



Advantage matrix: two novel multi-attribute decision-making methods and their applications

Bin Yu^{1,2} · Zeshui Xu¹

Published online: 17 January 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

By comparing attributes of objects in an information system, the advantage matrix on the object set is established in this paper. The contributions can be identified as follows: (1) The advantage degree is proposed by the accumulation of the advantage matrix. (2) Based on the advantage matrix, the advantage (disadvantage) neighborhood approximation operator and the advantage (disadvantage) correlation approximation operator are defined and studied. Based on these two new operators, the neighborhood degree and the correlation degree are presented. The relationships between them are also investigated to demonstrate the value of the proposed method. (3) Finally, based on the above three degrees, new algorithms are designed, in which the effectiveness and robustness of the algorithms are analyzed by practical examples.

Keywords Advantage matrix · Advantage (disadvantage) neighborhood approximation operator · Advantage (disadvantage) correlation approximation operator · Advantage (neighborhood · correlation) · Decision-making

1 Introduction

Many decision-making situations in real life need to consider several criteria (Liang et al. 2015; Lin et al. 2013; Pearman 2014; Qian et al. 2014; Sun et al. 2017, 2017). Multiple-attribute decision making (MADM) is an important branch of decision-making. MADM methods attempt to select the best alternative(s) from a set of alternatives with respect to several criteria. Some of the classic methods of MADM have been proposed, such as ELECTRE method (Figueira et al. 2009) and PROMETHEE method (Brans and Vincke 1986) which are based on “Outranking Relationship”, AHP method constructed with the concept of Hierarchy (Belton 1986; Tian et al. 2018), MAVT method based on “Attribute

✉ Bin Yu
yu7bin@hotmail.com

✉ Zeshui Xu
xuzeshui@263.net

¹ Business School, Sichuan University, Chengdu 610064, Sichuan, People’s Republic of China

² Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410084, Hunan, People’s Republic of China

Value function" (Ferretti et al. 2014) and TOPSIS method (Hwang and Yoon 1981; Fan et al. 2013) and EDAS (Evaluation Based On Distance from Average Solution) method (Ghorabae et al. 2015; Canciglieri et al. 2015). These methods with clear concept and strong practicability, are known as effective ranking approaches and have been widely used in the research of various MADM problems. For example: Dragan and Fatih (Pamučar and Ecer 2020) propose a novel subjective weighting method called the Fuzzy Full Consistency Method (FUCOM-F) for determining weights as accurately as possible under fuzziness, and apply this method to the green supplier evaluation problem. In order to meet the challenge of the COVID-19 (COroNaVIrus Disease-2019) pandemic, health systems must adjust to new circumstances and establish separate hospitals exclusive for patients infected with SARS-CoV-2 virus. Žižović et al. (2021) put forward a multiple-criteria model for the evaluation and selection of nurses for COVID-19 hospitals and so on.

Rough set theory (RST) (Pawlak 1982) is a mathematical approach to imprecision, vagueness and uncertainty in data analysis. In RST, lower and upper approximations were defined to characterize a concept, namely, a subset of the universe. The main advantage of RST in data analysis is that it does not need any preliminary or additional information about data. RST has been widely applied to many fields, e.g., forecasting (Yu et al. 2020, 2019, 2019), medical diagnosis (Pattaraintakorn and Cercone 2008), machine learning (Hong et al. 2008, 2002), decision-making (Salamó and López-Sánchez 2011; Son et al. 2012; Swiniarski and Skowron 2003; Tian et al. 2011; Xiong et al. 2012; Zeleny 1976; Ye et al. 2021; Wang et al. 2021), pattern recognition (Wang and Wang 2009), case-based reasoning (Huang and Tseng 2004), and data mining (Lingras and Yao 1998; Yamaguchi 2009).

RST (Pawlak 1997; Pawlak and Slowinski 1994) is a valid mathematical approach to deal with the decision analysis, and has attracted much of attentions. In the past few years, many rough set-based MADM methods were presented. Based on covering (I, T)-fuzzy rough sets, Zhang et al. (2019) defined four kinds of fuzzy β -covering models, and designed a fuzzy rough TOPSIS method. By triangular norms and fuzzy implication operators, Jiang et al. (2018) defined four types of fuzzy β -covering variable precision (I, T)-fuzzy rough set models, and designed another kind of fuzzy rough TOPSIS method. In the intuitionistic fuzzy environment, Zhang et al. (2019) defined covering (I, T)-intuitionistic fuzzy rough set model based on D'eer fuzzy domain operator, and designed intuitionistic fuzzy rough TOPSIS method. At the same time, from the perspective of classical indiscernibility relation and big data, Yu et al. (2019) gave the attribute oriented λ -indiscernibility relation, then constructed the λ -rough set model, and finally successfully integrate the λ -rough set model and TOPSIS method to establish a kind of rough TOPSIS method, which is applied to student performance evaluation. Yu et al. (2019) used the α -indiscernibility relation to construct an α -rough fuzzy set model from the perspective of granular computing. Then they successfully integrated the α -rough fuzzy set model with PROMETHEE, and established a class of rough fuzzy PROMETHEE method, and applied it to the analysis of enterprise development. Based on the (γ, δ) -rough fuzzy set model, Yu et al. (2020) designed a multi-attribute decision-making method (the prediction method based on the minimum deviation), and discussed the application of this method.

Based on the review conducted above, some genuine challenges are identified which are presented in a nutshell below:

1. In traditional methods, dimension and attribute weight are used as a priori information, and different processing methods lead to great differences in decision-making results.

Table 1 An information system $S = \{U, Q, f, V\}$

S	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
x_1	0.1341	0.4417	0.4561	0.2982	0.0809	0.3381	0.6035	0.6925	0.6079
x_2	0.2126	0.0133	0.1017	0.0464	0.7772	0.2940	0.5261	0.5567	0.7413
x_3	0.8949	0.8972	0.9954	0.5054	0.9051	0.7463	0.7297	0.3965	0.1048
x_4	0.0715	0.1967	0.3321	0.7614	0.5338	0.0103	0.7073	0.0616	0.1279
x_5	0.2425	0.0934	0.2973	0.6311	0.1092	0.0484	0.7814	0.7802	0.5495
x_6	0.0538	0.3074	0.0620	0.0899	0.8258	0.6679	0.2880	0.3376	0.4852

Next, the TOPSIS method is illustrated as an example. In Table 1, the TOPSIS based on 3 determining attribute weight methods (Entropy method (Zou et al. 2006), Coefficient of variation method (Faber and Korn 1991), Cumulation method) and 4 normalization methods (Max method, Sum method, Max-Min method, Vector method) carries out the decision analysis, and the decision results are shown in Fig. 1a, b. In Fig. 1a, the influence of weight on sorting results is analyzed, the distance is European distance, and the vector method is used for normalization. In Fig. 1b, the influence of normalization on sorting results is analyzed, the distance is European distance, and all attribute weights are set to the same.

- In the process of decision method design, the distance (Euclidean distance, City block metric, Chebyshev distance, Mahalanobis distance, etc.) is usually used to measure the quality of objects. Choosing different distances also affects the decision results. Next, the TOPSIS method is illustrated as an example, too. In Table 1, the TOPSIS based on 4 distance carries out the decision analysis, and the decision results are shown in Fig. 1c. In terms of decision semantics, although weight is used to adjust the importance of attributes in the distance, it do not think about the internal influence between attributes, such as incompatibility and non operability.
- In the existing methods, the advantage relations between objects are established from various angles, but these relations often have nonlinear relationships, which is difficult

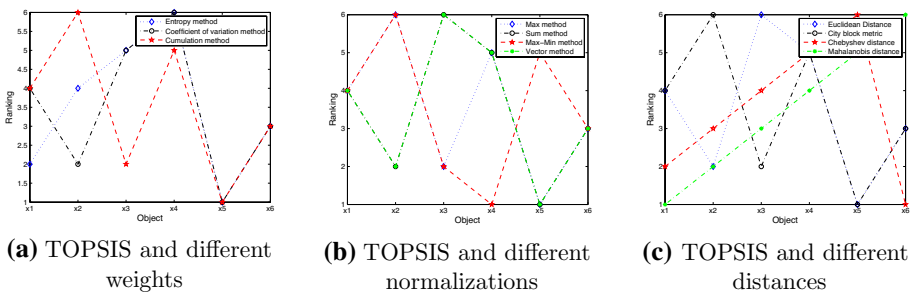


Fig. 1 Sorting results based on TOPSIS

to meet the needs of optimal ranking.

Motivated by these challenges, the major contributions of this research are threefold:

1. Through the comparison between attributes one by one, the advantage matrix is established to intuitively judge the advantage attributes among objects. From the perspective of decision semantics, it only compares the same attributes of different objects without considering the impact of attributes. The decision-making method based on advantage matrix design avoids dimensional processing, so as to weaken the influence of prior information on decision-making results.
2. The approximation operator is the core of domain rough set, and domain rough set does not need any prior information other than the data set required by the problem when describing or processing the problem, so the description or processing of the problem is more objective. The decision-making method based on approximation operators is designed to avoid the determination of attribute weight, so as to weaken the influence of a prior information on the decision-making results.
3. The non-linear advantage relation between objects is constructed based on the advantage matrix, the approximation operator of the advantage relation is established, and the neighborhood degree and the correlation degree are proposed to global analysis, further analysis of the object sequencing, so as to achieve the optimal ranking of objects.

The rest of this paper is organized as follows. In Sect. 2, based on the advantage matrix of objects, the neighborhood approximation operators and correlation approximation operators are constructed, together with investigating the properties of four types of approximations. In Sect. 3, two new MADM methods are designed according to the advantage matrix of objects. The experiments demonstrate that the synthetical results are effective and robust.

2 Main results

An information system (or knowledge representation system) is a finite table. The rows are labeled by objects, the columns are labeled by attributes, and the entries of the table are in their attribute-values. An information system is a 4-tuple $S = \{U, Q, f, V\}$, where (1) U is a non-empty finite set of objects; (2) Q is a non-empty finite set of attributes; (3) $V = \bigcup_{q \in Q} V_q$ where V_q is the set of values of the attribute q ; (4) $f : U \times Q \rightarrow V$ is a map such that $f(x, q) \in V_q$ for every $x \in U$ and $q \in Q$. For our purpose, throughout this study, let $V = [0, 1]$. The cardinality of a finite set X in this study is written as $|X|$.

Definition 2.1 (*Advantage matrix*) Given an information system S , its advantage matrix $D = [D(x_i, x_j)]$ is a $|U| \times |U|$ matrix, in which the element $D(x_i, x_j)$ for an object pair $(x_i, x_j) \in U \times U$ is defined by

$$D(x_i, x_j) = \{a \in Q \mid f(x_i, a) > f(x_j, a)\}. \quad (1)$$

The physical meaning of the matrix element $D(x_i, x_j)$ is that the object x_i is superior to x_j with respect to any attribute in $D(x_i, x_j)$. An advantage matrix D is nonsymmetric, i.e., $D(x_i, x_j) \neq D(x_j, x_i)$, $\exists x_i, x_j \in U$. Therefore, it is sufficient to consider the lower triangle and the upper triangle of the matrix.

Table 2 An information system $S = \{U, Q, f, V\}$

S	a_1	a_2	a_3	a_4
x_1	0.10	0.20	0.10	0.40
x_2	0.70	0.25	0.45	0.70
x_3	0.45	0.15	0.65	0.85
x_4	0.30	0.50	0.40	0.90
x_5	0.55	0.60	0.70	0.25

Example 2.2 Let $S = \{U, Q, f, V\}$ be an information system given in Table 2.

The advantage matrix D of Table 2 is shown below. For the object pair (x_3, x_1) , the entry $\{a_1, a_3, a_4\}$ indicates that the object x_3 is superior to x_1 by the attribute a_1, a_3 and a_4 .

$$D = \begin{bmatrix} \emptyset & \emptyset & \{a_2\} & \emptyset & \{a_4\} \\ Q & \emptyset & \{a_1, a_2\} & \{a_1, a_3\} & \{a_1, a_4\} \\ \{a_1, a_3, a_4\} & \{a_3, a_4\} & \emptyset & \{a_1, a_3\} & \{a_4\} \\ Q & \{a_2, a_4\} & \{a_2, a_4\} & \emptyset & \{a_4\} \\ \{a_1, a_2, a_3\} & \{a_2, a_3\} & \{a_1, a_2, a_3\} & \{a_1, a_2, a_3\} & \emptyset \end{bmatrix}$$

Definition 2.3 For an information system S and $x, y \in U$, if $D(x, y) = Q$ or $D(y, x) = \emptyset$, then x is completely superior to y , and denoted by $x \ggg y$ or $y \lll x$; if $|D(x, y)| > |D(y, x)|$, then x is relatively superior to y , and denoted by $x \gg y$ or $y \ll x$.

Examples are given below to elaborate Definition 2.3.

Example 2.4 (Continuation of Examples 2.2) By Definition 2.3, $x_2 \ggg x_1, x_4 \ggg x_1, x_3 \gg x_1, x_5 \gg x_1, x_5 \gg x_3$, and $x_5 \gg x_4$.

Definition 2.5 For an information system S . Then

1. If $\forall x, y \in U, x \neq y, D(x, y) \cup D(y, x) = Q$, then S is called a complementary information system (CIS).
2. If $\forall x, y \in U, x \neq y, D(x, y) \cup D(y, x) \neq Q$, then S is called a non-complementary information system (NCIS).
3. If $\forall x, y \in U, x \neq y, |D(x, y)| \neq |D(y, x)|$, then S is called a linear-comparable information system (LCIS).

Theorem 2.6 For an information system S and $x, y \in U$. Then

1. If S is a CIS, then $x \gg y \Leftrightarrow |D(x, y)| > \frac{|Q|}{2}$.
2. S is a CIS $\Leftrightarrow \forall x, y \in U, x \neq y, |D(x, y)| + |D(y, x)| = |Q|$.
3. S is an NCIS $\Leftrightarrow \exists a \in Q / (D(x, y) \cup D(x, y)), f(x, a) = f(y, a)$.
4. S is an LCIS $\Leftrightarrow \forall x, y \in U, x \gg y$ or $x \ll y$.

Proof According to Definitions 2.3 and 2.5, obviously. □

Proposition 2.7 For an information system S and $x, y, x_i, x_j \in U$. Then

1. $y \ggg x \Leftrightarrow f(y, a) > f(x, a), \forall a \in Q$.
2. $|D(y, x)| > \frac{|Q|}{2} \Rightarrow y \ggg x$ and $y \ggg x \Rightarrow |D(x, y)| < \frac{|Q|}{2}$.
3. $y \ggg x \Rightarrow y \gg x$.
4. $X = \bigcap_{j=1,2,\dots,|U|, x_j \neq x} D(x, x_j) \neq \emptyset \Rightarrow \forall a \in X, f(x, a) = \max_{j=1,2,\dots,|U|} \{f(x_j, a)\}$. In particular, if $X = Q$, then $x \gg x_j, \forall x_j \in U/\{x\}$, x is called a maximum object.
5. $X = \bigcap_{i=1,2,\dots,|U|, x_i \neq x} D(x_i, x) \neq \emptyset \Rightarrow \forall a \in X, f(x, a) = \min_{i=1,2,\dots,|U|} \{f(x_i, a)\}$. In particular, if $X = Q$, then $x_i \ggg x, \forall x_i \in U/\{x\}$, x is called a minimum object.

Proof According to Definitions 2.1 and 2.3, obviously. □

Definition 2.8 (Advantage degree) The advantage degree (AD) of an object x is defined by

$$AD(x) = \frac{\sum_{i=1}^{|U|} |D(x, x_i)|}{(|U| - 1) * |Q|}, \tag{2}$$

$x, x_i \in U$.

Note

1. In an information system, the advantage degree is used to describe by the advantage attribute of one object to other objects. By comparing the advantage degrees of each object, the advantages and disadvantages of the object are analyzed. This description method not only compares each object (by $[D(x, x_i)]$), but also compares all objects comprehensively (by \sum). In particular, $\frac{\sum_{i=1}^{|U|} |D(x, x_i)|}{(|U|-1)*|Q|} = \frac{\sum_{i=1}^{|U|} |D(x, x_i)| - |D(x, x)|}{(|U|-1)*|Q|}, |D(x, x)| \equiv 0$.
2. In Table 2, based on the total utility value ($\sum_{j=1}^4 f(x_i, a_j)$), $\sum_{j=1}^4 f(x_1, a_j) = 0.80, \sum_{j=1}^4 f(x_2, a_j) = 2.10, \sum_{j=1}^4 f(x_3, a_j) = 2.10, \sum_{j=1}^4 f(x_4, a_j) = 2.10,$ and $\sum_{j=1}^4 f(x_5, a_j) = 2.10$. The results allow us to believe that the objects $x_2, x_3,$ and x_4 are better than x_1 , and that x_2, x_3 and x_4 are not comparable. To achieve a more efficient sorting result in Table 2, the advantage degrees based on the advantage matrix are as follows: $AD(x_1) = 0.13, AD(x_2) = 0.63, AD(x_3) = 0.50, AD(x_4) = 0.56,$ and $AD(x_5) = 0.69$. Immediately, x_5 is superior to $x_2; x_2$ is superior to $x_4; x_4$ is superior to $x_3; x_3$ is superior to x_1 .

For the above two methods, incomparable situations will appear for two objects. In order to solve this kind of problem, two kinds of approximate operators on the basis of the advantage matrix are proposed.

Definition 2.9 Let $S = \{U, Q, f, V\}$ be an information system and $x, y \in U$.

1. The advantage and disadvantage neighborhood approximation operators are respectively defined as:

$$\underline{\uparrow}x = \{y \in U \mid |D(x, y)| < |D(y, x)|\} \tag{3}$$

and

$$\underline{\downarrow}x = \{y \in U \mid |D(x, y)| > |D(y, x)|\}. \tag{4}$$

2. The advantage and disadvantage correlation approximation operators are respectively defined as:

$$\overline{\uparrow}x = \{y \in U \mid D(y, x) \neq \emptyset\} \tag{5}$$

and

$$\overline{\downarrow}x = \{y \in U \mid D(x, y) \neq \emptyset\}. \tag{6}$$

Examples are given below to elaborate Definition 2.9.

Example 2.10 (Continuation of Examples 2.2) By Definition 2.9, we have $\overline{\uparrow}x_1 = \{x_2, x_3, x_4, x_5\}$, $\overline{\downarrow}x_1 = \emptyset$, $\overline{\uparrow}x_2 = \emptyset$, $\overline{\downarrow}x_2 = \{x_1\}$, $\overline{\uparrow}x_3 = \{x_5\}$, $\overline{\downarrow}x_3 = \{x_1\}$, $\overline{\uparrow}x_4 = \{x_5\}$, $\overline{\downarrow}x_4 = \{x_1\}$, $\overline{\uparrow}x_5 = \emptyset$, $\overline{\downarrow}x_5 = \{x_1, x_3, x_4\}$, $\overline{\uparrow}x_1 = \{x_2, x_3, x_4, x_5\}$, $\overline{\downarrow}x_1 = \{x_3, x_5\}$, $\overline{\uparrow}x_2 = \{x_3, x_4, x_5\}$, $\overline{\downarrow}x_2 = \{x_1, x_3, x_4, x_5\}$, $\overline{\uparrow}x_3 = \{x_1, x_2, x_4, x_5\}$, $\overline{\downarrow}x_3 = \{x_1, x_2, x_4, x_5\}$, $\overline{\uparrow}x_4 = \{x_2, x_3, x_5\}$, $\overline{\downarrow}x_4 = \{x_1, x_2, x_3, x_5\}$, $\overline{\uparrow}x_5 = \{x_1, x_2, x_3, x_4\}$, $\overline{\downarrow}x_5 = \{x_1, x_2, x_3, x_4\}$.

Note

- For an information system, the advantages and disadvantages of the objects are analyzed by combining the advantage and disadvantage neighborhood. They reflect the advantages and disadvantages of objects in information systems from different perspectives. Assuming that we have six objects x_1, x_2, x_3, x_4, x_5 , and x_6 , in which x_2 is superior to x_1 ; x_3 is superior to x_2 ; x_4 is superior to x_3 ; x_5 is superior to x_4 ; and x_6 is superior to x_5 . Then $\overline{\uparrow}x_1 \supseteq \overline{\uparrow}x_2 \supseteq \overline{\uparrow}x_3 \supseteq \overline{\uparrow}x_4 \supseteq \overline{\uparrow}x_5 \supseteq \overline{\uparrow}x_6$ and $\overline{\downarrow}x_1 \subseteq \overline{\downarrow}x_2 \subseteq \overline{\downarrow}x_3 \subseteq \overline{\downarrow}x_4 \subseteq \overline{\downarrow}x_5 \subseteq \overline{\downarrow}x_6$. By contrast, if $\overline{\uparrow}x_1 \supseteq \overline{\uparrow}x_2 \supseteq \overline{\uparrow}x_3 \supseteq \overline{\uparrow}x_4 \supseteq \overline{\uparrow}x_5 \supseteq \overline{\uparrow}x_6$ and $\overline{\downarrow}x_1 \subseteq \overline{\downarrow}x_2 \subseteq \overline{\downarrow}x_3 \subseteq \overline{\downarrow}x_4 \subseteq \overline{\downarrow}x_5 \subseteq \overline{\downarrow}x_6$, then x_2 is superior to x_1 ; x_3 is superior to x_2 ; x_4 is superior to x_3 ; x_5 is superior to x_4 ; and x_6 is superior to x_5 . The less objects in the advantage neighborhood of the object x , the result in the higher priority x achieves. The more objects in the disadvantage neighborhood of the object x , the result in the higher priority x achieves.
- In an information system, $\overline{\uparrow}x$ represents the set of objects that have an advantage attribute for the object x , $\overline{\downarrow}x$ represents the set of objects that have a disadvantage attribute for the object x . They explain the correlation between objects in the information system from different views. The semantic expression is as follows: The less objects in the advantage correlation of the object x , the result in the higher priority x achieves. The more objects in the disadvantage correlation of the object x , the result in the higher priority x achieves.

According to the semantic expression of neighborhood and correlation approximation operators, we propose the following two new degrees.

Definition 2.11 (Neighborhood degree and correlation degree) Let $S = \{U, Q, f, V\}$ be an information system and $x \in U$.

- The neighborhood degree (ND) of an object x is defined as

$$ND(x) = \frac{|\overline{\downarrow}x|}{|\overline{\uparrow}x|}. \tag{7}$$

2. The correlation degree (CD) of an object x is defined as

$$CD(x) = \frac{|\bar{\downarrow}x|}{|\uparrow x|}. \tag{8}$$

Proposition 2.12 *Let $S = \{U, Q, f, V\}$ be an information system and $x, y \in U$. Then*

1. *If x is a maximum object, then $\uparrow x = \emptyset$, $\downarrow x = U \setminus \{x\}$, $\bar{\uparrow}x = \emptyset$, and $\bar{\downarrow}x = U \setminus \{x\}$.*
2. *If x is a minimum object, then $\bar{\downarrow}x = \emptyset$, $\bar{\uparrow}x = U \setminus \{x\}$, $\downarrow x = \emptyset$, and $\uparrow x = U \setminus \{x\}$.*
3. *If x is a maximum object, then $\forall y \in U, \overline{AD}(x) \geq AD(y)$.*
4. *If x is a minimum object, then $\forall y \in U, AD(x) \leq AD(y)$.*

Proof According to Proposition 2.7 and Definition 2.9, obviously. □

Theorem 2.13 *Let $S = \{U, Q, f, V\}$ be an information system and $x, y \in U$. Then*

1. $\uparrow x \cap \downarrow x = \emptyset$ and $\bar{\uparrow}x \cap \bar{\downarrow}x = \emptyset$.
2. $\bar{\uparrow}x \subseteq \uparrow x$ and $\downarrow x \subseteq \bar{\downarrow}x$.
3. $\frac{|\bar{\uparrow}x|}{|\uparrow x|} \leq ND(x) \leq \frac{|\bar{\downarrow}x|}{|\downarrow x|}$ and $\frac{|\downarrow x|}{|\uparrow x|} \leq CD(x) \leq \frac{|\bar{\downarrow}x|}{|\bar{\uparrow}x|}$.

Proof

1. According to Definition 2.9, obviously.
2. If $y \in \uparrow x$, then $|D(x, y)| < |D(y, x)|$. Since $|D(x, y)| \geq 0$, we have $|D(y, x)| > 0$, that is, $D(y, x) \neq \emptyset$. So, $y \in \bar{\uparrow}x$, thus, $\bar{\uparrow}x \subseteq \uparrow x$. Similarly, $\downarrow x \subseteq \bar{\downarrow}x$.
3. According to (1), (3) is straightforward.

Theorem 2.14 *Let $S = \{U, Q, f, V\}$ be a CIS and $x \in U$. Then*

1. $\uparrow x = \{y \in U \mid |D(y, x)| \geq \frac{|Q|}{2}\}$ and $\downarrow x = \{y \in U \mid |D(x, y)| > \frac{|Q|}{2}\}$.
2. $\bar{\uparrow}x \cup \downarrow x = U \setminus \{x\}$ and $\bar{\uparrow}x \cup \downarrow x = U \setminus \{x\}$.
3. $ND(x) = \frac{|U|-1}{|\uparrow x|} - 1 = \frac{|\downarrow x|}{|U|-1-|\downarrow x|}$.
4. $CD(x) = \frac{|U|-1}{|\uparrow x|} - 1 = \frac{|\bar{\downarrow}x|}{|U|-1-|\bar{\downarrow}x|}$.

Proof

1. Since S is a CIS, we have $|D(x, y)| + D(y, x) = |Q|$ by Theorem 2.6. By Definition 2.9 (1), we have $|D(y, x)| - |D(x, y)| = |D(y, x)| - |Q| + |D(y, x)| > 0$, so $|D(y, x)| > \frac{|Q|}{2}$. Thus, $\uparrow x = \{y \in U \mid |D(y, x)| > \frac{|Q|}{2}\}$. Similarly, we can proof that $\downarrow x = \{y \in U \mid |D(x, y)| > \frac{|Q|}{2}\}$.
2. According to Definition 2.1, $D(x, x) = \emptyset, \forall x \in U$. By Definition 2.9, we have $x \in \uparrow x, \downarrow x, \bar{\uparrow}x, \bar{\downarrow}x$. Then $\bar{\uparrow}x \cup \downarrow x \subseteq U \setminus \{x\}$. Conversely, suppose $y \notin \bar{\uparrow}x \cup \downarrow x, \forall y \in U \setminus \{x\}$. Then, $y \notin \bar{\uparrow}x$ and $y \notin \downarrow x$. If $y \notin \bar{\uparrow}x$, then $|D(y, x)| < \frac{|Q|}{2}$. Since S is a CIS, we have $|D(x, y)| + D(y, x) = |Q|, \bar{D}(x, y) > \frac{|Q|}{2}$. That is, $y \in \downarrow x$, which contradicts $y \notin \downarrow x$. Thus, $y \in \bar{\uparrow}x \cup \downarrow x, \bar{\uparrow}x \cup \downarrow x \supseteq U \setminus \{x\}$. In summary, $\bar{\uparrow}x \cup \downarrow x = U \setminus \{x\}$. Similarly, $\bar{\uparrow}x \cup \downarrow x = U \setminus \{x\}$.

3. According to (2) and Theorem 2.13 (2), we have $|\underline{\uparrow}x| + |\underline{\downarrow}x| = |U| - 1$. Then,

$$ND(x) = \frac{|\underline{\downarrow}x|}{|\underline{\uparrow}x|} = \frac{|U|-1-|\underline{\uparrow}x|}{|\underline{\uparrow}x|} = \frac{|U|-1}{|\underline{\uparrow}x|} - 1 = \frac{|\underline{\downarrow}x|}{|U|-1-|\underline{\downarrow}x|}.$$

4. It is similar to the proof of (3).

□

3 Application

3.1 Algorithms and examples

In the sequel, we present the general problem with the best selection from m objects $(x_i, i = 1, 2, \dots, m)$. Here we evaluate and compare with the other objects on the basis of attribute $(a_j, j = 1, 2, \dots, n)$ whose values are known to us. Given a decision information system $S = \{U, Q, f, V\}$, where $U = \{x_1, x_2, \dots, x_m\}$, $Q = \{a_1, a_2, \dots, a_n\}$, and $f : U \times Q \rightarrow V, f : (x_i, a_j) \mapsto [0, 1]$. follows.

Algorithm 3.1 (Multi attribute decision making based on the advantage matrix)

Input: An information system $S = \{U, Q, f, V\}$

Output: A ranking result of all objects.

Step 1: Compute the advantage matrix $D = [D(x_i, x_j)]$, the object $x_i, x_j \in U, \setminus \setminus$ according to Definition 2.1;

Step 2: Compute the advantage degree $AD(x_i), x_i \in U, \setminus \setminus$ according to Definition 2.8;

Step 3: Rank the objects according to the AD of each object. The bigger the $AD(x_i)$ is, the better the object x_i will be,

$x_i < x_j \iff AD(x_i) < AD(x_j)$, or $x_i \sim x_j \iff AD(x_i) = AD(x_j), i, j = 1, 2, \dots, |U|$;

Step 4: End.

Algorithm 3.2 (Multi attribute decision making based on the neighborhood approximation operator and the correlation approximation operator)

Input: A information system $S = \{U, Q, f, V\}$ and an increase Δ .

Output: A ranking result of all objects.

Step 1: Compute the advantage matrix $D = [D(x_i, x_j)]$, the object $x_i, x_j \in U, \setminus \setminus$ according to Definition 2.1;

Step 2: Compute the neighborhood degree $ND(x_i), x_i \in U, \setminus \setminus$ according to Definition 2.11 (1);

Step 3: Compute the correlation degree $CD(x_i), x_i \in U, \setminus \setminus$ according to Definition 2.11 (2);

Step 4: If $x_i, x_j \in U, ND(x_i) = ND(x_j)$ and $CD(x_i) > CD(x_j)$, then $NDCD(x_i) = ND(x_i) + \Delta$; If $x_i, x_j \in U, ND(x_i) = ND(x_j)$ and $CD(x_i) < CD(x_j)$, then $NDCD(x_j) = ND(x_j) + \Delta$; If $x_i, x_j \in U, ND(x_i) < ND(x_j)$, then $NDCD(x_i) = ND(x_i)$ and $NDCD(x_j) = ND(x_j)$, where ND and CD are mixed to establish NDCD;

Step 5: Rank the objects according to the NDCD of each object. The bigger the $NDCD\{x_i\}$ is, the better the object x_i will be,

$x_i < x_j \iff NDCD\{x_i\} < NDCD\{x_j\}, i, j = 1, 2, \dots, |U|$;

Step 6: End.

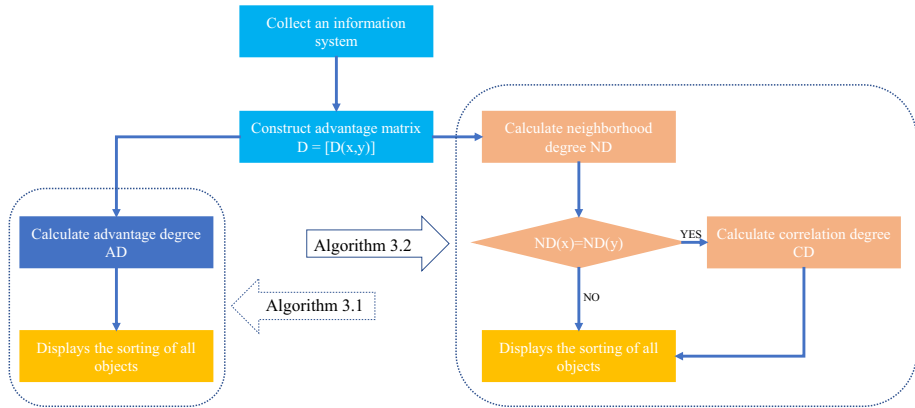


Fig. 2 The flowchart of Algorithms 3.1 and 3.2

Table 3 An information system $S = \{U, Q, f, V\}$

S	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
x_1	0.1080	0.8722	0.4519	0.4272	0.8291	0.8776	0.5811	0.5289	0.4445	0.3453
x_2	0.4599	0.0522	0.3334	0.1687	0.6266	0.0144	0.6372	0.6944	0.0854	0.9468
x_3	0.4509	0.2197	0.0591	0.7517	0.5387	0.2943	0.6513	0.2124	0.0573	0.5202
x_4	0.5511	0.4596	0.7409	0.3684	0.6505	0.1799	0.8646	0.5433	0.6295	0.9538
x_5	0.8054	0.9585	0.5068	0.9418	0.7266	0.9263	0.0560	0.7025	0.7962	0.0736
x_6	0.7009	0.7900	0.1999	0.0172	0.0945	0.0682	0.8169	0.9564	0.6912	0.2070

To understand Algorithms 3.1 and 3.2 better, they are described by a flowchart as Fig. 2. Next, an example is introduced to demonstrate the two algorithms.

Example 3.3 Let $S = \{U, Q, f, V\}$ be an information system given in Table 3.

(1) Based on Algorithm 3.1,

Step 1: Compute the advantage matrix $D = [D(x_i, x_j)]$.

\emptyset	$[a_2, a_3, a_4, a_5, a_6, a_9]$	$[a_2, a_3, a_5, a_6, a_8, a_9]$
$[a_1, a_7, a_8, a_{10}]$	\emptyset	$[a_1, a_3, a_5, a_8, a_9, a_{10}]$
$[a_1, a_4, a_7, a_{10}]$	$[a_2, a_4, a_6, a_7]$	\emptyset
$[a_1, a_3, a_7, a_8, a_9, a_{10}]$	$[a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_9, a_{10}]$	$[a_1, a_2, a_3, a_5, a_7, a_8, a_9, a_{10}]$
$[a_1, a_2, a_3, a_4, a_6, a_8, a_9]$	$[a_1, a_2, a_3, a_4, a_5, a_6, a_8, a_9]$	$[a_1, a_2, a_3, a_4, a_5, a_6, a_8, a_9]$
$[a_1, a_7, a_8, a_9]$	$[a_1, a_2, a_6, a_7, a_8, a_9]$	$[a_1, a_2, a_3, a_7, a_8, a_9]$
	$[a_2, a_4, a_5, a_6]$	$[a_5, a_7, a_{10}]$
	$[a_8]$	$[a_7, a_{10}]$
	$[a_4, a_6]$	$[a_7, a_{10}]$
	\emptyset	$[a_3, a_7, a_{10}]$
	$[a_1, a_2, a_4, a_5, a_6, a_8, a_9]$	\emptyset
	$[a_1, a_2, a_8, a_9]$	\emptyset

Step 2: Compute the advantage degree, $AD(x_1) = 0.5000$, $AD(x_2) = 0.3400$, $AD(x_3) = 0.3200$, $AD(x_4) = 0.6400$, $AD(x_5) = 0.7400$, and $AD(x_6) = 0.4600$.

Step 3: Rank the objects according to the AD of each object, $x_3 < x_2 < x_6 < x_1 < x_4 < x_5$

Step 4: End.

(2) Based on Algorithm 3.2,

Step 1: Compute the advantage matrix D , it is the same as (1).

Step 2: Compute the neighborhood degree, $ND(x_1) = 1.5000$, $ND(x_2) = 0.2500$, $ND(x_3) = 0$, $ND(x_4) = 4.0000$, $ND(x_5) = \infty$, and $ND(x_6) = 0.6667$.

Step 3: According to Step 2, skip.

Step 4: Skip.

Step 5: Rank the objects according to the $NDCD(NDCD = ND)$, $x_3 < x_2 < x_6 < x_1 < x_4 < x_5$

Step 6: End.

3.2 Case analysis

For the sake of proving the effectiveness and robustness of our new algorithms, Student Performance Data Set Student performance (2014) is selected from the UC Irvine Machine Learning Repository (UCI). In this data approach student achievement in secondary education of two Portuguese schools. The data attributes include student grades, demographic, social and school related features) and it was collected by using school reports and questionnaires, as shown in Tables 4 and 5 . Two data sets are provided regarding the performance in two distinct subjects: Mathematics (Mat) and Portuguese language (Por). In Cortez and Silva (2008), the two data sets are modeled under 5-level classification and regression tasks. Important note: the target attribute d_3 has a strong correlation with attributes d_2 and d_1 .

It is generally believed that $a_4, a_7, a_8, a_{13}, a_{14}, a_{16}, a_{17}, a_{18}, a_{20}, a_{21}, a_{24}, a_{25}, a_{27}, a_{28}, a_{29}$, and a_{30} are related to students' grades. Based on Student Performance Data Set Student performance (2014), 4 new information systems are set up, as shown in Tables 6, 7, 8 and 9. In Tables 7 and 9 , d_1, d_2 and d_3 represent the grades of 3 stages and actual results, as

Table 4 Attribute information

Attribute	Description (Domain)
a_1 school	Student's school (binary: Gabriel Pereira or Mousinho da Silveira)
a_2 sex	Student's sex (binary: female or male)
a_3 age	Student's age (numeric: from 15 to 22)
a_4 address	Student's home address type (binary: 1 - urban or 0 - rural)
a_5 famsize	Family size (binary: less or equal to 3 or greater than 3)
a_6 Pstatus	Parent's cohabitation status (binary: living together or apart)
a_7 Medu	Mother's education (numeric: 0 - none, 0.2 - primary education(4th grade), 0.4 - 5th to 9th grade, 0.6 - secondary education or 0.8 - higher education)
a_8 Fedu	Father's education (numeric: 0 - none, 0.2 - primary education(4th grade), 0.4 - 5th to 9th grade, 0.6 - secondary education or 0.8 - higher education)
a_9 Mjob	Mother's job (nominal: 'teacher', 'health' care related, civil 'services')
a_{10} Fjob	Father's job (nominal: 'teacher', 'health' care related, civil 'services')
a_{11} reason	Reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')
a_{12} guardian	Student's guardian (nominal: 'mother', 'father' or 'other')
a_{13} travelttime	Home to school travel time (numeric: 0.8 - less than 15 min., 0.6 - 15 to 30 min., 0.39 - 30 min. to 1 hour, or 0.19 - greater than 1 hour)
a_{14} studytime	Weekly study time (numeric: 0.2 - less than 2 hours, 0.4 - 2 to 5 hours, 0.6 - 5 to 10 hours, or 0.8 - greater than 10 hours)
a_{15} failures	Number of past class failures (numeric: n , if $1 \leq n < 3$, else 4)
a_{16} schoolsup	Extra educational support (binary: 1 - yes or 0 - no)
a_{17} famsup	Family educational support (binary: 1 - yes or 0 - no)
a_{18} paid	Extra paid classes within the course subject (Math or Portuguese) (binary: 1=yes or 0=no)
a_{19} activities	Extra-curricular activities (binary: yes or no)
a_{20} nursery	Attended nursery school (binary: 1 - yes or 1 - no)
a_{21} higher	Wants to take higher education (binary: 1 - yes or 0 - no)
a_{22} internet	Internet access at home (binary: yes or no)
a_{23} romantic	With a romantic relationship (binary: yes or no)
a_{24} famrel	Quality of family relationships (numeric: 0.2 - very bad, 0.4 - bad, 0.6 - good, 0.8 - very good or 1 - excellent)
a_{25} freetime	Free time after school (numeric: 0.2 - very low, 0.4 - low, 0.6 - general, 0.8 - high, 1 - very high)
a_{26} goout	Going out with friends (numeric: from 1 - very low to 5 - very high)
a_{27} Dalc	Workday alcohol consumption (numeric: 0.9 - very low, 0.6 - low, 0.4 - general, 0.2 - high, 0 - very high)
a_{28} Walc	Weekend alcohol consumption (numeric: 0.9 - very low, 0.6 - low, 0.4 - general, 0.2 - high, 0 - very high)
a_{29} health	Current health status (numeric: 0.2 - very bad, 0.4 - bad, 0.6-good, 0.8 - very good or 1 - excellent)
a_{30} attendance	(Rate of) attendance (numeric: $\frac{n}{93}$, n number of school attendances)

Table 5 Decision attribute information

Target attribute	Description (Domain)
d_1	First period grade (numeric: 5 - very good (16-20), 4 - good (14-15), 3 - satisfactory (12-13), 2 - sufficient (10-11) , 1 - fail (0-9)
d_2	Second period grade (numeric: 5 - very good (16-20), 4 - good (14-15), 3 - satisfactory (12-13), 2 - sufficient (10-11) , 1 - fail (0-9)
d_3	Final grade (numeric: 5 - very good (16-20), 4 - good (14-15), 3 - satisfactory (12-13), 2 - sufficient (10-11) , 1 - fail (0-9)

In the target item, the higher the level of the object, the better the object is

Table 6 An information system M based on Mathematics course

M	a_4	a_7	a_8	a_{13}	.	a_{17}	a_{18}	.	a_{24}	a_{25}	.	a_{29}	a_{30}
x_1	1.00	0.80	0.80	0.60	.	0	0	.	0.80	0.60	.	0.60	0.94
x_2	1.00	0.20	0.20	0.80	.	1.00	0	.	1.00	0.60	.	0.60	0.96
.
x_{100}	1.00	0.80	0.60	0.80	.	1.00	1.00	.	1.00	0.60	.	0.60	1.00
.
x_{200}	1.00	0.80	0.80	0.80	.	1.00	1.00	.	0.80	1.00	.	0.60	1.00
.
x_{395}	1.00	0.20	0.20	0.80	.	0	0	.	0.60	0.40	.	1.00	0.95

Table 7 Three period achievements based on Mathematics course

M	d_1	d_2	d_3
x_1	1	1	1
x_2	1	1	1
.	.	.	.
x_{100}	1	1	1
.	.	.	.
x_{200}	1	1	2
.	.	.	.
x_{395}	1	1	1

Table 8 An information system P based on Portuguese language course

P	a_4	a_7	a_8	a_{13}	.	a_{17}	a_{18}	.	a_{24}	a_{25}	.	a_{29}	a_{30}
y_1	1.00	0.80	0.80	0.60	.	0	0	.	0.80	0.60	.	0.6000	0.96
y_2	1.00	0.20	0.20	0.80	.	1.00	0	.	1.00	0.60	.	0.60	0.98
.
y_{200}	1.00	0.60	0.60	0.80	.	1.00	0	.	0.80	0.60	.	0.8000	0.94
.
y_{400}	1.00	0.60	0.40	0.80	.	0	0	.	1.00	0.8000	.	0.2000	0.96
.
y_{649}	0	0.60	0.40	0.39	.	0	0	.	0.80	0.80	.	1.00	0.96

Table 9 Three period achievements based on Portuguese language course

P	d_1	d_2	d_3
y_1	1	2	2
y_2	1	2	2
.	.	.	.
y_{200}	2	3	4
.	.	.	.
y_{400}	4	5	5
.	.	.	.
y_{649}	2	2	2

shown. Tables 7 and 9 are used as comparison indicators to verify whether the sorting results are correct.

3.3 Results and discussions

3.3.1 Sorting results and Spearman rank correlation coefficient

To demonstrate the application and effect of Algorithms 3.1 and 3.2, these two new algorithms are used to deal with two data sets (Mat and Por), $\Delta = 0.0001$. To show the experimental results, 3 object sets are randomly selected from Tables 6 and 8, respectively. $\{x_{56}, x_{362}, x_{132}, x_{17}, x_{302}\}$, $\{x_{277}, x_{309}, x_1, x_{19}, x_{235}\}$, $\{x_{207}, x_{204}, x_{326}, x_{100}, x_{365}\}$ are selected from Table 6. $\{y_{426}, y_{62}, y_{201}, y_{366}, y_{625}\}$, $\{y_{598}, y_{401}, y_{193}, y_{103}, y_{411}\}$, $\{y_{311}, y_{27}, y_{141}, y_{226}, y_{629}\}$ are selected from Table 8. Algorithms 3.1 and 3.2 are used to analyze the above 6 object sets, and the results are shown in Figs. 3 and 4.

In Fig. 3a, Algorithms 3.1 and 3.2 are used to sort $x_{56}, x_{362}, x_{132}, x_{17}, x_{302}$, and compare sorting results with Table 7 (d_1). In Fig. 3b, Algorithms 3.1 and 3.2 are used to sort $x_{277}, x_{309}, x_1, x_{19}, x_{235}$, and compare sorting results with Table 7 (d_2). In Fig. 3c, Algorithms 3.1 and 3.2 are used to sort $x_{207}, x_{204}, x_{326}, x_{100}, x_{365}$, and compare sorting results with Table 7 (d_3).

In Fig. 4a, Algorithms 3.1 and 3.2 are used to sort $y_{426}, y_{62}, y_{201}, y_{366}, y_{625}$, and compare sorting results with Table 9 (d_1). In Fig. 4b, Algorithms 3.1 and 3.2 are used to sort $y_{598}, y_{401}, y_{193}, y_{103}, y_{411}$, and compare sorting results with Table 9 (d_2). In Fig. 4c,

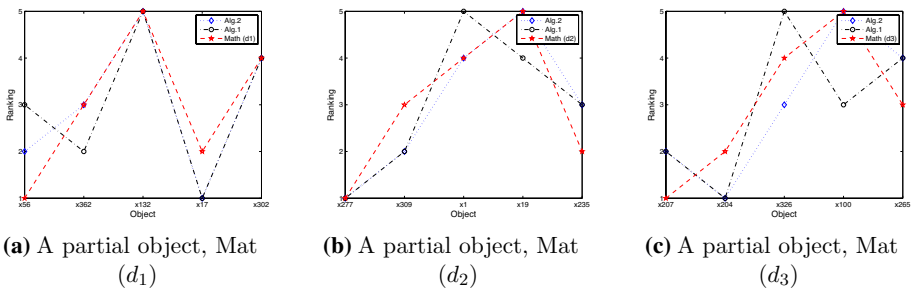


Fig. 3 The sort based on Mathematics

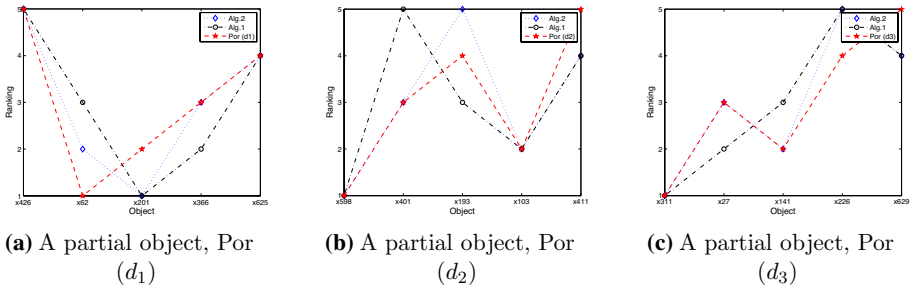


Fig. 4 The sort based on Portuguese

Table 10 Spearman ρ based on Fig. 3a

ρ	Alg.3.2	Alg.3.1	Mat(d_1)
Alg.3.2	1	0.9	0.9
Alg.3.1	0.9	1	0.7
Mat(d_1)	0.9	0.7	1

Table 11 Spearman ρ based on Fig. 3b

ρ	Alg.3.2	Alg.3.1	Mat(d_2)
Alg.3.2	1	0.9	0.9
Alg.3.1	0.9	1	0.8
Mat(d_2)	0.9	0.8	1

Table 12 Spearman ρ based on Fig. 3c

ρ	Alg.3.2	Alg.3.1	Mat(d_3)
Alg.3.2	1	0.6	0.8
Alg.3.1	0.6	1	0.6
Mat(d_3)	0.8	0.6	1

Table 13 Spearman ρ based on Fig. 4a

ρ	Alg.3.2	Alg.3.1	Por(d_1)
Alg.3.2	1	0.9	0.9
Alg.3.1	0.6	1	0.7
Por(d_1)	0.9	0.7	1

Algorithms 3.1 and 3.2 are used to sort $y_{311}, y_{27}, y_{141}, y_{226}, y_{629}$, and compare sorting results with Table 9 (d_3).

To better describe the consistency between the sorting results and the actual results, the Spearman rank correlation coefficient or Spearman ρ (Myers et al. 2013) is used to analyze the results in Figs. 3 and 4, the results are shown in Tables 10, 11, 12, 13, 14

Table 14 Spearman ρ based on Fig. 4b

ρ	Alg.3.2	Alg.3.1	Por(d_2)
Alg.3.2	1	0.6	0.9
Alg.3.1	0.6	1	0.7
Por(d_2)	0.9	0.7	1

Table 15 Spearman ρ based on Fig. 4c

ρ	Alg.3.2	Alg.3.1	Por(d_3)
Alg.3.2	1	0.9	0.9
Alg.3.1	0.9	1	0.8
Por(d_3)	0.9	0.8	1

and 15. It is widely known that if the Spearman ρ between sorting results of two methods belongs to $[0, 1]$, then the two methods are positively correlated. Moreover, as a commonly accepted notion, if the Spearman ρ between sorting results of two methods is greater than 0.6, then the correlation degree between the two methods is high.

In Figs. 3 and 4, the broken line trend of each subgraph is basically consistent, which also shows that the results obtained by Algorithms 3.1 and 3.2 are basically consistent with the actual sorting results. In Tables 10, 11, 12, 13, 14 and 15, the Spearman ρ among Algorithms 3.1, 3.2 and the real sorting results are greater than or equal to 0.6. In particular, the most results are greater than or equal to 0.8. In the summary, there is a high correlation among Algorithms 3.1, 3.2 and the real results, and Algorithms 3.1 and 3.2 are effective.

3.3.2 Random results and accuracy

The number of objects of the two data sets (Mat and Por) is relatively large, and the real ranking is 5 levels, so it is unreasonable to sort all objects. This paper uses the random sampling method to verify the accuracy of Algorithms 3.1 and 3.2. The specific process is as follows:

For Fig. 5a, the specific process is as follows: (1) two objects are randomly selected from 395 objects in the Mathematical data set; (2) the objects are sorted by comparing the NDCD values and Mat (d_1), if the two sorting results are consistent, the result of Algorithm 3.2 is accurate; (3) Repeat (1) and (2) 1000 times, and count the proportion of consistent sorting results as the accuracy of Algorithm 3.2; (4) Repeat (1), (2) and (3) 1000 times to obtain the accuracy of 1000 times, and draw the diagram. Other subgraphs of Fig. 5 are similarly processed. Overall, the accuracies of Algorithms 3.2 and 3.1 are distributed between 0.60 and 0.72, and the whole random process is randomly selected $1000 * 1000 = 1 * 10^6$ times (Select 2 of the 395 objects, $C_{395}^2 = 77815 < 1 * 10^6$).

Comparing the process in Fig. 5, the Mathematic data set is replaced by the Portuguese data set. Other processes are the same, and the results are shown in Fig. 6, the accuracies of Algorithms 3.2 and 3.1 are distributed between 0.62 and 0.74, and the whole random process is randomly selected $1000 * 1000 = 1 * 10^6$ times (Select 2 of the 649 objects, $C_{649}^2 = 210276 < 1 * 10^6$).

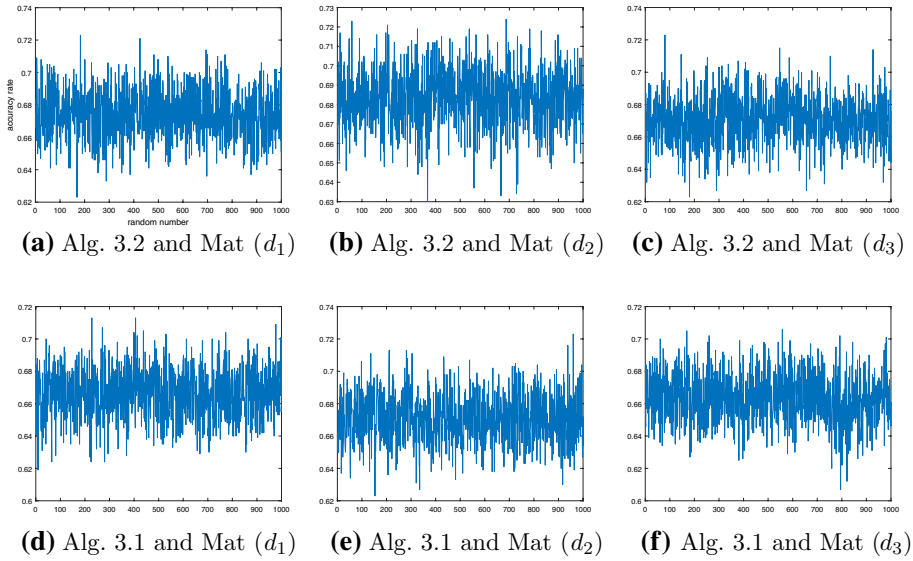


Fig. 5 Stochastic analysis based on the Mathematic data set

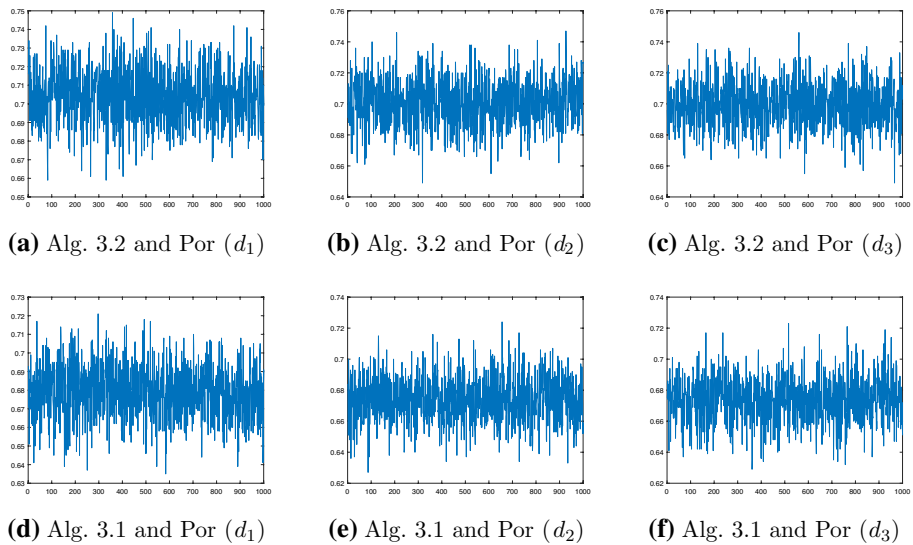
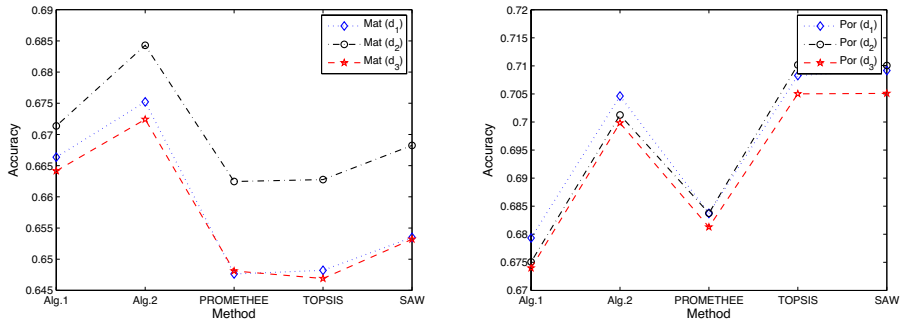


Fig. 6 Stochastic analysis based on the Portuguese data set

3.3.3 Accuracy comparison and stability analysis

In order to show the accuracy of Algorithms 3.2, 3.1, PROMETHEE, TOPSIS, and SAW, the detailed test process is as follows:

For Fig. 7a, the specific process is as follows: (1) Sort 395 objects in the Mathematical data set by Algorithms 3.2, 3.1, PROMETHEE, TOPSIS, and SAW; (2) Generate all



(a) Accuracy comparison based on Mat (b) Accuracy comparison based on Por

Fig. 7 Accuracy comparison

binary combinations, $C^2_{395} = 77815$; (3) Based on all binary combinations, the proportion that the sorting of Algorithm 3.2 (Algorithm 3.1, PROMETHEE, TOPSIS, SAW) is consistent with Mat (d_1) is recorded as the accuracy of Algorithm 3.2 (Algorithm 3.1, PROMETHEE, TOPSIS, SAW), and the results are shown in Fig. 7a (Mat (d_1)). Other broken lines in Fig. 7a shall be subject to the same test process. The accuracy analysis based on the Portuguese data set is shown in Fig. 7b.

Overall, the accuracy of Algorithms 3.2 and 3.1 is higher than that of PROMETHEE, TOPSIS, and SAW in the mathematical data set; In the Portuguese data set, Algorithm 3.2 is exactly the same as PROMETHEE, TOPSIS, and SAW, but Algorithm 3.1 is lower than PROMETHEE, TOPSIS, and SAW.

For the two data sets (Mat and Por), the sensitivity of 5 algorithms (Algorithms 3.2, 3.1, PROMETHEE, TOPSIS, SAW) is analyzed by deleting attributes. The specific test process is as follows:

For Fig. 8, by deleting the attributes one by one, the influence of attribute change on the accuracy of 5 algorithms (Algorithms 3.2, 3.1, PROMETHEE, TOPSIS, SAW) is verified. Combined with Figs. 7 and 8, the accuracy change is shown in Fig. 9.

For Fig. 9a–c, PROMETHEE and TOPSIS have the worst stability, followed by SAW, and Algorithms 3.1 and 3.2 have the best stability. For Fig. 9 (d),(e),(f), PROMETHEE has the worst stability, and other methods have similar stability. Overall, SAW is linear and more stable, but its scope of application is limited; PROMETHEE and TOPSIS are non-linear and poor stable, but their scope of application are extensive. Algorithms 3.1 and 3.2 are non-linear, the stability of them is similar to SAW, or even better.

In summary, (1) From the experimental results: Algorithms 3.1 and 3.2 are used for student ranking based on two data sets (Mat and Por). The sorting results of Algorithms 3.1 and 3.2 are compared with the real sorting results. Algorithms 3.1 and 3.2 are highly correlated with the actual results; By comparing Algorithms 3.1 and 3.2 with PROMETHEE, TOPSIS, and SAW, the accurate pair of display Algorithms 3.1 and 3.2 are higher than or equal to other methods; Through random process and deleting attributes, it shows that Algorithms 3.1 and 3.2 are robust and better than other methods. (2) For management implications: On the one hand, in the process of decision analysis, Algorithms 3.1 and 3.2 not only analyze the relationship between objects, but also analyze the impact of all objects on each object. In the process of Algorithm 3.2 design, a new rough set model is introduced. This model does not simply use the rough set

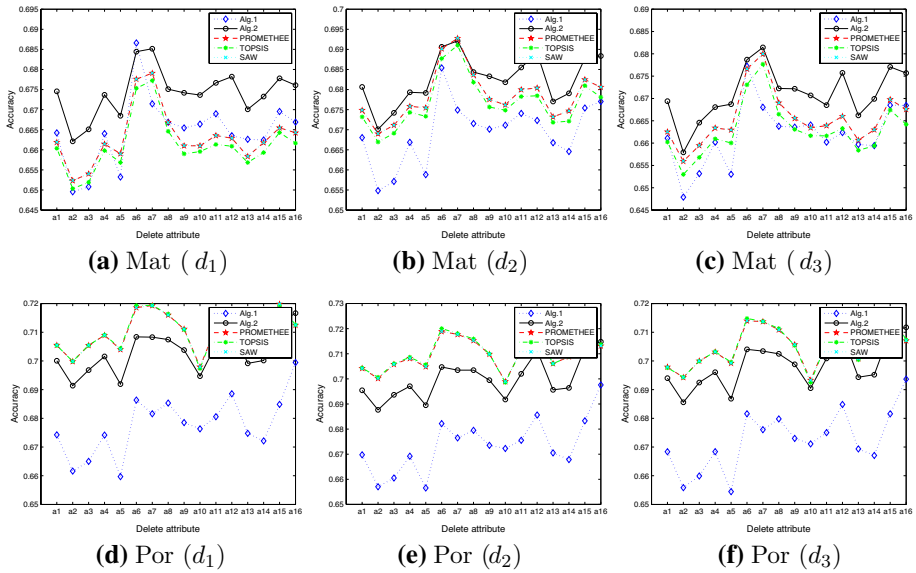


Fig. 8 Accuracy when deleting attributes based on Mat and Por

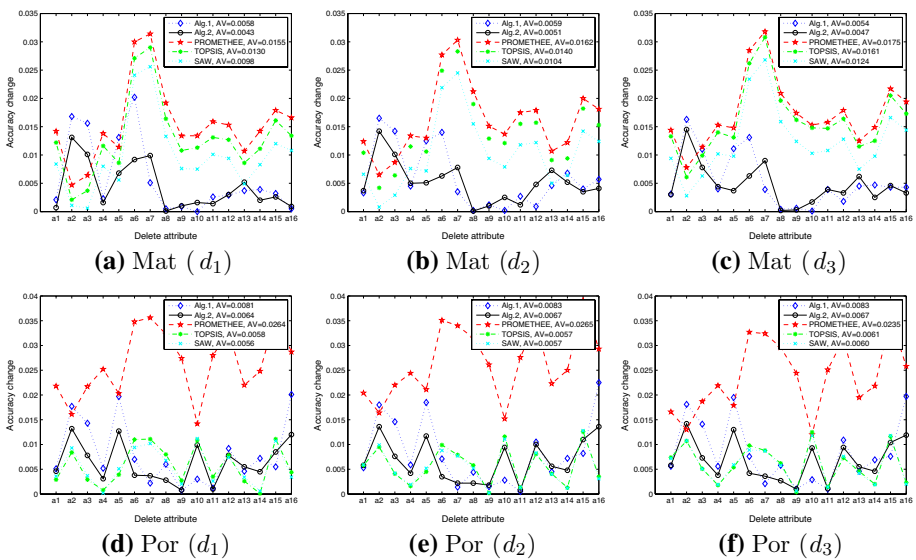


Fig. 9 Accuracy changes when deleting attributes based on Mat and Por

model to analyze the decision results, but uses the principle of rough set and combines with the purpose of decision analysis to establish a new rough model with management semantics; On the other hand, most of the existing decision-making methods rely on subjective experience or prior information, but different subjective experience or prior information leads to great differences in decision-making results. Algorithms 3.1 and

3.2 weaken the dependence on subjective experience and prior information, making the decision-making results relatively objective.

4 Conclusions

The present study is designed to design two multi-attribute decision-making methods based on less subjective experience or prior information. In this study, two new methods based on the advantage matrix are proposed to handle MADM problems. Compared with the real decision results and other methods, the accuracy and robustness of two new multi-attribute decision-making methods are analyzed. For the two new multi-attribute decision-making methods, the two methods rely on all objects, so the accuracy is high. At the same time, the change of a few attributes does not affect the decision result, that is, strong robustness. A number of limitations to our study should be noted, including that the element diversity of advantage matrix is needed. In other words, in Algorithm 3.1, the more diverse the modules of the elements of the advantage matrix, the better the decision-making effect; In Algorithm 3.2, the disadvantage and advantage neighborhood (correlation) approximations of each object are different. On the contrary, if a large number of empty or complete sets appear in the advantage matrix, Algorithms 3.1 and 3.2 fail. Our work clearly has some limitations. Despite this we believe our work could be a new perspective for decision analysis. For specific applications, the two methods are non-linear algorithms, which have a wider scope of application. Specifically, as long as an information system is established (see the previous information system description), and the attribute information is comparable, no other subjective experience or a prior information is required, and finally the decision result is obtained.

As for the directions of future research, we will investigate complementary issues as follows.

1. According to the characteristics of the advantage matrix, We will extend the advantage matrix to the intuitionistic fuzzy environment, linguistic environment, and the Hesitant fuzzy environment and generalize our methods in such environments.
2. According to the characteristics of the advantage matrix, the object is analyzed only from the perspective of the advantages of the attributes, and the disadvantages of the attributes are missing. We will establish the advantage-disadvantage matrix and define a new degree based on the advantage-disadvantage matrix for decision analysis.
3. For an order information system (Greco et al. 2002), the advantage matrix and directed graph between objects will be established, and the Graph Neural Networks (Wu et al. 2020) will be used to sort and cluster the objects.
4. According to Definition 2.9, the neighborhood operator is aimed at individual objects and is used for individual decision analysis of individual objects. We will define the neighborhood operator of the object group and use it for group decision analysis, such as comparison among multiple sales teams, asset oriented portfolio investment, etc.

Acknowledgements This work is supported by the Hunan Provincial Natural Science Foundation of China (Nos. 2019JJ50490, 2020JJ5346), and Scientific Research Foundation of Hunan Province Education Department (No. 18C0455).

Declarations

Conflict of interest The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

- Belton V (1986) A comparison of the analytic hierarchy process and a simple multi-attribute value function. *Eur J Oper Res* 26(1):7–21
- Brans J, Vincke MB (1986) How to select and how to rank projects: the promethee method. *Eur J Oper Res* 24(2):228–238
- Canciglieri O Jr (2015) Multi-attribute method for prioritization of sustainable prototyping technologies. *Clean Technol Environ Policy* 17(5):1–9
- Cortez P, Silva A (2008) Using data mining to predict secondary school student performance. *EUROSIS* 13:5–12
- Faber D, Korn H (1991) Applicability of the coefficient of variation method for analyzing synaptic plasticity. *Biophys J* 60(5):1288–1294
- Fan ZP, Zhang X, Liu Y, Zhang Y (2013) A method for stochastic multiple attribute decision making based on concepts of ideal and anti-ideal points. *Appl Math Comput* 219(24):11438–11450
- Ferretti V, Bottero M, Mondini G (2014) Decision making and cultural heritage: An application of the multi-attribute value theory for the reuse of historical buildings. *J Cult Herit* 15(6):644–655
- Figueira JR, Greco S, Roy B (2009) Electre methods with interaction between criteria: an extension of the concordance index. *Eur J Oper Res* 199(2):478–495
- Ghorabae M, Zavadskas E, Olfat L, Turskis Z (2015) Multi-criteria inventory classification using a new method of evaluation based on distance from average solution (EDAS). *Informatica* 26(3):435–451
- Greco S, Matarazzo B, Slowinski R (2002) Rough approximation by dominance relations. *Int J Intell Syst* 17(2):153–171
- Hong TP, Tseng LH, Wang SL (2002) Learning rules from incomplete training examples by rough sets. *Expert Syst Appl* 22(4):285–293
- Hong TP, Lin CE, Lin JH, Wang SL (2008) Learning cross-level certain and possible rules by rough sets. *Expert Syst Appl* 34(3):1698–1706
- <https://archive.ics.uci.edu/ml/datasets/Student+Performance>. Student performance data set. (2014)
- Huang C, Tseng T (2004) Rough set approach to case-based reasoning application. *Expert Syst Appl* 26(3):369–385
- Hwang CL, Yoon K (1981) Multiple attribute decision making. Springer, Berlin
- Jiang H, Zhan J, Chen D (2018) Covering-based variable precision (I, T) -fuzzy roughsets with applications to multiattribute decision-making. *IEEE Trans Fuzzy Syst* 27(8):1558–1572
- Liang D, Pedrycz W, Liu D, Hu P (2015) Three-way decisions based on decision-theoretic rough sets under linguistic assessment with the aid of group decision making. *Appl Soft Comput* 29:256–269
- Lin TY, Yao YY, Zadeh LA (eds) (2013) Data mining, rough sets and granular computing, vol 95. Physica
- Lingras PJ, Yao Y (1998) Data mining using extensions of the rough set model. *J Am Soc Inf Sci Technol* 49(5):415–422
- Myers J, Well A, Lorch R Jr (2013) Research design and statistical analysis. Routledge, Milton Park
- Pamučar D, Ecer F (2020) Prioritizing the weights of the evaluation criteria under fuzziness: the fuzzy full consistency method-FUCOM-F. *Facta Univ Ser Mech Eng* 18(3):419–437
- Pattaraintakorn P, Cercone N (2008) Integrating rough set theory and medical applications. *Appl Math Lett* 21(4):400–403
- Pawlak Z (1982) Rough set. *Int J Comput Inf Sci* 11(5):341–356
- Pawlak Z (1997) Rough set approach to knowledge-based decision support. *Eur J Oper Res* 99(1):48–57
- Pawlak Z, Slowinski R (1994) Decision analysis using rough sets. *Int Trans Oper Res* 1(1):107–114
- Pearman A (2014) Models and methods in multiple criteria decision making. Elsevier, Amsterdam
- Qian Y, Zhang H, Sang Y, Liang J (2014) Multigranulation decision-theoretic rough sets. *Int J Approx Reason* 55(1):225–237
- Salamó M, López-Sánchez M (2011) Rough set based approaches to feature selection for case-based reasoning classifiers. *Pattern Recogn Lett* 32(2):280–292

- Son CS, Kim YN, Kim HS, Park HS, Kim MS (2012) Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches. *J Biomed Inform* 45(5):999–1008
- Sun B, Ma W, Qian Y (2017) Multigranulation fuzzy rough set over two universes and its application to decision making. *Knowl Based Syst* 123:61–74
- Sun B, Ma W, Xiao X (2017) Three-way group decision making based on multigranulation fuzzy decision-theoretic rough set over two universes. *Int J Approx Reason* 81:87–102
- Swiniarski RW, Skowron A (2003) Rough set methods in feature selection and recognition. *Pattern Recogn Lett* 24(6):833–849
- Tian D, Zeng XJ, Keane JA (2011) Core-generating approximate minimum entropy discretization for rough set feature selection in pattern classification. *Int J Approx Reason* 52(6):863–880
- Tian G, Zhang H, Feng Y, Wang D, Peng Y, Jia H (2018) Green decoration materials selection under interior environment characteristics: a grey-correlation based hybrid mcdm method. *Renew Sustain Energy Rev* 81:682–692
- Wang H, Wang S (2009) Discovering patterns of missing data in survey databases: an application of rough sets. *Expert Syst Appl* 36(3):6256–6260
- Wang W, Zhan J, Zhang C (2021) Three-way decisions based multi-attribute decision making with probabilistic dominance relations. *Inf Sci* 559:75–96
- Wu Z, Pan S, Chen F et al (2020) A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst* 32(1):4–24
- Xiong W, Su Q, Li J (2012) The group decision-making rules based on rough sets on large scale engineering emergency. *Syst Eng Procedia* 4:331–337
- Yamaguchi D (2009) Attribute dependency functions considering data efficiency. *Int J Approx Reason* 51(1):89–98
- Ye J, Zhan J, Ding W et al (2021) A novel fuzzy rough set model with fuzzy neighborhood operators. *Inf Sci* 544:266–297
- Yu B, Cai M, Li Q (2019) A λ -rough set model and its applications with TOPSIS method to decision making. *Knowl Based Syst* 165:420–431
- Yu B, Guo L, Li Q (2019) A characterization of novel rough fuzzy sets of information systems and their application in decision making. *Expert Syst Appl* 122:253–261
- Yu B, Cai M, Dai J, Li Q (2020) A novel approach to predictive analysis using attribute-oriented rough fuzzy sets. *Expert Syst Appl* 161:113644
- Zeleny M (ed) (1976) The theory of the displaced ideal. In: *Multiple criteria decision making Kyoto 1975. Lecture Notes in Economics and Mathematical Systems (Operations Research)*, vol 123. Springer, Berlin, Heidelberg
- Zhang K, Zhan J, Wu W (2019) Fuzzy β -covering based (I, T) -fuzzy rough set models and applications to multi-attribute decision-making. *Comput Ind Eng* 128:605–621
- Zhang L, Zhan J, Xu Z (2019) Covering-based generalized IF rough sets with applications to multi-attribute decision-making. *Inf Sci* 478:275–302
- Žižović M, Pamučar D, Miljković B, Karan A (2021) Multiple-criteria evaluation model for medical professionals assigned to temporary SARS-CoV-2 hospitals. *Decis Mak Appl Manag Eng* 4(1):153–173
- Zou ZH, Yi Y, Sun JN (2006) Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *J Environ Sci* 18(5):1020–1023

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.