Information Sciences 700 (2025) 121860

Contents lists available at ScienceDirect

Information Sciences

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

Feature selection and information fusion based on preference ranking organization method in interval-valued multi-source decision-making information systems

Weihua Xu [a,b,c,*], Zhenyuan Tian [a]

[a] College of Artificial Intelligence, Southwest University, Chongqing, 400715, PR China
[b] National & Local Joint Engineering Research Center of Intelligent Transmission and Control Technology, Chongqing, 400715, PR China
[c] Chongqing Key Laboratory of Brain-inspired Computing and Intelligent Chips, Chongqing, 400715, PR China

ARTICLE INFO

Keywords:
Multi-source
Interval value
Feature selection
Information fusion

ABSTRACT

Multi-source Decision-Making Information Systems (DMSs) demonstrate superior capabilities in integrating and analyzing a diverse array of information sources, providing enhanced functionality over single-source systems. Within these systems, feature selection is crucial for identifying key attributes, which reduces information and enhance the efficiency of the decision-making process. However, current established information fusion techniques in multi-source DMSs, which integrate various sources into a unified framework, tend to be computationally intensive and are not adept at handling interval-valued data. This paper introduces an innovative feature selection model specifically developed for multi-source DMSs, employing the Preference Ranking Organization Method for Enrichment Evaluations (PROMETHEE). The model initiates by establishing the neighbourhood relationships among objects across different attributes. It then utilizes the PROMETHEE algorithm to rank these attributes based on their comparative strengths and weaknesses, facilitating the pinpointing of the most valuable features. The model further refines the selection process by quantifying the consensus level, thereby discovering the most reliable information sources. Our some experiments, performed utilizing a broad and comprehensive dataset, have validated both the model and its underlying algorithm. The results obtained provide compelling evidence of the model's effectiveness, especially highlighting its proficiency in handling interval-valued data. Furthermore, the outcomes illustrate the model's significance to the enhancement of decision-making processes within multi-source Decision-Making Information Systems (DMSs).

1. Introduction

A Decision Information System (DIS) stands as an indispensable technological aid for organizations and managers, engineered to facilitate effective decision-making in the face of today's complex and uncertain environments [25]. The challenges of the information age, such as burgeoning data volumes, swiftly changing market conditions, and the intricacies of business structures, have made the role of DIS more critical than ever. By integrating a variety of data sources, providing comprehensive modelling and analytical tools, and presenting an intuitive user interface, DIS enhances the clarity and manageability of decision-making processes. The utility of DIS

* Corresponding author.
E-mail addresses: chxuwh@swu.edu.cn (W. Xu), clearlove0530@email.swu.edu.cn (Z. Tian).

https://doi.org/10.1016/j.ins.2024.121860
Received 26 August 2024; Received in revised form 19 December 2024; Accepted 29 December 2024

Available online 7 January 2025

is not only in its technical prowess but also in its practical applications that extend into various domains. It is particularly effective in management information systems [21], streamlining operations and strategic planning. In the health care coverage sector [12], DIS aids in making informed policy decisions and optimizing resource allocation. Additionally, it addresses geographic issues [1] by supporting detailed spatial analysis and decision-making. These applications highlight the system's ability to articulate decision options clearly and manage them efficiently, which is invaluable for systematic analysis and optimization across different conditions.

In the realms of scientific research and engineering practice, multi-source analysis has emerged as a pivotal approach for examining complex systems and problems [13]. Distinguished from traditional single-source analysis methods, multi-source analysis possesses the capability to concurrently account for the system's attributes across various contexts. This holistic consideration enables a more comprehensive and precise comprehension of the system, allowing for a more revealing portrayal of its behaviours and characteristics. The adaptability of multi-source analysis lies in its dynamic scalability, aligning the scope of investigation with the specific demands and nuances of the problem at hand. This tailored approach enhances the method's flexibility and adaptability, equipping it to cater more effectively to the diverse needs of research and applications.

Multi-source decision-making information systems (DMSs) have become indispensable in the complex landscape of modern decision-making by seamlessly integrating a wealth of data from various internal and external sources. This multi-source integration provides a more comprehensive, diverse, and reliable information base that is crucial for decision-makers to enhance the accuracy and reliability of their decisions. With such systems, decision-makers gain a more holistic understanding of problems, access to more precise data analytics, comprehensive risk assessments, and a flexible array of decision-making options, thereby significantly improving the quality and effectiveness of decision-making processes [9]. The challenge of synthesizing vast amounts of data into coherent and useful information has driven significant interest in information fusion within multi-source systems [39]. Early contributions in this field, such as those by Cai et al. [4], introduced Bayesian networks to enhance information fusion, while Zhang et al. [44] applied rough set theory to achieve similar objectives. In the information age, feature selection [16] has emerged as a critical technique for approximating data by filtering out the unimportant and retaining the essential, a method that has since been integrated into multi-source information systems [17]. The field of feature selection is rich with diverse methodologies, including the use of causal features [11], statistical methods [15], and neural networks [32], all aimed at refining the analytical capabilities of these systems.

In scenarios involving multiple data sources, the potential for divergent observations or information across these sources can introduce data inconsistency and uncertainty [41][35]. To mitigate these challenges within decision-making information systems, the incorporation of interval values proves to be an effective strategy. In some cases, the data may only have an interval range rather than a specific value, e.g., in the process of drug treatment, the plasma concentration of a drug is an important indicator for judging the effectiveness and safety of the treatment. However, due to individual differences (e.g., age, weight, metabolic rate, etc.), the ideal range of drug concentration is usually an interval rather than a fixed value; also, for example, in practical applications, the accuracy of many measuring devices is limited, and thus they cannot provide a precise value, assuming that you are using a thermometer to measure the temperature of a certain environment, the instrument may only be accurate to $\pm$ 0.5 °C, rather than providing a specific precise value. Interval values offer a robust mechanism for capturing the inherent uncertainty and ambiguity of data, allowing decision-makers to contemplate a broader spectrum of possibilities and to tailor their decisions accordingly [2][26]. This approach is particularly crucial in practical decision-making, where the ability to express and manage uncertainty is essential for informed and strategic choices.

With the Evolution of Multi-Source Decision-Making information Systems, there has been a burgeoning interest in their theoretical foundations. This paper aims to extend the PROMETHEE method within the context of multi-source decision-making information systems, introducing an innovative framework for feature selection and information fusion. Our approach is designed to discern and prioritize the source exhibiting optimal classification efficacy amidst a constellation of interval-valued data sources.

This study is propelled by several key motivations:

1. **Innovation in Feature Selection:** A plethora of effective methods for multi-source feature selection has been documented [22][40]. Traditional approaches to multi-source feature selection are predicated on source fusion with selection occurring post-fusion within a consolidated data table. Herein, we integrate the PROMETHEE algorithm [3] directly into the feature selection process of multi-source decision-making information systems, devising a novel method that leverages dominance relationships for enhanced selection efficacy.
2. **Relevance to Modern Decision-Making:** Multi-source decision-making information systems are integral to the fabric of modern decision-making science and are omnipresent in various facets of real-life scenarios, underscoring the need for robust theoretical and practical frameworks to support complex decision-making processes.
3. **Addressing Real-World Data Characteristics:** Real-life data predominantly manifest as interval values, which inherently introduce elements of inconsistency and uncertainty [7]. Recognizing the significance of interval values in decision-making, this study advocates for their incorporation into the feature selection process to ensure a more accurate reflection of real-world complexities.

Building upon the preceding discussion, our methodology for interval-valued multi-source decision-making information systems introduces several innovative aspects:

1. **Direct Feature Selection:** Unlike the majority of existing studies that rely on information fusion for feature selection in multi-source decision-making information systems-a process that often struggles to directly identify the most relevant features, leading to increased computational costs and potential noise in the data-we introduce a novel approach that bypasses this indirect

Table 1
The description of mathematical symbols.

mathematical symbols	meaning
$U = \{x_1, x_2, ..., x_m\}$	the set of objects
$C = \{C_1, C_2, ..., C_n\}$	the set of attributes
$d = \{d_1, d_2, ..., d_t\}$	the set of decision attributes
$D_r \in d/R_t$	the class of decision
$dis(a, b)$	Distance between object a and object b
$(x_a^k)_j$	Neighbourhood class of object x_a with respect to the j-th attribute in the k-th source
$POS_j^k(D_r)$	the lower approximation of the j-th attribute in the decision class, D_r at the k-th source.
$NEG_j^k(D_r)$	the upper approximation of the j-th attribute in the decision class, D_r at the k-th source.
$CON_j^k(D_r)$	the confidence level of the j-th attribute in the decision class, D_r at the k-th source.
$F_r^k(a, b)$	Preference function between object a and object b of the r-th decision class at the k-th source.
$r(a, b)$	Difference in confidence level between object a and object b of the r-th decision class
lap_r^k	Difference between maximum and minimum values in confidence level of the r-th decision class at the k-th source.
$pd(a, b)$	the preference difference between object a and object b
$\varphi^+(a)$	the leaving flow of object a
$\varphi^-(a)$	the entering flow of object a
$\varphi(a)$	the net flow of object a
DC_k	the degree of consensus
C_{aj}^k	the intermediate values of interval data
C_{aj}^*	the source leader

selection process, selecting the best source directly from the available sources. We introduce a method capable of performing feature selection in a direct and cost-effective manner.
2. **Integration of Interval Values:** To address the limitations in Zhan et al.'s study of uncertain multi-source information systems, where interval values are often overlooked, we propose integrating interval values into the PROMETHEE algorithm. This novel approach tackles the categorization challenges specific to interval-valued systems while preserving and enhancing the strengths of the PROMETHEE method.
3. **Enhanced Object and Feature Management:** While Zhan et al.'s method [5] sequentially ranks objects based on a single criterion, our approach provides a more robust solution by simultaneously handling both feature ranking and optimal source selection. This dual functionality not only streamlines the decision-making process but also enhances the accuracy and efficiency of feature selection in complex, multi-source systems. By integrating these two critical steps, our method significantly expands its applicability to real-world scenarios.

This paper is organized as follows. In the next sections, we will review some of the most representative literature as well as recent developments, introduce some basic concepts of multi-source decision information systems. In Section 3 we provide a preliminary introduction to neighbourhood rough set and introduce the core ideas of the PROMETHEE method to interval-valued multi-source decision-making information systems and propose a new method for feature selection and information fusion in interval-valued multi-source decision-making information systems. Experimental results are given in Section 4. The conclusions of the experiments and future work are given in Section 5.

2. Literature review

This section presents a review of multi-source methods and related extended applications, as well as studies on feature selection and information fusion. It then revisits several basic concepts about multi-source decision-making information systems, with the aim of facilitating subsequent discussions. Table 1 provides a summary of the key mathematical symbols utilised in this study.

2.1. Multi-source information systems

In examining theories pertaining to multi-source information systems, numerous effective classification techniques have been enhanced and introduced to the multi-source problem. For instance, Srinivasan et al. [27] employed a knowledge-based approach to assess the viability of a novel scheme for integrating multiple sources into the classification process, proposing two novel classification methodologies. Tso et al. [30] put forward classification of multi-source remote sensing imagery using a genetic algorithm and Markov random fields. Watanachaturaporn et al. [33] used Support Vector Machines (SVM) as an alternative to classify remote sensing data into a multi-source information system, proposed a new classification method for multi-source data. Xu et al. [37] performed multi-source remote sensing data classification based on convolutional neural networks, which were then combined with other data features extracted from the cascade network.

All multi-source classification methods have their own characteristics and can approach the classification problem from different perspectives, such as Analytic Hierarchy Process (AHP) [31] method and PROMETHEE method. In specific, the AHP method is a multi-criteria decision-making method for dealing with complex decision-making problems. It does this by decomposing the problem into a hierarchy, determining the importance or priority between the levels in the hierarchy, then using pairwise comparisons to determine the weights of each level, and finally evaluating and ranking the options. The promotion and extension of the AHP method has also attracted the attention of many scholars. Alonso et al. introduced consistency into AHP and give the system the flexibility to adjust acceptance requirements to different scope and conformance requirements. Saaty et al. [24] proposed the negative priorities in the analytic hierarchy process, which can be used to deal with combinatorial priorities in opposite directions. In addition to the AHP, the PROMETHEE method is also quite effective in the classification of multi-source. The PROMETHEE method can sort all the data advantageously according to a given preference function and calculate the preference index to get the desired sorting result. Moreover, the promotion and optimization with the PROMETHEE method are also rich. Liao et al. [18] extended PROMETHEE to the intuitionistic fuzzy environment by augmenting PROMETHEE with intuitionistic fuzzy sets, considering intuitionistic fuzzy preferences and intuitionistic fuzzy weights. Macharis et al. [19] endeavoured to refine the PROMETHEE method by synthesizing the merits and drawbacks of both PROMETHEE and AHP methodologies. The method proposed by Hyde et al. [14] is based on reliability and incorporates a generalised criterion function to account for uncertainty in the standard performance values.

Nevertheless, the principal objective of the aforementioned approach is the categorisation of multi-source information systems. The quantity of data can result in an increase in the cost of classification and a reduction in the effectiveness of classification. Consequently, the objective is to investigate the potential of information fusion in the context of multi-source information systems. In the following section, we will review the current state of knowledge regarding feature selection and information fusion.

2.2. Feature selection and information fusion

The objective of feature selection and information fusion is to identify and retain the most pertinent features while eliminating those that are less significant, thereby reducing costs and enhancing accuracy. The following section will present the relevant knowledge pertaining to feature selection and information fusion.

Feature selection is a method that reduces the dimensionality of data, thereby improving the efficiency and performance of algorithms by selecting the most relevant and informative features. Zeng et al. [42] considered feature interaction in feature selection and proposed a feature selection method based on interaction weight factors. Farahat et al. [6] proposed an efficient greedy method for unsupervised feature selection. Zhang et al. used tabu search method for feature selection to select the optimal subset from the original large feature set.

Information fusion [8], which entails the integration of information from disparate sources or types of information, can enhance the predictive accuracy and resilience of a model. The integration of predictions from disparate models or feature sets can result in a reduction in prediction error and an enhancement of model stability. Information fusion is used in many ways [38], such as biometric identification technology [23], wireless sensor networks [20] and automotive sensors [28]. In recent years, information fusion has received a great deal of scholarly attention. Zhang et al. [43] combined information fusion with Bayesian networks to solve dynamic problems. Sun et al. [29] gave a generalised multi-sensor optimal information fusion decentralised Kalman filter with a two-layer fusion structure for multi-sensor optimal information fusion.

A considerable number of studies have sought to apply feature selection and information fusion to multi-source information systems [8]. The majority of current methods [10][36] compare each feature in each source according to a specific relationship, select the superior source features, and aggregate all the selection results into a single source, which retains the same number of features as the original data set. Subsequent feature selection is then performed on this individual source, removing the less effective attributes, and the remaining features constitute the final simplified result. The fundamental principle underlying this methodology is the fusion of features prior to selection. This involves the combination of all features present in each source, followed by a comparison to identify the source that contains the optimal attributes. This process is inherently time-consuming. In the following section, we will present a novel approach to feature selection and information fusion. This approach eliminates the need for feature combination, thus reducing the time required for the selection process.

3. PROMETHEE-driven feature selection and information fusion

In this section, we revisit several basic concepts about multi-source decision-making information systems with interval-valued data, neighbourhood rough sets and the PROMETHEE algorithm, then describe the process of transforming the data within the original interval-valued multi-source decision information system into data that can be attribute-sorted by the PROMETHEE algorithm. Subsequently, a preference function is constructed, enabling the exact confidence of each attribute to be determined under different decisions. Subsequently, the preference matrix is employed to derive the net flow and thus the attribute ranking. Subsequently, the consensus degree of each source is employed to calculate the optimal source.

3.1. Neighbourhood rough sets

Let U be a finite and nonempty set, contains a finite number of objects, each of which has a corresponding score under all attributes. Several objects are grouped into the same decision class.

Neighbourhood rough set is a branch of rough set theory. Compared with the traditional rough set theory, neighbourhood rough set considers the neighbourhood relationship between objects, which is more in line with the characteristics of practical problems.

Let $A = \{U, \mathbb{C} \cup \{d\}\} = \{U, \{C_j | j = 1, 2, ...n\} \cup \{d\}\}$ be a multi-source decision-making information system, where $U = \{x_1, x_2, ..., x_m\}$ is a universe, $d = \{d_1, d_2, ..., d_t\}$ is a special attribute called the decision attribute and $C = \{C_1, C_2, ..., C_n\}$ is the set of attributes.

Table 2
A multi-source decision information system with interval-values.

A	C_1^1	...	C_1^i	C_2^1	...	C_2^i	...	C_n^1	...	C_n^i	d
x_1	$[C_{11}^{1-}, C_{11}^{1+}]$	...	$[C_{11}^{i-}, C_{11}^{i+}]$	$[C_{12}^{1-}, C_{12}^{1+}]$	...	$[C_{12}^{i-}, C_{12}^{i+}]$	...	$[C_{1n}^{1-}, C_{1n}^{1+}]$	...	$[C_{1n}^{i-}, C_{1n}^{i+}]$	d_1
x_2	$[C_{21}^{1-}, C_{21}^{1+}]$	...	$[C_{21}^{i-}, C_{21}^{i+}]$	$[C_{22}^{1-}, C_{22}^{1+}]$	...	$[C_{22}^{i-}, C_{22}^{i+}]$	...	$[C_{2n}^{1-}, C_{2n}^{1+}]$	...	$[C_{2n}^{i-}, C_{2n}^{i+}]$	d_2
...	...	...	...	...	...	...	...	...	...	...	...
x_m	$[C_{m1}^{1-}, C_{m2}^{1+}]$	...	$[C_{m1}^{i-}, C_{m1}^{i+}]$	$[C_{m2}^{1-}, C_{m2}^{1+}]$	...	$[C_{m2}^{i-}, C_{m2}^{i+}]$	...	$[C_{mn}^{1-}, C_{mn}^{1+}]$	...	$[C_{mn}^{i-}, C_{mn}^{i+}]$	d_m

Let $X \subseteq U$, the neighbourhood relationship between x_m and other x_n can be defined as satisfying $dis(x_m, x_n) < \theta$. In this equation, $dis(x_m, x_n)$ is an arbitrary form of the distance formula, and different choices can be made depending on the situation. This equation represents the selection of two different objects, calculate some distance between the two objects, if this distance is less than a given value θ, the two objects are considered to be in each other's neighbourhood, it can be assumed that there is a neighbourhood relationship between these two objects. Typically, $dis(x_m, x_n)$ can be chosen as either the Euclidean distance or the Manhattan distance.

The object x_m can be calculated as a distance using the distance formula dis for every other object. Putting all objects that satisfy the condition $dis(x_a, x_b) < \theta$ for x_a into the neighbourhood set (x_a), the neighbourhood set (x_m) can be defined as

$$(x_a) = \bigcup_{dis(a,b)<\theta, b\in U} \{x_b\}. \tag{1}$$

It was mentioned above that several objects x in a multi-source decision-making information system all belong to the same decision class $d_r \in \{d_1, d_2, ..., d_t\}$. We use D to denote the set of all x, belonging to the same decision class. D can be expressed by the following equation

$$D_r = \bigcup_{d_a = d_r} \{x_a\}, \tag{2}$$

where d_a denotes the decision attribute of object x_a, $D_r \in d/R_t$ and R_t is the number of classes after categorizing all decisions, d_r denotes a decision.

Then we can lead to the definitions of lower and upper approximation. Let x in U, the lower and upper approximation are defined as

$$POS(D_r) = \bigcup_{x_a \in U} \{x_a | (x_a) \subseteq D_r\}, \tag{3}$$

$$NEG(D_r) = \bigcup_{x_a \in U} \{x_a | (x_a) \cap D_r \neq \emptyset\}. \tag{4}$$

3.2. Multi-source decision-making information system with interval-values

The Multi-source decision-making information system with interval-values is a decision-making model that integrates multiple sources of information. The system is designed to address the uncertainty and ambiguity inherent in the decision-making process. The system represents uncertainty by utilising interval values and combines data and information from multiple sources in order to enhance the accuracy and reliability of decisions.

The main concept of the multi-source approach is to use data from various channels or sources. Let $A = \{U, \mathbb{C} \cup \{d\}\}$ be a multi-source decision information system, where $U = \{x_1, x_2, ..., x_m\}$ is a universe, $d = \{d_1, d_2, ..., d_t\}$ is a special attribute called the decision attribute and $\mathbb{C} = \{\mathbb{C}_1, \mathbb{C}_2, ...\mathbb{C}_n\}$ is a finite set of standards which is called attribute. For each $x \in U$, if $\mathbb{C}_j$ have i sources, each $\mathbb{C}_j = \{C_j^1, C_j^2, ..., C_j^i\} \subseteq \mathbb{C}$, then A can be written as $A = \{U, \mathbb{C} \cup \{d\}\} = \{U, \{C_j^k | k = 1, 2, ..., i; j = 1, 2, ...n\} \cup \{d\}\}$.

A multi-source decision-making information system is a pair $A = \{U, \mathbb{C} \cup \{d\}\} = \{U, \{C_j^k | k = 1, 2, ..., i; j = 1, 2, ...n\} \cup \{d\}\}$. If each element in A is an independent interval value, it can be expressed as$[C_j^{k-}, C_j^{k+}]$. For $\{x_a | a = 1, 2, ..., m\}$, the data in the multi-source decision information system with interval-values table is represented as $[C_{aj}^{k-}, C_{aj}^{k+}]$

Example 1. Let $A = (U, \mathbb{C} \cup \{d\}) = (U, \{C_{aj}^{k-}, C_{aj}^{k+} | k = 1, 2, ..., i; j = 1, 2, ...n; a = 1, 2, ..., m\} \cup \{d\})$ be a multi-source decision information system with interval-values.

In Table 2, m denotes the number of objects in the universe U, n is the corresponding attribute in the multi-source decision information system and i denotes the corresponding source. Next we build a preference function on this information system. Relative advantage relationships derived from preference differences between objects.

3.3. The PROMETHEE algorithm

The PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations) algorithm is a decision-making tool designed for ranking multiple attributes across candidate items. It assesses these candidates based on a set of decision criteria to identify the optimal choice. In PROMETHEE, a decision-maker defines a preference function for each criterion, which quantifies the relative preference of each candidate. These functions allow for calculating a preference index that highlights the strengths and weaknesses of each candidate and establishes a comparative ranking.

A preference function in PROMETHEE expresses the degree of preference or relative importance of each option across various attributes, such as cost, benefit, or risk. Higher values from the preference function indicate a stronger preference for a particular option. Each criterion has an associated preference function, which enables a detailed evaluation of candidates by comparing their respective advantages and disadvantages. This systematic approach improves decision-making by providing a clear, quantitative basis for ranking alternatives.

For any two objects in U, the generalized preference function can be expressed by the following equation

$$F : U \times U \longrightarrow [0,1]. \tag{5}$$

The preference function can be used to evaluate the difference in preferences between any two objects a and b under the same attribute. The difference in preferences can be used to reflect the relationship between objects a and b, which can be categorized as dominance and inferiority. The difference in preferences under the j-th marking scheme can be expressed as γ_j.

• The dominance of object a over b under the j-th attribute can be expressed as

$$\gamma_j(a,b) = F(C_j(a) - C_j(b)). \tag{6}$$

• The inferiority of object a over b under the j-th attribute can be expressed as

$$\gamma_j(b,a) = F(C_j(b) - C_j(a)), \tag{7}$$

where γ_j denotes the relationship between the advantages and disadvantages of object a and object b on a scale. If there are many scoring criteria, it is generally necessary to weight each merit relationship to combine all scoring criteria. Depending on the preference function, the merit relationships between the same objects are generally different. We then need to summarise the advantages and disadvantages of the relationship between the two objects under each attribute separately, which leads to the concepts of the leaving flow and the entering flow.

• The leaving flow represents the sum of an object's advantages over other all objects under all sources and all attributes.

$$\varphi^+(a) = \sum_{b\in U} \gamma_j(a,b). \tag{8}$$

• The entering flow represents the sum of the other all object's advantages over this object under all sources and all attributes, and can also be interpreted as the sum of the object's disadvantages relative to the other objects.

$$\varphi^-(a) = \sum_{b\in U} \gamma_j(b,a). \tag{9}$$

• The net flow is defined as the global superiority of an object

$$\varphi(a) = \varphi^+(a) - \varphi^-(a). \tag{10}$$

Net flow is expressed as leaving flow minus entering flow. The leaving flow represents the advantage of the current object over other objects, and the entering flow represents the advantage of other objects over the current object. Net flows are meant to reflect the sum of an object's advantages over other objects in the evaluation criteria. According to the PROMETHEE algorithm, the greater the net flow, the greater the importance of the object in the overall information system, which will be mentioned in detail in the subsequent ranking of advantages. The steps of PROMETHEE algorithm can be divided into

(1) Designing a suitable preference function.
(2) Calculate preference relationships between all elements.
(3) Derive the preference matrix through the preference relationships between the elements.
(4) Calculation of leaving flow and entering flow.
(5) Calculation of net flow.
(6) Derive the final sorting result by the size of the net flows.

3.4. Feature selection in multiple sources

Definition 1. Based on the preference function, with the value of each interval, we can judge the superiority relationship between any two objects a and b under the same attribute. But if we want to evaluate the preference difference between any two attributes C_1 and C_2, we can't just utilize the values in the decision information table alone because the intervals in the decision information table are scores for object x under a certain attribute, they can be used for evaluating object x_a but cannot be used to evaluate attribute C in turn.

Table 3
Table for feature selection.

	d_1^1	...	d_1^i	...	d_t^1	...	d_t^i
C_1	$CON_1^1(D_1)$	...	$CON_1^i(D_1)$	...	$CON_1^1(D_t)$	...	$CON_1^i(D_t)$
...	...	...	...	...	...	...	...
C_n	$CON_n^1(D_1)$	...	$CON_n^i(D_1)$	...	$CON_n^1(D_t)$	...	$CON_n^i(D_t)$

In order to identify the most appropriate attributes through preference differences, it is necessary to perform some processing of the interval values in the table. The objective of the processing is to ensure that the magnitude of the processed numbers accurately reflects the attribute's preference, thereby enabling the feature to be selected.

We use the confidence level of an attribute under different decisions to evaluate the merit of the attribute. First, we have to compute the neighbourhood class of each object under a single attribute. For interval-valued data, we can use the Jaccard distance to get the equivalence relation, which is given by the following formula

$$dis(a,b) = 1 - \frac{|A \cap B|}{|A \cup B|}, \tag{11}$$

where A and B denote the interval values of the corresponding objects a and b under a certain attribute.

With the Jaccard distance, we can easily derive the neighbourhood class for a given attribute with respect to the object x_a, which are defined as

$$(x_a^k)_j = \bigcup_{dis(a,b)<\theta, b\in U} \{x_b\}, \tag{12}$$

where a denotes the a-th object, k denotes the k-th source, and j denotes the j-th attribute and θ is a threshold value given according to the actual situation. Given a decision class $D_r \in d/R_t$ contains all objects x with the same decision type, R_t is the number of classes after categorizing all decisions, then the scaled lower approximation of x_a is defined as

$$POS_j^k(D_r) = \bigcup_{x_a\in U} \{x_a | (x_a^k)_j \subseteq D_r\}. \tag{13}$$

The upper approximation of x_a is defined as

$$NEG_j^k(D_r) = \bigcup_{x_a\in U} \{x_a | (x_a^k)_j \cap D_r \neq \emptyset\}. \tag{14}$$

So the confidence level can be defined as

$$CON_j^k(D_r) = \frac{\left|POS_j^k(D_r)\right|}{\left|NEG_j^k(D_r)\right|}. \tag{15}$$

In Equation (15), we have derived confidence levels for different decision classes under a single attribute by using the neighbourhood class of all objects under a single attribute. Next, we can view the decision classes as attribute conditions and the original attributes as objects. The purpose of doing so is to change the original attribute-to-object scoring to decision-to-attribute scoring, and to view the confidence level as the result of the decision-to-attribute scoring, and to utilize the confidence level to judge the attribute's dominance and inferiority. As shown in Table 3.

Definition 2. After converting the raw interval-valued data into data that can be judged on the merits of the attributes, the next step is to derive the preference differences using the preference function and transforming them into the form of a preference matrix. We use a linear preference function as follows

$$F_r^k(C_a, C_b) = \begin{cases} 0, & r(C_a, C_b) < 0, \\ \frac{r(C_a, C_b)}{lap_r^k}, & 0 < r(C_a, C_b) < lap_r^k, r \in \{1,2,...,t\}, k \in \{1,2,...,i\}, \\ 1, & r(C_a, C_b) > lap_r^k, \end{cases} \tag{16}$$

where

$$r(C_a, C_b) = CON_a^k(D_r) - CON_b^k(D_r), a, b \in \{1,2,...,n\}, k \in \{1,2,...,i\}, r \in \{1,2,...,t\}, \tag{17}$$

and the scale-dependent parameters are calculated as follows

$$lap_r^k = \max_{D_r\in U} CON_{j1}^k(D_r) - \min_{D_r\in U} CON_{j2}^k(D_r), j1, j2 \in \{1,2,...,n\}, k \in \{1,2,...,i\}. \tag{18}$$

Table 4
A preference matrix.

	C_1	C_2	...	C_n	φ^+
C_1	0	$pd(C_1, C_2)$	...	$pd(C_1, C_n)$	$\varphi^+(C_1)$
C_2	$pd(C_2, C_1)$	0	...	$pd(C_2, C_n)$	$\varphi^+(C_2)$
...	...	...	...	...	...
C_n	$pd(C_n, C_1)$	$pd(C_n, C_2)$	...	0	$\varphi^+(C_n)$
φ^-	$\varphi^-(C_1)$	$\varphi^-(C_2)$	...	$\varphi^-(C_n)$	

Then the preference difference between a and b can be expressed as

$$pd(a, b) = \sum_{r=1}^{t} \sum_{k=1}^{i} F_r^k(a, b) \cdot \omega_r^k, \tag{19}$$

where ω_r^k represents the weight of each decision class, which is generally taken to be equal. The meaning of preference difference is to calculate the advantages and disadvantages of object a and object b sequentially under all sources and all decision class, to derive the relative advantages and disadvantages of the two objects under this one scoring attribute, and then weight and sum them up to derive the overall advantage of object a over object b.

Having obtained the object-to-object preference differences, we can then derive a preference matrix about all objects. Assuming that there is an information system containing n objects, each object can compute n preference differences including itself, and its own preference difference is 0. Then we can get a $n \times n$ matrix with diagonal elements equal to 0.

Definition 3. In the PROMETHEE method, the leaving flow and entering flow are defined as follows:
• The leaving flow represents the sum of an object's advantages over other all objects under all sources and all attributes,

$$\varphi^+(a) = \sum_{b \in U} pd(a, b). \tag{20}$$

• The entering flow represents the sum of the other all object's advantages over this object under all sources and all attributes, and can also be interpreted as the sum of the object's disadvantages relative to the other objects,

$$\varphi^-(a) = \sum_{b \in U} pd(b, a). \tag{21}$$

The net flow is defined as the global superiority of an object

$$\varphi(a) = \varphi^+(a) - \varphi^-(a). \tag{22}$$

There is a preference matrix in Table 4. It is obvious that the larger the leaving flow, the greater the sum of an object's advantages over all other objects. For the same reason, the smaller the entering flow, the less the sum of an object's disadvantages relative to all other objects. Responding to the relationship between the advantages and disadvantages of an object through its net flow, we can say that the larger the net flow, the better the object. Fig. 1 and Algorithm 1 show the specific flow of the algorithm. The time complexity of Algorithm 1 is $O(k \cdot n \cdot t + m^2 + n^2)$ and the space complexity is $O(k + m^2 + n^2)$.

Example 2. A brief arithmetic example is presented to illustrate the process of feature selection in a multi-source decision information system with interval-valued inputs.

In the context of the modern lifestyle, gout is increasingly being recognised as a metabolic disease. Gout is a disease that results from disturbances in uric acid metabolism, which significantly impacts the quality of life and health of patients. A company implements regular medical check-ups for its employees. The initial physical examination identifies three indicators of blood uric acid, uric acid and C-reactive protein (CRP), and determines whether the employee has gout based on these three physical examination results. Data from three medical examiners can be considered as two sources. In order to reduce the medical examination time, the original three indicators are now approximated into two. A single ranking of the original indicators is sufficient; the least advantageous test should be removed.

Data from all medical examinations are shown in Table 5.

Table 5 can be referred to $A = \{U, \mathbb{C} \cup \{d\}\} = \{U, \{C_j^k | k = 1, 2, 3; j = 1, 2, 3\} \cup \{d = 1, 2\}\}$. In this example, all the data has been initialized. Then we compute the neighbourhood relationship. Utilizing the Jaccard distance function, take the parameter $\theta = 0.5$.

We then use the Jaccard distance function to compute the neighbourhood relationship for each sample interval value, and then Equation (15) to derive the degree of certainty, which is used to subsequently carry over to the linear preference function to compute the preference differences.

The preference differences for each attribute under different decisions are shown in Table 6.

After obtaining the preference differences for each attribute, the leaving flow and the entering flow are calculated by Equation (20) and Equation (21). The preference matrix is shown in Table 7.

Algorithm 1: Feature selection method based on the PROMETHEE algorithm in multi-source decision information system with interval-values.

Input: A multi-source decision-making information system with interval-values $A=\{U,\mathbb{C}\cup\{d\}\}$.;
Output: Attributes ordering;
$CM \longleftarrow \{\}$.
for $i=1$ *to* k **do**
  $CM_i \longleftarrow \{\}$.
  for $j=1$ *to* n **do**
    for $r=1$ *to* t **do**
      calculate the confidence level $PM_i \longleftarrow CON_j^k(D_r)=\frac{\left|POS_j^k(D_r)\right|}{\left|NEG_j^k(D_r)\right|}$.
    end
  end
  $CM \longleftarrow CM_i$.
end
the preference matrix: $PM \longleftarrow \{\}$.
for a *in* CM **do**
  Calculation of preference difference through preference functions: $F_r^k(C_a,C_b)$.
  Weighting $F_r^k(C_a,C_b)$ to obtain $pd(a,b)$.
  $PM \longleftarrow pd(a,b)$.
end
$\varphi=0, \varphi^+ \longleftarrow \{\}, \varphi^- \longleftarrow \{\}$.
for $x=1$ *to* n **do**
  $pd_x \longleftarrow \{\}$.
  for $y=1$ *to* n **do**
    $pd_x += pd(C_x,C_y)$.
  end
  $\varphi^+ \longleftarrow pd_x$.
end
for $y=1$ *to* n **do**
  $pd_y \longleftarrow \{\}$.
  for $x=1$ *to* n **do**
    $pd_y += pd(C_x,C_y)$.
  end
  $\varphi^- \longleftarrow pd_y$.
end
$\varphi=\varphi^+-\varphi^-$.
Sorting is done according to φ.
return: Attributes ordering.

Table 5
A multi-source decision information system with interval-values.

A	C_1			C_2			d
	C_1^1	C_1^2	C_1^3	C_2^1	C_2^2	C_2^3	
x_1	[0.256, 0.312]	[0.094, 0.114]	[0.302, 0.37]	[0.542, 0.662]	[0.401, 0.491]	[0.731, 0.893]	1
x_2	[0.248, 0.304]	[0.253, 0.309]	[0.112, 0.136]	[0.405, 0.495]	[0.431, 0.527]	[0.424, 0.518]	1
x_3	[0.166, 0.202]	[0.089, 0.109]	[0.389, 0.475]	[0.711, 0.869]	[0.659, 0.805]	[0.716, 0.875]	1
x_4	[0.274, 0.336]	[0.154, 0.188]	[0.233, 0.285]	[0.358, 0.438]	[0.440, 0.538]	[0.689, 0.842]	2
x_5	[0.259, 0.317]	[0.150, 0.184]	[0.069, 0.085]	[0.515, 0.629]	[0.506, 0.618]	[0.276, 0.338]	2

A	C_3			d
	C_3^1	C_3^2	C_3^3	
x_1	[0.747, 0.913]	[0.297, 0.363]	[0.417, 0.509]	1
x_2	[0.819, 1.001]	[0.675, 0.825]	[0.589, 0.719]	1
x_3	[0.603, 0.737]	[0.306, 0.374]	[0.48, 0.586]	1
x_4	[0.531, 0.649]	[0.774, 0.946]	[0.625, 0.764]	2
x_5	[0.711, 0.869]	[0.441, 0.539]	[0.517, 0.631]	2

Table 6
Table for feature selection.

	d_1^1	d_1^2	d_1^3	d_2^1	d_2^2	d_2^3
C_1	0.25	0.25	1	1	1	1
C_2	0.5	0.333	0.5	0.333	0.25	0.25
C_3	0.5	0.333	1	1	0.5	0.333

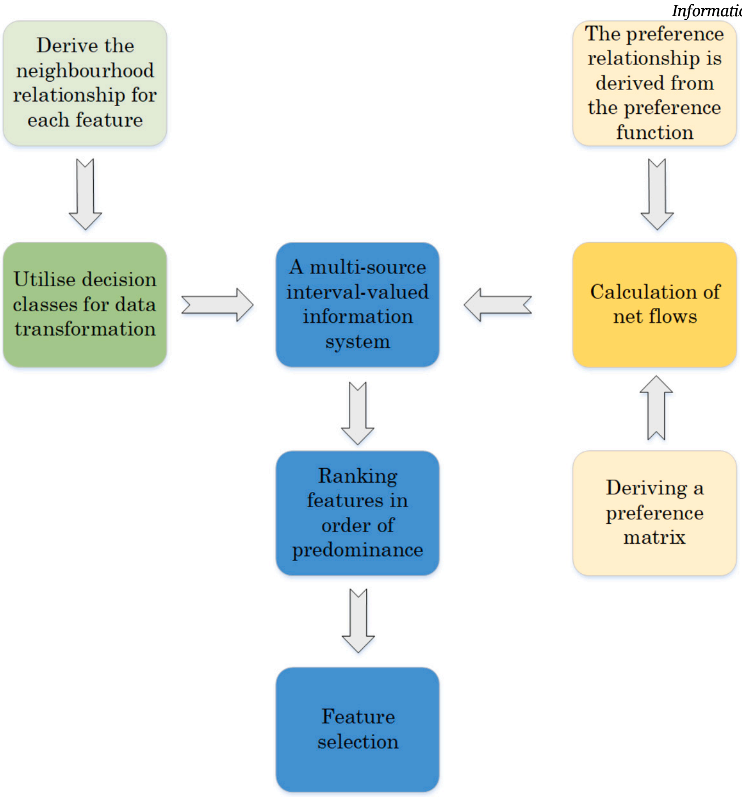


Fig. 1. Feature selection process for multi-source interval-valued information system.

Table 7
The preference matrix.

	C_1	C_2	C_3	φ^+
C_1	0	0.2916	0	0.2916
C_2	0.0833	0	0	0.0833
C_3	0.0833	0.2916	0	0.375
φ^-	0.1666	0.5833	0	

Calculating the net flow through Equation (22), it is obvious to see that $\varphi(C_3) > \varphi(C_1) > \varphi(C_2)$. Based on the previous analysis, attributes C_3 and attributes C_1 are more dominant than attribute attributes C_2, so it can be judged that attributes C_2 should be replaced.

3.5. *Information fusion*

In the previous section, each attribute has been ranked for the purpose of feature selection using the decision class-based degree of Confidence in the multi-source decision-making information system containing interval-valued data. Within this section, we need to use the attributes that have been selected to pick out the best one among all the sources.

Definition 4. Let $A' = \{U, \mathbb{C}' \cup \{d\}\} = \{U, \{C_j^k | k = 1, 2, ..., i; j = 1, 2, ...n'\} \cup \{d\}\}$ be a multi-source decision-making information system with interval-values that has been subjected to feature selection using the methods described above. We average all interval values in the system, which is defined as follows Equation (23),

$$C_{aj}^k = \frac{C_{aj}^{k-} + C_{aj}^{k+}}{2}. \tag{23}$$

After converting all interval-valued data to real-valued data, we need to go through all the sources and determine a leader. We consider each source to have the same percentage of weight, so the leader is the average of the data from all sources, reflecting the average of the data from the entire decision-making information system. Then the best one is selected through the leader for the purpose of optimal source selection.

We denote the data in the elected leaders by C^*_{aj}, which can be defined as

$$C^*_{aj} = \frac{\sum_{k=1}^{i} C^k_{aj}}{i}, \tag{24}$$

where i denotes the total number of sources, k denotes the current source, a denotes the current element, and j denotes the current attribute. In this step, we summed and averaged the corresponding values within so sources, thus obtaining an average data table with which to represent the average of the entire multi-source data set. Next, we will utilize the degree of consensus between the data within each source and the source leader's data to select the source that is most similar to the source leader, and ultimately to select the optimal source.

We define the degree of consensus as follows

$$DC_k = \sum_{a=1}^{m} \sum_{j=1}^{n'} |C^k_{aj} - C^*_{aj}|, \tag{25}$$

where j contains all the left-behind attributes, a denotes all the objects, k denotes the current source, and C^*_{aj} denotes the source leader's data. The information is fused to give us one of the most representative results. Here, he can reflect the data distribution of the whole multi-source information system, so we think he is the optimal one source and Algorithm 2 shows the specific flow. The time complexity of Algorithm 2 is $O(m \cdot n' \cdot i)$ and the space complexity is $O(m \cdot n')$.

Algorithm 2: Information fusion method based on the PROMETHEE algorithm in multi-source decision information system with interval-values.

Input: A multi-source decision-making information system with interval-values $A' = \{U, \mathbb{C} \cup \{d\}\}$ that has been subjected to feature selection.;
Output: An optimal source ;
create zero matrix: FP.
for C^{k-}_{aj} *and* C^{k+}_{aj} *in* A' **do**
    $FP \longleftarrow C^k_{aj} = \frac{C^{k-}_{aj} + C^{k+}_{aj}}{2}$.
end
create zero matrix: A^*.
for $a = 1$ *to* m ***do*** **do**
    for $j = 1$ *to* n' **do**
        for $k = 1$ *to* i **do**
            $C^*_{aj} = \frac{C^k_{aj}}{i}$.
        end
        $A^* \longleftarrow C^*_{aj}$.
    end
end
for $k = 1$ *to* i **do**
    Calculation of DC_k.
end
Select the optimal source based on DC_k.
return: A optimal source.

Example 3. In Example 2, we ranked the three attributes in order of the net flow, and based on the relationship between the advantages and disadvantages of each attribute, we removed attribute C_2, retained C_1 and C_3. Now we need to pick the most representative of the three medical examiners. By Equation (23), we start by converting all interval-valued data to floating-point numbers. The result of the floating-pointisation we show in Table 8. By Equation (24), we can construct a source leader in Table 9. By calculating the degree of consensus between each source and the source leader, we can conclude that source 3 is the best source.

Table 8
The result of the floating-pointisation.

	C_1			C_3			
A'	C_1^1	C_1^2	C_1^3	C_3^1	C_3^2	C_3^3	d
x_1	0.284	0.104	0.336	0.830	0.330	0.463	1
x_2	0.276	0.281	0.124	0.910	0.750	0.654	1
x_3	0.184	0.099	0.432	0.670	0.340	0.533	1
x_4	0.305	0.171	0.259	0.590	0.860	0.695	2
x_5	0.288	0.167	0.077	0.790	0.490	0.574	2

Table 9
The source leader.

A^*	C_1	C_3	d
x_1	0.241	0.541	1
x_2	0.227	0.771	1
x_3	0.238	0.514	1
x_4	0.245	0.715	2
x_5	0.177	0.618	2

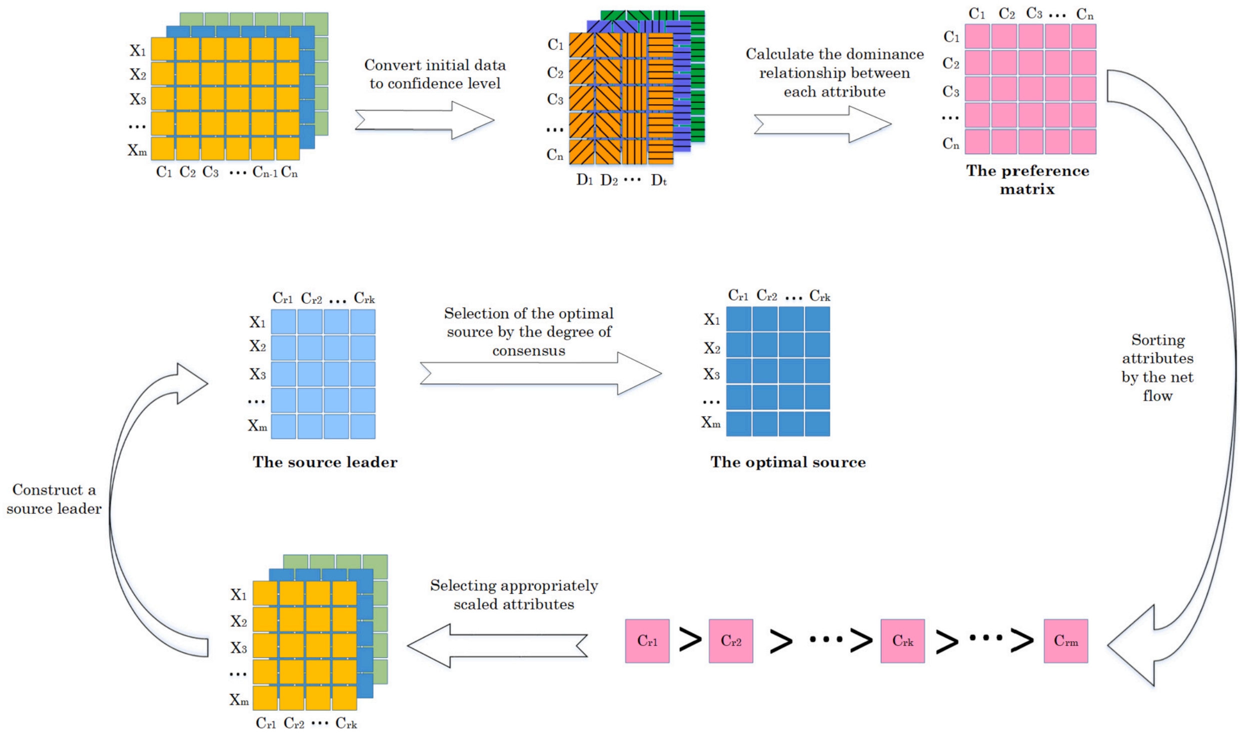


Fig. 2. Steps of the constructed method.

3.6. Steps of the constructed method

Fig. 2 illustrates the steps of the constructed method. The initial step is to convert the interval-valued data into a confidence level that can be evaluated by the PROMETHEE algorithm. This is done in order to evaluate the degree of attribute superiority or inferiority. This is done through decision classes and neighbourhood relations in a multi-source decision-making information system that contains interval-valued data. The linear preference function is then employed to derive a preference matrix for the data, which in turn yields the departure flow, the entry flow, and the net flow. The attributes are then sorted by the magnitude of the net flow, and finally feature selection is achieved.

Feature selection is performed by a certain ratio, and subsequently the approximated data containing multiple sources is obtained. An artificially created source leader is then calculated by calculating the average value based on the previous step. Subsequently, the consensus degree between each source and the leader source was calculated, and an optimal source was selected among all sources. The specific steps are as follows Table 10.

4. Experimental analysis

In order to verify the validity of the modelling algorithm, we conducted numerical experiments on some substantial datasets, selected from UCI, with the specific information shown in Table 11.

4.1. Multi-sourcing and interval initialization for data within a dataset

In practical applications, the data of multi-source decision-making information systems typically originate from a multitude of testing sources or a combination of multiple testing results. Despite the inherent discrepancies between data sets derived from disparate testing sources, these discrepancies are often not significant and tend to adhere to a normal distribution. Prior to the feature

Table 10
Steps of the constructed method.

	The constructed method
step1	Calculate the neighbourhood of interval-valued data under an attribute.
step2	The neighbourhood relationship of each object under a certain attribute is brought into the decision class to get the exact confidence of each attribute.
step3	Designing a preference function to derive a preference relation from the degree of certainty.
step4	Compute the preference matrix.
step5	Compute the departure flow, the entry flow, and the net flow.
step6	Sorting for feature selection.
step7	Replace interval data with average data.
step8	Find the average source leader.
step9	The sources of information fusion is derived from the degree of consensus between the source and the source leader.

Table 11
Description of data sets.

No	Data sets	Instances	Features	Classes
1	Wine	178	13	3
2	Parkinson's Disease Classification	756	754	2
3	Statlog (German Credit Data)	1000	20	2
4	Period Changer	90	1177	2
5	Predict Students' Dropout and Academic Success	4424	36	3
6	Statlog (Landsat Satellite)	6435	36	6
7	Seoul Bike Sharing Demand	8760	13	2
8	Online Shoppers Purchasing Intention Dataset	12330	17	2
9	Occupancy Detection	20560	6	2

Table 12
The result of data multi-sourcing.

A'	C_1	C_1^1	C_1^2	C_1^3	C_1^4	...	C_1^{19}	C_1^{20}
x_1	0.842	0.869	0.879	0.685	0.860	...	0.997	0.883
x_2	0.571	0.781	0.759	0.683	0.553	...	0.728	0.386
x_3	0.560	0.693	0.372	0.673	0.645	...	0.695	0.601
x_4	0.878	0.938	0.936	0.577	0.823	...	0.618	0.872
x_5	0.581	0.720	0.672	0.472	0.629	...	0.182	0.612

selection and information fusion, it is necessary to process the UCI dataset in order to render it applicable to real-life situations. Firstly, several related data tables must be generated based on the original data through normal distribution in order to simulate multiple sources of real-life data detection. Then, the data within each source must be intervalised. In this section, a simple strategy is employed to generate a multi-source decision information system with interval-values. The specific steps are as follows.

(1) First, each data set can be viewed as a decision table $A = \{U, C \cup \{d\}\} = \{U, \{C_j | j = 1, 2, ...n\} \cup \{d\}\}$. To generate multiple related sources, each containing a similar table of data, we need to make changes to the original data table. A total of 20 sources will be generated by first creating 20 random arrays that follow a normal distribution with a mean of 0 and a standard deviation of 0.1. These 20 normal distributions are then used in conjunction with the original data to produce 20 correlated sources. Modelling multiple sources in this way enables correlation between each source while maintaining their individuality, the step are follow as

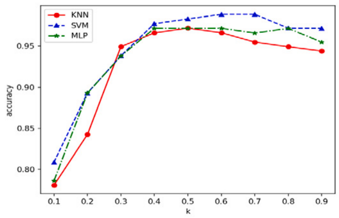
$$C_{aj}^k = C_{aj} \times (1 - \mathcal{N}_k(0, 0.1^2)), \tag{26}$$

where j denotes the corresponding attribute, k denotes the k-th source, and $\mathcal{N}_k$ is the corresponding k-th normal distribution.

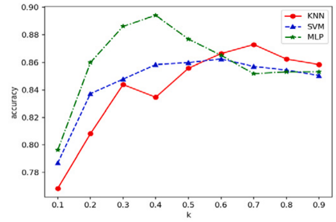
Example 4. In accordance with the aforementioned approach, the initial attribute of the initial five objects of the normalised Wine dataset was multisourced, resulting in the generation of 20 related sources. As shown in Table 12.

(2) Now we have 20 interrelated sources. To ensure consistency, we need to max-min normalise all the data since each attribute is measured on a different scale. Normalising all data has the advantage of converting data from different scales to decimals to $[0, 1]$. This ensures that each attribute score has the same influence in subsequent preference relationships, the normalisation process can be shown as

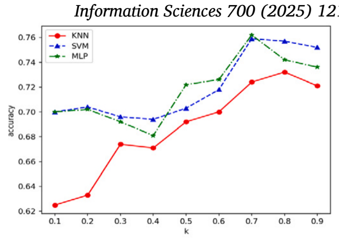
$$C_{aj}^k = \frac{C_{aj}^k - min(C_{aj}^k)}{max(C_{aj}^k) - min(C_{aj}^k)}. \tag{27}$$



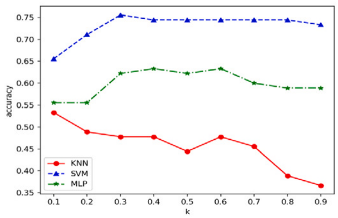
(a) Classification accuracies of Wine for different k



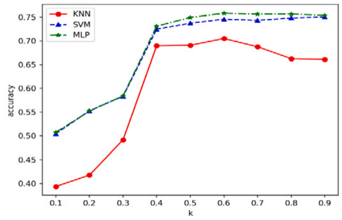
(b) Classification accuracies of Parkinson's Disease for different k



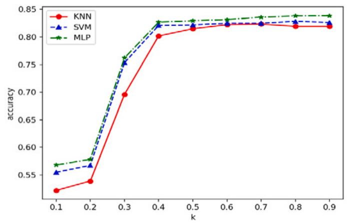
(c) Classification accuracies of Statlog (German Credit Data) for different k



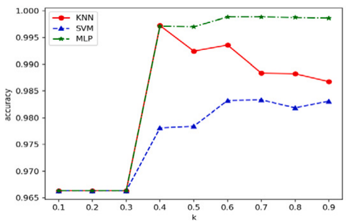
(d) Classification accuracies of Period Changer for different k



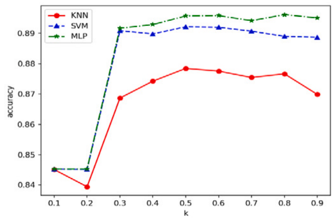
(e) Classification accuracies of Predict Students' Dropout for different k



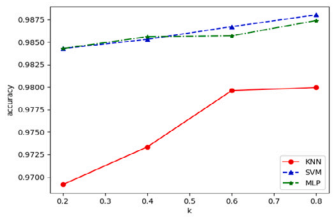
(f) Classification accuracies of Statlog (Landsat Satellite) for different k



(g) Classification accuracies of Seoul Bike Sharing Demand for different k



(h) Classification accuracies of Online Shoppers Dataset for different k



(i) Classification accuracies of Occupancy Detection for different k

Fig. 3. Classification accuracy of the optimal data.

(3) After normalising the data, it is necessary to convert it into intervals to simulate real-life interval-valued data. The interval values are calculated by adding or subtracting 0.1 times the value to or from itself. Let C^k_{aj}. The interval of C^k_{aj} is $[C^k_{aj} - 0.1 \times C^k_{aj}, C^k_{aj} + 0.1 \times C^k_{aj}]$.

Up to this point, we have divided a dataset into 20 related sources, each containing normalised interval data.

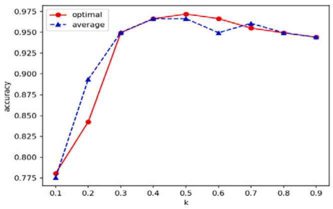
4.2. Classification accuracy of the optimal data

In these experiments, we selected all the datasets shown in Table 8 and varied the proportion k used in feature selection to observe the effect of this technique after optimal sorting. The step size for selecting parameter k differs due to variations in the number of attributes across datasets.

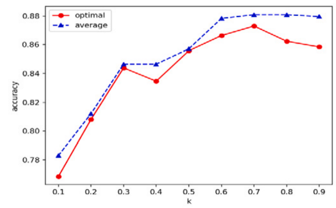
In this section, three machine learning classifiers, K Nearest Neighbours (KNN), Support Vector Machines (SVM) and Multilayer Perceptron (MLP), are used to evaluate the classification performance when different proportions of attributes are retained. All data are taken from the optimal source selected by the above methodology. Fivefold cross-validation is used to obtain the average classification accuracy as the final evaluation metric. Because Occupancy Detection contains fewer attributes, we chose a scale ratio k of $\{0.2, 0.4, 0.6, 0.8\}$, and $\{0.1 + 0.1k, k = 0, 1, ..., 8\}$ for the other datasets.

To demonstrate the variation of classification accuracy with attribute selection ratios, we present feature selection results for the UCI dataset in Fig. 3. The results are selected from the best source data. The x-axis represents the attribute selection scale k, and the y-axis represents the corresponding data's classification accuracy.

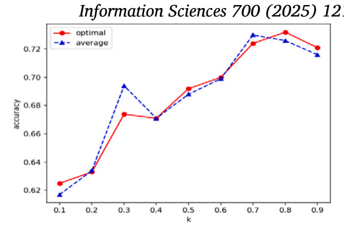
Fig. 3 shows that as the value of k increases, the proportion of selected attributes increases, resulting in smoother classification accuracy and decreased growth value. Based on our previous analyses, we have determined that attributes in the information system are sorted according to preference differences. The more advantageous attributes are listed first and are more likely to be selected when choosing attributes with the proportion of k. Conversely, the less advantageous attributes are listed last and are less likely to be selected when choosing attributes with the proportion of k. Therefore, a smaller proportion k of selected attributes results in better attribute selection and greater improvement in classification accuracy. In contrast, as the value of k increases, the advantage of the selected attributes diminishes, resulting in a relatively small improvement in classification accuracy. As a result, the classification



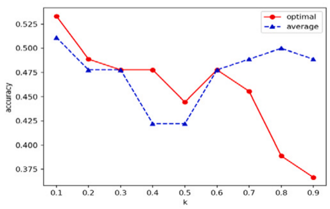
(a) Comparison of classification accuracy of Wine with source leader



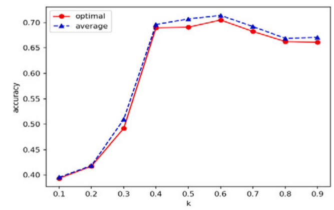
(b) Comparison of classification accuracy of Parkinson's Disease with source leader



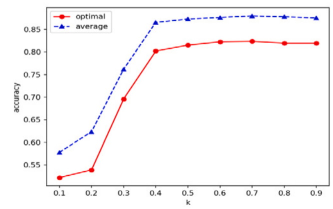
(c) Comparison of classification accuracy of Statlog(German) with source leader



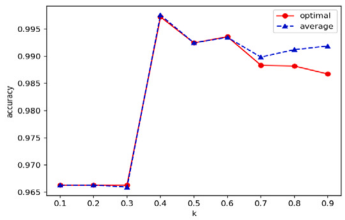
(d) Comparison of classification accuracy of Period Changer with source leader



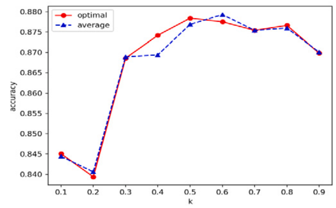
(e) Comparison of classification accuracy of Predict Students with source leader



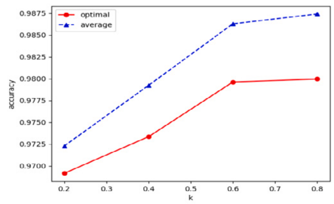
(f) Comparison of classification accuracy of Statlog (Landsat) with source leader



(g) Comparison of classification accuracy of Seoul Bike Sharing Demand with source leader



(h) Comparison of classification accuracy of Online Shoppers Purchasing Intention Dataset with source leader



(i) Comparison of classification accuracy of Occupancy Detection with source leader

Fig. 4. Comparison of the optimal data.

accuracy becomes smoother as we move further back, and in some cases, it may even decrease. From Fig. e, Fig. f, Fig. h and Fig. i, it can be seen that when more instances are included in the dataset, the classification results regarding attribute ordering are usually better, which indicates that our method can effectively handle big data datasets.

In the Jaccard distance function, the threshold value θ can be modified to some extent according to the actual situation, in the dataset used in this experiment, θ are taken as 0.1 as the threshold for determining the neighbourhood relationship of the interval values. In this case, when operating on the dataset Seoul Bike Sharing Demand Dataset, we removed the attributes Data, Seasons and Holiday and selected the remaining 10 attributes for experimentation. On the dataset Online Shoppers Purchasing Intention Dataset, we discarded three of the character-based data structures which is TrafficType, VisitorType and Weekend, then retained the other numerical data. On dataset Occupancy Detection, we removed the character attribute Data and kept the other numeric attributes.

4.3. Comparison of the optimal data

As previously stated, a source leader is defined as the optimal source among a multitude of sources. The source leader is the arithmetic mean of the data within all sources and is employed to reflect the overall mean of the sources. In selecting the optimal source, it is desirable to identify a source exhibiting moderate overall data, which confers the advantage of enhanced stability and objectivity. Consequently, all sources are compared with the previously constructed source leader, with the aim of selecting the one that is closest to the source leader. This is believed to be the optimal source.

To illustrate the efficacy of our approach to information fusion, we contrast the classification accuracy of the optimal sources with that of the source leaders, employing the K-nearest neighbour (KNN) algorithm.

Fig. 4 illustrates the relationship between classification accuracy and the optimal source and source leader under KNN, as the proportion of selected attributes k increases. The figure indicates that the optimal source, obtained after distance selection, is highly similar to the source leader in terms of classification accuracy. The data was averaged from all sources to obtain the source leader. The source closest to the source leader, which is the most similar to it, was then selected using the distance function. To retain the attributes, a suitable attribute proportion k was chosen based on the actual situation. The most representative source was then selected from the original multiple sources by constructing a source leader. Regardless of the proportion of attribute k taken, the classification accuracy of the optimal source and the source leader does not differ significantly. This demonstrates that our method functions as a means of information fusion. As can be observed from the structure of the experiment, our algorithm demonstrates superior performance with increased data. As illustrated in Fig.a, the limited number of attributes in the dataset results in a smaller number of fusion conditions for information fusion, which in turn leads to discrepancies between the source leaders and the optimal sources that have been selected. It is crucial to emphasise that even after attribute ranking and the selection of superior attributes, if the dataset has a limited number of attributes and a large number of objects, and k is relatively small, the classification accuracy may decline as k increases, as demonstrated in Fig.h. Nevertheless, the theoretical feasibility of the proposed approach can still be demonstrated.

4.4. Comparison with other methods of information fusion

In this subsection, we compare our method with other interval-valued multi-source information fusion methods under the above dataset. Five information fusion algorithms are adopted as comparison algorithms, as shown in the following.

(1) MinF: $C^*_{aj} = min\{C^{1-}_{aj}, ..., C^{k-}_{aj}\}$, where C^{k-}_{aj} denotes the lower bound on the value of the interval under the j-th attribute of the object a under the k-th source.

(2) MeanF: $C^*_{aj} = mean\{\frac{C^{1-}_{aj}+C^{1+}_{aj}}{2}, ..., \frac{C^{k-}_{aj}+C^{k+}_{aj}}{2}\}$, which means taking the average of the upper and lower bounds under each source

(3) MaxF: $C^*_{aj} = max\{C^{1+}_{aj}, ..., C^{k+}_{aj}\}$, where C^{k+}_{aj} denotes the upper bound on the value of the interval under the j-th attribute of the object a under the k-th source.

(4) The fusion approach is introduced by Zhang et al. [45] (written as DIFIV)

(5) The fusion approach is introduced by Xu et al. [34] (written as IFIEM)

We compared the newly proposed information fusion method based on the PROMETHEE algorithm with five other fusion methods on nine datasets. Table 13, Table 14 and Table 15 shows the classification accuracy of the six information fusion algorithms under three different classifiers, KNN, SVM, MLP. To facilitate the comparison with algorithm PRIFS, the percentage of selected attributes in PRIFS was set to 40%, 60%, and 80% in advance. The data in the table are the classification accuracy results of the five-fold cross-validation.

From the figure, it can be seen that the PRIFS algorithm has the highest average classification accuracy when 80% of the features are selected in both the classifiers, KNN and MLP, whereas under the SVM classifier, the PRIFS has the highest classification accuracy when 60% proportion of the features are selected.

It is worth noting that the small number of samples in the Period Changer dataset results in many algorithms having the same classification accuracy. This is a problem with the dataset itself, and we will not study it too much. However, on most of the other datasets, the classification accuracy of the PRIFS method is higher than the remaining five algorithms.

From Table 13, in terms of the average classification accuracies of the six information fusion methods under all datasets, under the KNN classifier, the average classification accuracy of the PRIFS method is significantly higher than that of MinF, MaxF and DIFIV, and slightly higher than that of MeanF and IFIEM when a 40% proportion of the features are selected. Whereas the average classification accuracy of the PRIFS method is significantly higher when a 60% proportion of the features are selected, which is the highest and outperforms all other methods. However, when 80% of the features are selected, the PRIFS method is only higher than MinF and MaxF. From Table 14, it can be seen that the three methods MinF, MaxF and MeanF perform poorly under the SVM classifier, while DIFIV and IFIEM perform better and are slightly higher than PRIFS when 40% and 60% of the proportion of the features are selected, but PRIFS when 80% of the proportion of the features are selected is still the one with the highest average classification accuracy. Table 15 shows the classification accuracies under MLP, from which it can be seen that the three methods MinF, MaxF and MeanF still perform poorly, while DIFIV and IFIEM exceed the PRIFS when selecting 40% and 60% of the proportion of the features, but the PRIFS when selecting 80% of the proportion of the features is still the one with the highest average classification accuracy, significantly higher than the other five methods. Therefore, in practical applications, choosing an appropriate threshold is crucial for improving the classification accuracy of information fusion.

Under all the datasets, we selected the PRIFS algorithm with one of the highest classification accuracies under all the feature selection ratios, and performed the WILCOXON test with the remaining five algorithms under the three classifiers, KNN, SVM, and MLP, to demonstrate the significance of our proposed approach. Let the null hypothesis be $H0$: No significant differences between PRIFS and comparison methods. When the significance level is 0.05, if the P value is greater than 0.05, it means that the null hypothesis is not rejected. If the P value is less than 0.05, it means that the null hypothesis is rejected. Table 16 shows the p value of the WILCOXON test. As can be seen in the figure, all the original hypotheses are rejected under KNN, indicating that there is a significant difference between PRIFS and these five compared methods. While in SVM, the p-value of PRIFS and DIFIV is greater than 0.05, which means that the difference is not significant. Similarly, under MLP, the p-value for both DIFIV and IFIEM is greater than 0.05, which may be caused by the small sample size. The PRIFS method is significant and scalable on the experimental dataset. This result provides a theoretical basis for real-life application of the model, which is expected to be able to show similar advantages in real data from different domains.

Table 13
Classification accuracy of information fusion based on KNN.

	MinF	MeanF	MaxF	DIFIV	IFIEM	PRIFS(40%)	PRIFS(60%)	PRIFS(80%)
Data1	0.92714	0.94952	0.95508	0.93285	**0.95539**	0.94920	0.95507	0.93253
Data2	0.80823	0.80823	0.80693	0.86502	0.85842	0.8346	**0.86635**	0.86238
Data3	0.722	0.719	0.719	0.711	0.723	0.671	0.7	**0.732**
Data4	0.45556	**0.52222**	0.47778	0.42222	0.42222	0.47778	0.47778	0.38889
Data5	0.65687	0.66094	0.6623	0.65263	0.65416	0.6896	**0.70502**	0.66252
Data6	0.7855	0.832	0.8315	0.8435	**0.8355**	0.8020	0.8225	0.8195
Data7	0.97728	0.97591	0.97705	0.98287	0.98778	**0.99725**	0.99360	0.98835
Data8	0.86156	0.86123	0.86099	0.86253	0.86285	0.87421	**0.87753**	0.87664
Data9	0.93979	0.92768	0.92845	0.977626	0.97699	0.97335	0.97962	**0.97996**
Average	0.79266	0.80630	0.80212	0.80558	0.80737	0.81032	**0.81971**	0.80475

Table 14
Classification accuracy of information fusion based on SVM.

	MinF	MeanF	MaxF	DIFIV	IFIEM	PRIFS(40%)	PRIFS(60%)	PRIFS(80%)
Data1	0.96651	0.96651	0.96079	**0.9719**	**0.9719**	0.97175	0.96063	0.94565
Data2	0.83736	0.83868	0.83999	0.84789	0.84525	0.85847	**0.86244**	0.8545
Data3	**0.761**	0.757	0.754	0.749	0.756	0.694	0.718	0.757
Data4	0.7	0.7	0.7	0.72222	0.7222	**0.74444**	**0.74444**	**0.74444**
Data5	0.75927	0.75881	**0.75972**	0.7561	0.75197	0.72469	0.74571	0.74819
Data6	0.8175	0.828	0.8285	**0.838**	0.8325	0.821	0.825	0.8286
Data7	0.98105	0.98082	0.9807	0.98298	0.98325	0.97808	**0.98322**	0.98185
Data8	0.88308	0.88491	0.88573	0.88215	0.88532	0.88978	**0.892**	0.88897
Data9	0.98025	0.97872	0.9786	0.98592	0.98398	0.98531	0.98572	**0.98804**
Average	0.85402	0.85493	0.85422	0.85957	0.85916	0.85195	0.85757	**0.85969**

Table 15
Classification accuracy of information fusion based on MLP.

	MinF	MeanF	MaxF	DIFIV	IFIEM	PRIFS(40%)	PRIFS(60%)	PRIFS(80%)
Data1	0.96079	**0.97762**	0.96651	0.97206	0.96634	0.96048	0.96587	0.96032
Data2	0.81094	0.83071	0.81881	0.83599	0.84257	**0.89417**	0.86512	0.85318
Data3	0.69	0.688	0.684	0.732	0.712	0.681	0.726	**0.742**
Data4	0.58889	0.6	0.61111	0.62222	**0.68**	0.63333	0.63333	0.58889
Data5	0.75474	0.7561	0.74525	0.73937	0.75746	0.73102	**0.75814**	0.75633
Data6	0.8045	0.8265	0.825	**0.858**	0.81	0.8265	0.831	0.838
Data7	0.99623	0.99646	0.99726	0.99783	0.97631	0.99715	**0.99886**	0.99874
Data8	0.89376	0.89294	0.88573	0.89502	0.88807	0.89286	0.89578	**0.89611**
Data9	0.9875	0.98745	**0.98833**	0.987	0.98711	0.9856	0.9857	0.9875
Average	0.83192	0.83954	0.83577	0.84883	0.84887	0.84467	**0.85109**	0.84679

Table 16
P value of the WILCOXON test.

	MinF	MeanF	MaxF	DIFIV	IFIEM
KNN	<0.05	<0.05	<0.05	<0.05	<0.05
SVM	<0.05	<0.05	<0.05	0.12	<0.05
MLP	<0.05	<0.05	<0.05	0.09	0.07

4.5. Parametric experiments

In order to further analyze the influence of parameters on the designed results, we change the threshold value θ in the Jaccard distance function and reorder the attributes of the Wine data set. A larger threshold indicates that the neighbourhood relationship contains more objects. Table 17 and Fig. 5 shows the parameters and sorting of the attributes.

It can be seen from Table 17 that the first three optimal objects are constant as the threshold value θ changes. In addition, from the ranking results of each group, the ranking results of all schemes are relatively stable, and only a few objects have slight changes in their rankings. A more intuitive ranking result can be obtained from Fig. 5. Overall, the ranking results of our method are still less affected by the threshold value θ.

Table 17
The comparison of different θ in Wine.

The threshold value θ	Ranking result	Optimal object
0.1	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_2 \succ C_{11} \succ C_6 \succ C_8 \succ C_9 \succ C_5 \succ C_4 \succ C_1 \succ C_3$	C_7
0.2	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_{11} \succ C_2 \succ C_6 \succ C_9 \succ C_8 \succ C_5 \succ C_4 \succ C_1 \succ C_3$	C_7
0.3	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_{11} \succ C_2 \succ C_6 \succ C_9 \succ C_5 \succ C_8 \succ C_1 \succ C_4 \succ C_3$	C_7
0.4	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_{11} \succ C_2 \succ C_6 \succ C_9 \succ C_5 \succ C_4 \succ C_1 \succ C_8 \succ C_3$	C_7
0.5	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_{11} \succ C_2 \succ C_9 \succ C_6 \succ C_5 \succ C_4 \succ C_1 \succ C_8 \succ C_3$	C_7
0.6	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_{11} \succ C_2 \succ C_9 \succ C_6 \succ C_5 \succ C_4 \succ C_1 \succ C_3 \succ C_8$	C_7
0.7	$C_7 \succ C_{13} \succ C_{10} \succ C_{12} \succ C_{11} \succ C_2 \succ C_6 \succ C_9 \succ C_5 \succ C_4 \succ C_1 \succ C_3 \succ C_8$	C_7
0.8	$C_7 \succ C_{13} \succ C_{10} \succ C_{11} \succ C_{12} \succ C_2 \succ C_9 \succ C_4 \succ C_6 \succ C_5 \succ C_1 \succ C_3 \succ C_8$	C_7
0.9	$C_7 \succ C_{13} \succ C_{10} \succ C_{11} \succ C_{12} \succ C_2 \succ C_4 \succ C_5 \succ C_9 \succ C_6 \succ C_1 \succ C_3 \succ C_8$	C_7

Table 18
The result of the FRIEDMAN test.

Friedman Value	χ_F^2	P value
3.841584	0.62783	0.87112

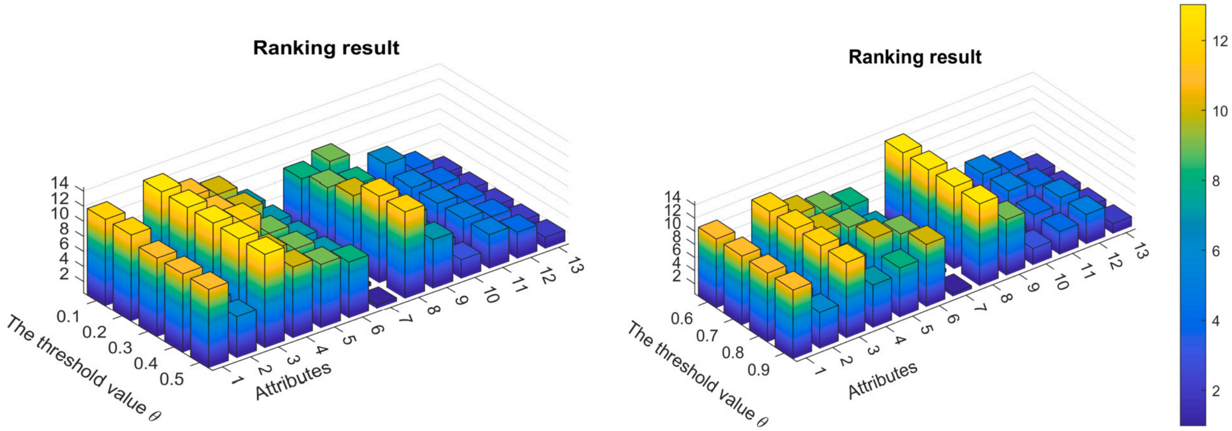


Fig. 5. The influence of the change of threshold value θ on attribute sorting.

Then we did a FRIEDMAN test on the above sorted results. Table 18 shows the P value is 0.87112, which is larger than 0.05. So we can think that there is no significant difference in the ordering of these groups, the threshold value θ has little effect on the ordering results of feature selection.

5. Conclusions

In this paper, we have applied the PROMETHEE algorithm to the realm of feature extraction, crafting a method that employs both the algorithm and confidence levels to rank attributes and synthesize data effectively. Our work culminates in the following principal contributions:

(1) A systematic feature extraction method based on the PROMETHEE algorithm was developed and subsequently applied to attribute ranking. This approach enables the effective assessment and measurement of the relevance of individual features to the classification task. By employing the PROMETHEE algorithm, we devised a structured framework for the selection of the most pertinent features, thus establishing a robust basis for subsequent data analysis and modelling.

(2) In the field of information fusion, we introduce the concept of the “source leader”, an innovative strategy that circumvents the conventional multi-step fusion process by selecting features directly from a multitude of information sources. By reducing the number of data processing steps, this strategy reduces the complexity of information processing and significantly improves the efficiency of feature selection, particularly in the context of datasets with multiple information sources. This innovation is not only unique in theory but also demonstrates a notable enhancement in practical performance.

(3) A comprehensive analysis is conducted to evaluate the stability and effectiveness of the proposed method in comparison with other prevalent feature selection techniques. The experimental results demonstrate that the proposed method exhibits high performance consistency across a multitude of datasets and application scenarios, thereby exhibiting strong generality and reliability. In particular, the present method demonstrates significant advantages when applied to datasets with a high number of features. The experimental results validate the feasibility of the proposed method and provide support for its use in a wider range of applications.

As we look to the future, the path for further research is clear. There is potential to explore more sophisticated methods for establishing attribute ranking criteria and to employ more precise techniques for identifying source leaders. Extending our approach to other complex environments for feature extraction and information fusion presents an exciting prospect. Additionally, investigating the integration of our method with other methodologies could yield innovative solutions within the broader context of decision-making systems.

CRediT authorship contribution statement

Weihua Xu: Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Zhenyuan Tian:** Writing – review & editing, Writing – original draft, Visualization, Software, Investigation, Formal analysis, Data curation.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant 62376229, and Natural Science Foundation of Chongqing, under Grant CSTB2023NSCQ-LZX0027.

Data availability

No data was used for the research described in the article.

References

[1] T. Bernhardsen, Geographic Information Systems: An Introduction, John Wiley & Sons, 2002.
[2] E. Borgonovo, M. Marinacci, Decision analysis under ambiguity, Eur. J. Oper. Res. 244 (2015) 823–836.
[3] J.P. Brans, P. Vincke, B. Mareschal, How to select and how to rank projects: the promethee method, Eur. J. Oper. Res. 24 (1986) 228–238.
[4] B. Cai, Y. Liu, Q. Fan, Y. Zhang, Z. Liu, S. Yu, R. Ji, Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network, Appl. Energy 114 (2014) 1–9.
[5] J. Deng, J. Zhan, W.Z. Wu, A ranking method with a preference relation based on the promethee method in incomplete multi-scale information systems, Inf. Sci. 608 (2022) 1261–1282.
[6] A.K. Farahat, A. Ghodsi, M.S. Kamel, An efficient greedy method for unsupervised feature selection, in: 2011 IEEE 11th International Conference on Data Mining, IEEE, 2011, pp. 161–170.
[7] W. Gao, C. Song, F. Tin-Loi, Probabilistic interval analysis for structures with uncertainty, Struct. Saf. 32 (2010) 191–199.
[8] D. Guo, W. Xu, W. Ding, Y. Yao, X. Wang, W. Pedrycz, Y. Qian, Concept-cognitive learning survey: mining and fusing knowledge from data, Inf. Fusion 109 (2024) 102426.
[9] D. Guo, W. Xu, Y. Qian, W. Ding, Fuzzy-granular concept-cognitive learning via three-way decision: performance evaluation on dynamic knowledge discovery, IEEE Trans. Fuzzy Syst. (2023).
[10] D. Guo, W. Xu, Y. Qian, W. Ding, M-fccl: memory-based concept-cognitive learning for dynamic fuzzy data classification and knowledge fusion, Inf. Fusion 100 (2023) 101962.
[11] I. Guyon, C. Aliferis, et al., Causal feature selection, in: Computational Methods of Feature Selection, Chapman and Hall/CRC, 2007, pp. 79–102.
[12] R. Haux, Health information systems–past, present, future, Int. J. Med. Inform. 75 (2006) 268–281.
[13] Q. Hu, Z. Li, L. Wang, Y. Huang, Y. Wang, L. Li, Rainfall spatial estimations: a review from spatial interpolation to multi-source data merging, Water 11 (2019) 579.
[14] K. Hyde, H.R. Maier, C. Colby, Incorporating uncertainty in the promethee mcda method, J. Multi-Criteria Decis. Anal. 12 (2003) 245–259.
[15] K. Kira, L.A. Rendell, The feature selection problem: traditional methods and a new algorithm, in: Proceedings of the Tenth National Conference on Artificial Intelligence, 1992, pp. 129–134.
[16] J. Li, H. Liu, Challenges of feature selection for big data analytics, IEEE Intell. Syst. 32 (2017) 9–15.
[17] Y. Li, T. Li, H. Liu, Recent advances in feature selection and its applications, Knowl. Inf. Syst. 53 (2017) 551–577.
[18] H. Liao, Z. Xu, Multi-criteria decision making with intuitionistic fuzzy promethee, J. Intell. Fuzzy Syst. 27 (2014) 1703–1717.
[19] C. Macharis, J. Springael, K. De Brucker, A. Verbeke, Promethee and ahp: the design of operational synergies in multicriteria analysis.: strengthening promethee with ideas of ahp, Eur. J. Oper. Res. 153 (2004) 307–317.
[20] E.F. Nakamura, A.A. Loureiro, A.C. Frery, Information fusion for wireless sensor networks: methods, models, and classifications, ACM Comput. Surv. 39 (2007) 9–es.
[21] J.A. O'brien, G.M. Marakas, Management Information Systems, vol. 6, McGraw-Hill Irwin New York, NY, USA, 2006.
[22] P. Qiu, Z. Niu, C. Zhang, Research on the multi-source causal feature selection method based on multiple causal relevance, Knowl.-Based Syst. 265 (2023) 110334.
[23] A. Ross, A. Jain, Information fusion in biometrics, Pattern Recognit. Lett. 24 (2003) 2115–2125.
[24] T.L. Saaty, M. Ozdemir, Negative priorities in the analytic hierarchy process, Math. Comput. Model. 37 (2003) 1063–1075.
[25] R. Santhanam, J. Kyparisis, A multiple criteria decision model for information system project selection, Comput. Oper. Res. 22 (1995) 807–818.
[26] S.P. Shary, A new technique in systems analysis under interval uncertainty and ambiguity, Reliab. Comput. 8 (2002) 321–418.
[27] A. Srinivasan, J. Richards, Knowledge-based techniques for multi-source classification, Remote Sens. 11 (1990) 505–525.
[28] C. Stiller, F.P. León, M. Kruse, Information fusion for automotive applications–an overview, Inf. Fusion 12 (2011) 244–252.
[29] S. Sun, Z. Deng, Multi-sensor optimal information fusion Kalman filter, Automatica 40 (2004) 1017–1023.
[30] B.C. Tso, P.M. Mather, Classification of multisource remote sensing imagery using a genetic algorithm and Markov random fields, IEEE Trans. Geosci. Remote Sens. 37 (1999) 1255–1260.
[31] O.S. Vaidya, S. Kumar, Analytic hierarchy process: an overview of applications, Eur. J. Oper. Res. 169 (2006) 1–29.

[32] A. Verikas, M. Bacauskiene, Feature selection with neural networks, Pattern Recognit. Lett. 23 (2002) 1323–1335.
[33] P. Watanachaturaporn, M.K. Arora, P.K. Varshney, Multisource classification using support vector machines, Photogramm. Eng. Remote Sens. 74 (2008) 239–246.
[34] W. Xu, K. Cai, D.D. Wang, A novel information fusion method using improved entropy measure in multi-source incomplete interval-valued datasets, Int. J. Approx. Reason. 164 (2024) 109081.
[35] W. Xu, D. Guo, Y. Qian, W. Ding, Two-way concept-cognitive learning method: a fuzzy-based progressive learning, IEEE Trans. Fuzzy Syst. 31 (2022) 1885–1899.
[36] W. Xu, Y. Pan, X. Chen, W. Ding, Y. Qian, A novel dynamic fusion approach using information entropy for interval-valued ordered datasets, IEEE Trans. Big Data 9 (2022) 845–859.
[37] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, B. Zhang, Multisource remote sensing data classification based on convolutional neural network, IEEE Trans. Geosci. Remote Sens. 56 (2017) 937–949.
[38] Y. Xu, C. Wang, J. Lai, Weighted multi-view clustering with feature selection, Pattern Recognit. 53 (2016) 25–35.
[39] R.R. Yager, A framework for multi-source data fusion, Inf. Sci. 163 (2004) 175–200.
[40] K. Yu, L. Liu, J. Li, W. Ding, T.D. Le, Multi-source causal feature selection, IEEE Trans. Pattern Anal. Mach. Intell. 42 (2019) 2240–2256.
[41] K. Yuan, D. Miao, W. Pedrycz, W. Ding, H. Zhang, Ze-hfs: zentropy-based uncertainty measure for heterogeneous feature selection and knowledge discovery, IEEE Trans. Knowl. Data Eng. (2024).
[42] Z. Zeng, H. Zhang, R. Zhang, C. Yin, A novel feature selection method considering feature interaction, Pattern Recognit. 48 (2015) 2656–2666.
[43] P. Zhang, T. Li, G. Wang, C. Luo, H. Chen, J. Zhang, D. Wang, Z. Yu, Multi-source information fusion based on rough set theory: a review, Inf. Fusion 68 (2021) 85–117.
[44] P. Zhang, T. Li, Z. Yuan, Z. Deng, G. Wang, D. Wang, F. Zhang, A possibilistic information fusion-based unsupervised feature selection method using information quality measures, IEEE Trans. Fuzzy Syst. 31 (2023) 2975–2988.
[45] X. Zhang, X. Chen, W. Xu, W. Ding, Dynamic information fusion in multi-source incomplete interval-valued information system with variation of information sources and attributes, Inf. Sci. 608 (2022) 1–27.