Applied Intelligence (2022) 52:15177–15192
https://doi.org/10.1007/s10489-022-03240-w

A heterogeneous multi-attribute case retrieval method based on neutrosophic sets and TODIM for emergency situations

Kai Zhang[1] · Jing Zheng[2,3] · Ying-Ming Wang[3,4]

Accepted: 12 January 2022 / Published online: 11 March 2022

Abstract
Heterogeneous multi-attribute case retrieval is a crucial step in generating emergency alternatives during the course of emergency decision making (EDM) by referring to historical cases. This paper develops a heterogeneous multi-attribute case retrieval method for EDM that considers five attribute formats: crisp numbers, interval numbers, intuitionistic fuzzy numbers, single-valued neutrosophic numbers (SvNNs), and interval-valued neutrosophic numbers (IvNNs). First, we propose a similarity measurement of IvNNs and calculate the attribute similarities for the five attribute formats. The attribute weights are established using an optimal model. Next, the case similarities are calculated and the set of the similar historical cases is constructed. Furthermore, the evaluated information based on heterogeneous multi-attribute from similar historical cases is provided, and the calculation method for the evaluation of utility based on TODIM (an acronym for interactive and multi-criteria decision-making in Portugese) is proposed. The most suitable historical case is determined based on the case similarity and the evaluated utility. From this, the emergency alternative is generated. Finally, we demonstrate the efficacy of the proposed method with a case study and conduct comparisons against the performance of existing methods to assess the validity and superiority of the proposed method.

Keywords Heterogeneous multi-attribute case retrieval · Interval-valued neutrosophic numbers · Emergency decision making · Case similarity · TODIM

1 Introduction

Emergency events (EEs), such as the novel coronavirus pneumonia in 2019, the missing Malaysia Airlines Flight MH370, earthquakes, and hurricanes, continue to adversely affect society [1]. When an EE occurs, quickly providing an emergency alternative is a key part of emergency decision making (EDM), and finding these alternatives has received significant research attention [2–5]. To better handle EEs, emergency departments have established emergency plans [6]; however, these plans are usually determined based on hypothetical situations, which are often very different from the actual EE that must be addressed. In such a case, the emergency plan may be ineffective [7]. Therefore, given the potential ineffectiveness of the emergency plan, quickly generating an emergency alternative is vital.

Case-based reasoning (CBR) is an artificial intelligence method that retrieves relevant historical cases and generates an alternative strategy for a particular problem by referring to the alternatives chosen from the historical cases. Moreover, this method is convenient and easy to use. Because of these advantages, researchers have recently begun to focus on the application of CBR to EDM [6, 8–12]. The basic process of using CBR to solve problems can be summarized into four main steps: retrieve, re-use, revise, and retain [13]. In previous research, retrieval has mainly been applied

✉ Kai Zhang
zhangkai@fjcpc.edu.cn

Jing Zheng
zhengjing80@qq.com

Ying-Ming Wang
msymwang@hotmail.com

1 College of Information and Intelligent Transportation, Fujian Chuanzheng Communications College, Fuzhou, 350007, Fujian, China

2 College of Electronics and Information Science, Fujian Jiangxia University, Fuzhou, 350108, Fujian, China

3 Institute of Decision Science, Fuzhou University, Fuzhou, 350116, Fujian, China

4 Key Laboratory of Spatial Data Mining & Information Sharing of Ministry of Education, Fuzhou University, Fuzhou, 350116, Fujian, China

to obtain the most suitable historical case and to make decisions based on that emergency alternative. Therefore, case retrieval plays a crucial role in these applications. If the retrieved historical cases are similar to the target case, then the generated alternatives are effective; otherwise, they are ineffective. Therefore, to increase the efficacy of the case retrieval results, it is crucial to study case similarity measurement.

In EDM based on CBR, the representation of emergencies often deploys case attributes, making it a multi-attribute decision-making problem. As the decision problems assessed by these methods become increasingly complex, multi-attribute decision problems based on fuzzy theory have become more popular, such as probabilistic linguistic term sets [14, 15], intuitionistic fuzzy sets (IFSs) [16], and Pythagorean fuzzy sets [17]. However, it is difficult to express the present complex decision-making problems with only a single format of information [18]. Thus, heterogeneous multi-attribute decision-making problems have drawn increasing research attention [19–21]. In particular, an EE is characterized by its abruptness, uncertainty, and time urgency, which means that heterogeneous multi-attribute decision making can better express the case information than methods using a single format for information. Meanwhile, many case similarity measurements have been developed to handle heterogeneous case information [6, 7, 9, 16, 18, 22]. Among these measurements, case attributes are based on crisp numbers, interval numbers, fuzzy linguistic variables, random variables, intuitionistic fuzzy numbers (IFNs), among others. However, emergency case information has become increasingly complex, and a great deal of information cannot be expressed accurately. Instead, the information is often expressed as fuzzy, vague, incomplete, indeterminate, and/or inconsistent. Existing attribute formats cannot handle independent components and some dependent components. Therefore, methods for dealing with the above situation have become a research priority for EDM based on CBR.

In an emergency, the most similar historical case alternative is not necessarily the most suitable alternative. Therefore, some researchers instead choose the most suitable historical cases by evaluating a set of similar historical cases [22–25]. For example, Zheng et al. [22] applied crisp numbers to assess the emergency alternatives. Similarly, Zhang et al. [23] applied linguistic variables and crisp numbers to evaluate similar historical cases. Wang et al. [24] used semantic data to express the evaluated data. However, the EDM problem has become increasingly complicated owing to limitations in emergency information, the inherent uncertainty and complexity of emergencies, and the fuzzy nature of human thinking [26]. Furthermore, the selection of the most suitable historical case is also a process of MADA (multi-attribute decision analysis). Hence, the representation of the evaluated attributes requires further discussion.

The neutrosophic set, which was developed by Florantin Smarandache [27], is an extension of IFSs and considers the uncertainty, imprecision, inconsistency, and vagueness of data. A neutrosophic set consists of three parts, truth membership, indeterminacy membership, and falsity membership, where the indeterminacy membership is particularly helpful in the representation of uncertain information. Neutrosophic sets have been widely used in multi-criteria decision making, with a high degree of success [28–31]. Bolturk and Kahraman [28] proposed a new analytic hierarchy process based on interval-valued neutrosophic numbers (IvNNs). Wang et al. [29] presented a weight determination method based on single-valued neutrosophic numbers (SvNNs) preference relations. Altun et al. [30] used probabilistic simplified neutrosophic sets to improve the PROMETHEE method. Zhang et al. [31] proposed a new group decision-making method that considers the IvNNs of the data. However, in EDM based on CBR, the uncertain information about the emergency represented by a neutrosophic set is rarely considered. Because of the suddenness that characterizes EEs, the information that can be collected is often limited and sometimes cannot be accurately expressed. It involves not only the degree of truth and uncertainty but also the degree of fallacy. In this scenario, a neutrosophic set can better represent the case information. Furthermore, neutrosophic sets can also accurately express decision-makers' evaluations of similar case set in complex emergencies. Therefore, this study investigates the case retrieval method with heterogeneous multi-attributes, wherein the case information and the evaluated information includes a neutrosophic set.

During EEs, decision makers do not always act completely rationally when selecting the most suitable historical case and may present inappropriate psychological behaviors. Prospect theory and TODIM (which is a Portuguese acronym for interactive and multi-criteria decision making) are the two most commonly used methods to describe the psychological behaviors that can occur during the decision-making process. However, prospect theory has some limitations compared to TODIM. For example, within prospect theory, the aspiration levels of the attributes should be determined in advance, but in practice this can be difficult to determine. Furthermore, TODIM has been used in several EDM methods to solve decision-making problems with consideration of the psychological behavior of decision makers [22, 32, 33]. Moreover, little attention has been paid to the TODIM in a neutrosophic set environment. Therefore, this paper applies TODIM to discuss the psychological behaviors during the selection of the most suitable historical case from heterogeneous multi-attribute evaluated information.

Through the above analysis, in our case representation, we not only refer to several traditionally used attribute formats, such as crisp numbers, interval numbers, IFNs, but also introduce SvNNs and IvNNs. In the evaluation of similar historical case set, we also use heterogeneous multi-attribute to represent the evaluated information, such as SvNN, IvNN and linguistic variable. In addition, TODIM is integrated into the evaluation utility value determination process. Thus, we propose a heterogeneous multi-attribute case retrieval method based on neutrosophic numbers and TODIM to select the most suitable historical case to generate an emergency alternative in the case of an EE. However, there are three challenges to be resolved: (1) how to measure the attribute similarities of heterogeneous multi-attribute with extended use range; (2) how to determine the attribute weights easily and scientifically; (3) how to calculate the evaluated utilities of similar historical cases under the heterogeneous multi-attribute environment, and with consideration of the decision makers' psychological behavior.

To overcome these challenges, this paper makes the following major contributions: (1) The heterogeneous multi-attribute case retrieval method is extended to consider attribute values that are SvNNs or IvNNs, which can better express the information of incomplete, indeterminate, and inconsistent problems. (2) An optimization model based on attribute distance is constructed to determine attribute weights, which is not only easy to train, but also more objective. (3) The most suitable historical case selection stage not only considers the heterogeneous evaluated information, but also considers the psychological behavior of the decision makers, which makes the result of case retrieval more consistent with the actual situation and therefore more applicable.

The remainder of this paper is structured as follows: Section 2 briefly reviews the method of case retrieval and the basic concepts of interval-valued neutrosophic sets (IvNSs) and IFSs, Section 3 describes the problem, Section 4 proposes a similarity measurement based on IvNS, Section 5 introduces the case retrieval method, Section 6 demonstrates the efficiency of the proposed method through a case study, and Section 7 concludes the paper.

2 Preliminaries

In this section, we descript the method of case retrieval and review the basic concepts of IvNSs and IFSs.

2.1 Description of case retrieval

Before generating alternatives, similar historical cases are retrieved to assist decision makers in their decisions. There are three main steps to this process: calculating attribute similarity, determining attribute weights, and calculating the case similarity. The specific process of case retrieval is shown in Fig. 1. Calculating attribute similarity is mainly done to determine the corresponding attribute distance/similarity calculation formulae for different formats of attributes, for example, the Manhattan distance, Euclidean distance, Gaussian distance, and grey correlation degree. The second step mainly determines the attribute weights based on case information. There are two main methods: building an optimization model and machine learning. Machine learning requires a large amount of data, whereas optimization models do not. Furthermore, case similarity should be extracted to then retrieve the case, based on which the similar historical cases will be obtained. Finally, the decision makers can make decisions according to the alternatives of similar historical cases.

2.2 Basic concepts of interval-valued neutrosophic sets

Definition 1 [34] Suppose X is an arbitrary universe of discourse. An IvNS A is provided by the following mathematical expression:

$A = \{< x, [T_A^-(x), T_A^+(x)], [I_A^-(x), I_A^+(x)], [F_A^-(x), F_A^+(x)] > | \, x \in X\}$,

where $T_A(x)$, $I_A(x)$, and $F_A(x) : X \to int[0, 1]$.There also exists the restriction of $0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3$.

Definition 2 [35] Let A and B be two IvNSs,$A \subseteq B$ if

$T_A^-(x) \leq T_B^-(x)$, $T_A^+(x) \leq T_B^+(x)$,
$I_A^-(x) \leq I_B^-(x)$, $I_A^+(x) \leq I_B^+(x)$,
$F_A^-(x) \leq F_B^-(x)$, $F_A^+(x) \leq F_B^+(x)$

Definition 3 [35] Let A and B be two IvNSs. Then the similarity between two IvNSs being satisfied is defined as $S(A, B)$, which satisfies the following properties:

(1) $0 \leq S(A, B) \leq 1$;
(2) $S(A, B) = 1 \leftrightarrow A = B$;
(3) $S(A, B) = S(B, A)$;
(4) If $A \subseteq B \subseteq C$, then $S(A, C) \leq S(A, B)$ and $S(A, C) \leq S(B, C)$.

Definition 4 Let A be an IvNS, the score function of an IvNS can be defined as

$Sc(A) = \frac{1}{6}(2 + T_{ij}^- + T_{ij}^+ - I_{ij}^- - I_{ij}^+ - F_{ij}^- - F_{ij}^+)$.

where a larger value of $Sc(A)$ indicates a larger IvNS.

From definition 4, when A is an SvNN, the score function is

$Sc(A) = \frac{1}{3}(2 + T_{ij} - I_{ij} - F_{ij})$.

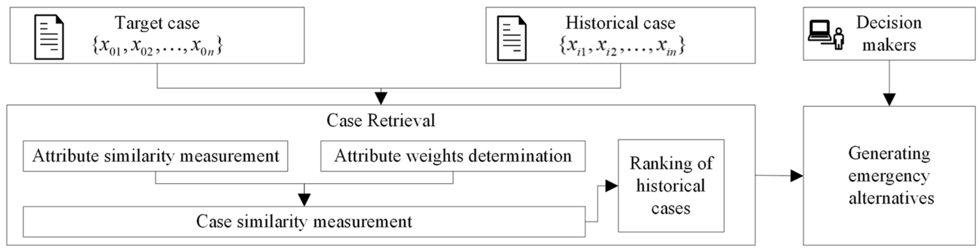

Fig. 1 The specific process of case retrieval

2.3 Basic concepts of intuitionistic fuzzy sets

Definition 5 [36] Let Y be a fixed set, then the IFS can be defined as:

$D = \{< y, \mu_D(y), \upsilon_D(y) > | \ y \in Y\}$,

where $\mu_D(y)$ is the membership function, and $\upsilon_D(y)$ is the non-membership function. $\mu_D(y)$ and $\upsilon_D(y)$ satisfy the condition: $\mu_D(y) \in [0, 1]$, $\upsilon_D(y) \in [0, 1]$, and $0 \leq \mu_D(y) + \upsilon_D(y) \leq 1, y \in Y$. The hesitant degree of the IFS can be described as $\pi_D(y) = 1 - \mu_D(y) - \upsilon_D(y)$ and $0 \leq \pi_D(y) \leq 1, y \in Y$.

For convenience, an IFN is defined as $\alpha = (\mu_\alpha, \upsilon_\alpha)$, satisfying the condition:$\mu_\alpha \in [0, 1]$, $\mu_\alpha + \upsilon_\alpha \leq 1$, where a score function is defined as $s(\alpha) = \mu_\alpha - \upsilon_\alpha$.

3 Problem description

This section briefly describes the problem of the heterogeneous multi-attribute case retrieval method with five different attribute value formats.

Suppose there exist m historical cases denoted by $C = \{C_1, C_2, \ldots, C_m\}$, where C_i denotes the ith historical case, $i \in \{1, 2, \ldots, m\}$, and suppose C_0 is the target case. Let $X = \{X_1, X_2, \ldots, X_n\}$ be a finite attribute set concerning the case problem of both historical cases and the target case, where X_j denotes the jth problem attribute, $j \in \{1, 2, \ldots, n\}$. Let $W = \{w_1, w_2, \ldots, w_n\}$ be a vector of problem attribute weights, where w_j denotes the jth problem attribute weight, such that $\sum\limits_{j=1}^{n} w_j = 1$ and $0 \leq w_j \leq 1$. Let $P_i = \{x_{i1}, x_{i2}, \ldots, x_{in}\}$ be a vector of the problem attribute value corresponding to the problem of historical case C_i, where x_{ij} denotes the jth problem attribute value of historical case C_i. Let $P_0 = \{x_{01}, x_{02}, \ldots, x_{0n}\}$ be a vector of the problem attribute value corresponding to the problem of target case C_0, where x_{0j} denotes the jth problem attribute value of target case C_0.

In this study, the problem attribute values of the case problem, x_{ij} and x_{0j} are considered to be expressed in five formats: crisp numbers, interval numbers, IFNs, SvNNs, and IvNNs. Based on this, we use Ω^1, Ω^2, Ω^3, Ω^4 and Ω^5 to represent these five formats. For the sake of convenience, the five formats of x_{ij} and x_{0j} are expressed as follows:

$$x_{ij} = \begin{cases} x_{ij}, & (X_j \in \Omega^1) \\ [x_{ij}^-, x_{ij}^+], & (X_j \in \Omega^2) \\ < \mu_{ij}, \upsilon_{ij} >, & (X_j \in \Omega^3) \\ < T_{ij}, I_{ij}, F_{ij} >, & (X_j \in \Omega^4) \\ < [T_{ij}^-, T_{ij}^+], [I_{ij}^-, I_{ij}^+], [F_{ij}^-, F_{ij}^+] >, & (X_j \in \Omega^5) \end{cases}$$

4 Similarity measurement based on interval-valued neutrosophic sets

Theorem 1 *Assume A and B are two IvNSs, namely,*

$A =< [T_A^-, T_A^+], [I_A^-, I_A^+], [F_A^-, F_A^+] >$ *and* $B =< [T_B^-, T_B^+], [I_B^-, I_B^+], [F_B^-, F_B^+] >$.

Let

$M_1 = \frac{1}{2}(1 + \sqrt{\frac{1}{2}[(T_A^- - T_B^-)^2 + (T_A^+ - T_B^+)^2]})$,

$M_2 = \frac{1}{2}(1 + \sqrt{\frac{1}{2}[(F_A^- - F_B^-)^2 + (F_A^+ - F_B^+)^2]})$,

$M_3 = \frac{1}{2}(1 - \sqrt{\frac{1}{2}[(I_A^- - I_B^-)^2 + (I_A^+ - I_B^+)^2]})$,

$M_4 = \frac{1}{2}(1 + \sqrt{\frac{1}{2}[(I_A^- - I_B^-)^2 + (I_A^+ - I_B^+)^2]})$,

$M_5 = \frac{1}{2}(1 - \sqrt{\frac{1}{2}[(T_A^- - T_B^-)^2 + (T_A^+ - T_B^+)^2]})$,

and

$M_6 = \frac{1}{2}(1 - \sqrt{\frac{1}{2}[(F_A^- - F_B^-)^2 + (F_A^+ - F_B^+)^2]})$,

then the similarity between A and B is $S(A, B) = \frac{1}{n}\sum_{i=1}^{n} \frac{min\{M_1, M_2, M_5, M_6\} + min\{max\{M_1, M_2\}, max\{M_5, M_6\}\} + |M_4 - M_3|}{max\{min\{M_1, M_2\}, min\{M_5, M_6\}\} + max\{M_1, M_2, M_5, M_6\} + |M_4 - M_3|}$

Proof A proof of the theorem is obtained according to Definition 3.

(1)$0 \leq M_1 \leq 1, 0 \leq M_2 \leq 1, 0 \leq M_3 \leq 1, 0 \leq M_4 \leq 1$, $0 \leq M_5 \leq 1$, and $0 \leq M_6 \leq 1$, so it is obvious that $0 \leq S(A, B) \leq 1$.

(2)$S(A, B) = 1 \Leftrightarrow min\{M_1, M_2\} = min\{M_5, M_6\}$, $max\{M_1, M_2\} = max\{M_5, M_6\}$, $M_4 = M_3 \Leftrightarrow T_A^- = T_B^-$,

$T_A^+ = T_B^+$, $F_A^- = F_B^-$, $F_A^+ = F_B^+$, $I_A^- = I_B^-$, $I_A^+ = I_B^+ \Leftrightarrow A = B$.

(3)$S(A, B)$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{min\{M_1,M_2,M_5,M_6\}+min\{max\{M_1,M_2\},max\{M_5,M_6\}\}+|M_4-M_3|}{max\{min\{M_1,M_2\},min\{M_5,M_6\}\}+max\{M_1,M_2,M_5,M_6\}+|M_4-M_3|}$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{min\{M_1,M_2,M_5,M_6\}+min\{max\{M_5,M_6\},max\{M_1,M_2\}\}+|M_4-M_3|}{max\{min\{M_1,M_2\},min\{M_5,M_6\}\}+max\{M_5,M_6,M_1,M_2\}+|M_4-M_3|}$$

$= S(B, A)$

(4)Because $A \subseteq B \subseteq C$,
then
$T_A^- \le T_B^- \le T_C^-$, $T_A^+ \le T_B^+ \le T_C^+$, $I_A^- \ge I_B^- \ge I_C^-$, $I_A^+ \ge I_B^+ \ge I_C^+$, $F_A^- \ge F_B^- \ge F_C^-$, $F_A^+ \ge F_B^+ \ge F_C^+$.
Clearly,
$min\{M_1, M_2\}(A, B) \le min\{M_1, M_2\}(A, C)$,
$max\{M_1, M_2\}(A, B) \le max\{M_1, M_2\}(A, C)$,
$min\{M_5, M_6\}(A, B) \ge min\{M_5, M_6\}(A, C)$,
$max\{M_5, M_6\}(A, B) \ge max\{M_5, M_6\}(A, C)$,
$min\{M_1, M_2\}(A, B) \le min\{M_5, M_6\}(A, B)$.
when
$max\{M_1, M_2\}(A, C) \le max\{M_5, M_6\}(A, C)$
and
$min\{M_1, M_2\}(A, C) \le min\{M_5, M_6\}(A, C)$
then
$min\{M_1, M_2\}(A, B) \le min\{M_1, M_2\}(A, C)$
$\le min\{M_5, M_6\}(A, C) \le min\{M_5, M_6\}(A, B)$,
$max\{M_1, M_2\}(A, B) \le max\{M_1, M_2\}(A, C)$
$\le max\{M_5, M_6\}(A, C) \le max\{M_5, M_6\}(A, B)$
then,
$min\{M_1, M_2, M_5, M_6\}(A, B)$
$\le min\{M_1, M_2, M_5, M_6\}(A, C)$
and
$min\{max\{M_1, M_2\}, max\{M_5, M_6\}\}(A, B)$
$\le min\{max\{M_1, M_2\}, max\{M_5, M_6\}\}(A, C)$,
and
$max\{M_1, M_2, M_5, M_6\}(A, B)$
$\ge max\{M_1, M_2, M_5, M_6\}(A, C)$
then
$S(A, B) \ge S(A, C)$.
Similarly, we obtain $S(B, C) \ge S(A, C)$. □

5 Case retrieval method for heterogeneous multi-attributes

To retrieve the most suitable historical case using heterogeneous multi-attributes, a problem attribute similarity measurement is presented. Next, the problem attribute weights are determined by constructing an optimization model. Based on this model, case similarities are attained and the similar historical case set is gain. Furthermore, the decision makers give the evaluation information for the similar historical case set, and the evaluated utilities are calculated based on TODIM. Finally, the most suitable historical case is retrieved based on the case similarities and evaluated utilities.

5.1 Problem attribute similarity measurement

When the heterogeneous problem attribute information includes crisp numbers, interval numbers, IFNs, SvNNs, and IvNNs, the problem attribute similarity measurement is presented as below.

If problem attributes x_{ij} and x_{0j} are crisp numbers, let $Sim_j(C_0, C_i)$ be the problem attribute similarity of problem attribute X_j between historical case C_i and target case C_0. According to Fan et al. [7], $Sim_j(C_0, C_i)$ is calculated as follows:

$$Sim_j(C_0, C_i) = exp[-\frac{\sqrt{(x_{0j} - x_{ij})^2}}{d_j^{max}}] \tag{1}$$

where $d_j^{max} = max\{\sqrt{(x_{0j} - x_{ij})^2} \mid i \in \{1, 2, \cdots, m\}\}$

Alternatively, if problem attributes x_{ij} and x_{0j} are interval numbers, that is, $x_{ij} = [x_{ij}^-, x_{ij}^+]$ and $x_{0j} = [x_{0j}^-, x_{0j}^+]$. According to Fan et al. [7], $Sim_j(C_0, C_i)$ is calculated as follows:

$$Sim_j(C_0, C_i) = exp[-\frac{\sqrt{(x_{0j}^- - x_{ij}^-)^2 + (x_{0j}^+ - x_{ij}^+)^2}}{d_j^{max}}] \tag{2}$$

where $d_j^{max} = max\{\sqrt{(x_{0j}^- - x_{ij}^-)^2 + (x_{0j}^+ - x_{ij}^+)^2} \mid i \in \{1, 2, \cdots, m\}\}$

Similarly, if problem attributes x_{ij} and x_{0j} are IFNs, that is, $x_{ij} = \langle u_{ij}, v_{ij}\rangle$ and $x_{0j} = \langle u_{0j}, v_{0j}\rangle$. According to Zheng et al. [16], $Sim_j(C_0, C_i)$ is calculated using

$$\begin{aligned} Sim_j(C_0, C_i) = &< min(1- \mid u_{ij} - u_{0j} \mid, 1- \mid v_{ij} - v_{0j} \mid), \\ &1 - max(1- \mid u_{ij} - u_{0j} \mid, \\ &1- \mid v_{ij} - v_{0j} \mid) > \end{aligned} \tag{3}$$

For convenience, following previous research [37], the score function can be utilized to transform the IFN into a scrip number, and the problem attribute similarity $Sim_j(C_0, C_i)$ is transformed using

$$\begin{aligned} Sim_j(C_0, C_i) = &\, min(1- \mid u_{ij} - u_{0j} \mid, 1- \mid v_{ij} - v_{0j} \mid) \\ &-[1 - max(1- \mid u_{ij} - u_{0j} \mid, 1- \mid v_{ij} \\ &-v_{0j} \mid)] \end{aligned} \tag{4}$$

If problem attributes x_{ij} and x_{0j} are SvNNs, that is, $x_{ij} = \langle T_{ij}, I_{ij}, F_{ij}\rangle$ and $x_{0j} = \langle T_{0j}, I_{0j}, F_{0j}\rangle$, let $Sim_j(C_0, C_i)$ be the problem attribute similarity of problem attribute X_j

between historical case C_i and target case C_0. According to Theorem 1, $Sim_j(C_0, C_i)$ is calculated using

$$Sim_j(C_0, C_i) = \frac{min\{m_{ij1}, m_{ij2}, m_{ij5}, m_{ij6}\} + min\{max\{m_{ij1}, m_{ij2}\}, max\{m_{ij5}, m_{ij6}\}\} + |m_{ij4} - m_{ij3}|}{max\{min\{m_{ij1}, m_{ij2}\}, min\{m_{ij5}, m_{ij6}\}\} + max\{m_{ij1}, m_{ij2}, m_{ij5}, m_{ij6}\} + |m_{ij4} - m_{ij3}|} \tag{5}$$

where
$m_{ij1} = \frac{1}{2}(1 + \sqrt{(T_{ij} - T_{0j})^2})$, $m_{ij2} = \frac{1}{2}(1 + \sqrt{(F_{ij} - F_{0j})^2})$,
$m_{ij3} = \frac{1}{2}(1 - \sqrt{(I_{ij} - I_{0j})^2})$, $m_{ij4} = \frac{1}{2}(1 + \sqrt{(I_{ij} - I_{0j})^2})$,
$m_{ij5} = \frac{1}{2}(1 - \sqrt{(T_{ij} - T_{0j})^2})$, and $m_{ij6} = \frac{1}{2}(1 - \sqrt{(F_{ij} - F_{0j})^2})$.

Lastly, if problem attributes x_{ij} and x_{0j} are IvNNs, that is, $x_{ij} = \left\langle [T_{ij}^-, T_{ij}^+], [I_{ij}^-, I_{ij}^+], [F_{ij}^-, F_{ij}^+] \right\rangle$ and $x_{0j} = \left\langle [T_{0j}^-, T_{0j}^+], [I_{0j}^-, I_{0j}^+], [F_{0j}^-, F_{0j}^+] \right\rangle$, let $Sim_j(C_0, C_i)$ be the problem attribute similarity of problem attribute X_j between historical case C_i and target case C_0. According to Theorem 1, $Sim_j(C_0, C_i)$ is calculated using

$$Sim_j(C_0, C_i) = \frac{min\{m_{ij1}, m_{ij2}, m_{ij5}, m_{ij6}\} + min\{max\{m_{ij1}, m_{ij2}\}, max\{m_{ij5}, m_{ij6}\}\} + |m_{ij4} - m_{ij3}|}{max\{min\{m_{ij1}, m_{ij2}\}, min\{m_{ij5}, m_{ij6}\}\} + max\{m_{ij1}, m_{ij2}, m_{ij5}, m_{ij6}\} + |m_{ij4} - m_{ij3}|} \tag{6}$$

where
$m_{ij1} = \frac{1}{2}(1 + \sqrt{\frac{1}{2}((T_{ij}^- - T_{0j}^-)^2 + (T_{ij}^+ - T_{0j}^+)^2)})$,
$m_{ij2} = \frac{1}{2}(1 + \sqrt{\frac{1}{2}((F_{ij}^- - F_{0j}^-)^2 + (F_{ij}^+ - F_{0j}^+)^2)})$,
$m_{ij3} = \frac{1}{2}(1 - \sqrt{\frac{1}{2}((I_{ij}^- - I_{0j}^-)^2 + (I_{ij}^+ - I_{0j}^+)^2)})$,
$m_{ij4} = \frac{1}{2}(1 + \sqrt{\frac{1}{2}((I_{ij}^- - I_{0j}^-)^2 + (I_{ij}^+ - I_{0j}^+)^2)})$,
$m_{ij5} = \frac{1}{2}(1 - \sqrt{\frac{1}{2}((T_{ij}^- - T_{0j}^-)^2 + (T_{ij}^+ - T_{0j}^+)^2)})$,
$m_{ij6} = \frac{1}{2}(1 - \sqrt{\frac{1}{2}((F_{ij}^- - F_{0j}^-)^2 + (F_{ij}^+ - F_{0j}^+)^2)})$.

5.2 Problem attribute weights determination

The determination of problem attribute weights plays a key role in case retrieval. Thus far, existing objective weight determination methods mainly include optimization models [38] and machine learning [39, 40]. The machine learning method requires large amounts of data. Because there is limited case information concerning emergencies, the problem attribute weights are determined using an optimal model in our proposed method. Wang and Fan [41] proposed a weight determination method based on absolute deviations, which has proven to be simple and effective. Inspired by this, the problem attribute weights in the present study were determined by an optimal model based on attribute distance. Let $d(x_{ij}, x_{kj})$ represent the distance between problem attributes x_{ij} and $x_{kj} (k \in \{1, 2, \cdots, m\})$. Then, the optimal model is constructed as follows:

$$max \quad d = \sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} w_j d(x_{ij}, x_{kj}) \tag{7}$$

$$s.t. \quad \sum_{j=1}^{n} w_j^2 = 1, w_j \geq 0, j = 1, 2, \cdots, n, \tag{8}$$

Corollary 1 *Let $w_o^* = (w_1^*, w_2^*, \cdots, w_n^*)$ be the optimal vector of model (7)-(8); then,*

$$w_j^* = \frac{\sum\limits_{i=1}^{m} \sum\limits_{k=1}^{m} d(x_{ij}, x_{kj})}{\sum\limits_{j=1}^{n} \sum\limits_{i=1}^{m} \sum\limits_{k=1}^{m} d(x_{ij}, x_{kj})} \tag{9}$$

Proof The Lagrange function of (7)-(8) can be written as

$$L(w_j^*, \lambda) = \sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} w_j d(x_{ij}, x_{kj}) - \lambda (\sum_{j=1}^{n} w_j^2 - 1) \tag{10}$$

where λ is the Lagrange multiplier. Let $\frac{\partial L}{\partial w_j} = 0, j = 1, 2, \cdots, n$, then,

$$\frac{\partial L}{\partial w_j} = \sum_{i=1}^{m} \sum_{k=1}^{m} d(x_{ij}, x_{kj}) - 2\lambda w_j = 0, j = 1, 2, \cdots, n \tag{11}$$

By solving (8) and (11) together, we obtain

$$\lambda^* = \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} d(x_{ij}, x_{kj}) \tag{12}$$

$$w_j^* = \frac{\sum\limits_{i=1}^{m} \sum\limits_{k=1}^{m} d(x_{ij}, x_{kj})}{\sum\limits_{j=1}^{n} \sum\limits_{i=1}^{m} \sum\limits_{k=1}^{m} d(x_{ij}, x_{kj})} \tag{13}$$

The problem attribute weights w_j^* can be obtained by (13), and $w_j^* \geq 0$, which satisfies the constraint of the problem attribute weights. Therefore, the proof is true. □

According to Corollary 1, the problem attribute weights $W = (w_1, w_2, \cdots, w_n)$ can be obtained.

Furthermore, we need to perform consistency checks on weighted attributes. Information consistency and attribute correlation play a critical role in heterogeneous multi-attribute data. The quality of data has an important effect in the generation of emergency response alternatives. Therefore, we should check the data firstly. Correlation coefficient is used to reflect the degree of correlation between variables, which reflects the consistency and validity of the data. It is necessary to calculate the correlation coefficient of attributes under emergencies to ensure the accuracy of the results. Thus, the steps for calculating the degree of correlation between attributes are as follows:

Step 1: Because of the different attribute formats, the information needs to be converted to the consistent. We use score function to transform the information as follows:

$$\tilde{x}_{ij} = \begin{cases} (x_{ij}^- + x_{ij}^+)/2, & (x_{ij} \in \Omega_2) \\ u_{ij} - v_{ij}, & (x_{ij} \in \Omega_3) \\ (T_{ij} + 1 - I_{ij} + 1 - F_{ij})/3, & (x_{ij} \in \Omega_4) \\ (T_{ij}^- + T_{ij}^+ + 2 - I_{ij}^- - I_{ij}^+ + 2 - F_{ij}^- - F_{ij}^+)/6, & (x_{ij} \in \Omega_5) \end{cases} \tag{14}$$

Step 2: Because $\tilde{x}_{ij}$ is a dimensional representation, it needs to be dimensionless when standardized. Suppose that $\Lambda_{benefit}$ and Λ_{cost} are the sets of benefits and costs, respectively, the standardization is as follows:

$$e_{ij} = \begin{cases} \frac{\tilde{x}_{ij} - \tilde{x}_j^{min}}{\tilde{x}_j^{max} - \tilde{x}_j^{min}}, \tilde{x}_{ij} \in \Lambda_{benefit} \\ \frac{\tilde{x}_j^{max} - \tilde{x}_{ij}}{\tilde{x}_j^{max} - \tilde{x}_j^{min}}, \tilde{x}_{ij} \in \Lambda_{cost} \end{cases} \tag{15}$$

where $\tilde{x}_j^{min} = \underset{1 \le i \le n}{min}\{\tilde{x}_{ij}\}$, $\tilde{x}_j^{max} = \underset{1 \le i \le n}{max}\{\tilde{x}_{ij}\}$

Step 3: Through the weighted summation of the attribute, the total value of each attribute can be calculated. Then the correlation coefficient between attribute X_j and $X_f (f \in \{1, 2, \cdots, n\}, f \neq j)$ can be calculated by

$$R_{j,f} = \frac{m \sum_{i=1}^{m} w_j e_{ij} w_j e_{if} - \sum_{i=1}^{m} w_j e_{ij} \sum_{i=1}^{m} w_j e_{if}}{\sqrt{m \sum_{i=1}^{m} (w_j e_{ij})^2 - \sum_{i=1}^{m} (w_j e_{ij})^2} \sqrt{n \sum_{i=1}^{m} (w_j e_{if})^2 - \sum_{i=1}^{m} (w_j e_{if})^2}} \tag{16}$$

where $\mid R_{j,f} \mid \in [0, 1]$. The greater $\mid R_{j,f} \mid$ is, the stronger the correlation between attribute X_j and X_f is. When $\mid R_{j,f} \mid \geq 0.7$, it indicates attribute X_j and X_f have a strong correlation, when $0.4 \leq \mid R_{j,f} \mid \leq 0.7$, it indicates attribute X_j and X_f have a general correlation, when $\mid R_{j,f} \mid \leq 0.3$, it indicates attribute X_j and X_f have a low correlation and attribute selection needs to be considered.

5.3 Identify similar historical cases

To retrieve similar historical cases, it is necessary to measure the case similarities. Let $Sim(C_0, C_i)$ denote the case similarity between historical case C_i and target case C_0. We use the simple additive weight method to aggregate the attribute similarities into case similarities through problem attribute weights. Then, the case similarity $Sim(C_0, C_i)$ can be derived using the following equations:

$$Sim(C_0, C_i) = \sum_{j=1}^{n} w_j Sim_j(C_0, C_j) \tag{17}$$

The case similarities can be calculated by (17), and $Sim(C_0, C_i) \in [0, 1]$. The greater $Sim(C_0, C_i)$ is, the more similar is the historical case C_i.

According to the case similarity $Sim(C_0, C_i)$, the set of similar historical cases can be constructed. The decision makers provide the case similarity threshold ξ based on the information obtained and their own personal experience, where $\xi \in [min\{Sim(C_0, C_i)\}, max\{Sim\ (C_0, C_i)\}]$. A larger value of ξ indicates that the decision makers have high expectations that the cases are similar. When the case similarity of historical case C_i meets the condition $Sim(C_0, C_i) \geq \xi$, the historical case C_i is selected. Then, the set of similar historical case is expressed as $S_v, v = \{1, 2, \cdots, h\}$, such that $S_v = \{C_i \mid Sim(C_0, C_i) \geq \xi\}$.

5.4 Determine the evaluated utility of the similar historical cases

To select the most suitable historical case, the decision maker provides an evaluation of the similar historical cases' alternatives, which will be used in the target case C_0. Let $R = \{R_1, R_2, \cdots, R_L\}$ be the evaluation attribute vector, where R_l represents the lth evaluation attribute, $l \in \{1, 2, \cdots, L\}$. Let $W^R = \{w_1^R, w_2^R, \cdots, w_L^R\}$ be the evaluation attribute weight vector, such that $\sum_{l=1}^{L} w_l^R = 1$, and $0 \leq w_l^R \leq 1$. Let q_{vl}^R denote the evaluation attribute value of the lth evaluation attribute with regard to the similar historical case S_v, and q_{vl}^R is an IvNN or SvNN. Considering the decision makers demonstrate bounded rationality, TODIM is used to represent their psychological behaviors and determine the evaluation utilities. Thus, the steps for determining the evaluated utilities of the similar historical cases based on TODIM are as follows.

Step 1: When the evaluated attribute is SvNN or IvNN, calculate the score function $Sc(q_{vl}^R)$ according to Definition 4; when the evaluated attribute is linguistic variable, calculate the normalization formula as follows:

$$Sc(q_{vl}^R) = 1/Seq(q_{vl}^R) \tag{18}$$

where $Seq(q_{vl}^R)$ is the subscript of the linguistic set $T = \{T_1, T_2, \cdots, T_t\}$.

Step 2: Calculate the relative weight $w_{lr}^R, l, r \in \{1, 2, \cdots, L\}$, *i.e.*,

$$w_{lr}^R = w_l^R / w_r^R \tag{19}$$

Step 3: Calculate the cumulative distribution function F_{vl}^R as follows:

$$F_{vl}^R(x) = \begin{cases} 0, & x < Sc(q_{vl}^R) \\ 1, & x \geq Sc(q_{vl}^R) \end{cases} \tag{20}$$

Step 4: Calculate the gain of similar historical case S_v over case $S_k (k \in \{1, 2, \cdots, h\})$ for the evaluated attribute R_l, i.e.,

$$G_{vkl}^R = \int_{\Omega_{vkl}^R} [F_{vl}^R(x) - F_{kl}^R(x)]dx \tag{21}$$

where $\Omega_{vkl}^R = \{x \mid F_{kl}^R(x) < F_{vl}^R(x), x \in [a_{vkl}^{R*}, b_{vkl}^{R*}]\}$, $a_{vkl}^{R*} = min\{Sc(q_{vl}^R), Sc(q_{kl}^R)\}$, $b_{vkl}^{R*} = max\{Sc(q_{vl}^R), Sc(q_{kl}^R)\}$.

Correspondingly, the loss of similar historical case S_v over case S_k for evaluated attribute R_l is calculated as follows:

$$L_{vkl}^R = -\int_{\Theta_{vkl}^R} [F_{kl}^R(x) - F_{vl}^R(x)]dx \tag{22}$$

where $\Theta_{vkl}^R = \{x \mid F_{vl}^R(x) < F_{kl}^R(x), x \in [a_{vkl}^{R*}, b_{vkl}^{R*}]\}$.

Step 5: Normalize G_{vkl}^R and L_{vkl}^R as follows:

$$\begin{aligned} Y_{vkl} &= \frac{G_{vkl}^R - \min_l\{G_{vkl}^R\}}{\max_l\{G_{vkl}^R\} - \min_l\{G_{vkl}^R\}} \\ Z_{vkl} &= \frac{L_{vkl}^R - \max_l\{L_{vkl}^R\}}{\max_l\{L_{vkl}^R\} - \min_l\{L_{vkl}^R\}} \end{aligned} \tag{23}$$

Step 6: Determine the evaluated attribute weights according to (9).

Step 7: Calculate the dominance degree φ_{vkl}

$$\varphi_{vkl} = \sqrt{\frac{v_l^R Y_{vkl}}{w_r^R \sum_{l=1}^{L} (w_l^R / w_r^R)}} - \frac{1}{\theta} \sqrt{\frac{-Z_{vkl}}{w_l^R} \sum_{l=1}^{L} (w_l^R / w_r^R)} \tag{24}$$

where $w_r^R = \max_l\{w_l^R\}$, and θ denotes the degree of loss aversion of the decision maker, $\theta > 0$.

Step 8: Calculate the comprehensive dominance degree φ_{vk}.

$$\varphi_{vk} = \sum_{l=1}^{L} \varphi_{vkl} \tag{25}$$

Step 9: Calculate the overall dominance degree $\xi(S_v)$

$$\xi(S_v) = \frac{\sum_{k=1}^{h} \varphi_{vk} - \min_v\{\sum_{k=1}^{h} \varphi_{vk}\}}{\max_v\{\sum_{k=1}^{h} \varphi_{vk}\} - \min_v\{\sum_{k=1}^{h} \varphi_{vk}\}} \tag{26}$$

The larger $\xi(S_v)$ is, the higher the evaluated utility is, indicating that the alternative of historical case S_v is better applied to the target case.

5.5 Generate the alternative

To generate the emergency alternative, it is necessary to refer to the alternative of the most suitable historical case, whereas the selection of the most suitable historical case needs to consider the case similarity and the evaluated utility. Based on this, the comprehensive utility is determined through product theory. Let U_v denote the comprehensive utility of the similar historical case S_v. The calculation is defined as follows:

$$U_v = \alpha \times Sim(C_0, S_v) + \beta \times \xi(S_v) \tag{27}$$

where $\alpha, \beta \in [0, 1]$, and $\alpha + \beta = 1$. Here, $U_v \in [0, 1]$ and the larger U_v is, the more suitable the alternative of the similar historical case S_v is for the target case.

According to the ranking of U_v, the most suitable historical case is selected, and the decision maker generates the alternative of the target case according to the alternative of this case.

In summary, the steps of the case similarity measurement method are as follows (see Fig. 2):

Step 1: Calculate the heterogeneous attribute similarity using (1)-(6), which include crisp numbers, interval numbers, IFNs, SvNNs, and IvNNs.

Step 2: Determine the problem attribute weights using (9), and calculated the correlation coefficient between attribute X_j and X_f using (14)-(16) to make sure the consistency and to reduce the redundancy of heterogeneous data.

Step 3: Calculate the case similarity $Sim(C_0, C_i)$ using (17).

Step 4: Construct the set of the similar historical cases $S_v, v = \{1, 2, \cdots, h\}$.

Step 5: Calculate the evaluated utility $\xi(s_v)$ using (18)-(26).

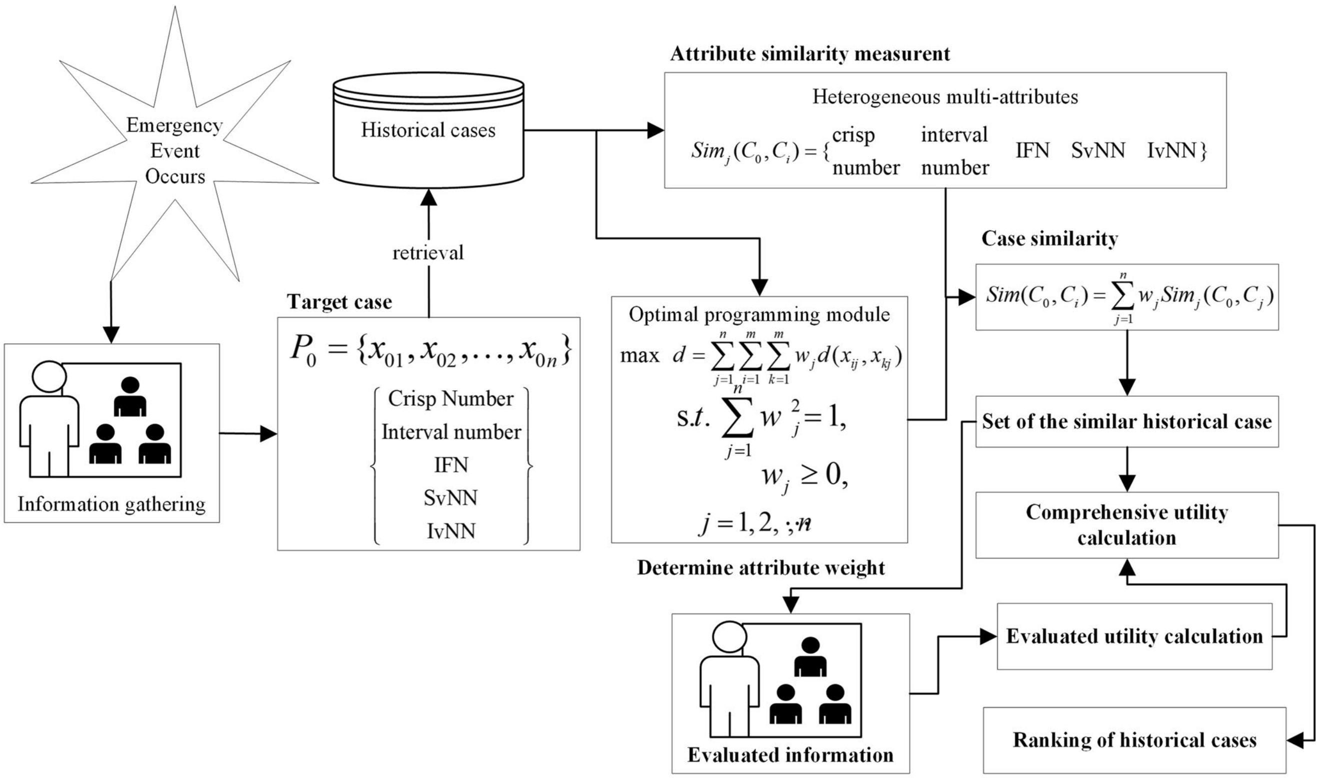

Fig. 2 Procedure of proposed method

Step 6: Calculate the comprehensive utility U_v using (27).

6 Case study

6.1 Problem description and analysis process

The rapid development of manufacturing economies and the increasing demand for energy have resulted in coal businesses being active on a regular basis. Therefore, methods for dealing with coal mine emergencies have attracted significant attention from both researchers and society at large.

In recent years, gas explosions have been one of the major types of coal mine emergencies in China, and the government and coal mining companies have attached great importance to finding ways to deal with these emergencies. Fortunately, gas explosion accidents are generally similar to one another. Because of this characteristic, a similar alternative can be chosen for new emergencies based on similar gas explosion accidents from the past. This means that coal mining companies can draw on historical cases to deal with new gas explosion accidents.

Company A is a coal company in China's Fujian province that refers to the emergency alternatives of historical cases to generate an emergency alternatives for contemporary gas explosions. In particular, Company A has collected ten historical cases $(C_1, C_2, \cdots, C_{10})$ and when a new gas explosion C_0 occurs, the company's decision makers generate alternatives based on the historical cases. The problem attributes of these gas explosion cases include eight attributes: the number of underground personnel (X_1 , unit: person), the scope of the explosion (X_2, unit: %), the degree of damage to the ventilation system (X_3), the degree of landslide (X_4), the scope of the fire (X_5), and the residual O_2 (X_6), CO(X_7), and CH_4 concentrations (X_8). Among these, X_1, X_6, X_7, and X_8 are crisp numbers, X_2 is an interval number, X_3 is an IFN, X_4 is an SvNN, and X_5 is an IvNN. The problem attribute values of the historical cases and the target case are shown in Table 1. Here, the objective of this study is to retrieve historical cases according to the target case and find the most similar historical case that can be used to make a decision for the target case.

To obtain the desirable historical case, the proposed method is used. Next, we discuss the steps in detail.

Step 1: Using (1)-(6), calculate the heterogeneous attribute similarities using, i.e., crisp numbers, interval numbers, IFNs, SvNNs, IvNNs. The results of this step are shown in Table 2.

Step 2: Using (9), the problem attribute weights are obtained, and the result is $W = (0.1843, 0.1396, 0.1507, 0.0503, 0.0516, 0.1681, 0.1106, 0.1448)$. Then, calculate the correlation coefficient between attribute X_j and X_k using (14)-(16), the results are shown in Fig. 3

Table 1 Information for historical cases and the target case

Cases	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
C_1	50	[26,30]	$\langle 0.89, 0.06\rangle$	$\langle 0.8, 0.2, 0.1\rangle$	$\langle [0.7, 0.8], [0.1, 0.3], [0.3, 0.5]\rangle$	12	31	7
C_2	72	[24,30]	$\langle 0.8, 0.14\rangle$	$\langle 0.9, 0.1, 0.05\rangle$	$\langle [0.8, 0.9], [0.1, 0.15], [0.2, 0.3]\rangle$	11	27	9
C_3	43	[19,27]	$\langle 0.55, 0.4\rangle$	$\langle 0.7, 0.3, 0.15\rangle$	$\langle [0.85, 0.95], [0.05, 0.1], [0.2, 0.4]\rangle$	28	32	11
C_4	78	[16,30]	$\langle 0.7, 0.25\rangle$	$\langle 0.8, 0.2, 0.15\rangle$	$\langle [0.85, 0.95], [0.05, 0.2], [0.1, 0.2]\rangle$	25	43	13
C_5	75	[32,40]	$\langle 0.8, 0.15\rangle$	$\langle 0.9, 0.1, 0.05\rangle$	$\langle [0.7, 0.9], [0.1, 0.15], [0.2, 0.4]\rangle$	19	20	8
C_6	39	[35,43]	$\langle 0.75, 0.15\rangle$	$\langle 0.75, 0.25, 0.1\rangle$	$\langle [0.75, 0.85], [0.2, 0.3], [0.1, 0.3]\rangle$	27	25	7
C_7	41	[25,30]	$\langle 0.8, 0.1\rangle$	$\langle 0.85, 0.15, 0.05\rangle$	$\langle [0.65, 0.8], [0.3, 0.4], [0.1, 0.2]\rangle$	30	27	15
C_8	35	[13,20]	$\langle 0.85, 0.1\rangle$	$\langle 0.95, 0.05, 0.1\rangle$	$\langle [0.8, 0.9], [0.1, 0.2], [0.2, 0.3]\rangle$	33	32	14
C_9	62	[15,23]	$\langle 0.85, 0.05\rangle$	$\langle 0.65, 0.35, 0.05\rangle$	$\langle [0.75, 0.9], [0.2, 0.3], [0.2, 0.4]\rangle$	28	35	10
C_{10}	68	[22,38]	$\langle 0.5, 0.45\rangle$	$\langle 0.75, 0.25, 0.15\rangle$	$\langle [0.75, 0.85], [0.2, 0.3], [0.3, 0.4]\rangle$	26	33	12
C_0	58	[20,31]	$\langle 0.7, 0.25\rangle$	$\langle 0.8, 0.2, 0.15\rangle$	$\langle [0.8, 0.9], [0.15, 0.25], [0.2, 0.4]\rangle$	18	29	13

It can be seen from Fig. 3 that most of the correlation coefficients are greater than 0.7, which indicates that there is a storing correlation between problem attributes. Further statistical analysis shows that the value of 79.7% results is greater than 0.7, the value of 96.9% results is greater than 0.4. It can be conclude that there are strong correlation between problem attributes, which also means there is a strong consistency among these heterogeneous information. Meanwhile, due to strong correlation between problem attributes and different data values, it can be explained that the data presented in this paper has less redundancy, and we can further apply these heterogeneous data.

Step 3: Using (17), the case similarities are obtained, i.e., $Sim(C_0, C_1) = 0.6478$, $Sim(C_0, C_2) = 0.6813$, $Sim(C_0, C_3) = 0.6632$, $Sim(C_0, C_4) = 0.7224$, $Sim(C_0, C_5) = 0.6245$, $Sim(C_0, C_6) = 0.5707$, $Sim(C_0, C_7) = 0.6505$, $Sim(C_0, C_8) = 0.5825$, $Sim(C_0, C_9) = 0.6506$, $Sim(C_0, C_{10}) = 0.6719$.

Step 4: The decision makers set the case similarity threshold at $\xi = 0.67$, then the similar historical case set can be attained $\{S_1, S_2, S_3\} = \{C_2, C_4, C_{10}\}$.

Step 5: The decision makers provide the evaluation information for the similar historical case set, which is presented in Table 3. The evaluated attributes were the reduction of casualties (R_1), the reduction of property loss (R_2), and the response time (R_3), and R_1 and R_2 were presented by SvNN and IvNN respectively, R_3 was expressed by linguistic variable, whose linguistic set is $S = \{VF : veryfast, F : fast, G : general, S : slow, VS : veryslow\}$. Then, according to Definition 4, the evaluated information was transformed into crisp numbers and the evaluated attribute weights were determined as $W^R = \{0.4121, 0.2747, 0.3132\}$ using (13). Furthermore, the evaluated utilities $\xi(S_v)$ were calculated using (18)-(26), and the following results are obtained:

$\xi(S_1) = 1, \xi(S_2) = 0.7113, \xi(S_3) = 0.$

Step 6: The comprehensive utilities U_v were calculated using (13), where the parameters were set as $\alpha = \beta = 0.5$, and the following results were obtained:

$U_1 = 0.8407, U_2 = 0.7169, U_3 = 0.3360.$

The greater the comprehensive utility, the more suitable the historical case is to the target case. We retrieved the most

Table 2 Problem attribute similarity

Cases	$Sim_1(C_0, C_i)$	$Sim_2(C_0, C_i)$	$Sim_3(C_0, C_i)$	$Sim_4(C_0, C_i)$	$Sim_5(C_0, C_i)$	$Sim_6(C_0, C_i)$	$Sim_7(C_0, C_i)$	$Sim_8(C_0, C_i)$
C_1	0.7062	0.7286	0.4300	0.9286	0.8261	0.6703	0.8669	0.3679
C_2	0.5441	0.8068	0.6900	0.8333	0.9078	0.6271	0.8669	0.5134
C_3	0.5209	0.8068	0.5500	0.8750	0.9363	0.5134	0.8071	0.7165
C_4	0.4191	0.8068	1.0000	1.0000	0.7881	0.6271	0.3679	1.0000
C_5	0.4775	0.4580	0.7000	0.8333	0.9078	0.9355	0.5258	0.4346
C_6	0.4378	0.3679	0.8000	0.9091	0.8478	0.5488	0.7515	0.3679
C_7	0.4775	0.7669	0.6500	0.8478	0.7700	0.4493	0.8669	0.7165
C_8	0.3679	0.5072	0.5500	0.8077	0.9054	0.3679	0.8071	0.8465
C_9	0.8404	0.6119	0.5000	0.7885	0.9511	0.5134	0.6514	0.6065
C_{10}	0.6474	0.6846	0.4000	0.9318	0.8831	0.5866	0.7515	0.8465

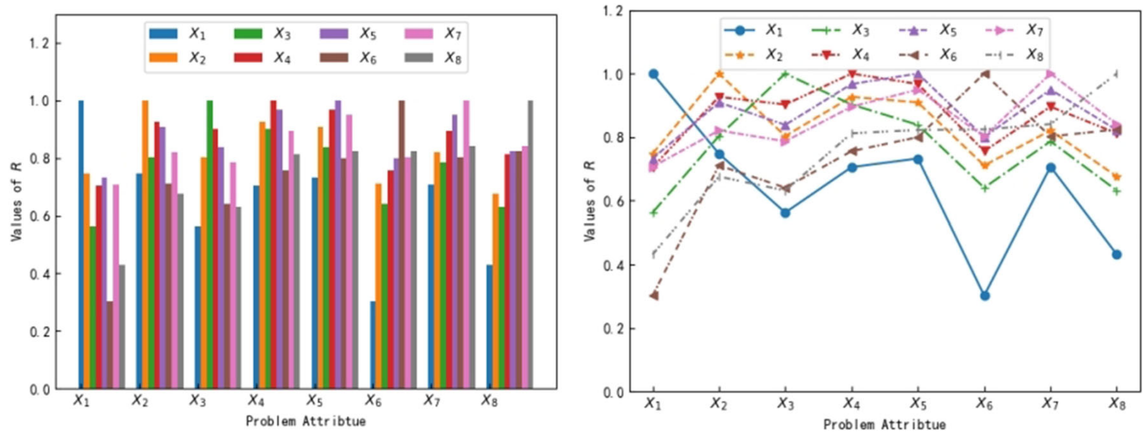

Fig. 3 The correlation coefficient between attribute X_j and X_k

suitable historical cases based on the comprehensive utility, and the result obtained was C_2. Therefore, the decision makers can refer to this historical case (C_2) to generate an alternative for the target case.

6.2 Comparative analysis and discussion

6.2.1 Comparison with the similarity measurement of neutrosophic numbers

To illustrate the characteristics of our method, as well as the effectiveness of the proposed similarity measurement of neutrosophic number, we compare some existing methods with the proposed approach.

Initially, we used the Euclidean distance according to SvNN [42] to calculate the attribute similarities of attribute X_4; the results and rank are shown in Table 4.

From Table 4, we can infer that the calculation results of the two methods are basically the same, except for the historical cases of C_2 and C_4, and the most similar historical case is C_{10}. Further statistical analysis shows that four-fifths (8/10) results from the proposed method are exactly the same as those from Euclidean distance. It indicates that the proposed method is consistent with the method of Euclidean distance. We believe that this happened because each part of an SvNN is a crisp number, the attribute similarities and the ranking of the two methods are similar. Therefore, the proposed method is effective.

Table 3 Evaluated information for the similar historical cases

	R_1	R_2	R_3
S_1	$\langle 0.7, 0.3, 0.2\rangle$	$\langle [0.7, 0.8], [0.15, 0.4], [0.2, 0.3]\rangle$	VF
S_2	$\langle 0.6, 0.4, 0.1\rangle$	$\langle [0.8, 0.9], [0.2, 0.3], [0.15, 0.3]\rangle$	G
S_3	$\langle 0.75, 0.25, 0.2\rangle$	$\langle [0.65, 0.8], [0.25, 0.4], [0.1, 0.2]\rangle$	S

Subsequently, we harnessed the similarity for IvNNs [43] to calculate the attribute similarity of attribute X_5. In a previous study by Ye and Du [43], there were two similarity measurements, which the authors called Yu-1 and Yu-2. Furthermore, we applied similarity for IvNNs based on Euclidean distance [44] to calculate the attribute similarity of attribute X_5. The results are shown in Table 5.

Table 5 demonstrates that there are differences among the calculation results of the four methods; however, the sort ranks are closer to one another. Further statistical analysis shows that RMSEs between the proposed method and Euclidean distance, Yu-1 and Yu-2 are 0.7746, 1.3784, and 1.3784 respectively, which are at low levels. It indicates that the proposed method is effective. In the process of the similarity calculation, the IvNN needs to be converted into a crisp number; consequently, there is a certain amount of information loss. Nevertheless, the method proposed in this paper always takes the uncertainty of the IvNNs into account in the calculation process and calculates the

Table 4 Comparison of the attribute similarities of SvNNs

	Euclidean distance		This paper	
Cases	$Sim_4(C_0, C_i)$	rank	$Sim_4(C_0, C_i)$	rank
C_1	0.8152	7	0.7321	7
C_2	0.9000	5	0.8333	4
C_3	0.9184	3	0.8750	3
C_4	0.9043	4	0.7917	5
C_5	0.6429	10	0.5286	10
C_6	0.9500	2	0.9091	2
C_7	0.6621	9	0.6250	9
C_8	0.7821	8	0.6724	8
C_9	0.8646	6	0.7885	6
C_{10}	0.9592	1	0.9318	1

Table 5 Comparison of the attribute similarities of IvNNs

	Euclidean distance		This paper		Yu-1		Yu-2	
Cases	$Sim_5(C_0, C_i)$	rank	$Sim_5(C_0, C_i)$	rank	$Sim_5(C_0, C_i)$	rank	$Sim_5(C_0, C_i)$	rank
C_1	0.7602	8	0.6154	9	0.9083	8	0.8000	10
C_2	0.9388	2	0.9545	1	0.9667	1	0.9625	3
C_3	0.8810	6	0.8261	6	0.9583	3	0.9000	7
C_4	0.8939	4	0.8298	4	0.9417	5	0.9250	4
C_5	0.8901	5	0.8298	5	0.9333	6	0.9250	4
C_6	0.7307	9	0.5273	10	0.8750	9	0.8250	9
C_7	0.8542	7	0.7864	7	0.9167	7	0.9250	4
C_8	0.9355	3	0.9091	2	0.9583	3	0.9750	1
C_9	0.7307	9	0.6393	8	0.8583	10	0.8500	8
C_{10}	0.9423	1	0.9091	2	0.9667	1	0.9750	1

distance between two attributes in the form of the IvNNs. This practice more effectively represents the uncertainty of the attributes.

6.2.2 Comparison with some existing case similarity measurements

In this section, we compare the proposed method with the existing case retrieval methods for heterogeneous case information [7, 16]. Because neither the IFNs nor the neutrosophic numbers were considered in the methods of Fan et al. [7], the IFNs here were converted into crisp numbers by a score function of IFNs [36], and the attribute distances of the neutrosophic numbers were calculated by Euclidean distance. Subsequently, the attribute and case similarities were calculated using the method described by Fan et al. [7]. The results of this comparison are shown in Table 5. Because Zheng et al.'s method [16] does not account for IFNs, we convert them into interval numbers using a score function of the IFNs [45]. The results are shown in Table 6.

From Table 6, we can deduce that the results presented in this paper are similar to that of Fan et al. [7] but quite different from that of Zheng et al. [16]. The similarity ranking obtained from the proposed method is close to those from the other methods. Further statistical analysis shows that RMSEs between the proposed method and Fan et al. [7], Zheng et al. [16] are 1.1831, and 1.3416 respectively, which are at low levels. With respect to the process of attribute similarity calculation, our method more closely resembles that of Fan et al. [7] in regards to the attributes of the crisp and interval numbers, except for the addition of IFNs and neutrosophic numbers. Zheng et al.'s approach [16] differs from our method, but not significantly. Therefore, our method is both feasible and effective. Furthermore, our proposed method enlarges the use rang for the EDM.

6.2.3 Comparison with some existing heterogeneous multi-attribute emergency decision-making methods

In this section, we compare the proposed method with several existing heterogeneous multi-attribute EDM methods [19–21]. The comparison is mainly conducted with respect to three aspects: heterogeneous attribute formats, decision-makers' psychology, and the pattern of identifying the results (see Table 7).

From Table 7, we conclude the following:

(1) We proposed a heterogeneous multi-attribute case retrieval method, which considers five types of problem attribute values, and can represent the information in complicated situations. Furthermore, the evaluated information is also represented by heterogeneous multi-attributes. In particular, SvNNs and IvNNs can handle incomplete, indeterminate, and inconsistent problems, whereas IFNs, linguistic variables

Table 6 Comparison of case similarity measurements

Proposed paper		Fan et al. [7]		Zheng et al. [16]	
Case similarity	rank	Case similarity	rank	Case similarity	rank
0.6347	7	0.5980	7	0.5054	7
0.7086	1	0.6749	1	0.6447	3
0.6852	2	0.6518	2	0.8175	1
0.6661	5	0.6369	4	0.5823	6
0.6032	10	0.5639	10	0.4090	10
0.6104	9	0.5750	8	0.4228	9
0.6247	8	0.5695	9	0.4451	8
0.6561	6	0.6265	5	0.61674	4
0.6687	3	0.6193	6	0.6020	5
0.6679	4	0.6496	3	0.6609	2

Table 7 Comparison with current heterogeneous multi-attribute EDM methods

Literature	Attribute formats	Decision-makers' psychology behavior	Pattern of identifying the results
Ref. [19]	Crisp numbers, interval numbers, linguistic variables	Completely rational	Sorting the values of the comprehensive optimal membership degrees
Ref. [20]	Real numbers, interval grey numbers, extended grey numbers	Completely rational	Sorting the case similarities
Ref. [21]	Intuitionistic fuzzy sets, triangular intuitionistic fuzzy numbers, trapezoidal intuitionistic fuzzy numbers, intervals, real numbers	Decision makers' preference	Sorting the adjusted normalized group priorities
Proposed method	Crisp numbers, interval numbers, intuitionistic fuzzy numbers, single-valued neutrosophic umbers, interval-valued neutrosophic numbers	Bounded rationality	Sorting the combination of case similarities and evaluation utilities

and others cannot. Our proposed method expands the application range of the heterogeneous multi-attribute retrieval method, and the effectiveness of the improved method is demonstrated by an example.

(2) The proposed method considers the decision-makers' psychological behavior and applies TODIM to describe the behavior of decision makers, which is consistent with real world situations. Previous research [19, 20] argued that decision makers are completely rational; however, other researchers [21] provided the preference information of decision makers for the alternative choices but did not express the psychological behavior during the decision-making process.The proposed method is more consistent with people's actual decision-making behavior, and the decision-making results are more reasonable and comprehensive.

(3) The proposed method comprehensively considers the case similarity and the evaluation utility, which makes retrieving the results more efficient. In some previous research [19], the most suitable historical case is obtained mainly according to the case similarity, whereas in other research [20, 21], the most suitable alternative is determined by the evaluation utility. The proposed method can make the retrieval results more effective by considering both objective data retrieval results and evaluation utilities of decision makers.

6.2.4 Sensitivity analysis of parameter in comprehensive utility

To illustrate the stability of selecting the most suitable historical case, sensitivity analysis was conducted in this paper on parameters α and β in (26). The results are shown in Fig. 4.

It can be seen from Fig. 4 that with the change of parameters (α, β) from (0.1, 0.9) to (0.9,0.1), the historical case S_2 ranks highest only at point (0.1,0.9), while historical case S_1 ranks highest in all the other points, therefore, the historical case S_1 was selected as a reference for generating an emergency alternative. Meanwhile, it indicates that the parameters have little influence on the results. This suggests that the selection of the most suitable historical case is stable and compatible.

6.2.5 Advantages of the proposed method

Based on the above analysis, the advantages of the proposed method can be summarized as follows:

(1) The proposed method considers heterogeneous case information, which involves crisp numbers, interval numbers, IFNs, IvNNs, and SvNNs and enlarges the application scope of case retrieval. Furthermore, the evaluated information is also expressed by heterogeneous multi-attributes. This is more consistent with the representation of data in real life.

(2) The problem attribute weights are determined objectively, and different historical cases with the same attribute have different problem attribute weights. Therefore, the results of case similarities are more objective and accurate than those obtained by existing methods. Moreover, the calculation process for problem attribute weights is relatively simple.

(3) The case retrieval method considers the evaluated information from similar historical cases to improve the effectiveness of emergency responses. The evaluated information is expressed as heterogeneous multi-attribute, and the decision-making process considers the decision-makers' psychological behaviors,

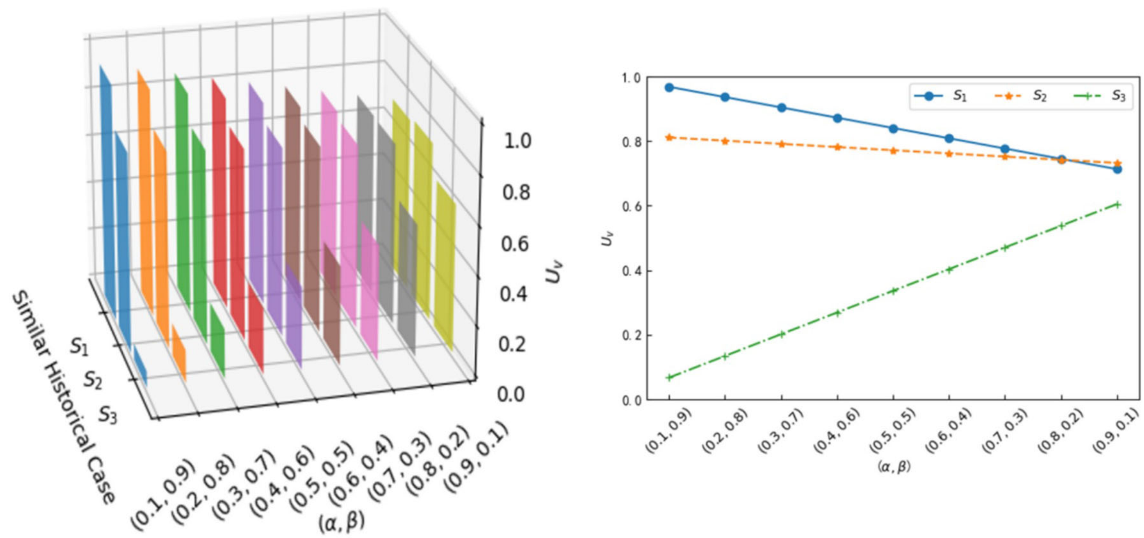

Fig. 4 Comprehensive utilities of the similar historical case for different parameters α and β

which makes the result representative of real-world scenarios.

7 Conclusions

Information about emergencies is becoming increasingly complex, and some emergencies exist as incomplete, indeterminate, and inconsistent problems. To address this issue, we proposed a heterogeneous multi-attribute case retrieval method based on neutrosophic sets and TODIM for emergency situations. A novel similarity measurement considering five formats attributes was proposed, which enlarges the application scope of heterogeneous multi-attribute case retrieval. The attribute weights are determined by an optimal model. Then, the case similarities between the target case and historical cases are calculated, and the set of similar historical cases is constructed. Based on this, the evaluated information based on heterogeneous multi-attribute is provided. The evaluated method based on TODIM was proposed to calculate the evaluated utilities. The most suitable historical case was selected, based on the case similarities and evaluated information. The results of a comparative application and analysis showed that the proposed method is feasible, effective, and practical.

The proposed method can provide decision makers in complex environments with a reference that they can use to effectively retrieve similar historical cases. However, there are several limitations pertaining to the proposed approach. For example, the case base is relatively limited; however, the development of the Internet of Things can offer more detailed case information about emergencies that exist in big data [46, 47]. Therefore, future research can focus on the case retrieval of big data based on CBR, in which the case information is uncertain and very large.

Acknowledgements This work was partly supported by the National Natural Science Foundation of China (No. 71371053), the Humanities and Social Sciences Foundation of the Chinese Ministry of Education (No. 20YJC630229), and the Natural Science Foundation of Fujian Province, China (Nos. 2021J01339, 2021J011223).

Declarations

Conflict of Interests The authors declare that they have no conflict of interest regarding the publication of this paper.

References

1. Liu B, Zhao X, Li Y (2016) Review and prospect of studies on emergency management. Procedia Engineering 145:1501–1508
2. Wang L, Zhang ZX, Wang YM (2015) A prospect theory-based interval dynamic reference point method for emergency decision making. Exp Syst Appl 42(23):9379–9388
3. Ding XF, Liu HC, Shi H (2019) A dynamic approach for emergency decision making based on prospect theory with interval-valued Pythagorean fuzzy linguistic variables. Comput Indust Eng 131:57–65
4. Gao J, Xu Z, Liang Z, Liao H (2019) Expected consistency-based emergency decision making with incomplete probabilistic linguistic preference relations. Knowl-Based Syst 176:15–28
5. Xu X, Yin X, Chen X (2019) A large-group emergency risk decision method based on data mining of public attribute preferences. Knowl Based Syst 163:495–509
6. Wang D, Wan K, Ma W (2020) Emergency decision-making model of environmental emergencies based on case-based reasoning method. J Environ Manag 262:1–10
7. Fan Z, Li Y, Wang X, Liu Y (2014) Hybrid similarity measure for case retrieval in CBR and its application to emergency response towards gas explosion. Exp Syst Appl 41(5):2526–2534

8. Fan ZP, Li YH, Zhang Y (2015) Generating project risk response strategies based on CBR: a case study. Exp Syst Appl 42(6):2870–2883
9. Yu F, Li XY (2018) Improving emergency response to cascading disasters: Applying case-based reasoning towards urban critical infrastructure. Int J Disas Reduct 30:244–256
10. Yu F, Li XY, Han XS (2018) Risk response for urban water supply network using case-based reasoning during a natural disaster. Saf Sci 106:121–139
11. Yu XB, Li CL, Zhao WX, Chen H (2020) A novel case adaptation method based on different evolution algorithm for disaster emergency. Appl Soft Comput 92:1–12
12. Zheng J, Wang YM, Chen L, Zhang K (2019) A new case retrieval method based on double frontiers data envelopment analysis. J Intell Fuzzy Syst 36:199–211
13. Kwon N, Cho J, Lee HS et al (2019) Compensation cost estimation model for construction noise claims using case-based reasoning. J Constr Eng Manag 145(8):1–15
14. Lin MW, Chen ZY, Xu ZS, Gou XJ, Herrear F (2021) Score function based on concentration degree for probabilistic linguistic term sets: an application to topsis and vikor. Inf Sci 551:270–290
15. Gou X, Xu Z, Liao H et al (2020) Probabilistic double hierarchy linguistic term set and its use in designing an improved vikor method: the application in smart healthcare. J Oper Res Soc 1:1–20
16. Zheng J, Wang YM, Lin Y et al (2019) Hybrid multi-attribute case retrieval method based on intuitionistic fuzzy and evidence reasoning. J Intell Fuzzy Syst 36(12):1–12
17. Ding XF, Liu HC (2019) A new approach for emergency decision-making based on zero-sum game with Pythagorean fuzzy uncertain linguistic variables. Int J Intell Syst 34(7):1667–1684
18. Yu GF, Li DF, Fei W (2018) A novel method for heterogeneous multi-attribute group decision making with preference deviation. Comput Indust Eng 124:58–64
19. Chen W, Wang X, Wang W et al (2021) A heterogeneous gra-cbr-based multi-attribute emergency decision-making model considering weight optimization with dual information correlation. Expert Syst Appl 182:1–13
20. Sun H, Dang Y, Mao W (2018) A decision-making method with grey multi-source heterogeneous data and its application in green supplier selection. Int J Environ Res Public Health 15(3):1–24
21. Dong J, Wan S (2018) A promethee-flp method for heterogeneous multi-attributes group decision making. IEEE Access 6:1–12
22. Zheng J, Wang YM, Zhang K (2020) Solution of heterogeneous multi-attribute case-based decision making problems by using method based on TODIM. Soft Comput 24:7081–7091
23. Zhang K, Wang YM, Zheng J (2021) Regret theory-based case-retrieval method with multiple heterogeneous attributes and incomplete weight information. Int J Comput Intell Syst 14(1):1022–1033
24. Wang D, Wan K, Ma W (2020) Emergency decision-making model of environmental emergencies based on case-based reasoning method. J Environ Manag 262:1–10
25. Jin Q, Jie H, Peng YH et al (2009) A case retrieval method combined with similarity measurement and multi-criteria decision making for concurrent design. Expert Syst Appl 36(7):10357–10366
26. Tan RP, Zhang WD (2020) Decision-making method based on new entropy and refined single-valued neutrosophic sets and its application in typhoon disaster assessment. Appl Intell 51(4):283–307
27. Smarandache F (1998) Neutrosophy neutrosophic probability: set, and logic. American Research Press, Rehoboth, pp 12–20
28. Bolturk E, Kahraman C (2018) A novel interval-valued neutrosophic AHP with cosine similarity measure. Soft Comput 22:4941–4958
29. Wang N, Meng F, Xu Y (2018) Deriving the priority weights from multiplicative consistent single-valued neutrosophic preference relations. Neural Comput Appl 31:6659–6683
30. Altun F, Sahin R, Guler G (2020) Multi-criteria decision making approach based on PROMETHEE with probabilistic simplified neutrosophic sets. Soft Comput 24:4889–4915
31. Zhang C, Li D, Kang X et al (2020) Multi-Attribute Group decision making based on multigranulation probabilistic models with Interval-Valued neutrosophic information. Mathematics 223:1–21
32. Ding Q, Wang YM, Goh M (2020) An extended todim approach for group emergency decision making based on bidirectional projection with hesitant triangular fuzzy sets. Comput Indust Eng 151(2):1–10
33. Li MY, Cao PP (2018) Extended todim method for multi-attribute risk decision making problems in emergency response. Comput Indust Eng 135(SEP.):1286–1293
34. Zhang X (2016) Multicriteria Pythagorean fuzzy decision analysis: A hierarchical QUALIFLEX approach with the closeness index-based ranking methods. Inf Sci 330:104–124
35. Du Y, Hou F, Zafar W, Yu Q (2017) A novel method for multiattribute decision making with Interval-Valued pythagorean fuzzy linguistic information. Int J Intell Syst 32:1085–1112
36. Atanassov K (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20:87–96
37. Xu ZS (2007) Intuitionistic fuzzy aggregation operators. IEEE Trans Fuzzy Syst 15(6):1179–1187
38. Zhao K, Yu X (2011) A case based reasoning approach on supplier selection in petroleum enterprises. Expert Syst Appl 38(6):6839–6847
39. Yan A, Shao H, Guo Z (2014) Weight optimization for case-based reasoning using membrane computing. Inf Sci 287:109–120
40. Wu D, Li J, Bao C (2018) Case-based reasoning with optimized weight derived by particle swarm optimization for software effort estimation. Soft Comput 22:5299–5310
41. Wang YM, Fan ZP (2007) Fuzzy preference relations: Aggregation and weight determination. Computers & Industrial Engineering 53(1):163–172
42. Biswas P, Pramanik S, Giri BC (2016) TOPSIS Method for multi-attribute group decision-making under single-valued neutrosophic environment. Neural Comput & Applic 27(3):727–737
43. Ye J, Du S (2017) Some distances, similarity and entropy measures for interval valued neutrosophic sets and their relationship. International Journal of Machine Learning & Cybernetics 10(3):1–10
44. Ye J (2014) Clustering Methods Using Distance-Based Similarity Measures of Single-Valued Neutrosophic Sets. Journal of Intelligent Systems 23(4):379–389
45. Zhang HY, Wang JQ, Chen XH (2014) Interval neutrosophic sets and their application in multicriteria decision-making problems. The Scientific World Journal 3:1–15
46. Guo Y, Wu K (2018) Research on case retrieval of Bayesian network under big data. Data & Knowledge Engineering 118(11):1–13
47. Guo Y, Chen W, Zhu YX, Guo YQ (2019) Research on the integrated system of case-based reasoning and Bayesian network. Isa Transactions 90:213–225

Kai Zhang is an professor at Fujian Chuanzheng Communications College. He has published approximately 20 papers and presided over ten scientific research projects. His current research interests include decision making theory and application and fuzzy information processing.

Jing Zheng received the Ph.D. degree at Fuzhou University. She is a Professor at at Fujian Jiangxia University. She has published over 20 papers and presided over eight scientific research projects . Her current research interests include decision theory and method.

Ying-Ming Wang received the M.Sc. degree in systems engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1987 and the Ph.D. degree in automatic control theory and application from Southeast University, Nanjing, China, in 1991. He is currently a Yangtze River scholar distinguished professor of Fuzhou University in China. He has published over 145 SCI and 39 SSCI indexed journal papers and has been one of the most cited Chinese researchers since 2014. His current research interests include multiple criteria decision analysis, data envelopment analysis, rule based inference, and quality function deployment.