Proceedings of the 2017 International Conference on Machine Learning and Cybernetics, Ningbo, China, 9-12 July, 2017

A WEIGHTED MULTI-GRANULATION DECISION-THEORETIC APPROACH TO MULTI-SOURCE DECISION SYSTEMS

YANTING GUO[1], ERIC C.C. TSANG[1], WEIHUA XU[2]

[1]Faculty of Information Technology, Macau University of Science and Technology, Taipa, Macau
[2]School of Science, Chongqing University of Technology, Chongqing 400054, PR China
Corresponding author: Tel.:+85388972136
E-MAIL: 1290237984@qq.com(YANTING GUO), cctsang@must.edu.mo(Eric C.C. Tsang)

Abstract:

Decision theoretic rough set is a typical generalization model of rough set, which has fault tolerance based on Bayes minimum decision risk. How to mine knowledge from the information collected from different sources is one of the focuses of current artificial intelligence. From a cognitive point of view, especially from the point of granulation, this paper studies decision theory of multi-source decision systems based on generalized multi-granulation and decision theoretic rough sets. It is because each granular structure is not equally important in practical issues. First of all, the method of granulation weight is proposed based on the internal uncertainty of systems and the external correlation between systems, namely the double weighted granulation (DGW) method. And then a weighted generalized multi-granulation decision-theoretic rough set (WGM-DTRS) model in multi-source decision systems is proposed. Finally, in order to verify the effectiveness of the (DGW) method, the approximation accuracy of decision classes under different weighted granulation methods is compared. The numerical results show that the proposed method is effective from the point of classification. Therefore, the WGM-DTRS model based on the DGW method is meaningful.

Keywords:

Decision-theoretic rough set; Multi-source decision syst ems; Granulation weight; Geighted generalized multi-granu lation

1. Introduction

With the development of information technology, information sources for research objects are increasing. Multi-source information can not only reflect research objects more comprehensively, but also make up for the deficiency of single source. At the same time, it brings new challenges to knowledge discovery [14]. As a new multi-view data analysis method, multi-granulation rough set theory can effectively mine knowledge from multi-source information systems, high-dimensional feature data, multi-intelligence agents and distributed information systems [15]. Therefore, this paper studies decision theory of multi-source decision systems from the perspective of multi-granulation.

The decision theoretic rough set (DTRS) proposed by Yao[1-2] is a representative decision model, in which thresholds can be calculated Bayes minimum decision risk, the conditional probability can be estimated by the naive Bayes model, and positive, negative and boundary regions of the probability rough set are regarded as the application of the three-way decision theory. The DTRS studies knowledge under a single granulation [3]. In many situations of real life, according to the goals or requirements of users to solve the problem [4], we need to use known knowledge produced by multiple granular structures to describe the target or request. Therefore, many scholars introduce multi-granulation into decision theoretic rough sets [4-12].

In particular, Qian et al. proposed optimistic, pessimistic and mean multi-granulation decision theoretic rough sets through combining multi-granulation with decision-theoretic rough sets [4]. Recently, considering that the approximation of targets in large data is extremely time-consuming, and in many cases it is almost impossible to label all the data, Qian et al. proposed local optimistic, pessimistic and mean multi-granulation decision theoretic rough sets [5]. Li et al. studied the optimistic, pessimistic multi-granulation decision theoretic rough sets in the distributed fuzzy condition decision information systems [6]. Lin et al. proposed optimistic, pessimistic fuzzy multi-granulation decision theoretic rough sets for multi-source information systems [7]. Sun et al. proposed a variable precision multi-granulation fuzzy decision-theoretic rough set over two universes based on three-way decisions and multi-granulation [8]. Feng et al. proposed two kinds of variable precision multi-granulation decision theoretic fuzzy rough sets through the maximum and minimum membership degrees of objects about fuzzy concepts under multi-granulation [9]. After this, Feng et al. studies the

978-1-5386-0408-3/17/$31.00 ©2017 IEEE

reduction of multi-granulation fuzzy information system based on variable precision multi-granulation decision-theoretic fuzzy rough sets. These researches mainly focus on multi-granulation decision-theoretic rough sets under optimistic and pessimistic circumstances [10]. This paper will study multi-granulation decision-theoretic rough sets under more general circumstances, namely weighted multi-granulation decision-theoretic rough sets for multi-source decision systems.

At present, methods of granulation weights are mainly summarized as follows: the definite support weighted method, the possible support weighted method, the determine average approximation support weighted method, the geometric average approximation support weighted method based on the upper and lower approximations; the approximate precision weighted method based on the upper and lower approximation ratio; the approximate rough weighted method based on roughness; the granulation entropy weighted method based on rough entropy and so on [13]. This paper proposed a new method of granulation weights through the internal uncertainty of systems and the external correlations between systems in multi-source decision systems.

The main contribution of this paper is listed as follows:

1) The weighted generalized multi-granulation is introduced into decision-theoretic rough sets, which provides a decision-making method for multi-source decision systems.

2) A new method of granulation weights is proposed based on the internal uncertainty of systems and the external correlations between systems, namely the double weighted granulation method.

3) Comparing approximation accuracy of decision classes under different weighted granulation methods in multi-source decision systems, we verified that the validity of the proposed weighted granulation method and embodied the significance of weighted generalized multi-granulation decision-theoretic rough sets.

2. Preliminaries

In this section, we briefly review some basic concepts such as multi-source decision systems, weighted multi-granulation rough sets (WMGRS) and decision-theoretic rough sets (DTRS).

2.1. Multi-source decision systems

With the development of information technology, means collected information are increasing, and information sources obtained the same information are increasing accordingly. Therefore, many signal decision system about the same research objects are called multi-source decision systems. The detailed description [15] is as follows:

Let U is a nonempty finite set of objects; C_i is a set of condition attributes and D_i is a set of decision attributes in the i^{th} decision system ; V_i is the domain of attribute values, namely, $V_i = \cup_{a_i \in C_i} V_{a_i}$; $F_i : U \times C_i \cup D_i \to V_i$ is an information function, i.e., $\forall a_i \in C_i \cup D_i$, $x \in U$, that $F(x, a_i) \in V_{a_i}$, where $F(x, a_i)$ is the value of the object x about the attribute a_i . Then $MI = \{I_i \mid I_i = (U, C_i \cup D_i, V_i, F_i)\}$ is called a multi-source decision system, where $I_i = (U, C_i \cup D_i, V_i, F_i)$ is the i^{th} decision system of MI .

Unless otherwise specified, all the multi-source decision systems and decision systems in this paper are analogous to that defined above. It is important to point out that multi-source decision systems studied in this paper have the same universe, attributes and decision classes but different information functions.

2.2. Weighted multi-granulation rough sets

Multi-granulation rough set theory (MGRS) provides a method for knowledge discovery of multi-source decision systems. Weighted multi-granulation rough set (WMGRS) theory is a generalized model of MGRS, in which the importance of different granular structures is different and the lower and upper approximations are approximated by granular structures induced by multiple binary relations [13].

Let $I = (U, C \cup D, V, F)$ be a decision system, and $A_i \subseteq C, i = 1, 2, \cdots, s\ (s \le 2^C), \varphi \in (0.5, 1]$. For an arbitrary subset $X \subseteq U$, the lower and upper approximations of X with respect to $\sum_{i=1}^{s} A_i$ can be defined as

$$\overline{GM}^{w}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U : (\sum_{i=1}^{s} \varpi_i (1 - S^{A_i}_{\sim X}(x))) > 1 - \varphi\} ;$$

$$\underline{GM}^{w}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U : (\sum_{i=1}^{s} \varpi_i S^{A_i}_{X}(x)) \ge \varphi\} ;$$

respectively, where $S^{A_i}_X(x)$ is the support characteristic function of $x \in U$ with respect to concept X under attribute set A_i ; if $[x]_{A_i} \subseteq X$, then $S^{A_i}_X(x) = 1$; else $S^{A_i}_X(x) = 0$. The concept X is called a definable set with respect to $\sum_{i=1}^{S} A_i$ if and only if $\overline{GM}_{\sum_{i=1}^{S} A_i}(X) = \underline{GM}_{\sum_{i=1}^{S} A_i}(X)$; otherwise X is called a rough set with respect to $\sum_{i=1}^{S} A_i$. φ is called the information level with respect to $\sum_{i=1}^{S} A_i$. ϖ_i is the weight of granulation A_i , and $\sum_{i=1}^{S} \varpi_i = 1$. Positive region $pos(X)$, negative region $neg(X)$, and boundary region $bnd(X)$ are defined as follows:

$$pos(X) = \underline{GM}^{\varpi}_{\sum_{i=1}^{S} A_i}(X); neg(X) = \sim \overline{GM}^{\varpi}_{\sum_{i=1}^{S} A_i}(X);$$

$$bnd(X) = \overline{GM}^{\varpi}_{\sum_{i=1}^{S} A_i}(X) - \underline{GM}^{\varpi}_{\sum_{i=1}^{S} A_i}(X) .$$

Let $I = (U, C \cup D, V, F)$ be a decision system, $A_i \subseteq C,\ i = 1, 2, \cdots, s\ (s \le 2^C),\ \ \varphi \in (0.5, 1]$. The partition of universe generated by D is $U / D_i = \{Y_1, Y_2, \cdots, Y_m\}\ (i = 1, 2, \cdots, s)$. The weight ϖ_i of granulation A_i can be calculated by the following methods.

1) In the definite support weighted method, the possible support weighted method and the determine average approximation support weighted method, the weight ϖ_i of granulation A_i is defined as

$$\varpi_i = \frac{\sum_{j=1}^{m} |\underline{R_{A_i}}(Y_j)|}{\sum_{i=1}^{s}\sum_{j=1}^{m} |\underline{R_{A_i}}(Y_j)|}; \quad \varpi_i = \frac{\sum_{j=1}^{m} |\overline{R_{A_i}}(Y_j)|}{\sum_{i=1}^{s}\sum_{j=1}^{m} |\overline{R_{A_i}}(Y_j)|};$$

$$\varpi_i = \frac{1}{2}\left(\frac{\sum_{j=1}^{m} |\overline{R_{A_i}}(Y_j)|}{\sum_{i=1}^{s}\sum_{j=1}^{m} |\overline{R_{A_i}}(Y_j)|} + \frac{\sum_{j=1}^{m} |\overline{R_{A_i}}(Y_j)|}{\sum_{i=1}^{s}\sum_{j=1}^{m} |\overline{R_{A_i}}(Y_j)|} \right).$$

2) In the approximate precision weighted method and the approximate rough weighted method, the weight ϖ_i of granulation A_i is defined as

$$\varpi_i = \frac{\sum_{j=1}^{m} \alpha_i(Y_j)}{\sum_{i=1}^{s}\sum_{j=1}^{m} \alpha_i(Y_j)}; \quad \varpi_i = \frac{\sum_{j=1}^{m} \rho_i(Y_j)}{\sum_{i=1}^{s}\sum_{j=1}^{m} \rho_i(Y_j)};$$

where $\alpha_i(Y_j)$ denotes the precision of Y_j with respect to A_i , which can be calculated by $\alpha_i(Y_j) = \frac{|\underline{R_{A_i}}(Y_j)|}{|\overline{R_{A_i}}(Y_j)|}$; and $\rho_i(Y_j) = 1 - \alpha_i(Y_j)$.

2.3. Decision-theoretic rough sets (DTRS)

Decision-theoretic rough sets proposed by Yao provide a feasible approach for decision makers [1]. Based on the idea of three-way decisions and Bayes minimum decision risk, the decision-making process of DTRS model is given as follows.

Let $\Omega = \{X, □ X\}$ be a state set indicating that an object is in a decision class X and not in X ; A be an action set consisting of three actions, namely $A = \{a_P, a_B, a_N\}$, where a_P , a_B , a_N denote three actions about deciding $x \in pos(X)$, $x \in bnd(X)$ and $x \in neg(X)$, respectively. Different actions under different conditions will cause different costs. Let λ_{PP} , λ_{BP} , λ_{NP} denote the costs caused by taking actions a_P , a_B , a_N , respectively, when an object belongs to X ; and λ_{PN} , λ_{BN} , λ_{NN} denote the costs caused by taking actions a_P , a_B , a_N , respectively, when an object belongs to $□ X$.

Given the cost function, the expected cost associated with taking different actions for the objects in the equivalence class $[x]_R$ can be expressed as:

$$R(a_P \mid [x]_R) = \lambda_{PP} P(X \mid [x]_R) + \lambda_{PN} P(□ X \mid [x]_R);$$

$$R(a_B \mid [x]_R) = \lambda_{BP} P(X \mid [x]_R) + \lambda_{BN} P(□ X \mid [x]_R);$$

$$R(a_N \mid [x]_R) = \lambda_{NP} P(X \mid [x]_R) + \lambda_{NN} P(□ X \mid [x]_R).$$

The conditional probability $P(X \mid [x]_R)$ can be calculated by $P(X \mid [x]_R) = \left| [x]_R \cap X \right| / \left| [x]_R \right|$.

Based on the Bayesian decision procedure, the followings are the minimum-risk decision rules:

(P) If $R(a_P \mid [x]_R) \le R(a_B \mid [x]_R)$, and $R(a_P \mid [x]_R) \le R(a_N \mid [x]_R)$, then decide $x \in pos(X)$;

(B) If $R(a_B \mid [x]_R) \le R(a_P \mid [x]_R)$, and $R(a_B \mid [x]_R) \le R(a_N \mid [x]_R)$, then decide $x \in bnd(X)$;

(N) If $R(a_N \mid [x]_R) \le R(a_P \mid [x]_R)$, and $R(a_N \mid [x]_R) \le R(a_B \mid [x]_R)$, then decide $x \in neg(X)$.

When $\lambda_{PP} \le \lambda_{BP} < \lambda_{NP}$ and $\lambda_{NN} \le \lambda_{BN} < \lambda_{NP}$, the above decision rules are further simplified as

(P) If $P(X \mid [x]_R) \ge \alpha$ and $P(X \mid [x]_R) \ge \gamma$, then decide $x \in pos(X)$;

(B) If $P(X \mid [x]_R) \le \alpha$ and $P(X \mid [x]_R) \ge \beta$, then decide $x \in bnd(X)$;

(N) If $P(X \mid [x]_R) \ge \beta$ and $P(X \mid [x]_R) \le \gamma$, then decide $x \in neg(X)$.

The expressions of above parameters α, β, γ are

$$\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})};\ \beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})};$$

$$\gamma = \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}.$$

When $(\lambda_{NP} - \lambda_{BP})(\lambda_{PN} - \lambda_{BN}) > (\lambda_{BP} - \lambda_{PP})(\lambda_{BN} - \lambda_{NN})$ is true, then $0 \le \beta < \gamma < \alpha \le 1$. Accordingly, the decision rules are simplified as

(P) If $P(X \mid [x]_R) \ge \alpha$, then decide $x \in pos(X)$;

(B) If $\beta < P(X \mid [x]_R) < \alpha$, then decide $x \in bnd(X)$;

(N) If $p(X \mid [x]_R) \le \beta$, then decide $x \in neg(X)$.

3. Weighted generalized multi-granulation decision -theoretic rough sets (WGM-DTRS)

In this section we introduce weighted generalized multi-granulation into decision-theoretic rough sets. First, a new method of weighted granulation is proposed. Then the WGM-DTRS model is constructed in multi-source decision systems. Finally, decision rules can be obtained.

3.1. The double weighted granulation method

In multi-source decision systems, because of the influence of external environment and technical means, the reliability of different sources is different, so the corresponding granular structures of different information sources are also different for overall decision support. Considering the internal uncertainty of information sources and the external correlations between different information sources, we proposed a new method of granulation weights.

Definition 3-1-1: Let $MI = \{I_i \mid I_i = (U, C_i \cup D_i, V_i, F_i)\}$ be a multi-source decision system and the partition of universe generated by the decision attributes be $U / D_i = \{Y_1, Y_2, \cdots, Y_m\} (i = 1, 2, \cdots, s)$. For each signal decision system $I_i \in MI$, the internal granulation weight of decision system I_i is defined as

$$IG(I_i) = \rho_i \times \left(1 - \frac{H_i(D_i \mid C_i)}{\sum_{i=1}^{s} H_i(D_i \mid C_i)}\right),$$

where $\rho_i = 1 - \sum_{j=1}^{m} \underline{R_{C_i}}(Y_j) / \sum_{j=1}^{m} \overline{R_{C_i}}(Y_j)$ denotes the accuracy of I_i, and $H_i(D_i \mid C_i) = -\sum_{i=1}^{|U|} \sum_{j=1}^{m} \frac{|[x_i]_{R_{C_i}} \cap Y_j|}{|U|} \log \frac{|[x_i]_{R_{C_i}} \cap Y_j|}{|[x_i]_{R_{C_i}}|}$ denotes the conditional entropy of decision attribute set D_i with respect to condition attribute set C_i, which can reflect the uncertainty of system I_i.

Definition 3-1-2: Let $MI = \{I_i \mid I_i = (U, C_i \cup D_i, V_i, F_i)\}$ be a multi-source decision system. For any two signal decision system $I_i, I_j \in MI$, the external difference between I_i and I_j can be defined as

$$d(I_i, I_j) = \sum_{x \in U} \sum_{a \in C} |[x]_a^i \cup [x]_a^j - [x]_a^i \cap [x]_a^j|,$$

where $[x]_a^i$ and $[x]_a^j$ denote equivalent classes of object x about attribute a in the i^{th} and j^{th} decision systems, respectively.

The external correlation between I_i and I_j can be defined as $ec(I_i, I_j) = 1 - \frac{d(I_i, I_j)}{|U|^2 \times |C|}$, where $|U|$ and $|C|$ denote the cardinality of the universe U and conditional attribute set C, respectively.

Definition 3-1-3: Let $MI = \{I_i \mid I_i = (U, C_i \cup D_i, V_i, F_i)\}$ be a multi-source decision system. For each signal decision system $I_i \in MI$, the external granulation weight of decision system I_i is defined as

$$EG(I_i) = \frac{1}{s} \sum_{j=1}^{s} ec(I_i, I_j) = \frac{1}{s} \sum_{j=1}^{s} \left(1 - \frac{d(I_i, I_j)}{|U|^2 \times |C|}\right).$$

Definition 3-1-4: Let $MI = \{I_i \mid I_i = (U, C_i \cup D_i, V_i, F_i)\}$ be a multi-source decision system. For each signal decision system $I_i \in MI$, the double granulation weight of decision system I_i in the MI can be defined as

$$\varpi_i = \frac{IG(I_i) \times EG(I_i)}{\sum_{j=1}^{s} IG(I_j) \times EG(I_j)}$$

3.2. The WGM-DTRS model

Considering the DTRS model has a certain tolerant ability during the evaluation and different granulations may not be equally important in practice, in this section we propose weighted generalized multi-granulation decision-theoretic rough sets (WGM-DTRS).

Definition 3-2-1: Let $MI = \{I_i \mid I_i = (U, C_i \cup D_i, V_i, F_i)\}$ be a multi-source decision system. For an arbitrary set $X \subseteq U$, the lower and upper approximations of X in the multi-source decision system MI can be defined as

$$\overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(X) = \{x \in U \mid \sum_{i=1}^{S} \varpi_i US_X^{C_i}(x) > 1 - \varphi\};$$

$$\underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(X)=\{x\in U \mid \sum_{i=1}^{S}\varpi_i LS_X^{C_i}(x)) \geq \varphi\};$$

respectively, where $\varphi \in (0.5,1]$ is called the information level with respect to $\sum_{i=1}^{S} C_i$. $US_X^{C_i}(x)$ and $LS_X^{C_i}(x)$ are the weighted upper and lower support characteristic functions of $x \in U$ with respect to concept X under C_i. $US_X^{C_i}(x)$ and $LS_X^{C_i}(x)$ are specifically defined as

$$US_X^{C_i}(x)=\begin{cases}1, & if\ P(X\mid[x]_{C_i})>\beta_i;\\ 0, & else.\end{cases} \quad LS_X^{C_i}(x)=\begin{cases}1, & if\ P(X\mid[x]_{C_i})\geq\alpha_i;\\ 0, & else.\end{cases}$$

And the granulation weight $\varpi_i (i=1,2,\cdots,s)$ can be calculated according to the proposed method of subsection 3.1.

X is called a definable set in the MI if and only if $\overline{GM}^{\varpi}_{\sum_{i=1}^{S} C_i}(X)=\underline{GM}^{\varpi}_{\sum_{i=1}^{S} C_i}(X)$; otherwise X is called a rough set in the MI. In the multi-source decision system MI, positive region $pos(X)$, negative region $neg(X)$, and boundary region $bnd(X)$ are defined as follows:

$$pos(X)=\underline{GM}^{\varpi}_{\sum_{i=1}^{S} C_i}(X); neg(X)=\sim\overline{GM}^{\varpi}_{\sum_{i=1}^{S} C_i}(X);$$

$$bnd(X)=\overline{GM}^{\varpi}_{\sum_{i=1}^{S} C_i}(X)-\underline{GM}^{\varpi}_{\sum_{i=1}^{S} C_i}(X).$$

Definition 3-2-2: Let $MI=\{I_i \mid I_i=(U,C_i\cup D_i,V_i,F_i)\}$ be a multi-source decision system and the partition of universe be $U/D_i=\{Y_1,Y_2,\cdots,Y_m\} (i=1,2,\cdots,s)$. The approximation accuracy of decision class Y_j in the multi-source decision system MI can be defined as

$$acc(Y_j)=\frac{|\underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j)|}{|\overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j)|}.$$

3.3. The decision rules of WGM-DTRS

Rule 3-1: Let $MI=\{I_i \mid I_i=(U,C_i\cup D_i,V_i,F_i)\}$ be a multi-source decision system and $\varphi\in(0.5,1]$. For an arbitrary subset $X\subseteq U$, the following decision rules for WGM-DTRS model are presented:

(P) If there is $\sum\{\varpi_i \mid P(X\mid[x]_{C_i})\geq\alpha_i\}\geq\varphi$, then decide $x\in pos(X)$;

(B) If there are $\sum\{\varpi_i \mid P(X\mid[x]_{C_i})>\beta_i\}>1-\varphi$ and $\sum\{\varpi_i \mid P(X\mid[x]_{A_i})\geq\alpha_i\}<\varphi$, then decide $x\in bnd(X)$;

(N) If there is $\sum\{\varpi_i \mid P(X\mid[x]_{A_i})>\beta_i\}\leq 1-\varphi$, then decide $x\in neg(X)$.

According to rule 3-1, if the sum of weight of granulations satisfying $P(X\mid[x]_{C_i})>\beta_i$ is not smaller than φ, then decide $x\in pos(X)$; if the sum of weight of granulations satisfying $P(X\mid[x]_{C_i})\geq\alpha_i$ is not smaller than φ and the sum of weight of granulations satisfying $P(X\mid[x]_{C_i})>\beta_i$ is smaller than φ, then decide $x\in bnd(X)$; if the sum of weight of granulations satisfying $P(X\mid[x]_{C_i})\geq\alpha_i$ is not greater than $1-\varphi$, then decide $x\in neg(X)$.

3.4. A multi-source decision algorithm

Different information sources can form different granular structures. Granulation weights reflect the significance of different sources. Based on the internal uncertainty and the external correlation, a new method of granulation weights is proposed. By combining weighted generalized multi-granulation rough sets with decision-theoretic rough sets, we provide a method for decision making in multi-source decision systems, namely the weighted generalized multi-granulation decision-theoretic rough set. The following **Algorithm 1** is designed to demonstrate the detailed decision process.

Algorithm 1: WGM-DTRS Algorithm for the decision-making of MI

Input: Multi-source Decision System MI, Information Level φ

Output: $\overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j)$ and $\underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j)$ of Decision Class Y_j

Step 1: For $i=1:s$

1.1) $H_i(D_i\mid C_i)=-\sum_{i=1}^{|U|}\sum_{j=1}^{m}\frac{|[x_i]_{R_{C_i}}\cap Y_j|}{|U|}\log\frac{|[x_i]_{R_{C_i}}\cap Y_j|}{|[x_i]_{R_{C_i}}|}$;

1.2) $\rho_i=1-\sum_{j=1}^{m}\underline{R_{C_i}}(Y_j)/\sum_{j=1}^{m}\overline{R_{C_i}}(Y_j)$;

End For

Step 2: **Compute** the internal granulation weight of I_i:

$$IG(I_i)=\rho_i\times\left(1-\frac{H_i(D_i\mid C_i)}{\sum_{i=1}^{s}H_i(D_i\mid C_i)}\right);$$

Step 3: **Compute** the external granulation weight of I_i :

3.1) $d(I_i, I_j) = \sum_{x \in U} \sum_{a \in C} | [x]_a^i \cup [x]_a^j - [x]_a^i \cap [x]_a^j |$;

3.2) $EG(I_i) = \frac{1}{s} \sum_{j=1}^{s} \left(1 - \frac{d(I_i, I_j)}{|U|^2 \times |C|} \right)$;

Step 4: $\varpi_i = \frac{IG(I_i) \times EG(I_i)}{\sum_{j=1}^{s} IG(I_j) \times EG(I_j)}$;

Step 5: For $j=1:m$

5.1) $\overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j) \leftarrow \varnothing$, $\underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j) \leftarrow \varnothing$;

5.2) For $x \in U$

5.2.1) If $\sum_{i=1}^{S} \varpi_i US_{Y_j}^{C_i}(x) > 1 - \varphi$

$\overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j) \leftarrow \overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j) \cup \{x\}$;

End If

5.2.2) If $\sum_{i=1}^{S} \varpi_i LS_{Y_j}^{C_i}(x)) \geq \varphi$

$\underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j) \leftarrow \underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j) \cup \{x\}$;

End If

End For

End For

Step 6: **Output** $\overline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j)$ and $\underline{GM}^{w}_{\sum_{i=1}^{S} C_i}(Y_j)$.

4. Case study

In order to verify the feasibility and superiority of the proposed double weighted granulation method, approximation accuracies of decision classes under different weighted granulation methods are compared in a multi-source decision system. For convenience of description, the double weighted granulation method, the definite support weighted granulation method, the possible support weighted granulation method, the determine average approximation support weighted granulation method and the approximate rough weighted granulation method are shortened by DGW, DSW, PSW, DAW, ARW in this section.

The detailed information about the multi-source decision system $MI = \{I_1, I_2, I_3, I_4, I_5, I_6\}$ is shown in **Table 1**. And $U / D_i = \{Y_1, Y_2\} = \{\{x_1, x_2, x_5, x_7, x_8\}, \{x_3, x_4, x_6, x_9, x_{10}\}\}$.

4.1. The double weighted granulation method

First of all, the internal granulation weight of each decision system can be calculated according to definition 3-3-1. Detailed results are shown in **Table 2**.

For any decision system I_i, $I_j \in MI$, the external difference between I_i and I_j can be calculated according to definition 3-1-2. Detailed results are shown in the matrix $D(d_{ij})$, where d_{ij} denotes the external difference between system I_i and I_j .

$$D(d_{ij}) = \begin{pmatrix} 0 & 74 & 80 & 80 & 120 & 104 \\ 74 & 0 & 118 & 94 & 114 & 118 \\ 80 & 118 & 0 & 88 & 96 & 100 \\ 80 & 94 & 88 & 0 & 120 & 88 \\ 120 & 114 & 96 & 120 & 0 & 128 \\ 104 & 118 & 100 & 88 & 128 & 0 \end{pmatrix}$$

Based on the external difference between systems and definition 3-1-2, the external correlations between systems can be calculated. Detailed results are shown in the matrix $EC(ec_{ij})$, where ec_{ij} denotes the external correlation between system I_i and I_j .

$$EC(ec_{ij}) = \begin{pmatrix} 1.0000 & 0.7533 & 0.7333 & 0.7333 & 0.6000 & 0.6533 \\ 0.7533 & 1.0000 & 0.6067 & 0.6867 & 0.6200 & 0.6067 \\ 0.7333 & 0.6067 & 1.0000 & 0.7067 & 0.6800 & 0.6667 \\ 0.7333 & 0.6867 & 0.7067 & 1.0000 & 0.6000 & 0.7067 \\ 0.6000 & 0.6200 & 0.6800 & 0.6000 & 1.0000 & 0.5733 \\ 0.6533 & 0.6067 & 0.6667 & 0.7067 & 0.5733 & 1.0000 \end{pmatrix}$$

TABLE 1. A multi-source decision system

	I_1			I_2			I_3		
	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3
x_1	1	3	1	1	2	1	1	2	2
x_2	1	3	1	1	3	1	1	2	2
x_3	2	1	3	1	1	3	2	1	3
x_4	2	1	3	1	1	3	2	1	3
x_5	1	3	1	1	3	1	1	3	2
x_6	3	3	2	3	3	2	3	3	3
x_7	2	1	3	1	2	1	3	3	3
x_8	2	1	3	1	1	3	2	2	1
x_9	3	3	2	3	3	2	2	2	1
x_{10}	1	3	3	1	3	1	1	3	1

	I_4			I_5			I_6		
	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3
x_1	1	3	1	1	3	2	1	3	1
x_2	1	1	1	3	1	3	1	3	1
x_3	2	1	3	2	1	1	1	1	1
x_4	2	1	3	3	1	3	3	3	3
x_5	1	1	1	3	3	2	1	1	1
x_6	3	1	2	3	3	2	3	3	3
x_7	2	2	2	3	3	2	2	2	3
x_8	2	2	2	2	1	1	2	2	2

x_9	2	1	3	2	2	1	2	2	3
x_{10}	1	1	1	1	3	2	1	3	1

TABLE 2. The information of internal granulation weight

Sources	I_1	I_2	I_3	I_4	I_5	I_6
$H_i(D_i \mid C_i)$	1.6000	1.6529	0.8000	0.8265	2.0265	1.6265
ρ_i	0.4286	0.2500	0.4286	0.5385	0.0526	0.1765
$IG(I_i)$	0.3482	0.2016	0.3884	0.4863	0.0401	0.1428

Therefore, the external granulation weight of each signal decision system $I_i \in MI$ can be obtained according to definition 3-1-3. Detailed results are listed as follows:

$$EG(I_1)=0.7456, EG(I_2)=0.7122, EG(I_3)=0.7322,$$
$$EG(I_4)=0.7389, EG(I_5)=0.6789, EG(I_6)=0.7011.$$

Based on the internal granulation weight and external granulation weight of each signal decision system, the double granulation weight of decision system I_i in the MI can be obtained according to definition 3-1-4. Detailed results are listed as follows:

$$\varpi_1=0.2211, \varpi_2=0.1223, \varpi_3=0.2422,$$
$$\varpi_4=0.3060, \varpi_5=0.0232, \varpi_6=0.0853.$$

4.2. The WGM-DTRS model

According to the double granulation weight of system $I_i (\mathrm{i}=1,2,\cdots,6)$ and weighted generalized multi-granulation decision-theoretic rough set theory, the upper and lower approximations of decision classes in multi-source decision systems can be calculated.

Considering the multi-source decision systems studied in this paper have the same universe, attributes and decision classes but different information functions, we find that the costs caused by different decisions should be the same under different source. According to the decision cost function given by experts, we can get different threshold parameter values. This paper focuses on the WGM-DTRS model of multi-source systems when $\alpha+\beta=1$, $\alpha+\beta>1$, $\alpha+\beta<1$. It is necessary to point out that the information level in the following experiment is selected as $\varphi=0.6$.

Next, we only discuss the decision making process under the case of $\alpha+\beta=1$, the other two cases are similar.

When $\alpha=0.6$ and $\beta=0.4$, the upper and lower approximations of decision classes Y_1, Y_2 in MI can be obtained as follows

$\overline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_1)=\{x_1,x_2,x_5,x_7,x_8,x_{10}\}$; $\underline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_1)=\{x_1,x_2,x_5\}$;

$\overline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_2)=\{x_3,x_4,x_6,x_7,x_8,x_9,x_{10}\}$; $\underline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_2)=\{x_3,x_4,x_6,x_9\}$.

When $\alpha=0.7$ and $\beta=0.3$, the upper and lower approximations of decision classes Y_1, Y_2 in MI can be obtained as follows

$\overline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_1)=\{x_1,x_2,x_3,x_5,x_7,x_8,x_{10}\}$;

$\overline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_2)=\{x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}$;

$\underline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_1)=\{x_1\}$; $\underline{GM}^{w}_{\sum_{i=1}^{6} C_i}(Y_2)=\{x_4,x_6,x_9\}$.

When $\alpha=0.8, \beta=0.2$, $\alpha=0.9, \beta=0.1$ and $\alpha=1, \beta=0$, the results of approximations are the same as the results of case $\alpha=0.7, \beta=0.3$.

4.3. Comparison of different weighted granulation methods

The feasibility and superiority of the proposed double weighted granulation method (DGW) are verified by comparing the approximate accuracy of decision classes under different weighted granulation methods in the multi-source decision system $MI=\{I_1,I_2,I_3,I_4,I_5,I_6\}$. Other weighted granulation methods are mainly DSW, PSW, DAW, and ARW. First, we calculate the approximation accuracies of decision classes under different weighted granulation methods. Detailed results are shown in Tables 3 and 4.

TABLE 3. Approximate accuracies of decision class Y_1 under five weighted granulation methods

(α,β)	(0.6,0.4)	(0.7,0.3)	(0.8,0.2)	(0.9,0.2)	(1.0,0.0)
PSW	0.4286	0.1250	0.1250	0.1250	0.1250
DSW	0.6667	0.1250	0.1250	0.1250	0.1250
DAW	0.4286	0.1250	0.1250	0.1250	0.1250
ARW	0.3000	0.0000	0.0000	0.0000	0.0000
DGW	0.5000	0.1429	0.1429	0.1429	0.1429

TABLE 4. Approximate accuracies of decision class Y_2 under five weighted granulation methods

(α,β)	(0.6,0.4)	(0.7,0.3)	(0.8,0.2)	(0.9,0.2)	(1.0,0.0)
PSW	0.4286	0.2222	0.2222	0.2222	0.2222
DSW	0.6667	0.2222	0.2222	0.2222	0.2222
DAW	0.4286	0.2222	0.2222	0.2222	0.2222
ARW	0.0000	0.2000	0.2000	0.2000	0.2000

DGW	0.5714	0.3333	0.3333	0.3333	0.3333

In order to show the advantages of the DGW more intuitively, the accuracies of different methods are compared. The detailed information is shown in Figs. 1 and 2.

From Fig. 1(a), 1(c), 1(d), it is obvious that the approximate accuracies of the first decision class under the DGW method are higher than the accuracies of PSW, DAW, ARW methods when selecting different thresholds. From Fig. 1(b), when α=0.6 and β=0.4, the approximate accuracies of the first decision class under the DGW method is lower than the accuracies of the DSW method. In other cases, the accuracies of the DGW method are higher than the accuracies of the DSW method.

From Fig. 2, the approximate accuracies of the second decision class under the DGW method are higher than the accuracies of PSW, DSW, DAW, ARW methods in either case.

According to the above comparison results, we know that the proposed DGW method is feasible. Moreover, the DGW has certain advantages from the view point of approximate precision. Also the significance of WGM-DTRS model based on the DGW method is embodied.

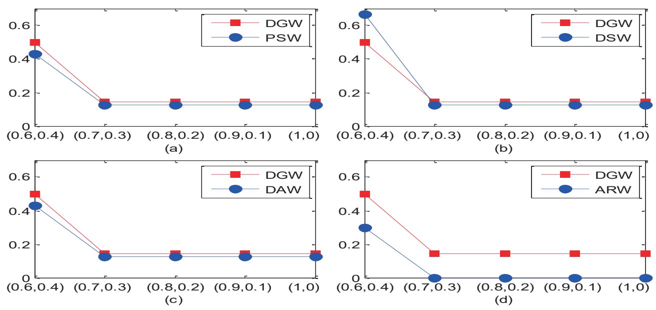

FIGURE 1. Approximate accuracies of decision class Y_1 under five weighted granulation methods

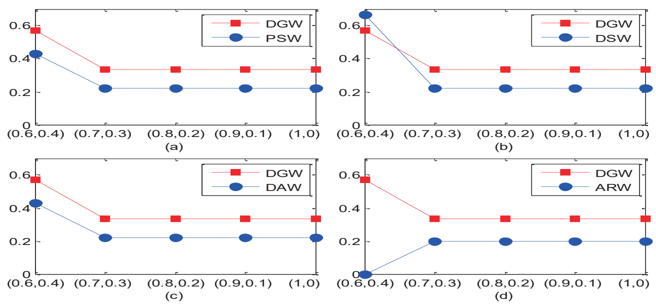

FIGURE 2. Approximate accuracies of decision class Y_2 under five weighted granulation methods

5. Conclusions

This paper introduces weighted generalized multi-granulation into decision-theoretic rough sets (WGM-DTRS) and then proposes WGM-DTRS in multi-source decision systems. First, the DGW is proposed based on the internal uncertainty of systems and the external correlation between systems. Second, the model and decision rules of WGM-DTRS in multi-source decision systems are studied. Finally, the validity of the double weighted granulation method is verified and the significance of WGM-DTRS model based on the DGW method is embodied through a case study and comparing the approximate accuracy of decision classes under different weighted granulation methods in a multi-source decision system.

Acknowledgments

This work is supported by the Macao Science and Technology Development Funds (100/2013/A3 and 081/2015/A3),

References

[1] Y Y Yao, Decision-theoretic rough set models, In: Proceeding of Rough Sets and Knowledge Technology, Lecture Notes in Artificial Intelligence, 2007, 4481: 1-12.

[2] Y Y Yao, The superiority of three-way decisions in probabilistic rough set models, Information Sciences. 2011, 181(6): 1080-1096.

[3] H Yu, G Y Wang, Y Y Yao, Current research and future perspectives on decision-theoretic rough sets, Chinese Journal of Computers, 2015, 38(8): 1628-1639.

[4] Y Qian, H Zhang, Y Sang, et al, Multigranulation decision-theoretic rough sets [J], International Journal of Approximate Reasoning, 2014, 55(1):225-237.

[5] Y Qian, X Y Liang, G P Lin, et al, Local multigranulation decision-theoretic rough sets [J], International Journal of Approximate Reasoning, 2017, 82:119-137.

[6] Z Li, X Liu, G Zhang, et al, A multi-granulation decision-theoretic rough set method for distributed fc -decision information systems: An application in medical diagnosis [J], Applied Soft Computing, 2017.

[7] G P Lin, J Y Liang, Y H Qian, et al, A fuzzy multigranulation decision-theoretic approach to multi-source fuzzy information systems [J], Knowledge-Based Systems, 2016, 91(C):102-113.

[8] B Sun, W Ma, X Xiao, Three-way group decision making based on multigranulation fuzzy decision-theoretic rough set over two universes [J], International Journal of Approximate Reasoning, 2016.

[9] T Feng, H T Fan, J S Mi, Uncertainty and reduction of variable precision multigranulation fuzzy rough sets based on three-way decisions [J], International Journal of Approximate Reasoning, 2017.

[10] T Feng, J S Mi. Variable precision multigranulation decision-theoretic fuzzy rough sets [J]. Knowledge-Based Systems, 2016, 91:93-101.

[11] W H Xu, Y T Guo. Generalized multigranulation double-quantitative decision-theoretic rough set [J]. Knowledge-Based Systems, 2016, 105:190-205.

[12] W T Li, W H Xu, Multigranulation decision-theoretic rough set in ordered information system, Fundamenta Informaticae, 2015, 139: 67-89.

[13] W H Xu. Ordered information systems and rough sets [M], Science Press, 2013. (in Chinese)

[14] J A Balazs, J D Squez, Opinion mining and information fusion [M]. Elsevier Science Publishers B. V. 2016.

[15] W H Xu, J H Yu, A novel approach to information fusion in multi-source datasets: A granular computing viewpoint [J]. Information Sciences, 2017, 378:410-423.