Postprint: A Comprehensive Method for Granularity Weight Determination in Multi-granulation Rough Sets

Authors: Peng Liangui, Yan Ruixia, Chen Zhaojun

Date: 2018-08-13T00:00:00+00:00

Abstract

To address the problem of strong subjectivity in existing granularity weight determination methods, this paper proposes a weight determination method based on granularity information content. First, information content is introduced into the lower approximation distribution of rough sets to define the information content of granularity sets within the rough set' s lower approximation distribution. Second, the importance of granularity is defined based on information content, and a comprehensive method for determining granularity weights based on information content is designed by using granularity importance as heuristic information. By introducing weight coefficients, decision makers can select the granularity weight determination approach according to actual circumstances: experience-dominated or objectivity-dominated. Finally, the effectiveness of the algorithm is verified through an example. The analysis results reveal that the experience-dominated determination method strengthens the importance of non-core granularities, while the objectivity-dominated determination method strengthens the importance of core granularities.

Full Text

A Synthetic Method for Determining Granularity Weights in Multi-Granularity Rough Sets

Peng Liangui, Yan Ruixia, Chen Zhaojun
School of Management, Shanghai University of Engineering Science, Shanghai 201620, China

Abstract

To address the strong subjectivity inherent in existing granularity weight determination methods, this paper proposes a novel approach based on granularity information quantity. First, information quantity is introduced into the lower approximation distribution of rough sets to define the information quantity of granularity sets within this distribution. Second, granularity importance is defined based on information quantity, which then serves as heuristic information for designing a comprehensive method to determine granularity weights. By introducing a weight coefficient, decision-makers can select the appropriate weight determination mode according to actual circumstances: empirically-oriented or objectively-oriented. Finally, the effectiveness of the algorithm is verified through an example. The analysis reveals that the empirically-oriented method strengthens the importance of non-core granularities, while the objectively-oriented method enhances the importance of core granularities.

Keywords: multi-granularity rough set; information quantity; granularity important degree; weight

0 Introduction

Since Pawlak introduced rough set theory, it has been widely applied and studied in decision analysis, artificial intelligence, pattern recognition, and other fields due to its advantage of not requiring prior information. In real-world decision-making processes, to evaluate decision objects reasonably, a comprehensive evaluation index system is typically established. After selecting the indicators, an effective method is needed to determine their weights, which reflect the importance of each indicator in the decision-making process. Consequently, weight determination methods have become a hot research topic in decision-making and evaluation.

Common weight determination methods include the Delphi method, Analytic Hierarchy Process (AHP) [?], grey relational analysis, attribute importance ranking, naïve Bayes method [?], 环比评分法, principal component analysis, and others. However, these methods generally rely on weights given by experts based on subjective experience, where differences in knowledge levels and preferences for specific granularities can affect the objectivity of decision results and increase evaluation costs, time complexity, and space complexity.

Following Zadeh' s initial proposal and discussion of fuzzy information granulation in 1979, granular computing has gradually developed within frameworks such as set theory and interval analysis, fuzzy sets [?], rough sets [?], probability theory [?], and entropy space theory [?]. Qian [?] analyzed the relationship between granular computing and rough set theory, combined them, and proposed the concept of dynamic granularity, subsequently extending single-granularity rough sets to multi-granularity rough sets. In multi-granularity rough set research, granularity weight has become a research hotspot, similar to attribute weights.

Xue et al. [?] redefined attribute certainty and reduction degree based on three-way decision theory and proposed a corresponding attribute weight construction method. Meng et al. [?] introduced information quantity into the lower approximation distribution reduction of pessimistic multi-granularity rough sets, defined granularity importance, and designed a heuristic granularity reduction algorithm using this importance as heuristic information, though they did not further discuss granularity weights. Wang et al. [?] explored how differences in granularity importance affect decision results but did not provide a clear method for determining granularity weights. Zhou [?] addressed multi-attribute decision problems by comprehensively analyzing attribute importance from multiple perspectives and levels of quotient spaces with different granularities, thereby determining attribute weights more accurately. Wang et al. [?] proposed a granular computing-based hesitant fuzzy multi-criteria decision method that determines attribute weights by constructing fuzzy preference matrices. While these methods utilize rough set and granular computing knowledge to determine attribute weights based on obtained data, considering only objective factors is insufficient given the presence of data noise and other influencing factors.

This paper proposes a granularity weight determination method based on granularity information quantity within the framework of rough set theory. By combining granularity weights obtained from data with expert experience, the method avoids excessive dependence on data and determines comprehensive weights through a combination of subjective and objective factors. By introducing a weight coefficient, decision-makers can adjust the proportion of subjective and objective weights according to their primary reference objects, thereby determining the direction of weight determination bias. Finally, an example demonstrates the method' s effectiveness, and a comparative analysis of weights determined under subjective and objective bias conditions reveals that the empirically-oriented method strengthens the importance of non-core granularities, while the objectively-oriented method strengthens the importance of core granularities, effectively avoiding the shortcomings of single-method approaches.

---

2 Synthetic Method for Determining Granularity Weights Based on Information Quantity

This method defines granularity importance based on the information quantity of granularity sets and designs a synthetic approach for determining weights in the lower approximation distribution of multi-granularity rough sets. The basic idea is to first compute the lower approximation distribution of the rough set and calculate the information quantity of different granularities. By removing granularities one by one and observing changes in information quantity, granularities causing large changes are deemed more important for decision-making, while those causing small changes are less important. Using granularity importance as heuristic information, the importance values are normalized to obtain each granularity' s weight.

Definition 1. An information system [?] is typically defined as a quadruple: $\langle U, AT, V, f\rangle$, where U is a non-empty finite set of all objects, AT is a non-empty finite set of attributes, V_a represents the value domain of attribute a, $V = \bigcup_{a\in AT} V_a$ is the set of all attribute value domains, and $f : U \times AT \to V$ is an information function indicating the value of object x on attribute a. If $AT = C \cup \{d\}$, where C is the condition attribute set and d is the decision attribute, then $\langle U, C, d, V, f\rangle$ is called a decision information system.

Definition 2. Let $\langle U, AT, V, f\rangle$ be an information system and $A \subseteq AT$. The indiscernibility relation on U is defined as [?, ?]:

$$IND(A) = \{(x, y) \in U \times U \mid \forall a \in A, f(x, a) = f(y, a)\}$$

From this definition, $IND(A)$ is an equivalence relation on U. From the perspective of granular computing, an equivalence relation corresponds to a granularity, $U/IND(A)$ corresponds to a granular structure or granularity space, and the equivalence class $[x]_A$ is called a knowledge granule.

Definition 3. Let $\langle U, AT, V, f\rangle$ be a complete decision information system [?], $A \subseteq AT$, and $X \subseteq U$ with $X \neq \emptyset$. The lower and upper approximations of X with respect to A are defined as:

$$\underline{A}(X) = \{x \in U \mid [x]_A \subseteq X\}, \quad \overline{A}(X) = \{x \in U \mid [x]_A \cap X \neq \emptyset\}$$

Definition 4. Let $\langle U, AT, D, V, f\rangle$ be a complete decision information system, $A_1, A_2, \dots, A_m \subseteq AT$, and $A = \{A_1, A_2, \dots, A_m\}$ be a granularity set. The information quantity of A in the lower approximation distribution is defined as:

$$I(A/D) = \sum_{j=1}^{r} \frac{|U - \underline{A}(Y_j)|}{|U|}$$

where $U/D = \{Y_1, Y_2, \dots, Y_r\}$ and $|\cdot|$ denotes set cardinality.

Definition 5. Let $\langle U, AT, D, V, f\rangle$ be a complete decision information system, $A_1, A_2, \dots, A_m \subseteq AT$, and $A = \{A_1, A_2, \dots, A_m\}$ be a granularity set. The importance of granularity A_i in A is defined as:

$$SGF(A_i, A) = I(A/D) - I(A - \{A_i\}/D)$$

A granularity $A_i \in A$ is necessary in A if and only if $SGF(A_i, A) > 0$.

Definition 6. The core of granularity set A is defined as [?]:

$$Core(A) = \{A_i \in A \mid SGF(A_i, A) > 0\}$$

Considering that this method neglects the role of expert experience, a weight parameter λ is introduced to scientifically evaluate and comprehensively consider granularity weights, thereby addressing the subjective-objective bias. The parameter λ is defined based on the dataset itself:

Let $|Core(A)|$ denote the cardinality of the core granularity set. When no core granularity exists, $|Core(A)| = 0$; when all granularities are core granularities, $|Core(A)| = |A|$; otherwise, $0 < |Core(A)| < |A|$. The comprehensive algorithm formula for determining granularity weights is:

a) Empirically-oriented:

$$\mu_\alpha(A_i) = \lambda\omega(A_i) + (1-\lambda)\mu(A_i)$$

b) Objectively-oriented:

$$\omega_\alpha(A_i) = \lambda\omega(A_i) + (1-\lambda)\mu(A_i)$$

where $\omega(A_i)$ represents the objective weight calculated from data, and $\mu(A_i)$ represents the subjective weight given by experts based on experience.

Algorithm 1: Objective Weight Calculation

Input: Complete decision information system $\langle U, AT, D, V, f\rangle$, granularity set $A = \{A_1, A_2, \dots, A_m\}$

Output: Objective weights $\omega(A_i)$ of each granularity in the decision information system

1. For each $A_i \in A$, compute U/A_i and U/D, and calculate $I(A_i/D)$ under U/A_i
2. Compute the information quantity $I(A/D)$ of the granularity set
3. For each $A_i \in A$, compute $I(A - \{A_i\}/D)$
4. Compute the importance of each granularity A_i: $SGF(A_i, A) = I(A/D) - I(A - \{A_i\}/D)$, and determine the core attribute set $Core(A) = \{A_i \mid SGF(A_i, A) > 0\}$
5. Compute the objective weight value $\omega(A_i)$ for each granularity A_i and the weight parameter $\lambda = \left|\frac{|Core(A)|}{|A|} - 0.5\right|$

Algorithm 2: Comprehensive Evaluation Weight Values

Input: Subjective weight values $\mu(A_i)$ given by experts and objective weights $\omega(A_i)$ from Algorithm 1

Output: Comprehensive evaluation weight values for each granularity

Empirically-oriented:

$$\mu_\alpha(A_i) = \lambda\omega(A_i) + (1-\lambda)\mu(A_i)$$

Objectively-oriented:

$$\omega_\alpha(A_i) = \lambda\omega(A_i) + (1-\lambda)\mu(A_i)$$

3 Example Analysis

A company faces an investment problem and has formulated eight investment schemes as alternatives. The company invites domain experts to evaluate these eight schemes based on evaluation indicators. Each expert assesses according to indicator $e = \{1, 2, 3, 4, 5\}$, where the evaluation levels are {excellent, good, average, slightly poor, poor}. Table 1 shows an evaluation table for company investment schemes provided by experts. This evaluation table is treated as a complete decision information system $\langle U, AT, D, V, f\rangle$, where $U = \{x_j \mid j = 1, \dots, 8\}$ represents the schemes, and equivalence relations construct each scheme' s equivalence classes. Each evaluation indicator' s assessment scheme forms a granularity space. Let $D = \{0, 1\}$ represent decision attributes, where 1 denotes "yes" and 0 denotes "no" decisions. In the multi-granularity space, different granularity weights are determined based on information quantity.

The granularity set is $A = \{A_1, A_2, A_3, A_4, A_5\}$, with decision attribute D. The equivalence classes are: - $U/A_1 = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}, \{x_6\}, \{x_7\}, \{x_8\}\}$ - $U/A_2 = \{\{x_1, x_7\}, \{x_2\}, \{x_3\}, \{x_4, x_6, x_8\}, \{x_5\}\}$ - $U/A_3 = \{\{x_1, x_2\}, \{x_3\}, \{x_4, x_5\}, \{x_6, x_7\}, \{x_8\}\}$ - $U/A_4 = \{\{x_1, x_7\}, \{x_2\}, \{x_3\}, \{x_4, x_5, x_6\}, \{x_8\}\}$ - $U/A_5 = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7, x_8\}\}$

The decision classes are $U/D = \{Y_1, Y_2\}$, where $Y_1 = \{x_1, x_2, x_3, x_7\}$ and $Y_2 = \{x_4, x_5, x_6, x_8\}$.

Step 1: From Algorithm 1, we compute: - $I(A/D) = 5/964$ - $I(A-\{A_1\}/D) = 5/964$ - $I(A-\{A_2\}/D) = 5/664$ - $I(A-\{A_3\}/D) = 5/164$ - $I(A-\{A_4\}/D) = 5/964$ - $I(A-\{A_5\}/D) = 5/664$

Step 2: From Algorithm 1(b)(c), we compute granularity importance: - $SGF(A_1, A) = SGF(A_4, A) = 0$ - $SGF(A_2, A) = SGF(A_5, A) = 3/64$ - $SGF(A_3, A) = 8/64$

Step 3: From Algorithm 1(d), the core attribute set is $Core(A) = \{A_2, A_3, A_5\}$.

Step 4: From Algorithm 1(f), we obtain objective weights: - $\omega(A_1) = \omega(A_4) = 0$ - $\omega(A_2) = \omega(A_5) = 0.214$ - $\omega(A_3) = 0.572$

Step 5: The weight parameter is $\lambda = 0.1$.

Step 6: Experts provide subjective weights: - $\mu(A_1) = 0.1$ - $\mu(A_2) = 0.15$ - $\mu(A_3) = 0.5$ - $\mu(A_4) = 0.1$ - $\mu(A_5) = 0.15$

Step 7: Using Algorithm 2, we compute comprehensive weights:

(a) Empirically-oriented: - $\mu_\alpha(A_1) = 0.09$ - $\mu_\alpha(A_2) = 0.1564$ - $\mu_\alpha(A_3) = 0.5072$ - $\mu_\alpha(A_4) = 0.09$ - $\mu_\alpha(A_5) = 0.1564$

(b) Objectively-oriented: - $\omega_\alpha(A_1) = 0.01$ - $\omega_\alpha(A_2) = 0.2076$ - $\omega_\alpha(A_3) = 0.5648$ - $\omega_\alpha(A_4) = 0.01$ - $\omega_\alpha(A_5) = 0.2076$

The comparison of differences (ε) shows: - Pure subjective method: $\varepsilon_\mu = 0.023$ - Pure objective method: $\varepsilon_\omega = 0.438$ - Empirically-oriented: $\varepsilon_{Sub} = 0.0245$ - Objectively-oriented: $\varepsilon_{Obj} = 0.0411$

Thus, $\varepsilon_\mu < \varepsilon_{Sub} < \varepsilon_{Obj} < \varepsilon_\omega$.

The analysis reveals that the pure subjective method yields the smallest differences among indicator weights, with high similarity that fails to adequately distinguish between indicators. The empirically-oriented method produces slightly greater weight fluctuations than the pure subjective method, but the differences among granularities remain insufficient compared to the objectively-oriented method. The objectively-oriented method generates larger weight fluctuations and greater differences among granularities, better expressing their distinctions and reflecting actual indicator differences. Although this approach appears more realistic, the appropriate method should be selected based on the decision context in practice. Moreover, while the pure objective method yields the largest difference value, it relies excessively on data and ignores expert experience, which is clearly unreasonable in domains like medicine where expert experience plays a crucial role. Overall, both empirically-oriented and objectively-oriented methods outperform pure subjective or pure objective methods, offering greater flexibility and more comprehensive consideration.

[Figure 1: see original paper] Comparison of granularity weights obtained by empirically-oriented and objectively-oriented methods

Through comparative analysis of the weight values in Figure 1, we observe:

---

4 Conclusion

In addressing decision problems, this paper proposes a comprehensive algorithm for determining granularity weights in multi-granularity rough sets based on information quantity to overcome limitations of single indicator weight determination methods. The algorithm uses granularity importance as heuristic information to determine objective weights. By introducing the weight coefficient λ, it solves the problem of over-reliance on a single weight source. This weight determination approach allows decision-makers to select formulas based on practical needs, offering significant flexibility and comprehensiveness. Additionally, the method effectively avoids zero-weight indicators.

A simple example demonstrates the algorithm' s effectiveness and compares different outcomes between empirically-oriented and objectively-oriented weight determination. Results show that the empirically-oriented method strengthens the importance of non-core granularities, while the objectively-oriented method enhances the role of core attributes. Moreover, the comprehensive weight method overcomes the incompleteness of single methods in weight determination.

Key findings: - In both methods, core granularity weights exceed non-core granularity weights, establishing a priority order based on weight magnitude - Empirically-oriented calculations strengthen non-core granularity importance, while objectively-oriented algorithms strengthen core granularity importance - Variance analysis shows: $\varepsilon_{Sub} < \varepsilon_{Obj}$, indicating the empirically-oriented method yields smaller weight dispersion, tending to assign similar weights to all granularities and weakening inter-granularity differences, which indirectly strengthens non-core and weakens core granularities - In practical decision-making, all selected indicators are irreplaceable, yet single weight determination methods often produce zero weights, which is unreasonable. The proposed comprehensive method effectively avoids zero weights for redundant indicators while preserving the fluctuation characteristics of single methods and providing more comprehensive information

---

References

[1] Wan Rong, Yan Ruixia. Method for determining customer demand weight based on rough set and FAHP [J]. Science and Technology Management Research, 2018, 38(4): 204-208.

[2] Zhang Wei, Wang Zhihai, Yuan Jidong, et al. A locally attribute weighted naïve Bayes classifier [J]. Journal of Beijing Jiaotong University, 2018, 42(02): 14-21.

[3] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356.

[4] Zadeh L A. Fuzzy sets and information granularity [C]//Advances in Fuzzy Set Theory and Applications, 1996: 3-18.

[5] Hu Qian, Mi Jusheng. Representation of multigranularity fuzzy rough sets and corresponding belief structure [J]. Computer Engineering and Applications, 2017, 53(19): 51-54.

[6] Luo Gongzhi, Qian Jiali. Neighborhood multi-granulation rough set based on multi-threshold for variable precision decisions [J/OL]. Application Research of Computers, 2018, 35(10). [2018-07-11]. http://kns.cnki.net/kcms/detail/51.1196.TP.20171010.1730.048.html.

[7] Zhang Qinghua, Guo Yonglong, Xue Yubin. Multi-granularity search algorithm based on probability statistics [J]. Pattern Recognition and Artificial Intelligence, 2015, 28(5): 422-428.

[8] Xiong Shenghua, Wu Sheng, Chen Zhensong, et al. Generalized intuitionistic fuzzy entropy and its application in weight determination [J]. Control and Decision, 2017, 32(5): 845-854.

[9] Qian Yuhua. Research on theory and method of granularity computing based on rough set [D]. Taiyuan: Shanxi University, 2005.

[10] Xue Zhanao, Zhu Tailong, Xue Tianyu, et al. Methodology of attribute weights acquisition based on three-way decision theory [J]. Computer Science, 2015, 42(08): 265-268, 272.

[11] Meng Huili, Ma Yuanyuan, Xu Jiucheng. The granularity reduction of pessimistic multi-granulation based on information quantity [J]. Journal of Nanjing University: Natural Sciences, 2015, 51(02): 343-348.

[12] Wang Xiaoyan, Shen Jialan, Shen Yuanxia. Graded multi-granulation rough set based on weighting granulations and dominance relation [J]. Journal of Shandong University: Natural Science, 2017, 52(03): 97-104.

[13] Zhou Danchen. A method for ascertaining the weight of attributes based on granular computing [J]. CAAI Transactions on Intelligent Systems, 2015, 10(2): 273-280.

[14] Wang Baoli, Ling Jiye, Hu Yunhong. Granular computing based hesitant fuzzy multi-criteria decision making [J]. Pattern Recognition and Artificial Intelligence, 2016, 29(3): 252-262.

[15] Zhang Ming, Cheng Ke, Yang Xibei, et al. Multi-granulation rough set based on weighted granulations [J]. Control and Decision, 2015, 30(2): 222-228.

[16] Xu W H, Sun W X, Zhang X Y, et al. Multiple granulation rough set approach to ordered information systems [J]. International Journal of General Systems, 2012, 41(5): 475-501.

[17] Huang Weihua. On a multi-granulation rough set model [J]. Journal of Southwest China Normal University: Natural Science Edition, 2017, 42(05): 137-143.

Note: Figure translations are in progress. See original paper for figures.