Complex & Intelligent Systems (2022) 8:5011–5026
https://doi.org/10.1007/s40747-022-00723-8

ORIGINAL ARTICLE

A novel method to estimate incomplete PLTS information based on knowledge-match degree with reliability and its application in LSGDM problem

Huimin Xiao[1] · Shouwen Wu[1] · Liu Wang[1]

Received: 27 October 2021 / Accepted: 5 March 2022 / Published online: 30 April 2022

Abstract

In recent years, large-scale group decision making (LSGDM) has been researched in various fields. Probabilistic linguistic term set (PLTS) is an useful tool to describe evaluation information of experts when solving the LSGDM problem. As decision-making becomes more complex, in most cases, decision makers are unable to give complete evaluations over alternatives, which leads to the lack of evaluation information. To estimate missing information, this paper proposes a new method based on knowledge-match degree with reliability that knowledge-match degree means the matching level between evaluation values provided by individual and ones from group. The possession of reliability associated with evaluation information depends on fuzzy entropy of PLTS. Compared with previous methods, this approach can enhance accuracy and reliability of estimated values of missing evaluation information. Based on this method, we develop a complete decision process of LSGDM including information collection, subgroup detecting, consensus reaching process (CRP), information aggregation and ranking alternatives. Subsequently, a case about pharmaceutical manufacturer selection is used to illustrate the proposed decision method. To verify effectiveness and superiority, we make a comparative analysis with other methods and finally draw a conclusion.

Keywords Large-scale group decision making · Probabilistic linguistic term set · Knowledge-match degree · Fuzzy entropy

Introduction

With the rapid development of society and technology, decision problem becomes increasingly complicated which contains usually much uncertain information. A decision problem may involve various fields. For instance, when a company makes a selection of investors, it will consider many factors of investor such as financial capacity, credit risks, industry familiarity and so on. In traditional group decision making, a small number of experts are hard to give comprehensive evaluations for schemes. Thus, large-scale group decision making (LSGDM) is introduced, in which a large number of decision makers (DMs) who come from distinct fields participate the process of decision making to ensure accuracy and efficiency of decision results. Normally, there are more than 20 (contain 20) DMs in LSGDM. Large-scale group decision making has attracted much attention of researchers [1, 2] and it is applied to many aspects, e.g., high-speed rail system [3], earthquake shelter selection [4], and healthcare service [5].

Due to the cognitive complexity of human, when conducting decision making, it is more appropriate to use fuzzy relations to depict people's evaluations [6]. Besides, compared with numeral rating, DMs tend to adopt linguistic labels because of its directness and simplicity. In fuzzy decision making, Zadeh [7] introduced the concept of fuzzy sets (FSs) and Torra [8] proposed hesitant fuzzy sets (HFSs) which well express the hesitancy and fuzziness of DMs. Cui et al. [9] utilized the HFSs to measure product similarity. Based on linguistic variables and HFSs, Rodriguez et al. [10] developed the theory of hesitant fuzzy linguistic term sets (HFLTSs) so that DMs can use a linguistic term set (LTS) to describe their preferences. On the basis of HFLTS, Chen et al. [11] proposed the proportional HFLTSs (PHFLTSs) and explored the aggregation of HFLTS possibility based on K-means clustering [12]. Nevertheless, the defect of HFLTSs is that the

✉ Shouwen Wu
wswlvs@163.com

[1] School of Computer and Communication Engineering, Henan University of Economics and Law, Zhengzhou 450046, China

weights or importance of terms are viewed as identical ones. To overcome the shortcoming, Pang et al. [13] developed the concept of probabilistic linguistic term sets (PLTSs) which reflect flexibly ambiguous information of DMs and importance distribution of LTs. PLTS is an useful tool in decision making, especially in LSGDM, because experts can utilize it to give more accurate and actual evaluations. In multi-attribute LSGDM problem, the decision matrix with PLTS is often used to represent evaluation information of DMs with respect to alternatives and attributes. However, in most cases, owing to limited knowledge and time urgency, experts are not able to provide a complete decision matrix. Thus, before conducting decision making, the incomplete matrix should be filled.

Many scholars have proposed various methods to estimate the missing information in the incomplete matrix and we have a brief review below. From the respective of estimation strategies, Ureña et al. [14] summarized as individual strategy and social strategy that the former only considers information of incomplete decision matrix while the latter needs to refer to the decision information of other experts. Based on additive consistency and intermediate value, Herrera-Viedma et al. [15] proposed an iterative process to estimate the missing preference values, which avoids the contradiction among preferences. Some researchers filled incomplete decision information according to consistency level of group [16]. Meng et al. [17] utilized the goal programming model to estimate missing values of decision matrix under fuzzy preference relation. Capuano et al. [18] developed an iteration process to solve the problem of incomplete evaluation information based on additive consistency and transitivity. Guo et al. [19] proposed the least deviation model to estimate missing preference, which achieved the maximum confidence under the mechanism of consistency and the range of missing values. Liang et al. [20] estimated incomplete information by means of collaborative filtering algorithm. Although various methods have been proposed to estimate incomplete decision matrix, few researches take the cognitive level of experts into account, which is an important factor to ensure the accuracy of estimated values. When estimating the missing information, we usually refer to the related information of other experts. If the referenced experts have low cognitive level, the accuracy of their evaluation information will be also poor. Thus, to solve this problem, a novel estimation method is proposed based on knowledge-math degree with reliability in this paper.

To apply the proposed method to LSGDM, we develop a decision process including estimating incomplete information, subgroup detecting, consensus reaching process (CRP) and selection process. With respect to subgroup detecting, some methods are proposed such as fuzzy C-means algorithm [23], K-means method [3], the approach based on supporting degree of alternatives [24] and variable

Table 1 Literature review

Literature	Whether involve PLTS?	Estimation method of missing information	Consensus analysis
Reference [14]	No	Summarize the individual strategy and social strategy	–
Reference [15–19]	No	Consider the preference relations or preference consistency	–
Reference [20]	No	Collaborative filtering algorithm	–
Reference [21]	No	–	Two-stage CRP model
Reference [22]	No	–	The CRP model based on conflict degree

grouping method [25]. In this paper, due to the efficiency and superiority of K-means, we adopt K-means method to cluster the subgroups. Besides, consensus reaching process has been the research focus in LSGDM and many CRP models are developed [26–28], e.g., two-stage CRP model [21], the CRP model based on conflict degree [22]. To intuitively analysis previous researches, the Table 1 is given. Due to the large amount of calculation in previous CRP of LSGDM, we construct a new CRP based on social trust relationship, which can also remain the original evaluation information to some extent. The main contributions in this paper are as follows.

(1) The new concepts of knowledge-match degree and PLTS reliability are introduced to measure the importance of referenced experts. On the basis of knowledge-match degree, we integrate the reliability into knowledge-match degree, namely knowledge-match degree with reliability, which is used to measure the cognitive level of experts.
(2) Based on knowledge-match degree with reliability, a novel method is developed to estimate the missing information in incomplete matrix represented by PLTS. To apply the method better, we design a algorithm to realize it.
(3) To solve the LSGDM problem, we construct a LSGDM process in term of the proposed estimation method. In the LSGDM process, a new CRP is constructed based on trust relationship.

The rest contents are arranged as follows. In the next section, some basic knowledge related to PLTS are reviewed including definitions, operations, normalization and so on.

Subsequently, we propose the novel method to estimate incomplete decision matrix with PLTS based on knowledge-match degree with reliability in the third section. In the fourth section, the LSGDM process is developed based on the proposed estimation method and a new CRP is construct based on trust relationship. A case about pharmaceutical manufacturer selection is used to illustrate the proposed decision method in the fifth section. To verify effectiveness and superiority of the proposed methods, we make the comparative analysis with other methods in the sixth section, and finally, draw a conclusion in the last section.

Preliminary

In this section, some basics about PLTS are introduced to prepare for latter theoretical method. The PLTS is proposed by Pang [13], and the comparison rule, the normalized method, and some operations are also given.

PLTS

Definition 1 [13] Suppose that S is a linguistic term set (LTS), $S = \{s_\alpha | \alpha = 0,1,2,\ldots,2\tau\}$. Let $L(p)$ be a PLTS:

$$L(p) = \left\{ L^k(p^k) | L^k \in S,\ p^k \geq 0,\ k = 1, 2, ..., \#L(p), \times \sum_{k=1}^{\#L(p)} p^k \leq 1 \right\},$$

where L^k and p^k denote the kth term and its probability in $L(p)$. $\#L(p)$ is the number of terms in $L(p)$. If $\sum_{k=1}^{\#L(p)} p^k < 1$, the PLTS is incomplete, which should be normalized before calculation or aggregation. According to Pang's method [13], the normalized PLTS is given by

$$\dot{L}(p) = \{L^k(\dot{p}^k) | k = 1, 2, ..., \#L(p)\}, \quad (1)$$

where $\dot{p}^k = p^k / \sum_{k=1}^{\#L(p)} p^k$.

Besides, when calculate two PLTSs, the cardinality of terms should be uniformed. Given any two PLTSs, $L_1(p)$ and $L_2(p)$, if $\#L_1(p) > \#L_2(p)$, add $\#L_1(p)$-$\#L_2(p)$ terms to $L_2(p)$ and the added terms are the least ones in $L_2(p)$. The probabilities of added elements are set 0. If $\#L_2(p) > \#L_1(p)$, it is the same as above.

Pang [13] also proposed some operations of PLTS, which are listed as follows".

(1) $(1)L_1(p) \oplus L_2(p) = \cup_{L_1^k \in L_1(p),\, L_2^k \in L_2(p)} \{p_1^k L_1^k \oplus p_2^k L_2^k\}$.

(2) $L_1(p) \otimes L_2(p) = \cup_{L_1^k \in L_1(p),\, L_2^k \in L_2(p)} \{(L_1^k)^{p_1^k} \otimes (L_2^k)^{p_2^k}\}$.

(3) $\lambda L(p) = \cup_{L^k \in L(p)} \lambda p^k L^k$, $\lambda \geq 0$.

(4) $L(p)^\lambda = \cup_{L^k \in L(p)} \{(L^k)^{\lambda p^k}\}$.

Definition 2 [13] Let $L(p)$ be a PLTS, $L(p) = \{L^k(p^k) | k = 1, 2, ..., \#L(p)\}$, then the score function of $L(p)$ is given by.

$$E(L(p)) = \sum_{k=1}^{\#L(p)} L^k p^k / \sum_{k=1}^{\#L(p)} p^k. \quad (2)$$

Definition 3 [13] Let $L(p)$ be a PLTS, $L(p) = \{L^k(p^k) | k = 1, 2, ..., \#L(p)\}$, then the deviation degree of $L(p)$ is given by.

$$\sigma(L(p)) = \frac{\sqrt{\sum_{k=1}^{\#L(p)} (p^k L^k - \overline{\alpha})^2)}}{\sum_{k=1}^{\#L(p)} p^k}, \quad (3)$$

where $\overline{\alpha} = \sum_{k=1}^{\#L(p)} L^k p^k / \sum_{k=1}^{\#L(p)} p^k$, $\#L(p)$ is the number of terms in $L(p)$.

The comparison rules of PLTSs based on score function and deviation degree are as follows:

(1) If $E(L_1(p)) > E(L_2(p))$, then $L_1(p) < L_2(p)$.
(2) If $E(L_1(p)) < E(L_2(p))$, then $L_1(p) > L_2(p)$.
(3) If $E(L_1(p)) = E(L_2(p))$, then.

1) If $\sigma(L_1(p)) > \sigma(L_2(p))$, then $L_1(p) < L_2(p)$.
2) If $\sigma(L_1(p)) < \sigma(L_2(p))$, then $L_1(p) > L_2(p)$.
3) If $\sigma(L_1(p)) = \sigma(L_2(p))$, then $L_1(p) = L_2(p)$.

To aggregate multiple PLTSs, the most used approach is aggregation operators. In Pang's research [13], some aggregation operators are introduced to integrate PLTSs and the definition of probabilistic linguistic weighted averaging (PLWA) operator is given as follows.

Definition 4 [13] Let $L_i(p) = \{L_i^k(p_i^k) | k = 1, 2, ..., \#L(p)\} (i = 1, 2, ..., n)$ be n PLTSs, then the PLWA operator is.

$$\begin{aligned} & PLWA(L_1(p), L_2(p), ..., L_n(p)) \\ & = w_1 L_1(p) \oplus w_2 L_2(p) \oplus ... \oplus w_n L_n(p) \\ & = \cup_{L_1^k \in L_1(p)} \{w_1 p_1^k L_1^k\} \oplus \cup_{L_2^k \in L_2(p)} \{w_2 p_2^k L_2^k\} \\ & \quad \oplus ... \oplus \cup_{L_n^k \in L_n(p)} \{w_n p_n^k L_n^k\}, \end{aligned} \quad (4)$$

where L_i^k and p_i^k denote the kth term and its probability in $L_i(p)$, and w_i $(i = 1,2,\ldots,n)$ is the weight of $L_i(p)$.

Incomplete decision matrix

In decision making, above of all, DMs give their evaluations for distinct alternatives associated with multiple attributes according to their knowledge and experience. Given PLTS is a useful and superior tool to express experts' opinions, especially in some emergency and vital decision cases, it becomes the preferred expression model. To describe experts' evaluation information, the decision matrices are used as follows:

$$P^1 = \begin{bmatrix} L_{11} & L_{12} & \ldots & L_{1n} \\ L_{21} & L_{22} & \ldots & L_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ L_{m1} & L_{m2} & \ldots & L_{mn} \end{bmatrix} P^2 = \begin{bmatrix} L_{11} & * & \ldots & L_{1n} \\ * & L_{22} & \ldots & L_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ L_{m1} & * & \ldots & L_{mn} \end{bmatrix}.$$

In decision matrix P^1 represented by PLTS,L_{ij} represents the evaluation value of alternative i over attribute j,$L_{ij} = \{L_{ij}^k(p_{ij}^k)|k = 1, 2, ..., \#L_{ij}(p_{ij})\}$. If all L_{ij} are known,P^1 is called the complete decision matrix. However, due to some reasons, DMs cannot give a complete decision matrix in most cases. In decision matrix P^2, '*' denotes the missing value and P^2 is called the incomplete decision matrix because there exist a values in P^2. The element with missing value is not equal to useless or valueless information, which should not be ignored [14]. Thus, before conducting the decision making, the process to estimate missing information in decision matrix is an indispensable segment. For the sake of convenience, some concepts and mathematical notions are given as follows:

a. Missing expert: the expert whose decision matrix is incomplete. The set of missing experts is denoted as $ME = \{e_k^*|k = 1, 2, ..., \#e^*\}$, where $\#e^*$ is the number of missing experts.

b. Referenced expert: the expert that they provide reference evaluation information for missing expert. The set of referenced experts is denoted as $RE = \{e_k^r|k = 1, 2, ..., \#e^r\}$, where $\#e^r$ is the number of referenced experts.

c. Missing element: the element information in decision matrix is unknown. Denote $L_{ijk}^*(p)$ as missing element, which means that the information of expert e_k over alternative i associated with attribute j is missing. The set of missing elements is denoted as $ML = \{L_{ij}^*|i = 1, 2, ..., m;\ j = 1, 2, ..., n\}$.

Estimating incomplete PLTS information based on knowledge-match degree with reliability

When estimating incomplete decision matrix, an important approach is to refer to evaluation information from other experts (namely referenced experts). Owing to distinct knowledge levels, referenced experts provide the evaluation information of different quality in which the information quality includes two indicators, accuracy and reliability. The higher the accuracy and reliability of evaluation information, the bigger the importance of referenced expert. Thus, in this section, a new method to estimate missing PLTS information in complete decision matrix is proposed based on knowledge-match degree with reliability and an algorithm is designed to work out the approach.

Knowledge-match degree with reliability

To measure the accuracy of evaluation information of referenced expert, a new indicator is proposed, namely knowledge-match degree. The main idea is that the collective evaluation is viewed as benchmark, then the knowledge-match degree can be obtained according to the distance between individual evaluation and the benchmark. The bigger the distance, the low is the knowledge-match degree. The distance is given by

$$d_{ij}^k = |E(L_{ijk}) - \frac{1}{\#e^r}\sum_{g=1}^{\#e^r} E(L_{ijg})|, \tag{5}$$

where d_{ij}^k denotes the distance between the referenced expert e_k^r and collective evaluation over alternative i associated with attribute j. $\#e^r$ is the number of referenced experts in RE.

Definition 5 Suppose that e_k^r is a referenced expert and RE is the set of referenced experts, $e_k^r \in RE$,$RE = \{e_k^r|k = 1, 2, ..., \#e^r\}$. With respect to the evaluation information of e_k^r $L_{ijk}(p)$, its knowledge-match degree is given by.

$$KN_{ij}^k = 1 - \frac{d_{ij}^k}{\sum_{g=1}^{\#e^r} d_{ij}^g},\ g = 1, 2, ..., \#e^r, \tag{6}$$

where KN_{ij}^k denotes the knowledge-match degree of expert e_k^r over alternative i associated with attribute j. $\#e^r$ is the number of referenced experts in RE.

Remarks As we see, the knowledge-match degree is used to measure the accuracy of referenced experts' evaluations. In the context of PLTS, experts provide their evaluation information by PLTS. When calculating the knowledge-match degree, the PLTS is transformed to a crisp value by score function (Eq. (2)). The benchmark is represented by the mean of collective evaluations, namely $\frac{1}{\#e^r}\sum_{g=1}^{\#e^r} E(L_{ijg})$ in Eq. (5). In Eq. (6), we conduct a normalization of distance so as to obtain the knowledge-match degree which belongs to between 0 and 1.

Example 1 Let $S = \{s_0, s_1, s_2, s_3, s_4\}$ be the LTS and the set of referenced experts is $RE = \{e_k^r|k = 1, 2, 3\}$, the evaluations with respect to alternative 1 associated with attribute 1 are:

$L_{11}^1(p) = \{s_0(0.3), s_1(0.7)\}$, $L_{11}^2(p) = \{s_1(0.5), s_2(0.5)\}$, $L_{11}^3(p) = \{s_1(1)\}$.

Then, by Eq. (5), we can get the distance:$d_{11}^1 = 0.43$, $d_{11}^2 = 0.49$, $d_{11}^3 = 0.08$.

According to Eq. (6), the knowledge-match degrees of e_k^r are.

$$KN_{11}^1 = 0.57,\ KN_{11}^2 = 0.51 KN_{11}^3 = 0.92.$$

Besides, reliability is also an essential indicator to measure the importance of evaluations. In some cases, expert may give the evaluations with high knowledge-match degree but the expert has great uncertainty. Thus, on the basis of knowledge-match degree, the reliability of evaluation information should be also considered. Evidently, the higher the reliability, the bigger the importance of the evaluation information provided by the referenced expert. In the context of PLTS, there exists fuzziness in a PLTS and the more the fuzziness, the more the uncertainty of referenced expert. Hence, the reliability of evaluation information can be obtained by calculating fuzzy entropy of PLTS. Evidently, the smaller the fuzzy entropy, the bigger the reliability. In this research, we adopt the method proposed by Liu [29] to measure fuzzy entropy of PLTS and give the definition of reliability.

Definition 6 [29] Suppose that S is a LTS, $S = \{s_\alpha | \alpha = 0, 1, 2, ..., 2\tau\}$. Let $L(p) = \{L^k(p^k) | k = 1, 2, ..., \#L(p)\}$ be a PLTS, then the fuzzy entropy of $L(p)$ is given by.

$$E_F(L(p)) = 1 - \sum_{i=1}^{\#L(p)} p_i |1 - 2\alpha_i|, \tag{7}$$

where E_F denotes the fuzzy entropy, $\alpha_i = L_i/(2\tau)$, and g represents the number of terms in $L(p)$.

Evidently, the value of fuzzy entropy (Eq. (7)) is between 0 and 1 (contain 0,1). With the aid of fuzzy entropy, we propose the definition of PLTS reliability.

Definition 7 Suppose that S is a LTS, and $S = \{s_\alpha | \alpha = 0, 1, 2, ..., 2\tau\}$. Let $L(p) = \{L^k(p^k) | k = 1, 2, ..., \#L(p)\}$ be a PLTS, then the reliability of $L(p)$ is given by.

$$R(L(p)) = 1 - E_F. \tag{8}$$

According to Eqs. (7), (8) is also written as

$$R(L(p)) = \sum_{i=1}^{\#L(p)} p_i |1 - 2\alpha_i|. \tag{9}$$

Proposition 1. In definition 7, the value of reliability is always between 0 and 1 (contain 0,1).

Proof. The formula of reliability is Eq. (9), and $R(L(p)) = \sum_{i=1}^{\#L(p)} p_i |1 - 2\alpha_i|$. For $\alpha_i = L_i/(2\tau)$, L_i is the subscript of the i th term in $L(p)$. Evidently, $0 \le L_i/(2\tau) \le 1$, namely $0 \le \alpha_i \le 1$. Thus, $0 \le |1 - 2\alpha_i| \le 1$. Since $\sum_{i=1}^{\#L(p)} p_i = 1$, $0 \le \sum_{i=1}^{\#L(p)} p_i |1 - 2\alpha_i| \le 1$. Hence, $R(L(p)) \in [0, 1]$.

Combining the knowledge-match degree with reliability, we can well measure the reliability level of evaluation information provided by referenced experts. The knowledge-match degree is used to identify the accuracy of evaluations while the reliability checks whether the evaluation information is reliable. The knowledge-match degree with reliability is given by

$$KR_{ij}^k = KN_{ij}^k \cdot R_{ij}^k, \tag{10}$$

where KR_{ij}^k represents the knowledge-match degree with reliability provided by expert e_k^r with respect to alternative i associated to attribute j. Since $KN_{ij}^k \in [0, 1]$ and $RE_{ij}^k \in [0, 1]$, KR_{ij}^k also belongs to the interval [0,1].

Example 2 According to example 1, $L_{11}^1(p) = \{s_0(0.3), s_1(0.7)\}$, $L_{11}^2(p) = \{s_1(0.5), s_2(0.5)\}$, and $L_{11}^3(p) = \{s_1(1)\}$, and their knowledge-match degrees are $KN_{11}^1 = 0.57$, $KN_{11}^2 = 0.51$, and $KN_{11}^3 = 0.92$.

By Eq. (9), the reliability is $R_{11}^1 = 0.65$, $R_{11}^2 = 0.25$, and $R_{11}^3 = 0.5$.

According to Eq. (10), we can obtain the knowledge-match degree with reliability:

$$KR_{11}^1 = 0.3705,\ KR_{11}^2 = 0.1275,\ KR_{11}^3 = 0.4600.$$

Estimating incomplete decision matrix

Different from Liang's method [20] to estimate the missing values, we consider the evaluation information of referenced experts and utilize the knowledge-match degrees with reliability to aggregate the referenced information. The knowledge-match degree with reliability can be obtained by Eq. (10), namely $KR_{ij}^k = KN_{ij}^k \cdot R_{ij}^k$, while it should be normalized first to become the aggregation operator. The normalized KR_{ij}^k is

$$NKR_{ij}^k = \frac{NKR_{ij}^k}{\sum_{\substack{g=1 \\ g \ne k}}^{\#e^r} NKR_{ij}^g}, \tag{11}$$

where $\#e^r$ is the number of referenced experts in RE.

Subsequently, according to Eq. (4), namely the PLWA operator, obtain the estimated value of missing information L_{ij}^* by aggregating all referenced evaluations of referenced experts. Among that, NKR_{ij}^k is viewed as the weight of PLWA operator. A missing expert e_k^* may have several missing values in the decision matrix. In this case, estimate all missing values of e_k^* and then change to the next missing expert. Thus, we design an algorithm to estimate the missing information.

Algorithm 1

Input: The incomplete decision matrix.

Output: The complete decision matrix.

Step 1. Find out the missing elements, missing experts and construct the corresponding sets: *ML* and *ME*. $ME = \{e_k^* \mid k = 1,2,...,\#e^*\}$, the missing experts are ordered in term of the subscripts from small to large. $ML = \{L_{ij}^* \mid i = 1,2,..., m; j = 1,2,..., n\}$ and $\#L_{ij}^*$ is the number of missing elements in *ML*. Let *z* be the number of iteration and its initial value is equal to 0.

Step 2. Select the missing expert in *ME* e_1^*, and search for its missing element in *ML* L_{ij}^*. According to L_{ij}^*, find out referenced experts of e_1^* and form the set of referenced experts $RE = \{e_k^r \mid k = 1,2,...,\#e^r\}$. Then, by equation (5)~(6) and (9)~(10), calculate the knowledge-match degrees with reliability of referenced experts respectively.

Step 3. According to equation (4) and (11), aggregate all evaluation information of referenced experts in $RE = \{e_k^r \mid k = 1,2,...,\#e^r\}$ and obtain the estimated value of L_{ij}^*.

Step 4. Let z=z+1. Subsequently, check whether $z = \#L_{ij}^*$. If $z < \#L_{ij}^*$, return step 2. If $z = \#L_{ij}^*$, conduct the next step.

Step 5. Output the complete decision matrix.

Step 6. End.

The application of estimation method in LSGDM

Problem description

In the LSGDM, the purpose is to achieve the optimal scheme in all alternatives. In our research, PLTS is used to express the evaluation information of DMs. If there exist incomplete PLTS information in decision matrix, the missing information will be first estimated. In the process of LSGDM, some notions are as follows:

The alternative set: $X = \{x_i | i = 1, 2, ..., m\}$.

The attribute set: $C = \{c_j | j = 1, 2, ..., n\}$.

The expert set: $E = \{e_k | k = 1, 2, ..., p\}$.

The set of missing experts: $ME = \{e_k^* | i = 1, 2, ..., \#e^*\}$.

The set of referenced experts: $RE = \{e_k^r | k = 1, 2, ..., \#e^r\}$.

The set of missing elements: $ML = \{L_{ij}^* | i = 1, 2, ..., m; j = 1, 2, ..., n\}$.

In this study, it contains four main steps to solve the problem of LSGDM including (1) estimating incomplete information, (2) subgroup detecting, (3) consensus reaching process (CRP), and (4) selection process. In step 1, to handle the incomplete PLTS information in decision matrices provided by experts, the Algorithm 1 is applied. In step

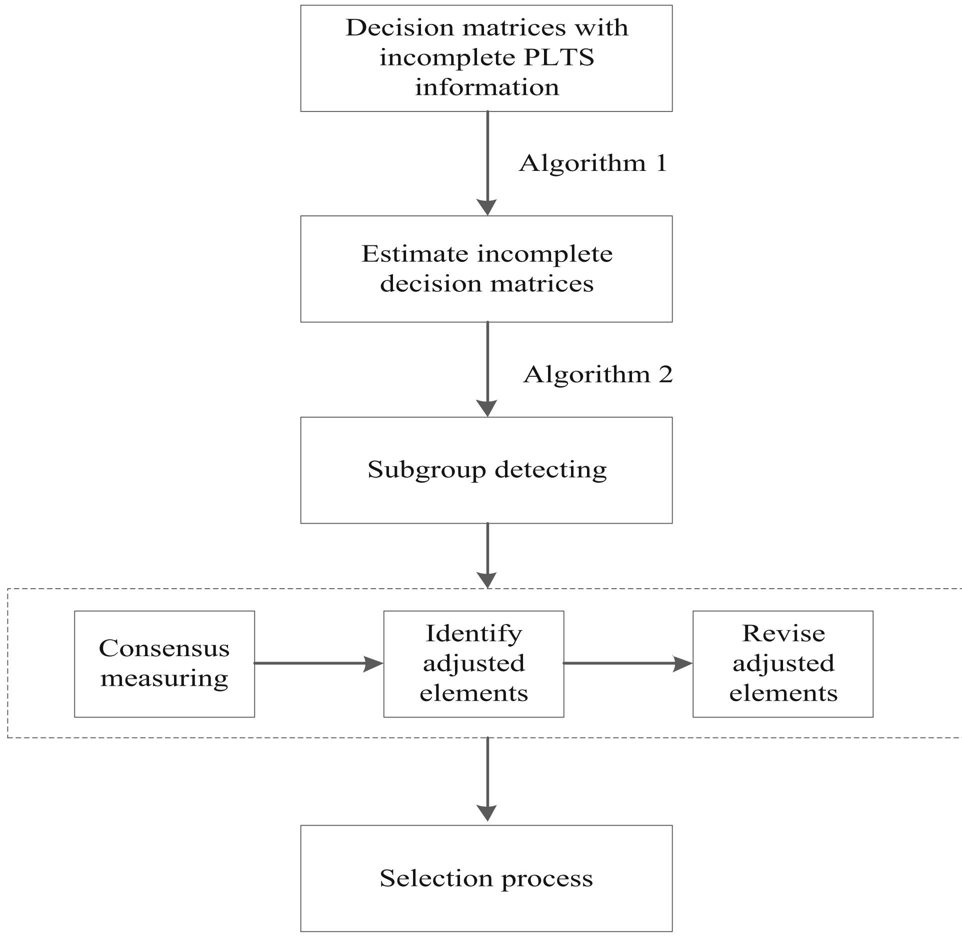

Fig. 1 Flowchart of LSGDM

2, subgroup detecting means that the experts are divided into some small subgroups in term of specific standards or methods to simplify the complexity of decision problem. In step 3, CRP is an important section to check whether the consensus of experts satisfies the expected level (consensus threshold). In step 4, all evaluation information from subgroups are aggregated to obtain the comprehensive evaluation value. According to the comprehensive values, rank the alternatives.

The process of LSGDM based on the estimation of incomplete decision matrix

In this subsection, based on the proposed estimation method, the decision process is developed to solve the problem of LSGDM. To see the whole framework intuitively, we give the flowchart of LSGDM (Fig. 1).

Estimating incomplete information

In the beginning of decision, experts give their evaluations over various alternatives associated to attributes in term of professional knowledge and experience. However, there always exist the incomplete evaluation information provided by DMs because of the limited time or the complexity of decision problem. To efficiently utilize the evaluation information, it is vital to fill the missing elements in decision matrices. According to the Algorithm 1, the missing values can be estimated with the aid of the related information of referenced experts. The details of estimation are omitted here.

Subgroup detecting

In traditional group decision making, the number of DMs may be less and it is often in the interval [3, 20]. Nevertheless, the number of experts participating decision making is large in LSGDM. The decision information given by them are very complicated and it is much difficult to analysis if we directly conduct the decision making. Thus, all experts are divided to some small subgroups to simplify the decision analysis. K-means clustering is a useful method to classify and it has been successfully applied in LSGDM to solve subgroup detecting [3]. In this research, we also adopt this method to divide experts.

With respect to K-means method, two important indicators that should be determined are K value and initial centroids For the K value, we mainly obtain it according to the error sum of squares SSE. By means of software, such as Python, draw graphic of relation between SSE and K value, and find out the point in which the degree of decline becomes smooth from sharp, which is the best K value. Then, we randomly choose K experts as the initial centroids and remaining experts are assigned to the nearest centroids. The process of subgroup detecting is summarized in Algorithm 2 as follows.

Algorithm 2

Input: The decision matrices of experts

Output: Several subgroups containing various experts

Step 1. Utilize the software Python to obtain the K value, namely the number of subgroup.

Step 2. Choose K experts as initial centroids randomly, the decision matrix P^k of expert is viewed as the initial point, then calculate the distances between experts and centroids and classify them to the nearest subgroup. According to equation (2) and Euclidean distance, the distance of two experts' matrices is given by

$$d(P^1,P^2)=\sqrt{\sum_{i=1}^{m}(\sum_{j=1}^{n}E(L_{ij}^1)-\sum_{j=1}^{p}E(L_{ij}^2))^2} \tag{12}$$

Step 3. Compute the mean of experts' decision matrices in every subgroup G_t, and it becomes the new centroid of each group in the next iteration, which is denoted as $\overline{P}^{G_t}$. The elements in $\overline{P}^{G_t}$ can be obtained by

$$\overline{L}_{ij}^{G_t}=\frac{1}{\#G_t}\sum_{k=1}^{\#G_t}\sum_{i=1}^{m}\sum_{j=1}^{n}L_{ij}^k \tag{13}$$

Where $\#G_t$ denotes the number of experts in subgroup G_t. $\overline{L}_{ij}^{G_t}$ is the element in $\overline{P}^{G_t}$ with respect to alternative i associated to attribute j. L_{ij}^k represents the PLTS information given by expert e_k.

Step 4. Conduct step 3 and step 4 repeatedly until the final subgroups do not change, which are denoted as $G_t, t\in\{1,2,...,K\}$.

Step 5. End.

Consensus reaching process based on trust relationship

Here, a consensus reaching process is developed to ensure that experts in each subgroups are able to reach the accepted consensus level. Since the complete consensus is difficult to achieve and it will consume large cost, the "soft consensus" is often used to measure whether the evaluations of experts satisfy the expected consensus. That is to say, we preset the consensus threshold ξ, and it reaches consensus if the consensus index $CI \geq \xi$. However, in most cases, it is difficult to reach consensus at the beginning as there always exist some conflicts of opinions given by DMs, which leads to low consensus degree. If $CI < \xi$, some decision matrices with low consensus will be revised to improve the consensus level. The consensus reaching process (CRP) contains three parts: consensus measuring, identification strategy and consensus improvement.

Consensus measuring

When measuring the consensus of subgroup, Li et al. [30] considered it from four aspects: element, alternative, decision matrix and subgroup. In Li's research, first, the collective opinion needs to be obtained by aggregating the information of all subgroups. However, it may be complicated in calculation by utilizing Li's method. Thus, we directly measure the consensus degree in terms of distance between any two elements, which does not need to calculate the collective opinion in advance. After measuring the consensus level of element, alternative, decision matrix and subgroup respectively, the collective consensus degree can be obtained. The process of consensus measure is as follows:

a. The consensus of element level

$$CI(L_{ij}^{k_1}) = \frac{1}{\#G_t - 1} \sum_{\substack{k_2=1 \\ k_2 \neq k_1}}^{\#G_t} \times (1 - |E(L_{ij}^{k_1}(p)) - E(L_{ij}^{k_2}(p))|),\ e_{k_1} \in G_t. \tag{14}$$

b. The consensus of alternative level

$$CI(L_i^k) = \frac{1}{n} \sum_{j=1}^{n} CI(L_{ij}^k). \tag{15}$$

c. The consensus of decision matrix level

$$CI(P^{k_1}) = \frac{1}{m} \sum_{i=1}^{m} CI(L_i^{k_1}) \tag{16}$$

d. The consensus of subgroup level

$$CI(G_t) = \frac{1}{\#G_t} \sum_{k=1}^{\#G_t} CI(P^k). \tag{17}$$

Finally, the collective consensus is

$$CI = \frac{1}{K} \sum_{t=1}^{K} CI(G_t). \tag{18}$$

Identification strategy

If $CI \geq \xi$, it satisfies the expected consensus and continue to perform to the next step, selection process. But if $CI < \xi$, we should identify the elements with low consensus and revise them to reach the consensus threshold. It will consume large cost and largely change original information of experts to modify all elements which do not satisfy the consensus threshold. Thus, with the aid of Li's research [30], we only need to revise the element with the lowest consensus. The identification processes are as follows:

a. Identify the subgroup with the lowest consensus

$$AL(G_t) = \{G_t | CI_t = \min CI(G_t)\}, \tag{19}$$

where $AL(G_t)$ represents the subgroup with the lowest consensus among all subgroups.

b. Identify the decision matrix with lowest consensus

$$AL(P^k) = \{P^k \Big| CI(P^k) = \min CI(P^k) \wedge e_k \in AL(G_t)\}. \tag{20}$$

c. Identify the alternative with the lowest consensus

$$AL(L_i^k) = \{L_i^k \Big| CI(L_i^k) = \min CI(L_i^k) \wedge L_i^k \in AL(P^k)\}. \tag{21}$$

d. Identify the element with the lowest consensus

$$AL(L_{ij}^k) = \{L_{ij}^k | CI(L_{ij}^k) = \min CI(L_{ij}^k) \wedge L_{ij}^k \in AL(L_{ij}^k)\}. \tag{22}$$

After conducting the above four steps, we can obtain the final element $AL(L_{ij}^k)$ which needs to be modified. Evidently, through the identification strategy, it only needs to adjust the $AL(L_{ij}^k)$ if the consensus level does not satisfy consensus threshold. The advantage of the method is to save the adjustment cost and remain the original evaluation information to the greatest extent.

Consensus improvement based on trust relationship

For the element $AL(L_{ij}^k)$, it is the evaluation information given by expert e_k with respect to alternative i associated to attribute j. To some extent, the initial information will be changed when modifying the element with the lowest consensus. To reduce the loss of original evaluation information, we develop a new method to modify the element based on trust relationship of experts. As we all know, in a social network, the higher the trust degree, the bigger the preference similarity of two experts. Here, we utilize social network analysis (SNA) and graph theory to represent the relationship of experts. The related basic knowledge are as follows.

Definition 8 [31] Let G (V, E, W) be a graph, where V is the set of vertices, $V = \{v^k | k = 1, 2, ..., p\}$, E denotes the set of edges, $E = \{e^{kl} | k \neq l \wedge k, l = 1, 2, ..., p\}$ and W represents the set of weights, $W = \{w^{kl} | k \neq l \wedge k, l = 1, 2, ..., p\}$, $w^{kl} \in [0, 1]$. The social network graph is shown in Fig. 2.

In this research, v^k denotes expert e_k and e^{kl} denotes the relationship between e_k and e_l. w^{kl} represents the trust degree between e_k and e_l. Here, we utilize the directional social work [22], namely $e^{kl} \neq e^{lk}$ and $w^{kl} \neq w^{lk}$. When analysis the trust relationship among experts, we denote the weight w^{kl} as the trust degree T_{kl} from e_k to e_l.

Based on SNA, we can find out the most trusted expert of e_k in the same subgroup, which is denoted as e_k^T,

$$e_k^T = \{e_l | T_{kl} = \max T_{kl} \wedge CI(P^l) \geq \xi \wedge e_k, e_l \in G_t\}. \tag{23}$$

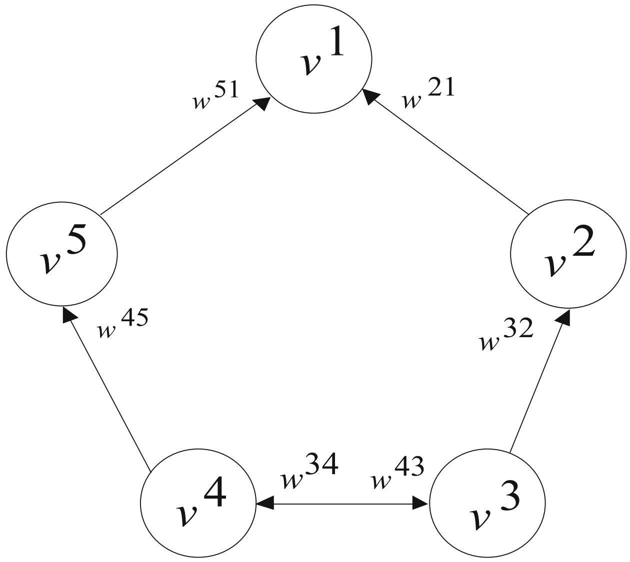

Fig. 2 Social network graph

In Eq. (23), the expert e_k^T needs satisfy three conditions. First, e_k and e_k^T are in the same subgroup. Second, the consensus degree of e_k^T must exceed the consensus threshold. Finally, the trust degree between e_k and e_k^T is the maximum one. According to the corresponding element of e_k^T, we modify the $AL(L_{ij}^k)$ and preset a parameter θ to control the adjustment proportion. The modification mechanism is given by

$$AL(\overline{L}_{ij}^k) = (1-\theta)L_{ij}^k(p) + \theta L_{ij}^l(p), \tag{24}$$

where $AL(\overline{L}_{ij}^k)$ represents the modified $AL(L_{ij}^k)$, $L_{ij}^l(p) = \{L_{ij}^l | e_l = e_k^T\}$.

By Eqs. (23) and (24), we can improve the consensus levels of subgroup and collective through modifying the elements with the low consensus iteratively. To realize the complete consensus reaching process, the Algorithm 3 is designed as follows.

Algorithm 3

Input: The original decision matrices of subgroups

Output: The modified decision matrices of subgroups which reach consensus

Step 1. Consensus measuring. We preset the consensus threshold ξ . By equation (14)~(18), obtain the consensus degree of collective *CI*, and check whether $CI \geq \xi$. If $CI \geq \xi$, skip to step 5.

Step 2. Identify the element with the lowest consensus. According to equation (19) and (22), the element $AL(L_{ij}^k)$ can be achieved.

Step 3. Modify the element $AL(L_{ij}^k)$. By equation (23), we can find out the most trusted expert of e_k ,namely, e_k^T . By equation (24), the $AL(L_{ij}^k)$ will be revised through the adjustment parameter θ .

Step 4. Calculate the collective consensus degree again. If $CI < \xi$, then return step 2 and conduct repeatedly. If $CI \geq \xi$, go to the next step.

Step 5. End.

Selection process

After the consensus reaching process, the evaluation information of all experts satisfy the consensus threshold. Subsequently, in the selection process, the decision matrices of experts will be aggregated into the collective decision matrix. Step by step, we first integrate the decision information in the subgroup by Eq. (4), where the weights of experts in the same subgroup are identical by default, e.g., the expert's weight $1/\#G_t$. We denote the decision matrix of subgroup as P^{G_t}:

$$P^{G_t} = PLWA(P^1, P^2, ..., P^{\#G_t}). \tag{25}$$

Then, according to each decision matrix of subgroup P^{G_t}, we can get the collective evaluation. The weights of subgroups λ_t are given by the size of persons in subgroup, e.g., $\lambda_t = \#G_t / \sum_{r=1}^{K} \#G_r$. Thus, the collective decision matrix P^c is

$$P^c = PLWA(P^{G_1}, P^{G_2}, ..., P^K). \tag{26}$$

As for the attribute weight ω_j, we preset the value in terms of the actual decision situation. Similarly, utilize PLWA and score function (Eq. (2)) to aggregate the attribute information of P^c into the comprehensive values of alternatives $E(L_i)$. According to the comprehensive values, we rank the alternative in descending order and obtain the optimal one.

In our research, the main steps of LSGDM are summarized as follows:

Step 1. Collect the experts' evaluation information and estimate the missing PLTS information in incomplete decision matrices. Using the Algorithm 1, the incomplete decision matrices can be transformed to the complete ones.

Step 2. Subgroup detecting. According to the Algorithm 2, the experts are divided into some distinct subgroups.

Step 3. Consensus reaching process. By utilizing the Algorithm 3, the evaluations of experts are able to satisfy anticipated consensus level.

Step 4. Selection process. According to Eqs. (25) and (26), obtain the collective decision matrix P^c. Calculate the comprehensive values of alternatives $E(L_i)$ and rank the alternatives. Finally, we obtain the optimal one.

In next section, the proposed process of LSGDM will be applied to a medical case.

Case study

In medical field, decision analysis is often used to solve some complex decision problem, such as selection of drug suppliers, drug procurement management, medicine classification based on AI [32] and so on. When coping with such decision problems, large number of human experts who come from different fields and have rich experience are invited to give evaluations over distinct alternatives. Although human are multi-modal thinking and can make decision intelligently and efficiently according to their knowledge and experience [33], there often exist cognitive fuzziness and hesitation when people face the complex problem and this uncertainty should be considered. Thus, PLTS is a very useful tool to express the evaluation information of experts. Now, suppose that a hospital is facing such a decision problem:

There are four kinds of drugs: medicine A, medicine B, medicine C and medicine D. Since the four drugs can treat the same disease while they have different efficacy and side effects, the hospital will choose the best one among the four drugs. Twenty experts from various medical fields are invited to participate in this decision making. They will give evaluations by means of PLTS according to three evaluation criteria: efficacy, side effects and cost. Given the importance of different criteria, we preset the weights of attributes to 0.5, 0.3 and 0.2, respectively. According to the proposed decision process in "The application of estimation method in LSGDM", we solve the problem step by step.

Evidently, this is a LSGDM. Experts use a linguistic term set with 5 granularity to give evaluations, namely $S = \{s_0, s_1, s_2, s_3, s_4\}$. From left to right, the terms mean: very bad, bad, medium, good, and very good. There are totally 20 DMs, 4 alternatives (4 kinds of drugs) and 3 attributes (3 evaluation criteria) which are denoted as.

The set of alternatives: $X = \{x_i | i = 1, 2, 3, 4\}$.

The attribute set: $C = \{c_j | j = 1, 2, 3\}$.

The weight vector of attribute: $\omega = (0.5, 0.3, 0.2)$.

The expert set: $E = \{e_k | k = 1, 2, ..., 20\}$.

Step 1. By means of PLTS, the experts give their evaluation information in the form of matrix as follows: ('*' represents missing information)

$$P^1 = \begin{bmatrix} \{s_2(1)\} & \{s_1(0.3), s_2(0.7)\} & \{s_3(1)\} \\ \{s_4(0.8)\} & \{s_2(1)\} & \{s_2(0.6), s_3(0.2)\} \\ \{s_0(0.8), s_1(0.2)\} & \{s_3(0.9)\} & \{s_1(0.5), s_2(0.5)\} \\ \{s_3(1)\} & \{s_2(1)\} & \{s_3(0.8)\} \end{bmatrix}$$

$$P^2 = \begin{bmatrix} * & \{s_2(1)\} & \{s_1(0.2), s_2(0.8)\} \\ \{s_2(0.5), s_3(0.5)\} & \{s_2(0.8)\} & \{s_3(1)\} \\ \{s_2(0.8)\} & \{s_4(1)\} & \{s_1(0.8), s_2(0.2)\} \\ \{s_3(1)\} & \{s_2(0.8), s_3(0.2)\} & \{s_2(1)\} \end{bmatrix}$$

$$P^3 = \begin{bmatrix} \{s_3(0.8)\} & \{s_2(1)\} & \{s_2(0.5), s_3(0.5)\} \\ \{s_3(0.9), s_4(0.1)\} & \{s_1(1)\} & \{s_3(1)\} \\ \{s_4(1)\} & \{s_1(0.3), s_2(0.5)\} & \{s_2(1)\} \\ \{s_2(1)\} & \{s_3(0.8)\} & \{s_1(0.5), s_2(0.5)\} \end{bmatrix}$$

$$P^4 = \begin{bmatrix} \{s_2(0.5), s_3(0.5)\} & * & \{s_3(1)\} \\ \{s_2(0.8)\} & \{s_2(0.3), s_3(0.7)\} & \{s_2(1)\} \\ \{s_1(0.4), s_2(0.6)\} & \{s_3(1)\} & \{s_3(0.2), s_4(0.8)\} \\ \{s_2(1)\} & \{s_3(1)\} & \{s_2(0.5), s_3(0.5)\} \end{bmatrix}$$

$$P^5=\begin{bmatrix}\{s_3(1)\} & \{s_2(0.5),\ s_3(0.5)\} & \{s_2(1)\}\\ \{s_1(0.5),\ s_2(0.3)\} & \{s_2(1)\} & \{s_3(0.9)\}\\ \{s_4(1)\} & \{s_1(0.6)\} & \{s_2(1)\}\\ \{s_2(1)\} & \{s_3(1)\} & \{s_2(0.6),\ s_3(0.4)\}\end{bmatrix}$$

$$P^6=\begin{bmatrix}\{s_1(1)\} & \{s_2(0.5),\ s_3(0.5)\} & \{s_3(1)\}\\ \{s_3(0.6),\ s_4(0.4)\} & * & \{s_2(1)\}\\ \{s_2(0.8)\} & \{s_3(1)\} & \{s_2(0.5),\ s_3(0.5)\}\\ \{s_1(0.5)\} & \{s_2(0.8),\ s_3(0.2)\} & \{s_2(1)\}\end{bmatrix}$$

$$P^7=\begin{bmatrix}\{s_1(0.8),\ s_2(0.2)\} & \{s_2(1)\} & \{s_3(1)\}\\ \{s_2(0.6),\ s_3(0.4)\} & \{s_3(1)\} & \{s_2(0.5),\ s_3(0.5)\}\\ \{s_1(0.8)\} & \{s_2(0.6),\ s_3(0.4)\} & \{s_2(1)\}\\ \{s_2(1)\} & \{s_2(1)\} & \{s_3(1)\}\end{bmatrix}$$

$$P^8=\begin{bmatrix}\{s_2(1)\} & \{s_1(0.5),\ s_2(0.5)\} & \{s_3(1)\}\\ \{s_2(0.8),\ s_3(0.2)\} & \{s_1(1)\} & \{s_1(0.3),\ s_2(0.7)\}\\ \{s_3(1)\} & \{s_2(0.6),\ s_3(0.4)\} & \{s_2(1)\}\\ \{s_2(1)\} & \{s_2(0.5),\ s_3(0.5)\} & *\end{bmatrix}$$

$$P^9=\begin{bmatrix}\{s_2(0.8)\} & \{s_3(1)\} & \{s_3(0.8),\ s_4(0.2)\}\\ \{s_1(1)\} & \{s_3(1)\} & \{s_0(0.9)\}\\ \{s_2(1)\} & \{s_3(0.6),\ s_4(0.4)\} & \{s_3(0.8)\}\\ \{s_1(0.3),\ s_2(0.7)\} & \{s_2(0.8),\ s_3(0.2)\} & \{s_3(1)\}\end{bmatrix}$$

$$P^{10}=\begin{bmatrix}\{s_3(1)\} & \{s_2(0.3),\ s_3(0.7)\} & \{s_4(1)\}\\ \{s_1(0.5),\ s_2(0.5)\} & \{s_2(1)\} & \{s_2(0.8),\ s_3(0.2)\}\\ \{s_1(0.3),\ s_2(0.7)\} & \{s_3(1)\} & \{s_4(1)\}\\ \{s_2(1)\} & \{s_2(0.3),\ s_3(0.7)\} & \{s_2(0.8)\}\end{bmatrix}$$

$$P^{11}=\begin{bmatrix}\{s_2(1)\} & \{s_2(0.5),\ s_3(0.5)\} & \{s_1(1)\}\\ * & \{s_4(0.8)\} & \{s_1(0.3),\ s_2(0.7)\}\\ \{s_3(0.3),\ s_4(0.7)\} & \{s_2(1)\} & \{s_3(1)\}\\ \{s_2(1)\} & \{s_3(0.6),\ s_4(0.4)\} & \{s_2(1)\}\end{bmatrix}$$

$$P^{12}=\begin{bmatrix}\{s_1(1)\} & \{s_2(0.6),\ s_3(0.4)\} & \{s_4(1)\}\\ \{s_1(0.6),\ s_2(0.4)\} & \{s_2(1)\} & \{s_2(0.5),\ s_3(0.5)\}\\ \{s_2(1)\} & \{s_3(1)\} & \{s_1(0.5),\ s_2(0.5)\}\\ \{s_3(0.6),\ s_4(0.4)\} & \{s_2(0.8),\ s_3(0.2)\} & \{s_2(1)\}\end{bmatrix}$$

$$P^{13}=\begin{bmatrix}\{s_2(1)\} & \{s_1(0.8),\ s_2(0.2)\} & \{s_3(1)\}\\ \{s_2(0.4),\ s_3(0.6)\} & \{s_3(1)\} & \{s_2(0.8)\}\\ \{s_1(1)\} & \{s_3(1)\} & \{s_2(0.6),\ s_3(0.4)\}\\ \{s_2(0.5),\ s_3(0.5)\} & \{s_1(1)\} & \{s_2(1)\}\end{bmatrix}$$

$$P^{14}=\begin{bmatrix}\{s_1(0.6),\ s_2(0.4)\} & \{s_2(1)\} & \{s_3(0.5),\ s_4(0.5)\}\\ \{s_3(1)\} & \{s_1(0.3),\ s_2(0.7)\} & \{s_2(1)\}\\ \{s_2(1)\} & \{s_3(0.8)\} & \{s_2(0.4),\ s_3(0.6)\}\\ \{s_4(1)\} & \{s_2(0.8),\ s_3(0.2)\} & \{s_3(1)\}\end{bmatrix}$$

$$P^{15}=\begin{bmatrix}\{s_1(1)\} & \{s_2(0.5),\ s_3(0.5)\} & \{s_2(1)\}\\ \{s_1(0.6),\ s_2(0.4)\} & \{s_3(0.5),\ s_4(0.3)\} & \{s_4(1)\}\\ \{s_2(1)\} & \{s_3(1)\} & \{s_2(0.5),\ s_3(0.5)\}\\ \{s_1(0.6),\ s_2(0.4)\} & \{s_3(0.8),\ s_4(0.2)\} & \{s_2(1)\}\end{bmatrix}$$

$$P^{16}=\begin{bmatrix}\{s_2(1)\} & \{s_3(0.5),\ s_4(0.5)\} & \{s_3(1)\}\\ \{s_3(1)\} & \{s_2(1)\} & \{s_2(0.8),\ s_3(0.2)\}\\ \{s_1(0.4),\ s_2(0.6)\} & \{s_3(1)\} & \{s_2(1)\}\\ \{s_1(0.8)\} & \{s_2(0.3),\ s_3(0.7)\} & \{s_4(1)\}\end{bmatrix}$$

$$P^{17}=\begin{bmatrix}\{s_1(0.5),\ s_2(0.5)\} & \{s_2(1)\} & \{s_3(1)\}\\ \{s_2(0.8)\} & \{s_1(0.6),\ s_2(0.4)\} & \{s_2(1)\}\\ \{s_3(1)\} & \{s_1(0.5),\ s_2(0.5)\} & \{s_2(0.6),\ s_3(0.4)\}\\ \{s_2(1)\} & \{s_3(1)\} & \{s_1(0.5),\ s_2(0.5)\}\end{bmatrix}$$

$$P^{18}=\begin{bmatrix}\{s_2(1)\} & \{s_1(0.6),\ s_2(0.4)\} & \{s_3(1)\}\\ \{s_2(0.3),\ s_3(0.7)\} & \{s_2(0.8)\} & \{s_2(1)\}\\ \{s_1(0.5),\ s_2(0.5)\} & \{s_3(1)\} & \{s_2(0.3),\ s_3(0.7)\}\\ \{s_2(1)\} & \{s_2(0.5),\ s_3(0.5)\} & \{s_2(1)\}\end{bmatrix}$$

$$P^{19}=\begin{bmatrix}\{s_3(1)\} & \{s_2(0.9),\ s_3(0.1)\} & \{s_2(1)\}\\ \{s_1(0.3),\ s_2(0.7)\} & \{s_2(1)\} & \{s_2(0.5),\ s_3(0.5)\}\\ \{s_3(0.8)\} & \{s_1(0.4),\ s_2(0.6)\} & \{s_2(1)\}\\ \{s_2(1)\} & \{s_2(0.9)\} & \{s_3(1)\}\end{bmatrix}$$

$$P^{20}=\begin{bmatrix}\{s_1(0.5),\ s_2(0.5)\} & \{s_3(1)\} & \{s_2(0.5),\ s_3(0.5)\}\\ \{s_2(1)\} & \{s_1(0.8)\} & \{s_1(0.6),\ s_2(0.4)\}\\ \{s_3(1)\} & \{s_1(0.7),\ s_2(0.3)\} & \{s_3(1)\}\\ \{s_2(0.5),\ s_3(0.5)\} & \{s_2(1)\} & \{s_1(1)\}\end{bmatrix}$$

In the above decision matrices, find out the incomplete decision matrices and estimate the missing information. The missing experts and missing elements are as follows:

The set of missing experts: $ME = \{e_2^*, e_4^*, e_6^*, e_8^*, e_{11}^*\}$.

The set of missing elements: $ML = \{L_{11}^*, L_{12}^*, L_{22}^*, L_{43}^*, L_{21}^*\}$.

We take e_2^* as an example. For e_2^*, the missing element is L_{11}^*, namely denoting L_{112}^*. According to the Algorithm 1, L_{112}^* can be estimated with the aid of PLTS information of referenced experts. In addition, the set of referenced experts is $MR = \{e_1^r, e_3^r, ..., e_{20}^r\}$. Subsequently, calculate the knowledge-match degree with reliability on the element L_{11} for each referenced expert. By Eqs. (5)–(11), we can obtain the normalized KR_{ij}^k:

$NKR_{11}^1 = 0$ $NKR_{11}^3 = 0.099$ $NKR_{11}^4 = 0.053$ $NKR_{11}^5 = 0.099$ $NKR_{11}^6 = 0.101$

$NKR_{11}^7 = 0.081$ $NKR_{11}^8 = 0$ $NKR_{11}^9 = 0$ $NKR_{11}^{10} = 0.099$ $NKR_{11}^{11} = 0$

$NKR_{11}^{12} = 0.101 NKR_{11}^{13} = 0 NKR_{11}^{14} = 0.064 NKR_{11}^{15} = 0.101 NKR_{11}^{16} = 0$

$NKR_{11}^{17} = 0.053 NKR_{11}^{18} = 0 NKR_{11}^{19} = 0.099 NKR_{11}^{20} = 0.053$

Then, according to Eq. (4) and the NKR_{ij}^k is used as the aggregation weight. We can get the estimated value of missing element $L_{112}^* = \{s_2(0.4), s_{2.1}(0.6)\}$.

Similarly, for the rest elements in *ML*, we have

$$L_{124}^* = \{s_{2.3}(0.5), s_{2.8}(0.5)\}\ L_{226}^* = \{s_3(1)\}$$

$$L_{438}^* = \{s_1(0.8), s_{1.6}(0.2)\}\ L_{2111}^* = \{s_{2.8}(0.5), s_3(0.5)\}$$

Step 2. Subgroup detecting. Utilize the software Python to make a diagram between *SSE* and *k* value (see Fig. 3), and determine the best *k* value.

From Fig. 3, we can see that the best k value is equal to 3. Thus, the number of subgroup should be taken 3. According to the Algorithm 2, choose randomly the experts e_1, e_2 and e_3 as initial centroids of subgroups. The result of subgroup detecting is.

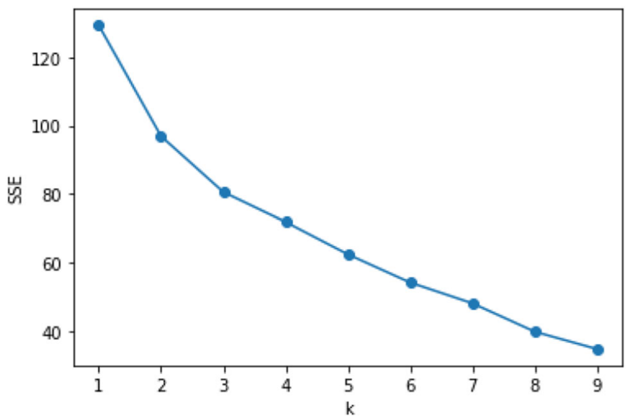

Fig. 3 The diagram between SSE and k value

Subgroup 1: e_1,e_4,e_5,e_9,e_{12}.
Subgroup 2: e_2,e_6,e_7,e_{11},e_{16},e_{17},e_{18}.
Subgroup 3: e_3,e_8,e_{10},e_{13},e_{14},e_{15},e_{19},e_{20}.

After three iterations, the subgroups are stable and the final result is as follows:

Subgroup 1: e_1,e_5,e_9, e_{11},e_{12},e_{13}.
Subgroup 2: e_2,e_4,e_6,e_7,e_{14},e_{16},e_{17},e_{18}.
Subgroup 3: e_3,e_8,e_{10},,e_{15},e_{19},e_{20}.

Step 3. Consensus reaching process according to the Algorithm 3. Given the importance of drug selection and the complexity of experts, we preset the consensus threshold $\xi = 0.9$. By Eqs. (14) ~ (18), compute the consensus degrees from four aspects: element level, alternative level, decision matrix level and subgroup level. The consensus degrees on decision matrix level and subgroup level are shown in Table 2.

According to the results in Table 1, we obtain the collective consensus degree $CI = 0.8991$. Since $CI < \xi$, the collective consensus needs to be improved by modifying some elements. In term of the identification strategy, using Eqs. (19)–(22), we find out the element with lowest consensus. First, the subgroup with lowest consensus is G_1 and the decision matrix level with lowest consensus is P^{12}. Then, in P^{12}, the alternative with lowest consensus is L_1. Finally, the element with lowest consensus is L_{13} in the decision matrix of expert e_{12}. Subsequently, based on the trust relationship, we revise the element $L_{13.}$ The social relationship among experts in G_1 is shown in Fig. 4.

In Fig. 4, we see the most trusted expert of e_{12} is e_9, namely $e_{12}^T = \{e_9\}$. In Eq. (24), θ is set 0.2. Thus, $AL(\overline{L}_{13}^{12}) = (1-0.2)L_{13}^{12}(p) + 0.2L_{13}^{9}(p) = \{s_{3.8}(0.8), s_4(0.2)\}$. Again, according to Eqs. (14) ~ (18), calculate the collective consensus degree $CI = 0.9026$. $CI > \xi$, the consensus reaching process ends.

Table 2 Consensus degrees on different levels

Subgroup	Consensus degrees on decision matrix level	Consensus degrees on subgroup level
G_1	$CI(P^1) =$ 0.8824$CI(P^5) =$ 0.9032$CI(P^9) = 0.9200$ $CI(P^{11}) = 0.8920$ $CI(P^{12}) =$ 0.8564$CI(P^{13}) =$ 0.9122	$CI(G_1) = 0.8944$
G_2	$CI(P^2) =$ 0.8640$CI(P^4) =$ 0.9200$CI(P^6) = 0.8994$ $CI(P^7) = 0.9300$ $CI(P^{14}) = 0.8848$ $CI(P^{16}) =$ 0.9012$CI(P^{17}) =$ 0.9120$CI(P^{18}) =$ 0.9420	$CI(G_2) = 0.9067$
G_3	$CI(P^3) =$ 0.8920$CI(P^8) =$ 0.9020$CI(P^{10}) =$ 0.8840 $CI(P^{15}) = 0.8960$ $CI(P^{19}) =$ 0.8624$CI(P^{20}) =$ 0.9400	$CI(G_3) = 0.8961$

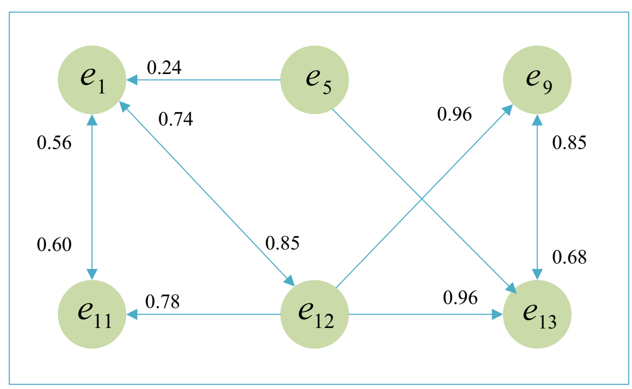

Fig. 4 Social relationship of experts in G_1

Step 4. Selection process. By Eq. (25), each decision matrix of subgroup can be obtained. According to the size of subgroup, the weights of subgroups are $\lambda_1 = 0.3$, $\lambda_2 = 0.4$ and $\lambda_3 = 0.3$. By Eq. (26), the collective decision matrix is

$$P^c = \begin{bmatrix} \{s_{2.1}(0.6), s_{2.4}(0.4)\} & \{s_{2.3}(0.5), s_{2.6}(0.5)\} & \{s_{2.7}(0.3), s_3(0.7)\} \\ \{s_2(0.2), s_{2.3}(0.8)\} & \{s_{1.3}(0.8), s_{1.6}(0.2)\} & \{s_2(0.7), s_{2.4}(0.3)\} \\ \{s_{1.5}(0.6), s_{1.9}(0.4)\} & \{s_2(0.7), s_{2.3}(0.3)\} & \{s_2(0.8), s_{2.3}(0.2)\} \\ \{s_{3.4}(0.5), s_{3.7}(0.5)\} & \{s_2(0.5), s_{2.4}(0.5)\} & \{s_3(0.4), s_{3.3}(0.6)\} \end{bmatrix}.$$

The weights of attributes are $\omega_1 = 0.5, \omega_2 = 0.3$ and $\omega_3 = 0.2$. By Eq. (4), the comprehensive evaluation of each alternative is

$$L_1 = \{s_{2.3}(0.3), s_{2.4}(0.36), s_{2.5}(0.2), s_{2.6}(0.14)\}$$

$$L_2 = \{s_{1.8}(0.112), s_{1.9}(0.524), s_2(0.316), s_{2.1}(0.048)\}$$

$$L_3 = \{s_{1.8}(0.564), s_{1.9}(0.036), s_2(0.376), s_{2.1}(0.024)\}$$

$$L_4 = \{s_{2.9}(0.1), s_3(0.25), s_{3.1}(0.4), s_{3.2}(0.25)\}.$$

According to Eq. (2), calculate the score functions of alternatives:

$$E(L_1) = 2.418,\ E(L_2) = 1.930,\ E(L_3) = 1.886,\ E(L_4) = 3.080$$

.

Thus, the priority of alternatives is: $x_4 \succ x_1 \succ x_2 \succ x_3$. Namely, the medicine D is the optimal drug.

Comparative analysis

In "Case study", we use a medical case to illustrate our research method. Besides, the proposed method possesses more superiority and rationality from three aspects.

First and foremost, the method to estimate missing values in decision matrix is more scientific and cautious. Different from previous researches, we consider the accuracy and reliability of referenced information when they are used to estimated missing values, i.e., the knowledge-match degree with reliability, which can guarantee the accuracy of estimated values. However, in some previous methods, it has been never taken into account. For instance, Liang et al. [20] utilized the collaborative filtering algorithm to estimate missing values while it only considers the similarity of nearest neighbors. In some cases, even though the similarity of referenced information is the highest, it may be low accuracy if we do not take the accuracy and reliability of referenced information into account. Thus, it is necessary to use knowledge-match degree with reliability to measure the accuracy and dependability of referenced information.

Besides, in the process of consensus measuring, we directly calculate the distance between any two objectives to obtain the consensus degree $CI(L_{ij}^{k_1})$(see Eqs. (14) ~ (18)), which can reduce the cost of time and calculation. However, in Li's research, first, the collective opinion needs to be obtained by aggregating the information of all subgroups. It needs to conduct complicated calculations including the integration of decision matrix of each subgroup and collective decision matrix. Especially, in LSGDM, there exist large number of evaluation information. If we adopt Li's method [30] to measure consensus level, the computed cost will be high. Hence, in this paper, the proposed method to measure consensus level is more simple and efficiency.

Finally, to improve the collective consensus, we design a new method to revise the elements based on trust relationship in the social network. Different from previous method, when modifying the element provided by the expert e_k, we refer to the evaluation information of the most trusted expert of e_k rather than all other trusted experts. That is to say, the proposed method can reduce the computation cost. Besides, the adjustment parameter θ is proposed to control the amplitude of adjustment. The smaller is the parameter θ, the more the reserved original information of e_k. According to actual situation and the trust degree, the parameter θ can be set flexibly.

Conclusion

In large-scale group decision making, due to various reasons, DMs may provide incomplete decision matrices. The missing information in those incomplete decision matrices also play an important role. In addition, the PLTS is a useful tool to describe the information uncertainty. Thus, under the environment of LSGDM, we design a new method to estimate the missing PLTS information in incomplete decision matrices based on knowledge-match degree with reliability. Compared with previous methods, the proposed estimation method can enhance the accuracy and reliability of estimated values. Besides, to apply the estimation method to LSGDM, a complete process of LSGDM is developed and we propose a new consensus reaching process based on trust degree. The new CRP in LSGDM can well solve the consensus problem. In this paper, the case of medicine selection is used to illustrate our decision method, which indicates that the proposed decision process can help managers and decision makers to solve some complicated multi-attribute problem, especially the decision case with much uncertain information. In addition to its application in the field of medicine, the proposed method can well be applied to bid evaluation [34], logistics provider selection [35], emergency decision and so on.

However, there are some limitations in this research. The first defect is that the computation of estimation method may be large and the reliability of PLTS is obtained by means of simple fuzzy entropy, which need to be improved in next work. Besides, with respect to the process of consensus improvement, the adjustment parameter is set randomly, which should be determined scientifically and reasonably. Therefore, in the next work, we will solve the above problems. We will have a deep research on incomplete decision matrix in an uncertain environment and explore the application related to the PLTS decision methods.

Declarations

Conflict of interests The authors disclose they have no financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

References

1. Chao X, Kou G, Li T, Peng Y (2018) Jie Ke versus AlphaGo: a ranking approach using decision making method for large-scale data with incomplete information. Eur J Oper Res 265:239–247
2. Liu B, Yu L, Ding R-X et al (2018) A decision-making method based on a two-stage regularized generalized canonical correlation analysis for complex multi-attribute large-group decision making problems. Journal of Intelligent & Fuzzy Systems 34:3941–3953
3. Chen Z-S, Liu X-L, Chin K-S et al (2021) Online-review analysis based large-scale group decision-making for determining passenger demands and evaluating passenger satisfaction: Case study of high-speed rail system in China. Information Fusion 69:22–39
4. Xu Y, Wen X, Zhang W (2018) A two-stage consensus method for large-scale multi-attribute group decision making with an application to earthquake shelter selection. Comput Ind Eng 116:113–129
5. Li S, Wei C. A large scale group decision making approach in healthcare service based on sub-group weighting model and hesitant fuzzy linguistic information. Computers & Industrial Engineering 2020; 144.
6. Liu P, Li Y (2019) Multi-attribute decision making method based on generalized maclaurin symmetric mean aggregation operators for probabilistic linguistic information. Comput Ind Eng 131:282–294
7. Zadeh LA (1965) Fuzzy Sets*. Inf Control 8:338–353
8. Torra V (2010) Hesitant fuzzy sets. Int J Intell Syst 25:529–539
9. Cui Chunsheng, Li Jielu, Zhenchun. Z. Measuring Product Similarity with Hesitant Fuzzy Set for Recommendation. Mathematics 2021; 9.
10. Rodriguez RM, Martinez L, Herrera F (2012) Hesitant fuzzy linguistic term sets for decision making. IEEE Trans Fuzzy Syst 20:109–119
11. Chen Z-S, Chin K-S, Li Y-L, Yang Y (2016) Proportional hesitant fuzzy linguistic term set for multiple criteria group decision making. Inf Sci 357:61–87
12. Chen Z-S, Zhang X, Pedrycz W *et al.* K-means clustering for the aggregation of HFLTS possibility distributions: N-two-stage algorithmic paradigm. Knowledge-Based Systems 2021; 227.
13. Pang Q, Wang H, Xu Z (2016) Probabilistic linguistic term sets in multi-attribute group decision making. Inf Sci 369:128–143
14. Ureña R, Chiclana F, Morente-Molinera JA, Herrera-Viedma E (2015) Managing incomplete preference relations in decision making: a review and future trends. Inf Sci 302:14–32
15. Herrera-Viedma E, Chiclana F, Herrera F, Alonso S (2007) Group decision-making model with incomplete fuzzy preference relations based on additive consistency. IEEE Trans Syst Man Cybern B Cybern 37:176–189
16. Chiclana F, Tapia García JM, del Moral MJ, Herrera-Viedma E (2013) A statistical comparative study of different similarity measures of consensus in group decision making. Inf Sci 221:110–123
17. Meng F, Chen X (2015) A new method for group decision making with incomplete fuzzy preference relations. Knowl-Based Syst 73:111–123
18. Capuano N, Chiclana F, Fujita H et al (2018) Fuzzy group decision making with incomplete information guided by social influence. IEEE Trans Fuzzy Syst 26:1704–1718
19. Guo W, Gong Z, Xu X, Herrera-Viedma E (2021) Additive and multiplicative consistency modeling for incomplete linear uncertain preference relations and its weight acquisition. IEEE Trans Fuzzy Syst 29:805–819
20. Liang D, He X, Xu Z. Multi-attribute dynamic two-sided matching method of talent sharing market in incomplete preference ordinal environment. Applied Soft Computing 2020; 93.
21. Xiao J, Wang X, Zhang H (2020) Managing personalized individual semantics and consensus in linguistic distribution large-scale group decision making. Information Fusion 53:20–34
22. Ding R-X, Wang X, Shang K, Herrera F (2019) Social network analysis-based conflict relationship investigation and conflict degree-based consensus reaching process for large scale decision making using sparse representation. Information Fusion 50:251–272
23. Tang M, Liao H, Xu J et al (2020) Adaptive consensus reaching process with hybrid strategies for large-scale group decision making. Eur J Oper Res 282:957–971
24. Wang Z, Rodríguez RM, Wang Y-m, Martínez L. A two-stage minimum adjustment consensus model for large scale decision making based on reliability modeled by two-dimension 2-tuple linguistic information. Computers & Industrial Engineering 2021; 151.
25. Wu Z, Xu J (2018) A consensus model for large-scale group decision making with hesitant fuzzy information and changeable clusters. Information Fusion 41:217–231
26. Zhang H, Dong Y, Chen X (2018) The 2-rank consensus reaching model in the multigranular linguistic multiple-attribute group decision-making. IEEE Transactions on Systems, Man, and Cybernetics: Systems 48:2080–2094

27. Dong Q, Cooper O (2016) A peer-to-peer dynamic adaptive consensus reaching model for the group AHP decision making. Eur J Oper Res 250:521–530
28. Zhang W, Xu Y, Wang H (2015) A consensus reaching model for 2-tuple linguistic multiple attribute group decision making with incomplete weight information. Int J Syst Sci 47:389–405
29. Liu Hongbin, Jiang Le, Zeshui X. Entropy Measures of Probabilistic Linguistic Term Sets. International Journal of Computational Intelligence Systems 2018; 11:45–57.
30. Li S, Wei C. A two-stage dynamic influence model-achieving decision-making consensus within large scale groups operating with incomplete information. Knowledge-Based Systems 2020; 189.
31. Tan X, Zhu J, Cabrerizo FJ, Herrera-Viedma E. A cyclic dynamic trust-based consensus model for large-scale group decision making with probabilistic linguistic information. Applied Soft Computing 2021; 100.
32. Hudec M, Mináriková E, Mesiar R *et al.* Classification by ordinal sums of conjunctive and disjunctive functions for explainable AI and interpretable machine learning solutions. Knowledge-Based Systems 2021; 220.
33. Holzinger A, Malle B, Saranti A, Pfeifer B (2021) Towards multimodal causability with graph neural networks enabling information fusion for explainable AI. Information Fusion 71:28–37
34. Chen Z-S, Zhang X, Rodríguez RM *et al.* Expertise-based bid evaluation for construction-contractor selection with generalized comparative linguistic ELECTRE III. Automation in Construction 2021; 125.
35. Chen Z-S, Zhang X, Govindan K *et al.* Third-party reverse logistics provider selection: A computational semantic analysis-based multi-perspective multi-attribute decision-making approach. Expert Systems with Applications 2021; 166.