|} \\ &= \sum_{i=1}^n \frac{|\sum_{i=1}^q \overline{K_i^p(X_i)}|}{|U|} - \sum_{i < j} \frac{|\sum_{i=1}^q \overline{K_i^p(X_i)} \cap \sum_{i=1}^q \overline{K_i^p(X_j)}|}{|U|} \pm \dots (-1)^{n+1} \frac{|\sum_{i=1}^q \overline{K_i^p(X_1)} \cap \dots \cap \sum_{i=1}^q \overline{K_i^p(X_n)}|}{|U|} \\ &= \sum_{i=1}^n \frac{|\sum_{i=1}^q \overline{K_i^p(X_i)}|}{|U|} - \sum_{i < j} \frac{|\sum_{i=1}^q \overline{K_i^p(X_i \cap X_j)}|}{|U|} \pm \dots (-1)^{n+1} \frac{|\sum_{i=1}^q \overline{K_i^p(X_1 \cap \dots \cap X_n)}|}{|U|}. \end{aligned}$$

Hence, $Bel_{\otimes}^p(X)$ is a belief function.

According to the duality of the belief and plausibility functions, one can prove that $Pl_{\otimes}^p(X)$ is a plausibility function. \square

It is noted that Theorem 3.8 offers an interpretation of the Dempster's combination rule.

Theorem 3.9. *Suppose Bel_1 and Bel_2 are two belief functions induced by two granular structures (or two distinct bodies of evidence) in a multigranulation approximation space (U, K_1, K_2) and $X \subseteq U$. Then a pessimistic combination belief function is defined as:*

$$Bel_{\otimes}^p(X) = \sum_{j(A) \subseteq X} m_{\otimes}^p(j(A)) - \frac{1}{2} \sum_{A \cap B = j(A)} m_{\otimes}^p(j(A)),$$

where

$$m_{\otimes}^p(B \cap C) = \begin{cases} \frac{1}{1 - \sum_{B, C \subseteq X} m_1(B)m_2(C)}, & K_1 \cap K_2 \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Similar to the proof of Theorem 3.2, it can be proved. \square

Theorem 3.9 gives a quantitative presentation for the qualitative pessimistic combination rule, which may be intuitional in the engineering field.

Example 3.4. Given a universe $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ and two granular structures K_1 and K_2 on U , where $K_1 = \{\{x_1, x_7\}, \{x_2, x_6\}, \{x_3, x_4, x_5\}, \{x_8\}\}$ and $K_2 = \{\{x_1, x_2\}, \{x_3, x_4, x_5\}, \{x_6, x_7, x_8\}\}$. For a target concept $X = \{x_1, x_2, x_6, x_8\}$, then, according to Theorem 3.8, $Bel_{\otimes}^p(X) = \frac{|\sum_{i=1}^2 \overline{K_i^p(X)}|}{|U|} = \frac{1}{8}$ and $Pl_{\otimes}^p(X) = \frac{|\sum_{i=1}^2 \overline{K_i^p(X)}|}{|U|} = \frac{5}{8}$. By Theorem 3.9, one has the same value as the above.

Corollary 3.3. *Suppose Bel_1 and Bel_2 are two belief functions induced by two granular structures (or two distinct bodies of evidence) in a multigranulation approximation space (U, K_1, K_2) and $X \subseteq U$. Then a pessimistic combination plausibility function is defined as:*

$$P_{\otimes}^p(\sim X) = 1 - \left(\sum_{j(A) \subseteq \sim X} m_{\otimes}^p(j(A)) - \frac{1}{2} \sum_{A \cap B = j(A)} m_{\otimes}^p(j(A)) \right).$$

If an approach used to combine the granular structures is based on the original Dempster's combination rule, the results of combination are unequal to that based on the quality method in the case that bodies of evidence are inconsistent. For this reason, we should reconsider how to represent the combination rule.

Theorem 3.10. Suppose $Bel_1, Bel_2, \dots, Bel_q$ are q belief functions induced by granular structures (bodies of evidence) in a multigranulation approximation space $(U, K_1, K_2, \dots, K_q), K_1 \preceq K_2 \preceq \dots \preceq K_q$, and $X \subseteq U$. Then

$$\underline{Q}^p(X) = \frac{|K_q(X)|}{|U|}.$$

From Theorem 3.10, we can find that when belief structures satisfy the generalized consonant property, the approximate quality obtained by the pessimistic combination rule is the combination belief function.

Theorem 3.11. Suppose $Bel_1, Bel_2, \dots, Bel_q$ are q belief functions induced by granular structures (bodies of evidence) in a multigranulation approximation space $(U, K_1, K_2, \dots, K_q), K_1 \preceq K_2 \preceq \dots \preceq K_q$, and $X_1 \subseteq X_2 \subseteq \dots \subseteq X_l \subseteq U$. Then

$$Bel_{\otimes}^p(X_1) \leq Bel_{\otimes}^p(X_2) \leq \dots \leq Bel_{\otimes}^p(X_l).$$

4. A two-grade fusion approach by combining Dempster–Shafer theory and MGRS theory

In order to differentiate conflict and reliable evidence, we first introduce the granulation distance for characterizing the difference among granular structures on the multigranulation space and then use the distance to cluster the conflict and reliable evidence.

In what follows, we give a definition of the distance among granular structures based on Liang's distance [12].

Definition 4.1. Let $F = (U, K_1, K_2, \dots, K_q)$ be a multigranulation space. Suppose $K_i, K_j \in F$ are two granular structures, where $K_i = \{G_i(x_l) | x_l \in U\}$ and $K_j = \{G_j(x_l) | x_l \in U\}$. Granulation distance between K_i and K_j is defined as

$$d(K_i, K_j) = \frac{1}{|U|^2} \sum_{l=1}^{|U|} |G_i(x_l) \oplus G_j(x_l)|,$$

where $|G_i(x_l) \oplus G_j(x_l)| = |G_i(x_l) \cup G_j(x_l)| - |G_i(x_l) \cap G_j(x_l)|$.

From the definition of distance, granulation distance obtains the maximum value $\frac{|U|-1}{|U|}$ if and only if $K_i = K(\omega) = \{\{x_1, x_2, \dots, x_n\}\}$ and $K_j = K(\delta) = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$ (or $K_i = K(\delta)$ and $K_j = K(\omega)$), which illustrates that evidence has complete conflict. Whereas granulation distance obtains the minimum value 0 if and only if $K_i = K_j$, which illustrates that evidence has complete reliable.

Definition 4.2. Let $F = (U, K_1, K_2, \dots, K_q)$ be a multigranulation space. Suppose $K_i, K_j \in F$ are two granular structures, where $K_i = \{G_i(x_l) | x_l \in U\}$ and $K_j = \{G_j(x_l) | x_l \in U\}$. Similarity between K_i and K_j is defined as

$$Sim(K_i, K_j) = 1 - d(K_i, K_j).$$

Definition 4.3. Let $F = (U, K_1, K_2, \dots, K_q)$ be a multigranulation space. Suppose $K_i, K_j \in F$ are two granular structures, where $K_i = \{G_i(x_l) | x_l \in U\}$ and $K_j = \{G_j(x_l) | x_l \in U\}$. Similarity matrix of F is defined as

$$S_{q \times q} = (s_{ij})_{q \times q},$$

where $s_{ij} = Sim(K_i, K_j)$.

Example 4.1. (Continued from Example 3.3.) In order to calculate the distance between two granular structures, we here use the uniform representation to rewrite K_1, K_2 , and K_3 as:

$$K_1 = \{\{x_1, x_3, x_4\}, \{x_2, x_6\}, \{x_1, x_3, x_4\}, \{x_1, x_3, x_4\}, \{x_5\}, \{x_2, x_6\}\},$$

$$K_2 = \{\{x_1, x_3\}, \{x_2, x_5, x_6\}, \{x_2, x_5, x_6\}, \{x_1, x_3\}, \{x_4\}, \{x_2, x_5, x_6\}, \{x_2, x_5, x_6\}\},$$

and

$$K_3 = \{\{x_1, x_3, x_4\}, \{x_2, x_5, x_6\}, \{x_1, x_3, x_4\}, \{x_1, x_3, x_4\}, \{x_2, x_5, x_6\}, \{x_2, x_5, x_6\}\}.$$

Then we get $d(K_1, K_2) = \frac{11}{36}$, $d(K_1, K_3) = \frac{4}{36}$, and $d(K_2, K_3) = \frac{7}{36}$. Accordingly, we have $Sim(K_1, K_2) = 1 - d(K_1, K_2) = \frac{25}{36}$, $Sim(K_1, K_3) = 1 - d(K_1, K_3) = \frac{32}{36}$, and $Sim(K_2, K_3) = 1 - d(K_2, K_3) = \frac{29}{36}$. Then the similarity matrix is

$$\begin{pmatrix} 1 & \frac{25}{36} & \frac{32}{36} \\ \frac{25}{36} & 1 & \frac{29}{36} \\ \frac{32}{36} & \frac{29}{36} & 1 \end{pmatrix}.$$

In light of the proposed evidence similarity as a criterion [6], we can justify whether two bodies of evidence are conflict or not. If $0 \leq Sim(K_i, K_j) \leq \frac{1}{2}$, the bodies of evidence K_i and K_j are conflict. If $\frac{1}{2} < Sim(K_i, K_j) \leq 1$, the bodies of evidence K_i and K_j are reliable.

Based on this criterion and the *k*-modes clustering algorithm [5], *q* granular structures K_1, K_2, \dots, K_q can be grouped into *k* clusters. Granular structures in the same cluster are reliable and those in different clusters are conflict.

Definition 4.4. Let $MS = \{IS_i | IS_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i)\}$ be a multi-source information system. K_1, K_2, \dots, K_q are *q* granular structures induced by AT_1, AT_2, \dots, AT_q , where $K_i = \{G_{i1}, G_{i2}, \dots, G_{i|U|}\}$. Then, a two-grade fusion function is formally defined as

$$f(K_1, K_2, \dots, K_q) = M_1^{\otimes} \oplus M_2^{\otimes} \oplus \dots \oplus M_k^{\otimes},$$

where M_i^{\otimes} is the combination result of every cluster and \oplus is the orthogonal sum for the *k* clusters.

Our combination rule is parallel to the quantitative combination rule proposed by Wong and Lingras [42], in which the combined belief function is derived by minimizing the entropy function.

Algorithm 4.1. A two-grade fusion algorithm

-
- Input: A multi-source information system.
 - Output: Belief and plausibility measures.
 - Step 1. A family of attribute sets induce multiple granular structures, denoted by K_1, K_2, \dots, K_q .
 - Step 2. Based on the relationship between the classical rough set theory and Dempster–Shafer theory, the assignment function can be obtained, which is defined as $m(G_j) = \frac{|G_j|}{|U|}$.
 - Step 3. Computing the evidence distance of granular structures and establishing evidence similarity matrix.
 - Step 4. According to the evidence distance of granular structures and the clustering algorithm, we classify multiple bodies of evidence into *k* groups: granular structures in the same class are reliable and those in different class are conflict.
 - Step 5. Based on the optimistic and pessimistic combination rules, we establish a combination fusion function.
 - Step 6. Output the belief and plausibility measures.
-

The architecture of the above algorithm is shown as Fig.1.

5. Example

In what follows, we employ an example to illustrate the effectiveness of the proposed combination rules.

Example 5.1. Let $MS = \{IS_i | IS_i = (U, AT_i, \{(V_a)_{a \in AT_i}\}, f_i, \{d\})\}$ be a multi-source decision information system, where $U = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ is a universe of six objects which are here regarded as patients. Suppose there are four hospitals ($H_i, i = \{1, 2, 3, 4\}$) providing us information regarding the attributes *A, B, C* of the objects. These attributes represent the patients’ physical examination indicators. *D* is the decision attribute with attribute values $\{+, -\}$, where + represents the patient has some disease and –the converse. Table 1 depicts the information provided by the four hospitals.

Firstly, by steps 1, 2, and 3 of Algorithm 4.1, we have the following results.

Each H_i gives rise to a granular structure K_i ($i \in \{1, 2, 3, 4\}$) as follows:

$$K_1 = \{\{e_1\}, \{e_2, e_5\}, \{e_3, e_4\}, \{e_6\}\}, K_2 = \{\{e_1, e_3\}, \{e_2, e_4, e_5\}, \{e_6\}\},$$

$$K_3 = \{\{e_1\}, \{e_2, e_3, e_5\}, \{e_4, e_6\}\}, \text{ and } K_4 = \{\{e_1, e_2, e_4, e_5, e_6\}, \{e_3\}\}.$$

In addition, we can get a granular structure induced by all attributes, which is

$K = \{\{e_1\}, \{e_2, e_5\}, \{e_3\}, \{e_4\}, \{e_6\}\}$. As a result, one obtains five belief structures based on the relationship between DS theory and rough set theory: $\vec{m}_1 = \{\frac{1}{6}, \frac{2}{6}, \frac{2}{6}, \frac{1}{6}\}$, $\vec{m}_2 = \{\frac{2}{6}, \frac{3}{6}, \frac{1}{6}\}$, $\vec{m}_3 = \{\frac{1}{6}, \frac{3}{6}, \frac{2}{6}\}$, $\vec{m}_4 = \{\frac{5}{6}, \frac{1}{6}\}$, and $\vec{m} = \{\frac{1}{6}, \frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\}$.

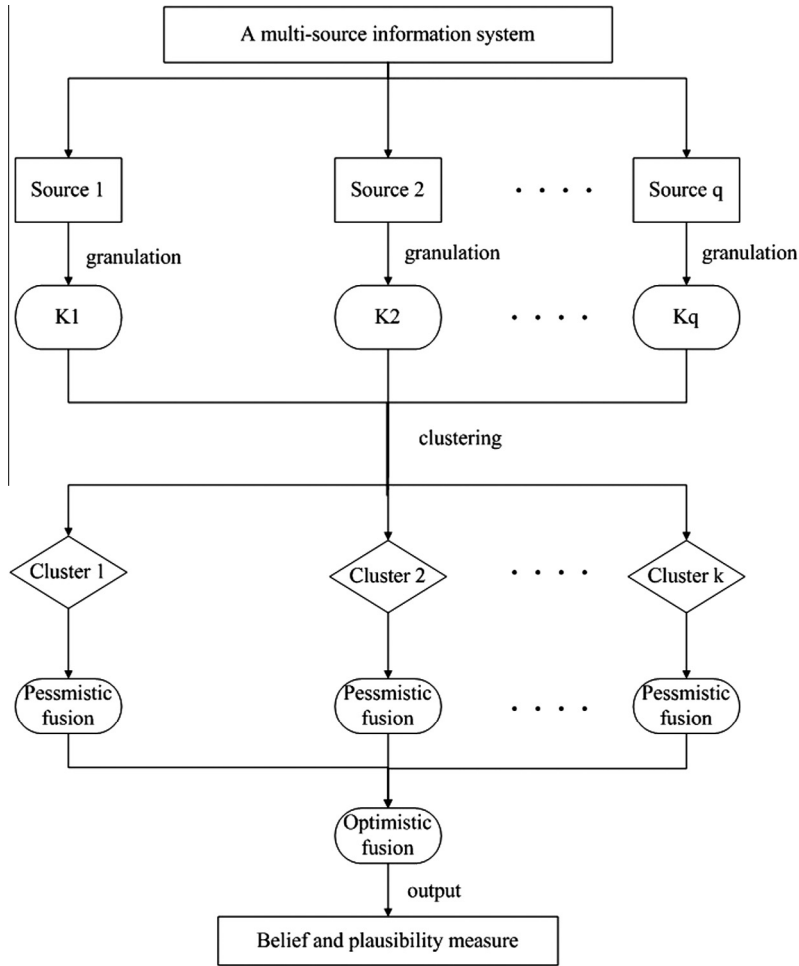


Fig. 1. The architecture of Algorithm 4.1.

Table 1
A multi-source decision information system.

	H_1			H_2			H_3			H_4			D
	A	B	C	A	B	C	A	B	C	A	B	C	
e_1	b	b	a	b	b	b	b	b	a	a	a	b	+
e_2	a	a	b	b	b	a	b	a	c	a	a	b	-
e_3	a	c	a	b	b	b	b	a	c	b	c	a	+
e_4	a	c	a	b	b	a	b	c	a	a	a	b	-
e_5	a	a	b	b	b	a	b	a	c	a	a	b	+
e_6	c	a	a	c	a	c	b	c	a	a	a	b	-

Furthermore, a family of decision classes $U/\{d\} = \{D_1, D_2\}$, where $D_1 = \{e_1, e_3, e_5\}$ and $D_2 = \{e_2, e_4, e_6\}$. In the following, we not only calculate a belief function value and a plausibility function value by Dempster's combination rule, Pawlak rough set, the pessimistic combination rule, and the proposed combination rule, respectively, but also we do that by the optimistic combination rule.

- (1) According to the Dempster's approach, one has $Bel_{\oplus}^o(D_1) = \frac{1}{17}$ and $Pl_{\oplus}^o(D_1) = \frac{10}{17}$.
- (2) According to the Pawlak rough set approach, one has $\underline{K}(D_1) = \{e_1, e_3\}$, $\overline{K}(D_1) = \{e_1, e_2, e_3, e_5\}$. Hence, $Bel(D_1) = \frac{2}{6}$ and $Pl(D_1) = \frac{4}{6}$.
- (3) According to the pessimistic combination rule, $\sum_{i=1}^4 K_i^p(D_1) = \emptyset$ and $\overline{\sum_{i=1}^4 K_i^p(D_1)} = \{e_1, e_2, e_3, e_4, e_5, e_6\}$. Hence, $Bel_{\otimes}^p(D_1) = 0$ and $Pl_{\otimes}^p(D_1) = 1$.

(4) According to the proposed combination rule, we first compute the granulation distance between them. They are as follows.

$$d(K_1, K_2) = \frac{8}{36}, d(K_1, K_3) = \frac{8}{36}, d(K_1, K_4) = \frac{20}{36}, d(K_2, K_3) = \frac{12}{36}, d(K_2, K_4) = \frac{16}{36}, \text{ and } d(K_3, K_4) = \frac{20}{36}.$$

Secondly, we employ a clustering algorithm based on the granulation distance to group the granular structures into two clusters: $\{K_1, K_2, K_3\}$ and $\{K_4\}$. We can utilize the pessimistic combination rule to combine the granular structures in one cluster and utilize optimistic combination rule to combine the granular structures in different clusters. Hence, we obtain $\sum_{i=1}^3 K_i^p(D_1) = \{e_1\}$ and $\overline{\sum_{i=1}^3 K_i^p(D_1)} = \{e_1, e_2, e_3, e_4, e_5\}$. Hence, $Bel_{\otimes}^p(D_1) = \frac{2}{6}$ and $Pl_{\otimes}(D_1) = \frac{5}{6}$, then combining with the cluster $\{K_4\}$, one has $\sum_{i=1}^3 G_i^p(D_1) \oplus K_4(D_1) = \{e_1, e_3\}$ and $\overline{\sum_{i=1}^3 K_i^p(D_1)} \oplus \overline{K_4(D_1)} = \{e_1, e_2, e_3, e_4, e_5\}$. Hence, the combination result is $Bel(K_1, K_2, K_3, K_4)(D_1) = \frac{2}{6}$, and $Pl(K_1, K_2, K_3, K_4)(D_1) = \frac{5}{6}$.

(5) According to the optimistic combination rule, $\underline{Q}^o(D_1) = \{e_1, e_3\}$ and $\overline{Q}^o(D_1) = \{e_1, e_2, e_3, e_5\}$. Hence, $\underline{Q}^o(D_1) = \frac{2}{6}$ and $\overline{Q}^o(D_1) = \frac{4}{6}$.

(6) For a new object $e = (a, a, b; c, a, c; b, a, c; b, c, a)$, according to K_1, K_2, K_3, K_4 , then we have the description of the object e by four granular structures: $[e]_{H_1} = \{e_2, e_5\}$ induced by the source H_1 , $[e]_{H_2} = \{e_6\}$ induced by the source H_2 , $[e]_{H_3} = \{e_2, e_3\}$ induced by the source H_3 , $[e]_{H_4} = \{e_3\}$ induced by the source H_4 . Based on the results of clustering, one can obtain $\sum_{i=1}^4 G_i(D_1)_e = (\emptyset \cap \emptyset \cap \{e_3\}) \cup \{e_3\} = \{e_3\}$, hence, $Bel_e(D_1) = \frac{1}{6}$. Similarly, $\sum_{i=1}^4 K_i(D_2)_e = (\emptyset \cup \{e_6\} \cup \{e_2\}) \cap \emptyset = \emptyset$, hence, $Bel_e(D_2) = 0$. Therefore, based on the maximization rule, the new object e belongs to the decision class D_1 .

$\sum_{i=1}^4 K_i(D_1)_e = (\{e_2, e_5\} \cap \{e_6\} \cap \{e_2, e_3\}) \cup \{e_3\} = \{e_3\}$, hence, $Bel_e(D_1) = \frac{1}{6}$. Similarly, $\sum_{i=1}^4 K_i(D_2)_e = (\{e_2, e_5\} \cup \{e_6\} \cup \{e_2, e_3\}) \cap \emptyset = \emptyset$, hence, $Bel_e(D_2) = 0$. Therefore, based on the maximization rule, the new object e belongs to the decision class D_1 .

Remark 2. From this example, we can find that: (1) if bodies of evidence are conflicting, the belief fusion value obtained by the proposed combination rule is no more than that computed by the optimistic combination rule, and no less than that done by the pessimistic combination rule. It seems to be more reasonable in human reasoning and problem solving; (2) the belief fusion value is close to that calculated by the single Pawlak rough set method, which illustrates that the proposed method may be used to approximate the Pawlak rough set method when it fails to deal with large scale data.

6. Conclusion

In this paper, we proposed the concept of granulation fusion which is one of important issues in the field of granular computing. Multigranulation rough set theory has provided a qualitative fusing method with no demand of prior information. The existing multigranulation fusion functions, optimistic and pessimistic multigranulation fusion functions, are too relaxed or restrictive for data analysis. Dempster's rule of combination and its improved rules have been employed as a major method for reasoning with multiple highly conflicting evidence. One may capture prior information as a mass function according to the relationship between the single granulation Pawlak rough set and the evidence theory. We have addressed the connection between the multigranulation rough set theory and Dempster–Shafer theory. A two-grade fusion approach involved in evidence theory and multigranulation rough set theory is proposed and it is based on a well-defined distance function between granulation structures. Finally, an example has been given to show the effectiveness of the proposed fusion method. The results of this study will be useful for pooling the uncertain data from a multi-source information system and significant for establishing a new direction of granular computing.

Acknowledgements

The authors would like to thank the anonymous reviewers and the editor for their constructive and valuable comments. This work is supported by grants from National Natural Science Foundation of China under Grant (Nos. 71031006, 61322211, 61379021, 11061004, 61303131, 61432011), National Key Basic Research and Development Program of China (973) (No. 2013CB329404), Doctoral Program of Higher Education (No. 20121401110013), Innovative Talents of Higher Learning Institutions of Shanxi, China (No. 20120301), and Education Committee of Fujian Province under Grant (Nos. JA13204, JA13202, JK2013027, 2013J01028, 2013J01029).

References

[1] S.B. Cho, J.H. Kim, Combining multiple neural networks by fuzzy integral for robust classification, *IEEE Trans. Syst. Man Cybernet.* 25 (2) (1995) 380–384.
 [2] A. Dempster, Upper and lower probabilities induced by a multivalued mapping, *Annals Math. Stat.* 38 (1967) 325–339.
 [3] D. Dubois, H. Prade, On the unicity of Dempster rule of combination, *Int. J. Intell. Syst.* 1 (1986) 133–142.
 [4] D. Dubois, H. Prade, A review of fuzzy set aggregation connectives, *Inform. Sci.* 36 (1985) 85–121.
 [5] Z.X. Huang, Extensions to the k-means algorithm for clustering large data sets with categorical values, *Data Mining Knowl. Discovery* 2 (1998) 283–304.
 [6] A.L. Jousselme, D. Grenier, E. Bosse, A new distance between two bodies of evidence, *Inform. Fusion* 2 (2001) 91–101.

- [7] M.A. Khan, M. Banerjee, Formal reasoning with rough sets in multiple-source approximation systems, *Int. J. Approx. Reason.* 49 (2008) 466–477.
- [8] M.A. Khan, M. Banerjee, Multiple-source approximation systems: membership functions and indiscernibility, *Lecture Notes Comput. Sci.* 5009 (2008) 80–87.
- [9] Y. Leung, N.N. Ji, J.H. Ma, An integrated information fusion approach based on the theory of evidence and group decision-making, *Inform. Fusion* 14 (2013) 410–422.
- [10] J.H. Li, C.L. Mei, Y.J. Lv, Knowledge reduction in real decision formal contexts, *Inform. Sci.* 189 (2012) 191–207.
- [11] J.H. Li, C.L. Mei, Y.J. Lv, Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction *International J. Approx. Reason.* 54 (1) (2013) 49–165.
- [12] J.Y. Liang, R. Li, Y.H. Qian, Distance: a more comprehensible perspective for measure in rough set theory, *Knowl.-Based Syst.* 27 (2012) 126–136.
- [13] 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 (6) (2012) 912–926.
- [14] G.P. Lin, Y.H. Qian, J.J. Li, NMGRS: Neighborhood-based multigranulation rough sets, *Int. J. Approx. Reason.* 53 (7) (2012) 1080–1093.
- [15] G.P. Lin, J.Y. Liang, Y.H. Qian, Multigranulation rough sets: from partition to covering, *Inform. Sci.* 241 (2013) 101–118.
- [16] G.P. Lin, J.Y. Liang, Y.H. Qian, Topology approach to multigranulation rough sets, *Int. J. Machine Learn. Cybernet.* 5 (2) (2014) 233–243.
- [17] Z. Pawlak, Rough sets, *Int. J. Comput. Inform. Sci.* 11 (1982) 341–365.
- [18] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, 1991.
- [19] W. Pedrycz, Relational and directional aspects in the construction of information granules, *IEEE Trans. Syst. Man Cybernet. Part A* 32 (5) (2002) 605–614.
- [20] W. Pedrycz, *Granular Computing: An Emerging Paradigm*, vol. 70, Springer, 2001.
- [21] W. Pedrycz, *Granular Computing: Analysis and Design of Intelligent Systems*, CRC Press/Francis Taylor, Boca Raton, 2013.
- [22] Y.H. Qian, J.Y. Liang, Rough set method based on multi-granulations, in: *Proceedings of the Fifth IEEE Conference on Cognitive Informatics*, vol. 1, 2006, pp. 297–304.
- [23] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: a multi-granulation rough set, *Inform. Sci.* 180 (2010) 949–970.
- [24] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, Incomplete multigranulation rough set, *IEEE Trans. Syst. Man Cybernet. Part A* 20 (2010) 420–430.
- [25] Y.H. Qian, S.Y. Li, J.Y. Liang, Z.Z. Shi, F. Wang, Pessimistic rough set based decisions: a multigranulation fusion strategy, *Inform. Sci.* 264 (2014) 196–210.
- [26] 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.
- [27] Y.H. Qian, J.Y. Liang, C.Y. Dang, Knowledge structure, knowledge granulation and knowledge distance in a knowledge base, *Int. J. Approx. Reason.* 19 (2) (2011) 263–264.
- [28] 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.
- [29] R. Raghavan, B.K. Tripathy, On some topological properties of multigranular rough sets, *Adv. Appl. Sci. Res.* 2 (3) (2011) 536–543.
- [30] M. Saerens, F. Fous, Yet another method for combining classifiers outputs: a maximum entropy approach, in: *Multiple Classifier Systems*, Springer, 2004, pp. 82–91.
- [31] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.
- [32] Y.H. She, X.L. He, On the structure of the multigranulation rough set model, *Knowl.-Based Syst.* 36 (2012) 81–92.
- [33] Y.H. She, X.L. He, Rough approximation operators on R_0 -algebras (nilpotent minimum algebras) with an application in formal logic L^* , *Inform. Sci.* 277 (2014) 71–89.
- [34] Y.H. She, G.J. Wang, An axiomatic approach of fuzzy rough sets based on residuated lattices, *Comput. Math. Appl.* 58 (1) (2009) 189–201.
- [35] A. Skowron, The rough sets theory and evidence theory, *Fundam. Inform.* XII (1990) 245–262.
- [36] P. Smets, The combination of evidence in the transferable belief models, *IEEE Trans. Pattern Anal. Machine Intell.* 12 (5) (1990) 447–485.
- [37] P. Smets, R. Kennes, The transferable belief model, *Artif. Intell.* 66 (2) (1994) 191–234.
- [38] B. Tessen, Approximations for efficient computation in the theory of evidence, *Artif. Intell.* 61 (1993) 315–329.
- [39] C.Z. Wang, Q. He, D.G. Chen, Q.H. Hu, A novel method for attribute reduction of covering decision systems, *Inform. Sci.* 254 (1) (2014) 181–196.
- [40] C.Z. Wang, D.G. Chen, Q.H. Hu, Fuzzy information systems and their homomorphisms, *Fuzzy Sets Syst.* 249 (2014) 128–138.
- [41] S.K.M. Wong, Y.Y. Yao, P. Bollmann, H.C. Birger, Axiomatization of qualitative belief structure, *IEEE Trans. Syst. Man Cybernet.* 21 (1991) 726–734.
- [42] S.K.M. Wong, P. Lingras, Combination of evidence using the principle of minimum information gain, in: *Proceedings of the Sixth International Conference on Uncertainty in Artificial Intelligence*, 1990, pp. 450–459.
- [43] W.Z. Wu, Y. Leung, Theory and applications of granular labeled partitions in multi-scale decision tables, *Inform. Sci.* 181 (2011) 3878–3897.
- [44] W.Z. Wu, Y. Leung, Optimal scale selection for multi-scale decision tables, *Int. J. Approx. Reason.* 54 (8) (2013) 1107–1129.
- [45] W.Z. Wu, Y. Leung, W.X. Zhang, Connections between rough set theory and Dempster–Shafer theory of evidence, *Int. J. General Syst.* 31 (4) (2002) 405–430.
- [46] W.Z. Wu, Y. Leung, J.S. Mi, On generalized fuzzy belief functions in infinite spaces, *IEEE Trans. Fuzzy Syst.* 17 (2009) 385–397.
- [47] W.Z. Wu, M. Zhang, H.Z. Li, J.S. Mi, Knowledge reduction in random information systems via Dempster–Shafer theory of evidence, *Inform. Sci.* 174 (2005) 143–164.
- [48] W.Z. Wu, Knowledge reduction in random incomplete decision tables via evidence theory, *Fundam. Inform.* 115 (2–3) (2012) 203–218.
- [49] W.H. Xu, Q.R. Wang, X.T. Zhang, Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space, *Int. J. Fuzzy Syst.* 13 (4) (2011) 46–259.
- [50] C.F. Xu, W. Gong, Y. Pan, Review of Dempster–Shafer method for data fusion, *Acta Electron Sinica* 29 (3) (2001) 393–396.
- [51] Y.Y. Yao, Granular computing basis issues and possible solutions, in: *Proceedings of the Fifth International Conference on Computing and Information* (1), 2000, pp. 186–189.
- [52] Y.Y. Yao, S.K.M. Wong, Representation, propagation and combination of uncertain information, *Int. J. General Syst.* 23 (1994) 59–83.
- [53] Y.Y. Yao, P.J. Lingras, Interpretations of belief functions in the theory of rough sets, *Inform. Sci.* 104 (1998) 81–106.
- [54] R.R. Yager, On the Dempster–Shafer framework and new combination rules, *Inform. Sci.* 41 (2) (1987) 93–138.
- [55] R.R. Yager, On the fusion of imprecise uncertainty measures using belief structures, *Inform. Sci.* 181 (15) (2011) 3199–3209.
- [56] X.B. Yang, Y.H. Qian, J.Y. Yang, Hierarchical structures on multigranulation spaces, *J. Comput. Sci. Technol.* 27 (6) (2012) 1169–1183.
- [57] X.B. Yang, Y.S. Qi, X.N. Song, J.Y. Yang, Test cost sensitive multigranulation rough set: model and minimal cost selection, *Inform. Sci.* 250 (2013) 184–199.
- [58] X.B. Yang, Y.H. Qian, J.Y. Yang, Hierarchical structures on multigranulation spaces, *J. Comput. Sci. Technol.* 27 (6) (2012) 1169–1183.
- [59] X.B. Yang, Y.H. Qian, J.Y. Yang, On characterizing hierarchies of granulation structures via distances, *Fundam. Inform.* 123 (3) (2013) 365–380.
- [60] L.A. Zadeh, Some reflections on soft computing, granular and their roles in the conception, design and utilization of information/intelligent systems, *Soft Comput.* 2 (1998) 23–25.
- [61] L.A. Zadeh, A simple view of the Dempster–Shafer theory and its implications for the rule of combination, *AI Mag.* 2 (7) (1986) 85–90.