Group Decision and Negotiation (2021) 30:1343–1372
https://doi.org/10.1007/s10726-020-09681-3

Hierarchical Punishment-Driven Consensus Model for Probabilistic Linguistic Large-Group Decision Making with Application to Global Supplier Selection

Sumin Yu[1] · **Zhijiao Du**[2] · **Xuanhua Xu**[3]

Published online: 3 August 2020

Abstract

Large-group decision making (LGDM) has attracted extensive attention and has been used to model complex decision problems. It is necessary to implement a consensus reaching process (CRP) due to the need to obtain a decision that is acceptable to the majority. The theory of probabilistic linguistic term sets (PLTSs) is very useful in addressing uncertain information in the decision-making process. In this paper, we develop a hierarchical punishment-driven consensus model for LGDM problems in the context of probabilistic linguistic information. The model has three stages. In the first stage, we define probabilistic linguistic large-group decision making. To improve the performance of PLTSs in the CRP, we redefine the rules governing their normalization and operations. In the second stage, the original large group is divided into several small subgroups by hierarchical clustering. In the third stage, we propose three levels of consensus measures and two adjustment strategies to refine the scope of measure and adjustment to the matrix element level. Then, a hierarchical punishment-driven consensus model is established that can provide guidance for adjustment and soften the human supervision of the CRP. Finally, a case study on global supplier selection illustrates the utility and applicability of the model, and a comparison with other linguistic models reveals its advantages.

Keywords Probabilistic linguistic large-group decision making (PL-LGDM) · Hierarchical punishment-driven consensus model (HPDCM) · Global supplier selection · Hierarchical clustering · Hard adjustment · Soft adjustment

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10726-020-09681-3) contains supplementary material, which is available to authorized users.

✉ Zhijiao Du
duzhj3@mail2.sysu.edu.cn; 131611141@csu.edu.cn

Sumin Yu
yusumin@szu.edu.cn

Xuanhua Xu
xuxh@csu.edu.cn

Extended author information available on the last page of the article

1 Introduction

Economic globalization and fierce market competition are driving companies to seek overseas suppliers (Awasthi et al. 2018; Kim et al. 2018). Global sourcing benefits companies, but it also increases the risks of supplier selection as the range of alternatives expands. Supplier selection represents a strategic managerial issue that is closely related to corporate performance and sustainability (Viswanadham and Samvedi 2013). Such a complex decision is difficult for a single person or a small organization to reach. In practice, a special committee (including relevant department heads and experts) or a special board meeting will be established to conduct in-depth assessments. This configuration is a typical large-group decision-making scenario where multiple decision makers (DMs) attempt to select a common solution from a set of alternatives. Large-group decision making (LGDM) has received much attention in the field of decision analysis. Generally, a GDM problem can be called an LGDM problem when the number of DMs is more than 20 (Ding et al. 2019; Gou et al. 2018). And yet in the studies of Palomares et al. (2014) and Wu et al. (2018), the illustrative examples included 50 experts. Situations involving large-scale DMs are considered to present new challenges for the decision process, such as scalability, constant preference supervision, different behaviors towards consensus, time cost (Labella et al. 2018; Rodríguez et al. 2018; Xu et al. 2019).

The DMs themselves play a central role in solving decision problems, but the form in which the evaluation information is expressed is also important for the decision output. Linguistic assessment information can express natural language well, and it has been one of the most commonly used forms of information expression in modeling realistic decision-making scenarios (Herrera et al. 1996, 2009; Liu et al. 2016; Merigó and Gil-Lafuente 2013; Zuheros et al. 2018). The characteristics of several existing linguistic models are described in Table 1. The traditional linguistic assessment information generally allows

Table 1 A review of the characteristics of several existing linguistic models

Linguistic formats	Number of linguistic terms	Probabilistic information	Sum of probability values
Linguistic term set (Herrera et al. 1996)	Single	N/A	N/A
Uncertain linguistic term set (Park et al. 2011)	A pair	N/A	N/A
Multi-granular linguistic term set (Morente-Molinera et al. 2019)	Single	N/A	N/A
Hesitant fuzzy linguistic term set (Rodríguez et al. 2012)	Several	N/A	N/A
Linguistic information based on discrete fuzzy numbers (Massanet et al. 2014)	Several	N/A	N/A
Possibility distribution-based HFLTS (Wu and Xu 2015)	Several	Completely known	1
Granulating linguistic information (Cabrerizo et al. 2018)	Single	N/A	N/A
Linguistic distribution assessment (Yu et al. 2018)	Several	Completely known	1
Probabilistic linguistic term set (Pang et al. 2016)	Several	Partially known	≤ 1

DMs to express their preferences with single linguistic terms. However, sometimes, it is difficult to depict complex qualitative information only by one linguistic term. For instance, a DM may evaluate an item as "very good", "good" or "somewhat good" but be unsure of how good the item is. To address such issues, Rodríguez et al. (2012) proposed the concept of hesitant fuzzy linguistic term sets (HFLTSs). HFLTSs collect all possible linguistic terms provided by the DMs, and all of these terms are assigned the same importance. In reality, a DM may have different preferences for multiple linguistic terms when judging an item. Therefore, Pang et al. (2016) extended the HFLTSs to a more general concept, named as probabilistic linguistic term sets (PLTSs). The PLTSs enable DMs to assign different weights to possible linguistic terms. To date, most studies on the GDM events with probabilistic linguistic information have focused on scenarios involving small-scale DMs, typically no more than 20. (e.g., Bai et al. 2018; Zhang et al. 2017).

A probabilistic linguistic large-group decision making (PL-LGDM) problem can be defined as a situation in which a large number of DMs seek to choose a common solution from a set of alternatives in the context of probabilistic linguistic information. We characterize its features as follows: (1) the number of DMs is large (usually more than 20); (2) most or all DMs provide their evaluation information by means of PLTSs; and (3) obtaining a high-consensus output requires the implementation of a consensus reaching process (CRP). Generally, solving a PL-LGDM problem requires the following four processes: the probabilistic linguistic representation of the evaluation information, clustering process, consensus reaching process and selection process.

To address the scalability issue in LGDM, we assume that among a large number of DMs, there will be subgroups of them with similar opinions (Rodríguez et al. 2018). A clustering process can be used to divide the large group into several small-scale subgroups based on opinion similarity. Many clustering methods have been proposed, such as the *K*-means algorithm (Wu and Xu 2018), fuzzy *c*-means algorithm (Bezdek et al. 1984), and hierarchical clustering algorithm (Dong et al. 2006; Johnson 1967). This study adopts the hierarchical clustering method to classify large-scale DMs.

The CRP is considered particularly important in LGDM because opinions among a large number of DMs can easily be controversial. In contrast to the ideal consensus that requires unanimity, the notion of soft consensus allows differences of opinion within a reasonable range (Kacprzyk and Fedrizzi 1988). Various consensus models have been proposed to address GDM problems under different situations. However, research on consensus in the context of probabilistic linguistic information has just begun. Zhang et al. (2017) designed a CRP with probabilistic linguistic preference relations. Wu and Liao (2019) proposed a consensus-based probabilistic linguistic gained and lost dominance score method. We find that the case studies in the above literature focus on GDM problems composed of three experts and do not involve large-scale DMs.

Faced with the challenges of classical consensus models for LGDM and the scarcity of research on consensus in PL-LGDM, this study develops a hierarchical punishment-driven consensus model and applies it to global supplier selection. The following efforts are made to address the consensus in PL-LGDM, which can overcome the scalability challenge and soften the human supervision of the CRP.

1 *Problem framework construction:* We define probabilistic linguistic large-group decision making and characterize its constituent elements. We present novel operational laws to preprocess probabilistic linguistic information.

2 *Clustering process:* A hierarchical clustering method is utilized to divide the large group into several small-scale subgroups. The cluster is regarded as the basic unit in the decision-making process.
3 *Consensus reaching process:* A hierarchical punishment-driven consensus model is employed to manage the differences among the opinions of clusters. According to the current group consensus index, the model can adopt different strategies to adjust the clusters' opinions.

The remainder of this paper is organized as follows: Sect. 2 reviews the basic concepts of PLTSs and CRPs and formulates a GDM event with PLTSs. In Sect. 3, a PL-LGDM problem is established and the hierarchical clustering method is used to classify large-scale DMs. In Sect. 4, a hierarchical punishment-driven consensus model is developed. Section 5 applies the proposed model to a case study on global supplier selection. Section 6 presents the comparative analysis and managerial implications. This paper ends with conclusions in Sect. 7.

2 Preliminaries

In this section, we first review some concepts related to PLTSs and present the construction of a GDM problem with probabilistic linguistic information. The CRP for LGDM problems is then described.

2.1 Group Decision Making with Probabilistic Linguistic Information

As an extension of HFLTSs, the concept of PLTSs was originally proposed by Pang et al. (2016), which enables DMs to express possible linguistic terms with different weights. Clearly, Pang et al. (2016) defined PLTSs on the basis of an additive linguistic evaluation scale, while Zhang et al. (2016) used PLTSs based on a subscript-symmetric linguistic evaluation scale.

Definition 1 (Pang et al. 2016) Let $S = \{s_\alpha | \alpha = -\tau, \ldots - 1, 0, 1, \ldots, \tau\}$ (τ is a positive integer) be a linguistic term set; then a PLTS is defined as

$$L(p) = \left\{ L^{(u)}\left(p^{(u)}\right) \middle| L^{(u)} \in S,\ p^{(u)} \geq 0,\ u = 1, 2, \ldots, \#L(p),\ \sum_{u=1}^{\#L(p)} p^{(u)} \leq 1 \right\}, \quad (1)$$

where $L^{(u)}\left(p^{(u)}\right)$ is a probabilistic linguistic element that includes the linguistic term $L^{(u)}$ and the associated probability $p^{(u)}$, and $\#L(p)$ is the number of different linguistic terms in $L(p)$.

A probabilistic linguistic element consists of two parts: a linguistic term and its corresponding probability. Based on this unique construct, a PLTS can be obtained in two ways: the statistical aggregation of single linguistic terms (see Pang et al. 2016) and the aggregation of multiple PLTSs (see Zhang et al. 2017). Formally, a multicriteria GDM problem with probabilistic linguistic information includes (1) a set of alternatives $X = \{x_1, x_2, \ldots, x_m\}$ $(m \geq 2)$, which are the possible solutions to the problem; (2) a set of criteria $A = \{a_1, a_2, \ldots, a_n\}$ $(n \geq 2)$ used to evaluate the alternatives in X; and (3) a set of DMs $E = \{e_1, e_2, \ldots, e_q\}$ $(q \geq 2)$ who provide judgments on the alternatives. Let $V_l =$

$\left(v_{l,ij}\right)_{m\times n}$ be the individual opinion given by DM $e_l(l = 1, 2, \ldots, q)$, where $v_{l,ij}$ represents the evaluation of DM e_l on alternative $x_i \in X$ with respect to criterion $a_j \in A$. $v_{l,ij}$ can be either a single linguistic term or a PLTS. The former expression represents a relatively precise evaluation value, while the latter reflects the DM's hesitation. To deal with the above different types of linguistic information, we attach a probability of 1 to the single linguistic term to convert it to a PLTS. For example, a given linguistic term s_3 can be changed to $\{s_3(1)\}$. In particular, if all individual opinions are expressed by means of single linguistic terms, a collective opinion can be obtained as a PLTS by aggregating linguistic terms (Song and Li 2019). In this study, we focus on a GDM in which most or all DMs provide evaluation information in the form of PLTSs. This is consistent with the fact that, due to the complexity of LGDM problems, it is difficult for DMs to express opinions by only using single linguistic terms.

2.2 Consensus Reaching Process in LGDM Problems

The increasing complexity of current decision-making scenarios requires the participation of large-scale DMs. Two main differences between classical GDM and LGDM are that (1) the latter case usually involves a larger number of DMs, and (2) implementing a CRP to achieve a high-consensus outcome is considered both more necessary and more difficult, as large-scale DMs tend to be controversial (Labella et al. 2018; Xu et al. 2019).

Consensus can be defined as "a state of mutual agreement among members of a group, where all legitimate concerns of individuals have been addressed to the satisfaction of the group'' (Saint and Lawson 1994). Obtaining a complete consensus is often unnecessary in practice and, as a consequence, the notion of soft consensus is introduced, which requires most (but perhaps not all) of the DMs to agree on the most important alternatives (Kacprzyk and Fedrizzi 1988; Palomares et al. 2014). Reaching a consensus is an iterative group discussion process in which some DMs must modify their opinions to bring them closer to the group opinion (Wu and Xu 2018). Two types of consensus measures have commonly been used: the first is based on the distance to the group opinion, and the second is based on the distances between individual opinions (Du et al. 2020; Labella et al. 2018; Wu and Xu 2016). A general scheme for a CRP is described as follows:

- *Problem framework configuration:* A GDM problem is established, including a finite set of alternatives, a set of DMs, and some pre-set important parameters (such as the consensus threshold).
- *Opinion gathering:* The opinions provided by DMs are gathered.
- *Consensus measure:* The group consensus index is calculated based on distance measures, and it reflects the level of agreement in the group.
- *Consensus control:* If the obtained consensus index is greater than the consensus threshold, this indicates that the desired consensus has been achieved and the selection process is concluded; otherwise, more consensus iterations should be carried out.
- *Consensus progress:* A procedure is adopted to identify the opinions that contribute less to the consensus and to guide the opinion adjustment process to increase the consensus index in the subsequent iterations. Then, another iteration starts by gathering opinions again.

To overcome the scalability challenge in LGDM problems, we consider that there will be DMs with similar opinions. In this study, a clustering algorithm is implemented to divide a large group and to weight subgroups before calculating the consensus measures. A hierarchical punishment-driven consensus model is then employed, by which the opinions of clusters are adjusted in the direction of the increasing consensus index.

3 Hierarchical Clustering Method for PL-LGDM Problems

In Sect. 3.1, we describe the establishment of a PL-LGDM problem and the preprocessing of decision information. To tackle the scalability challenge in PL-LGDM, Sect. 3.2 adopts a hierarchical clustering algorithm to classify the large group.

3.1 Problem Framework Configuration and Preprocessing of Decision Information

We first provide a description of the PL-LGDM scenario studied in this study. Let $X = \{x_1, x_2, \ldots, x_m\}$ be a finite set of alternatives, $A = \{a_1, a_2, \ldots, a_n\}$ be a set of criteria, $E = \{e_1, e_2, \ldots, e_q\}(q \geq 20)$ be a set of DMs, and S be a linguistic term set as before. This study considers a group to be a large group when the number of DMs exceeds 20. Let $V_l = (v_{l,ij})_{m\times n}$ be the individual decision matrix provided by DM $e_l(l = 1, 2, \ldots, q)$, where $v_{l,ij} = \left\{L_{l,ij}^{(u)}\left(p_{l,ij}^{(u)}\right)\middle|L_{l,ij}^{(u)} \in S, p_{l,ij}^{(u)} \geq 0,\ u = 1, 2, \ldots, \#v_{l.ij}, \sum_{u=1}^{\#v_{l,ij}} p_{l,ij}^{(u)} \leq 1\right\}$. $\#v_{l,ij}$ is the number of different linguistic terms in $v_{l,ij}$. Generally, there are benefit types and cost types in the criteria. By using the linguistic negation operator (Xu 2005), we convert the cost types to benefit types (for convenience, the converted result is still represented by $v_{l,ij}$) so that $L_{l,ij}^{(u)} = neg\left(L_{l,ij}^{(u)}\right)$ when a_j is a cost criterion.

To obtain the aggregation and distance measure of a PLTS, normalization is necessary. Generally, discrete linguistic terms are used to evaluate alternatives, while virtual linguistic terms only appear in operations and rankings (Xu 2009). Intuitively, in CRPs, discrete linguistic terms enable DMs to better understand the differences between individual opinions, which is conducive to making more reasonable adjustments. Therefore, we believe that before the selection process, the normalization and aggregation of PLTSs should be expressed in terms of discrete linguistic terms associated with probability values. Here, we redefine the normalization and operational laws of PLTSs. Normalization includes two steps:

1. *Granularity normalization:* All PLTSs contain the same linguistic terms;
2. *Probability normalization:* The sum of the probability values of all linguistic terms in each PLTS is 1.

Definition 2 (*Granularity normalization*) Let $L(p)_1$ and $L(p)_2$ be any two PLTSs, where $L(p)_1 = \left\{L_1^{(u)}\left(p_1^{(u)}\right)\middle|u = 1, 2, \ldots, \#L(p)_1\right\}$ and $L(p)_2 = \left\{L_2^{(u)}\left(p_2^{(u)}\right)\middle|u = 1, 2, \ldots, \#L(p)_2\right\}$. If there is a linguistic term $L_1^{(u)}$ in $L(p)_1$ that does not appear in $L(p)_2$, then add $L_1^{(u)}$ to $L(p)_2$. By this way, a new probabilistic linguistic element is obtained as $L_2^{(u)}\left(p_2^{(u)}\right)$, where $L_2^{(u)} = L_1^{(u)}$ and $p_2^{(u)} = 0$. Repeat the above steps until the two PLTSs have the same linguistic terms.

Definition 3 (*Probability normalization*) Given any PLTS $L(p)$ with $\sum_{u=1}^{\#L(p)} p^{(u)} < 1$, the associated PLTS $\dot{L}(p)$ is obtained as

$$\dot{L}(p) = \left\{ L^{(u)}\left(\dot{p}^{(u)}\right) \middle| L^{(u)} \in S,\ u = 1, 2, \ldots, \#L(p) \right\}, \tag{2}$$

where $\dot{p}^{(u)} = p^{(u)} \Big/ \sum_{u=1}^{\#L(p)} p^{(u)}$, $u = 1, 2, \ldots, \#L(p)$. $\dot{L}(p)$ is called the normalization form of $L(p)$.

For simplicity, the normalized PLTS (NPLTS) is still written as $L(p)$. The granularity of the NPLTS $L(p)$ is denoted as $\#\#L(p)$, where $\#\#L(p) \geq \#L(p)$.

Example 1 Given two PLTSs $L(p)_1 = \{s_{-1}(0.1), s_1(0.4)\}$ and $L(p)_2 = \{s_0(0.2), s_1(0.3), s_2(0.5)\}$, where the linguistic terms of both are drawn from $S_1 = \{s_{-2}, s_{-1}, s_0, s_1, s_2\}$, normalization is divided into two steps:

Step 1 Granularity normalization: $L(p)_1 = \{s_{-1}(0.1), s_0(0), s_1(0.4), s_2(0)\}$ and $L(p)_2 = \{s_{-1}(0), s_0(0.2), s_1(0.3), s_2(0.5)\}$. Thus, we have $\#\#L(p)_1 = \#\#L(p)_2$

Step 2 Probability normalization: $L(p)_1 = \{s_{-1}(0.2), s_0(0), s_1(0.8), s_2(0)\}$ and $L(p)_2 = \{s_{-1}(0), s_0(0.2), s_1(0.3), s_2(0.5)\}$

Next, we describe some basic operations.

Definition 4 Given any two NPLTSs, $L(p)_1 = \left\{ L_1^{(u)}\left(p_1^{(u)}\right) \middle| L_1^{(u)} \in S,\ u = 1, 2, \ldots, \#\#L(p)_1 \right\}$ and $L(p)_2 = \left\{ L_2^{(u)}\left(p_2^{(u)}\right) \middle| L_2^{(u)} \in S,\ u = 1, 2, \ldots, \#\#L(p)_2 \right\}$, $0 \leq \lambda_1, \lambda_2 \leq 1$. Then, we have

1. $L(p)_1 \oplus L(p)_2 = \left\{ L_3^{(u)}\left(p_3^{(u)}\right) \middle| L_3^{(u)} \in S,\ u = 1, 2, \ldots, \#\#L(p)_1 \right\}$,
2. $\lambda L(p)_1 = \left\{ L_1^{(u)}\left(\lambda p_1^{(u)}\right) \middle| u = 1, 2, \ldots, \#\#L(p)_1 \right\}$,
3. $\lambda_1 L(p)_1 \oplus \lambda_2 L(p)_2 = \left\{ L_4^{(u)}\left(p_4^{(u)}\right) \middle| L_4^{(u)} \in S,\ u = 1, 2, \ldots, \#\#L(p)_1 \right\}$,

where $L_3^{(u)} = L_1^{(u)}$, $p_3^{(u)} = p_1^{(u)} + p_2^{(u)} - p_1^{(u)} p_2^{(u)}$, $L_4^{(u)} = L_1^{(u)}$, $p_4^{(u)} = \lambda_1 p_1^{(u)} + \lambda_2 p_2^{(u)} - \lambda_1 \lambda_2 p_1^{(u)} p_2^{(u)}$.

Remark 1 Notice that sometimes the sum of the probability values of the PLTS obtained by Definition 4 is less than 1. In this case, we need to redo the process of probability normalization.

Example 2 (*Continued with Example 1*) Let $\lambda_1 = 0.4$ and $\lambda_2 = 0.6$; then, we have the following results:

Pang et al. (2016)'s operational law yields:

$$\begin{aligned} \lambda_1 L(p)_1 \oplus \lambda_2 L(p)_2 &= 0.4\{s_1(0.8), s_{-1}(0), s_{-1}(0.2)\} \oplus 0.6\{s_2(0.5), s_1(0.3), s_0(0.2)\} \\ &= \{0.4 \times 0.8 s_1 \oplus 0.6 \times 0.5 s_2, 0.4 \times 0 s_{-1} \oplus 0.6 \times 0.3 s_1, 0.4 \times 0.2 s_{-1} \oplus 0.6 \times 0.2 s_0\} \\ &= \{0.32 s_1 \oplus 0.3 s_2, 0 s_1 \oplus 0.18 s_1, 0.08 s_{-1} \oplus 0.12 s_0\} = \{s_{0.92}, s_{0.18}, s_{-0.08}\}, \end{aligned}$$

and our operational law yields:

$$\lambda_1 L(p)_1 \oplus \lambda_2 L(p)_2 = 0.4\{s_{-1}(0.2), s_0(0), s_1(0.8), s_2(0)\} \oplus 0.6\{s_{-1}(0), s_0(0.2), s_1(0.3), s_2(0.5)\}$$
$$= \{s_{-1}(0.0849), s_0(0.1273), s_1(0.4694), s_2(0.3184)\}.$$

From Example 2, we find that the probabilities of linguistic terms are missing if Pang et al.'s (2016) operational law is used. In this case, the unique feature of PLTSs compared with ordinary linguistic term sets disappears. More importantly, the virtual linguistic terms make it difficult for DMs to understand and determine the adjustments in the CRP. In our operational law, the ordinary linguistic terms remain unchanged, yet the associated probabilities always change.

To better apply PLTSs to GDM problems, the following aggregation operators are defined:

Definition 5 Let $L(p)_i = \left\{L_i^{(u)}\left(p_i^{(u)}\right) \middle| L_i^{(u)} \in S,\ k_i = 1,2,\ldots,\#\#L(p)_i\right\}$ $(i = 1,2,\ldots,m)$ be a set of m NPLTSs and θ_i be the weight of $L(p)_i$, such that $\theta_i \geq 0$, $i = 1,2,\ldots,m$, and $\sum_{i=1}^{m} \theta_i = 1$; then, the probabilistic linguistic weighted averaging (PLWA) operator is defined as

$$L(p) = PLWA\left(L(p)_1, L(p)_2, \ldots, L(p)_m\right) = \theta_1 L(p)_1 \oplus \theta_2 L(p)_2 \oplus \cdots \oplus \theta_m L(p)_m, \tag{3}$$

where $L(p) = \left\{L^{(u)}\left(p^{(u)}\right) \middle| L^{(u)} = L_1^{(u)},\ u = 1,2,\ldots,\#\#L(p)\right\}$, such that

$$p^{(u)} = \sum_{i=1}^{m} \theta_i p_i^{(u)} - \sum_{1 \leq i \leq j \leq m} \theta_i p_i^{(u)} \theta_j p_j^{(u)} + \sum_{1 \leq i \leq j \leq l \leq m} \theta_i p_i^{(u)} \theta_j p_j^{(u)} \theta_l p_l^{(u)} + \cdots + (-1)^{n-1} \prod_{i=1}^{m} \theta_i p_i^{(u)}. \tag{4}$$

Definition 6 Let $L(p)_1$ and $L(p)_2$ be any two NPLTSs; then, the distance between $L(p)_1$ and $L(p)_2$ is calculated by

$$d\left(L(p)_1, L(p)_2\right) = \frac{\left(\sum_{u=1}^{\#\#L(p)_1} \left(\left|I_1^{(u)}\right| \left|p_1^{(u)} - p_2^{(u)}\right|\right)^{\lambda}\right)^{1/\lambda}}{\#\#L(p)_1}, \tag{5}$$

where $\lambda > 0$ and $I_1^{(u)}$ is the subscript of the linguistic term $L_1^{(u)}$. In this study, we adopt the Euclidean distance, i.e., $\lambda = 2$.

It is clear that the distance measures have the following properties:

1 $0 \leq d\left(L(p)_1, L(p)_2\right) \leq 1$,
2 $d\left(L(p)_1, L(p)_2\right) = 0$, if $L(p)_1 = L(p)_2$,
3 $d\left(L(p)_1, L(p)_2\right) = d\left(L(p)_2, L(p)_1\right)$.

3.2 Hierarchical Clustering Algorithm

Clustering is a widely used methodology for analyzing and processing large-scales DMs, and it is considered to be effective in addressing the scalability challenge in LGDM problems. In this study, we adopt a hierarchical clustering method to classify the large-scale DMs into $K(1 \le K \le q)$ subgroups. Hierarchical clustering treats each data object as a separate cluster and then finds clusters with small distances to merge in each iteration. This process repeats until a designated number of clusters is reached or there is only one cluster. The procedure for the hierarchical clustering method is described in Algorithm 1.

Algorithm 1 Hierarchical clustering method with probabilistic linguistic information

Input: Normalized individual opinions $R_l\,(l=1,2,\ldots,q)$ and the designated number of clusters K.

Output: Clusters $c_1, c_2, \ldots, c_K$.

Step 1: Consider each DM as a cluster, denoted as $c_1, c_2, \ldots, c_{20}$.

Step 2: Use Eq. (5) to calculate the distances among clusters, denoted as $d(R_l, R_h)$, $l,h=1,2,\ldots,q$.

Step 3: Select the minimum distance, i.e., $d\left(R_{l^*}, R_{h^*}\right)=\min\left\{d\left(R_l, R_h\right)\right\}$, and merge clusters c_{l^*}, c_{h^*} into a single cluster c_{k^*}.

Step 4: Compute the distances between the new cluster and the old clusters, obtained by

$$d\left(R_{k^*}, R_l\right)=\max\left\{d\left(R_{l^*}, R_l\right), d\left(R_{h^*}, R_l\right)\right\}. \tag{6}$$

Step 5: Repeat steps 3-4 until the number of clusters equals to K.

Step 6: Output the clusters $c_1, c_2, \ldots, c_K$.

Theorem 1 *The time complexity of Algorithm 1 is* $O(q^3)$.

Proof First, it will take $O(q^2)$ time to calculate the distances between each pair of clusters. Once a new cluster is determined, the distance between the new cluster and each of the other clusters needs to be calculated. It can be seen that Algorithm 1 must be carried out for q iterations and that each iteration updates the distances, so the time complexity is the cube of the number of original clusters, i.e., $O(q^3)$. □

Calculating the weights of the clusters is a necessary step because it is closely related to the generation of group opinions. Inspired by Rodríguez et al. (2018), this study takes into account the size and cohesion. The size refers to the number of DMs in a cluster. Based on the majority principle, the larger the size of a cluster is, the more weight it should be given. However, the representation of the size should be adjusted according to the number of DMs involved in the LGDM problem. Based on computing with words (Quesada et al. 2015), Rodríguez et al. (2018) modeled the size with the fuzzy membership function μ_{size} shown in Fig. 1, where the universe of discourse is the number of DMs in a cluster and the membership degree reflects the cluster's influence on all the DMs in terms of the number of DMs. The points a and b of the membership function depend on the numbers of alternatives and DMs in the LGDM scenario. The highest membership degree is for values above b, the lowest membership degree is for values below a, and varying importance is assigned between a and b.

The cohesion represents the consensus index of the DMs' opinions in a cluster. The more coherent the individual opinions within a cluster are, the more weight we give the cluster.

Definition 7 For cluster c_k including n_k DMs, the cohesion of the cluster is defined as

$$cohesion(c_k) = 1 - \frac{1}{n_k(n_k-1)/2}\sum_{h=1}^{n_k-1}\sum_{l=h+1}^{n_k} d(R_l, R_h), \tag{7}$$

where $e_l, e_h \in c_k$. Clearly, $0 \leq cohesion(c_k) \leq 1$.

By integrating size and cohesion, Rodríguez et al. (2018) defined a function to calculate the weights of clusters, which could be flexibly adapted to a specific LGDM problem.

Definition 8 (Rodríguez et al. 2018) Let $Y_{c_k} = \{y_1, y_2\}$ be the values obtained for cohesion and size, respectively, where $y_1, y_2 \in [0, 1]$; then, the fusion value of the cohesion and size is calculated by

$$\varphi(Y_{c_k}) = (1 + y_2)^{y_1 \beta}, \tag{8}$$

where the parameter $\beta(\beta > 0)$ is used to adjust the effect of cohesion in calculating the weight of the cluster.

The aggregated value $\varphi(Y_{c_k})$ reflects the relevance of cluster c_k. Therefore, the weight vector of the clusters, $\theta = (\theta_1, \theta_2, \ldots, \theta_K)^T$, can be obtained from

$$\theta_k = \frac{\varphi(Y_k)}{\sum_{k=1}^{K}\varphi(Y_k)}. \tag{9}$$

Clearly, $0 \leq \theta_k \leq 1$ $(k = 1, 2, \ldots, K)$ and $\sum_{k=1}^{K}\theta_k = 1$. By using Eq. (3), the cluster opinion is obtained; i.e., $G_k = \left(g_{k,ij}\right)_{m \times n}$, where $g_{k,ij} = PLWA_\theta\left(r_{1,ij}, \ldots, r_{n_k,ij}\right)$.

We use the case in Sect. 5 to illustrate the weight calculation. In this study, we consider 10% of DMs to define the point a and the number of DMs divided by the number of alternatives to define the point b: $a = round(q \cdot 0.1) = 2$, $b = round(q/m) = 4$, where $round(\cdot)$ is the round function. Table 2 shows the values of size, membership degree, and cohesion for each cluster. Figure 2 shows the weights of clusters when different values of

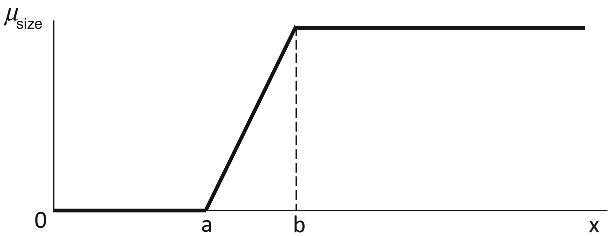

Fig. 1 Membership function for the cluster size

the parameter β are used to solve Eq. (8). We find that although clusters c_2 and c_3 have the same membership degree in terms of size, they are assigned different weights due to the different values of cohesion among their members. As shown in Fig. 2, when the value of β increases, the weight of cluster c_2 increases more than that of cluster c_3. This means that the parameter β can adjust the degree of the influence of cohesion on the weight.

4 Hierarchical Punishment-Driven Consensus Model in PL-LGDM Problems

We first describe the consensus measure. Section 4.2 presents punishment-driven consensus iterations. In Sect. 4.3, a hierarchical punishment-driven consensus model is given.

4.1 Consensus Measure

The consensus measure is designed to compute the differences among DMs' opinions. As previously discussed, the calculation of the consensus measure can be based on the distance to the group opinion or the distances among individual opinions. Xu et al. (2019) gave a detailed comparison of these two measures. The former measure is heavily influenced by the weights of the clusters. In this case, the CRP tends to target the opinions of low-weight clusters because their opinions contribute less to the group opinion. So that our model uses the latter measure, which is completely based on the differences among cluster's opinions, without considering the influence of the clusters' weights. We present the three-level consensus index measure below.

Level 1 *Individual consensus index at the matrix element level.* The consensus index of a cluster with respect to the others on alternative x_i under criterion a_j is

$$ICI_{k,ij}^{t} = 1 - \frac{1}{K-1}\sum_{h=1,h\neq k}^{K} d\left(g_{k,ij}^{t}, g_{h,ij}^{t}\right), \tag{10}$$

where $d\left(g_{k,ij}^{t}, g_{h,ij}^{t}\right)$ is the distance between $g_{k,ij}^{t}$ and $g_{h,ij}^{t}$. Clearly, $0 \le ICI_{k,ij}^{t} \le 1$. The greater the value of $ICI_{k,ij}^{t}$ is, the higher the consensus index of the matrix element $g_{k,ij}^{t}$.

Level 2 *Individual consensus index at the matrix level*. The consensus index of a cluster opinion with respect to the others is

$$ICI_{k}^{t} = \frac{1}{m \times n}\sum_{i=1}^{m}\sum_{j=1}^{n} ICI_{k,ij}^{t}. \tag{11}$$

Table 2 The values of size, membership degree, and cohesion for each cluster

	Size	Membership degree	Cohesion
c_1	4	1	0.8427
c_2	5	1	0.8446
c_3	4	1	0.8173
c_4	4	1	0.8196
c_5	2	0	0.8286
c_6	1	0	1

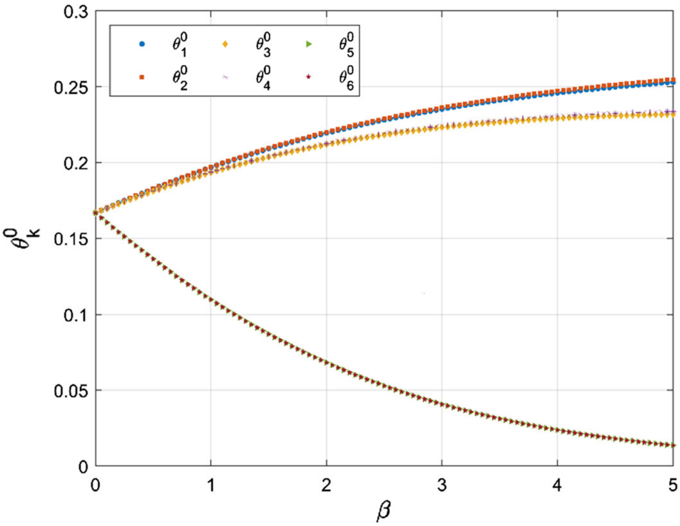

Fig. 2 Weights of the clusters as β varies. *Note*: set $a = 2$, $b = 4$

Clearly, $0 \leq ICI_k^t \leq 1$. The greater the value of ICI_k^t is, the higher the consensus index of cluster c_k.

Level 3 *Consensus index at the group level.* The group consensus index is computed by

$$GCI^t = \frac{1}{K}\sum_{k=1}^{K} ICI_k^t. \tag{12}$$

Clearly, $0 \leq GCI^t \leq 1$. The higher the value of GCI^t is, the higher the consensus among the clusters.

Usually, we need to define the consensus threshold (denoted as $\overline{GCI}$, where $0 \leq \overline{GCI} \leq 1$), which is used to determine whether the CRP can terminate. Calculating the threshold is an important matter, and many related studies have been published (see Labella et al. 2018; Xu et al. 2015; Zhang et al. 2017).

Remark 2 The three-level consensus index measure provides inspiration for constructing a hierarchical consensus model. We can either adjust all elements of the opinion (based on the individual consensus indexes at the matrix level) in each iteration or adjust only some elements of the opinion (based on the individual consensus index at the matrix element level).

4.2 Punishment-Driven Consensus Iterations

There are usually two ways to adjust opinions: the cluster opinion that has the largest distance from the group opinion is adjusted in each iteration (e.g. Wu and Liao 2019); or the cluster opinions that contribute less to the consensus are adjusted (e.g. Rodríguez et al. 2018). The former way sorts the individual consensus indexes and adjusts the opinions of clusters accordingly until the current group consensus index meets the consensus threshold.

Compared with the latter way, the former has the advantage in terms of reducing the number of adjusted opinions. This paper selects the first method. Since this study takes the cluster as the basic decision-making unit, the opinions of DMs have been merged into the opinions of clusters before the CRP begins. This means that the next step for DMs is to focus on discussing the adjustment of the clusters' opinions. Suppose cluster $c_{k^\cdot}$ has the lowest consensus index in the t-th iteration, i.e., $ICI_{k^\cdot}^t = \min\{ICI_k^t\}$. Cluster $c_{k^\cdot}$ provides the adjustment coefficient $\delta_{k^\cdot}^t$ $(0 \le \delta_{k^\cdot}^t \le 1)$ to make its opinion closer to the others. To provide guidance to the cluster regarding the adjustment coefficient and soften the human supervision of the CRP, we propose the concept of the punishment coefficient, which is the degree to which a cluster opinion is required to move towards the collective opinion.

Definition 9 Given the consensus threshold $\overline{GCI}$ and identified cluster consensus index $ICI_{k^\cdot}^t$ in the t-th iteration, the punishment coefficient is defined as

$$pc_{k^\cdot}^t = \left(\frac{\overline{GCI} - ICI_{k^\cdot}^t}{\overline{GCI}}\right)^\sigma, \tag{13}$$

where $\sigma(0 \le \sigma \le 1)$ is the power of the punishment coefficient and is used to indicate the urgency of reaching a consensus. Clearly, $0 \le pc_{k^\cdot}^t \le 1$.

Figure 3 shows the distribution of the punishment coefficient for different values of σ. We set $\sigma = 1$ as the benchmark. We find that the larger the value of σ is, the smaller the punishment coefficient $pc_{k^\cdot}^t$. $\sigma = 0$ is an extreme case, indicating that the identified cluster opinion is completely replaced by the collective opinion. Another important function of the punishment coefficient is that it can be used as the lower limit of the adjustment coefficient. If the given adjustment coefficient is less than the punishment coefficient, the cluster is required to use the punishment coefficient to adjust its opinion; i.e., if $\delta_{k^\cdot}^t < pc_{k^\cdot}^t$, then set $\delta_{k^\cdot}^t = pc_{k^\cdot}^t$. This is why our model is called a punishment-driven consensus model. In this way, the model not only provides guidance for opinion adjustment, but also designs a semi-automatic mandatory adjustment mechanism, which can soften the human supervision of the CRP.

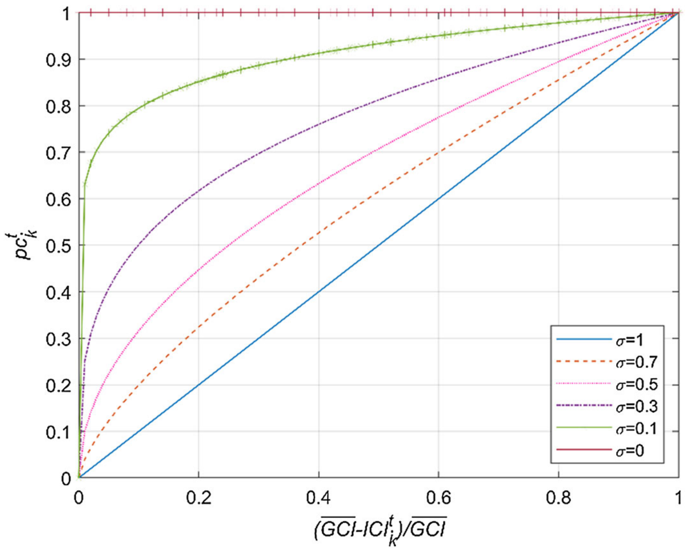

Fig. 3 Distribution of the punishment coefficient for different values of σ

Based on the operational laws proposed in this paper, the clusters' opinions are always expressed in the form of PLTSs before the selection process begins, which include discrete linguistic terms and their corresponding probability values. We define the adjusted opinion of the CRP as follow.

Definition 10 Let $G_{\dot{k}}^{t}$ be the identified opinion with the lowest consensus index in the t-th iteration, $\delta_{\dot{k}}^{t}$ be the given adjustment coefficient, and $pc_{\dot{k}}^{t}$ be the punishment coefficient. If $GCI^{t} < \overline{GCI}$, then the opinion $G_{\dot{k}}^{t}$ should be adjusted as $G_{\dot{k}}^{t+1} = \left(g_{\dot{k},ij}^{t+1}\right)_{m\times n}$, such that

$$g_{\dot{k},ij}^{t+1} = \begin{cases} \delta_{\dot{k}}^{t} g_{c\to\dot{k},ij}^{t} \oplus \left(1-\delta_{\dot{k}}^{t}\right) g_{\dot{k},ij}^{t} & \delta_{\dot{k}}^{t} \geq pc_{\dot{k}}^{t} \\ pc_{\dot{k}}^{t} g_{c\to\dot{k},ij}^{t} \oplus \left(1-pc_{\dot{k}}^{t}\right) g_{\dot{k},ij}^{t} & \delta_{\dot{k}}^{t} < pc_{\dot{k}}^{t} \end{cases}, \tag{14}$$

where $g_{c\to\dot{k},ij}^{t} = \oplus_{h=1,h\neq\dot{k}}^{K} g_{k,ij}^{t} \bar{\theta}_{h}^{t}$ is the collective opinion used to guide the adjustment of $G_{\dot{k}}^{t}$ and $\bar{\theta}_{h}^{t}$ is the remaining cluster c_h's weight, which satisfies $\bar{\theta}_{h}^{t} = \theta_{h}^{t} \Big/ \sum_{k=1,k\neq\dot{k}}^{K} \theta_{k}^{t}$. The addition operation "$\oplus$" is interpreted as

$$L_{\dot{k},ij}^{(u),t+1} = \begin{cases} \delta_{\dot{k}}^{t} L_{c\to\dot{k},ij}^{(u),t} \oplus \left(1-\delta_{\dot{k}}^{t}\right) L_{\dot{k},ij}^{(u),t} & \delta_{\dot{k}}^{t} \geq pc_{\dot{k}}^{t} \\ pc_{\dot{k}}^{t} L_{c\to\dot{k},ij}^{(u),t} \oplus \left(1-pc_{\dot{k}}^{t}\right) L_{\dot{k},ij}^{(u),t} & \delta_{\dot{k}}^{t} < pc_{\dot{k}}^{t} \end{cases}, \quad u = 1,\ldots,\#\#g_{\dot{k},ij}^{t}, \tag{15}$$

$$p_{\dot{k},ij}^{(u),t+1} = \begin{cases} \delta_{\dot{k}}^{t} p_{c\to\dot{k},ij}^{(u),t} + \left(1-\delta_{\dot{k}}^{t}\right) p_{\dot{k},ij}^{(u),t} & \delta_{\dot{k}}^{t} \geq pc_{\dot{k}}^{t} \\ pc_{\dot{k}}^{t} p_{c\to\dot{k},ij}^{(u),t} + \left(1-pc_{\dot{k}}^{t}\right) p_{\dot{k},ij}^{(u),t} & \delta_{\dot{k}}^{t} < pc_{\dot{k}}^{t} \end{cases}, \quad u = 1,\ldots,\#\#g_{\dot{k},ij}^{t}, \tag{16}$$

where $p_{c\to\dot{k},ij}^{(u),t} = \sum_{h=1,h\neq\dot{k}}^{K} p_{h,ij}^{(u),t} \bar{\theta}_{h}^{t}$, $L_{c\to\dot{k},ij}^{(u),t} = L_{\dot{k},ij}^{(u),t}$, $u = 1,\ldots,\#\#g_{\dot{k},ij}^{t}$, $i = 1,2,\ldots,m$, $j = 1,2,\ldots,n$.

4.3 Hierarchical Punishment-Driven Consensus Iterations

In Sect. 4.1, we divide the consensus measures into three levels. The group consensus index is used to determine whether the current consensus satisfies the threshold. The other consensus indexes can be used to determine which clusters' opinions or matrix elements need to be adjusted. Different matrix elements in a cluster opinion usually have different consensus indexes. Therefore, it is reasonable to use different adjustment coefficients (or punishment coefficients).

Definition 11 For the consensus threshold $\overline{GCI}$ and the identified cluster $c_{\dot{k}}$ in the t-th iteration, the punishment coefficient at the matrix element level is defined as

$$pc_{\dot{k},ij}^{t} = \begin{cases} \left(\dfrac{\overline{GCI} - ICI_{\dot{k},ij}^{t}}{\overline{GCI}}\right)^{\sigma} & ICI_{\dot{k},ij}^{t} < \overline{GCI} \\ 0 & ICI_{\dot{k},ij}^{t} \geq \overline{GCI} \end{cases}. \tag{17}$$

Let σ be the same as in Definition 9. Clearly, $0 \le pc^{t}_{\dot{k},ij} \le 1$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

Based on Eq. (14), we use the following formula to adjust the matrix elements in the opinion $G^{t}_{\dot{k}}$:

$$g^{t+1}_{\dot{k},ij} = \begin{cases} \delta^{t}_{\dot{k},ij} g^{t}_{c \to \dot{k},ij} \oplus \left(1 - \delta^{t}_{\dot{k},ij}\right) g^{t}_{\dot{k},ij} & \delta^{t}_{\dot{k},ij} \ge pc^{t}_{\dot{k},ij} \\ pc^{t}_{\dot{k},ij} g^{t}_{c \to \dot{k},ij} \oplus \left(1 - pc^{t}_{\dot{k},ij}\right) g^{t}_{\dot{k},ij} & \delta^{t}_{\dot{k},ij} < pc^{t}_{\dot{k},ij} \end{cases}, \tag{18}$$

where $\delta^{t}_{\dot{k},ij}\left(0 \le \delta^{t}_{\dot{k},ij} \le 1\right)$ is the adjustment coefficient given by cluster $c_{\dot{k}}$ that is used to modify the probability value of the matrix element of position (i,j) in $G^{t}_{\dot{k}}$.

In general, although the identified cluster has the lowest consensus index at the individual matrix level, it does not necessarily indicate that the consensus indexes of all elements in the cluster opinion do not satisfy the threshold. Using the same coefficient to adjust all the matrix elements may result in adjusting elements that should not be adjusted, or should not be adjusted by such a large amount, which will result in opinion distortion. We consider opinion distortion to have two parts: the adjustment amount of the opinion and the proportions of the adjusted matrix elements in the opinion. The latter part reflects the number of matrix elements that were adjusted but should not have been.

Definition 12 Based on Eq. (5), the adjustment amount of the identified opinion $G^{t}_{\dot{k}}$ after the t-th iteration is calculated by

$$\overline{AA}^{t+1}_{\dot{k}} = d\left(G^{t+1}_{\dot{k}}, G^{t}_{\dot{k}}\right). \tag{19}$$

Clearly, $0 \le \overline{AA}^{t+1}_{\dot{k}} \le 1$. A larger value of $\overline{AA}^{t+1}_{\dot{k}}$ indicates a greater adjustment in going from $G^{t}_{\dot{k}}$ to $G^{t+1}_{\dot{k}}$.

Definition 13 Let $IX\left(G^{t}_{\dot{k}}, G^{t+1}_{\dot{k}}\right)$ be the number of matrix elements that have been adjusted in the t-th iteration. The proportion of adjusted matrix elements is obtained by

$$\widehat{AA}^{t+1}_{\dot{k}} = \frac{IX\left(G^{t}_{\dot{k}}, G^{t+1}_{\dot{k}}\right)}{m \times n}. \tag{20}$$

Definition 14 Given the adjustment amount and the proportion of adjusted matrix elements, the distortion degree of the opinion $G^{t}_{\dot{k}}$ is defined as

$$DD^{t+1}_{\dot{k}} = \frac{\overline{AA}^{t+1}_{\dot{k}} + \widehat{AA}^{t+1}_{\dot{k}}}{2}. \tag{21}$$

Clearly, $0 \le DD^{t+1}_{\dot{k}} \le 1$. A larger value of $DD^{t+1}_{\dot{k}}$ represents a higher distortion degree in going from $G^{t}_{\dot{k}}$ to $G^{t+1}_{\dot{k}}$.

We propose the soft adjustment strategy, which allows different punishment coefficients to be used to adjust the matrix elements. Conversely, hard adjustment strategy requires that all matrix elements be adjusted with the same coefficient. Employing hard adjustment often leads to adjusting a larger number of matrix elements than employing soft adjustment in each consensus iteration. However, when the CRP begins, there are usually a large number of matrix elements in the identified opinion whose consensus indexes do not meet the requirements. Too much time will be spent reviewing the consensus index for each element and providing the adjustment coefficient if soft adjustment is used. Therefore, we consider that the hard adjustment strategy can be adopted first to improve the consensus indexes at different levels. When the group consensus index approaches the consensus threshold closely, the soft adjustment strategy is used to balance the relationship between opinion distortion and consensus improvement. To determine which adjustment strategy to adopt, we introduce a new parameter, the attainment rate of the consensus index, denoted as $ARoCI(0 \leq ARoCI \leq 1)$. This parameter elevates the punishment-driven consensus model to a hierarchical consensus model.

Definition 15 Given the group consensus index GCI^t and consensus threshold $\overline{GCI}$, the attainment rate of the consensus index is defined as

$$ARoCI^t = \frac{GCI^t}{\overline{GCI}}. \tag{22}$$

Clearly, $GCI^0 / \overline{GCI} \leq ARoCI^t \leq 1$. $ARoCI^t$ reflects the percentage of the current consensus index that satisfies the consensus threshold. The higher the group consensus index, the greater the value of $ARoCI^t$. DMs often set the threshold $\overline{ARoCI} \in \left[GCI^0 / \overline{GCI}, 1\right]$ for $ARoCI^t$.

The essential difference between hard adjustment and soft adjustment is whether a matrix element that satisfies the consensus threshold should be adjusted when the consensus index at the matrix level does not meet the consensus threshold. We define the following rule.

Definition 16 Let $IX\left(G^t_{k\cdot}\right)$ be the number of matrix elements that contribute less to the consensus. When the conditions $ARoCI^t \geq \overline{ARoCI}$ and $IX\left(G^t_{k\cdot}\right) / m \times n \leq \overline{IX}$ are met, the soft adjustment strategy is adopted; otherwise, the hard adjustment strategy is used. $\overline{IX}\left(0 \leq \overline{IX} \leq 1\right)$ is the threshold of $IX\left(G^t_{k\cdot}\right) / m \times n$.

If $ARoCI^t \geq \overline{ARoCI}$ and $IX\left(G^t_{k\cdot}\right) / m \times n \leq \overline{IX}$, this indicates that the group consensus index is high enough, and the number of elements in the identified opinion that need to be adjusted drops significantly. In this case, the soft adjustment strategy is used. Clearly, when we set $\overline{ARoCI} = 1$ or $\overline{IX} = 0$, the hard adjustment strategy will be used throughout the CRP. By combining the hard adjustment and soft adjustment, the hierarchical punishment-driven consensus model (HPDCM) is obtained. Algorithm 2 presents the implementation process for the model.

Algorithm 2 Hierarchical punishment-driven consensus model for PL-LGDM problems

Input: The initial individual opinions $V_l^0\left(l=1,2,\ldots,q\right)$, the consensus threshold $\overline{GCI}$, the threshold of the attainment rate of the consensus index $\overline{ARoCI}$, the weight vector of the criteria ω, and parameters $\overline{IX}$, a, b, and σ.

Output: The final number of iterations, the final group opinion, and the final ranking of the alternatives.

Process 1: Preprocessing of the decision information.

Implement granularity normalization and probability normalization to obtain the standardized individual opinions $R_l^0\left(l=1,2,\ldots,q\right)$.

Process 2: Clustering process.

Adopt Algorithm 1 to classify the initial large group into K clusters. Use the PLWA operator to obtain the clusters' opinions $G_k^0\left(k=1,2,\ldots,K\right)$, and compute the weight vector of the clusters with Eq. (9).

Process 3: Consensus reaching process.

Step 3.1: Set $t=0$.

Step 3.2: Consensus measure.

Compute the consensus measures via Eqs. (10)-(12). If $GCI^t \geq \overline{GCI}$, then proceed to Step 3.4; otherwise, proceed to the next step.

Step 3.3: Opinion adjustment.

Calculate the attainment rate of the consensus index $ARoCI^t$ and the rate of the number of matrix elements that have small consensus indexes $IX\left(G_k^t\right)/m\times n$. If $ARoCI^t \geq \overline{ARoCI}$ and $IX\left(G_k^t\right)/m\times n \leq \overline{IX}$, then use Eq. (18) to perform the adjustment; otherwise, use Eq. (14) to adjust the identified opinion. Let $t=t+1$, and then return to Step 3.2.

Step 3.4: Let $t^*=t$. Output the final number of iterations t^* and the final opinions of the clusters $G_k^{t^*}\left(k=1,2,\ldots,K\right)$.

Process 4. Selection process.

Step 4.1: Aggregate all the clusters' opinions to obtain the group opinion R_c via Eq. (3).

Step 4.2: Calculate the overall criterion values $Z\left(x_i\right)\left(i=1,2,\ldots,m\right)$ based on Eq. (3).

Step 4.3: Use the rule in Pang et al. (2016) to compare the overall criterion values and obtain the ranking of the alternatives.

5 Case Study

In this section, we apply the HPDCM to a PL-LGDM problem concerning global supplier selection.

5.1 Problem Description

ABC, an electronics manufacturer located in Quanzhou, China, is reassessing the suppliers it currently works with. Considering the rising cost of domestic labor and the policy advantages of the Belt and Road cooperation initiative, the company has decided to select a material supplier located along this route and add it to the supplier list. Five alternatives $X = \{x_1, x_2, x_3, x_4, x_5\}$ have been preselected for further evaluation, and their geographical locations are Alma-Ata, Ho Chi Minh City, Manila, Calcutta and Minsk, as shown in Fig. 4. A special committee of 20 relevant department heads and experts is formed to evaluate the alternative suppliers. Drawing from the research (Awasthi et al. 2018; Reefke and Sundaram 2018; Viswanadham and Samvedi 2013; Yücenur et al. 2011) and business practices, three comprehensive criteria are used: commercial factors (c_1), such as quality, price, quantity and delivery time; sustainability dimensions (c_2), including economic, environmental, and social bottom lines; and global risks (c_3), such as currency fluctuations, political instability, terrorism, and cultural incompatibility. The weight vector of the criteria is set as $\omega = (0.4, 0.3, 0.3)^T$. The DMs provide their evaluations by means of PLTSs, in which the linguistic evaluation scale

$$S = \left\{ \begin{array}{l} s_{-3} = Extremely\,poor,\ s_{-2} = Poor,\ s_{-1} = Somewhat\,poor,\ s_0 = Neutral, \\ s_1 = Somewhat\,good,\ s_2 = Good,\ s_3 = Extremely\,good \end{array} \right\}$$

is used. The original individual opinions are given in Online Resource.

Fig. 4 Geographical distribution of preselected suppliers along the Belt and Road. *Note*: The original source of this map is http://www.sohu.com/a/111613685_472018

5.2 Decision Process

Here, the HPDCM is used to manage the processes of consensus and selection. The following steps are involved. To save space, only the main results are listed. A detailed textual description of the decision process is given in Online Resource.

Input: The initial individual opinions $V_l^0 (l = 1, 2, \ldots, 20)$, the weight vector of the criteria ω, the consensus threshold $\overline{GCI} = 0.88$, and other parameters $\overline{ARoCI} = 0.95$, $\overline{IX} = 1$, and $\sigma = 1/3$

Output: The final ranking of the alternatives

Process 1 Preprocessing of the decision information.

Normalize the individual opinions. Here, we just give the normalized opinion of cluster c_1:

$$R_1^0 = \begin{pmatrix}
\left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0), s_0(0.5), \\ s_1(0), s_2(0.4), s_3(0.1) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(1), s_{-1}(0), s_0(0), \\ s_1(0), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.1), s_0(0), \\ s_1(0), s_2(0.9), s_3(0) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.5), s_0(0), \\ s_1(0), s_2(0.5), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0), s_0(0.4), \\ s_1(0.6), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.5), s_0(0), \\ s_1(0.5), s_2(0), s_3(0) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.3), s_0(0), \\ s_1(0.7), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0.2), s_{-1}(0.3), s_0(0), \\ s_1(0.5), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.7), s_0(0), \\ s_1(0.3), s_2(0), s_3(0) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.2), s_0(0.8), \\ s_1(0), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0), s_0(0), \\ s_1(1), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.6), s_0(0), \\ s_1(0.4), s_2(0), s_3(0) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0.5), s_0(0), \\ s_1(0.5), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0.2), s_{-1}(0), s_0(0.8), \\ s_1(0), s_2(0), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0), s_{-1}(0), s_0(0), \\ s_1(0), s_2(1), s_3(0) \end{array}\right\}
\end{pmatrix}$$

Process 2 Clustering process.

Use Algorithm 1 to divide the large group into six clusters by setting the minimum distance to 0.23, as shown in Fig. 5. As previously discussed, we set $\beta = 1$, $a = round(q \cdot 0.1) = 2$, and $b = round(q/m) = 4$. Table 3 presents the clustering results.

Process 3 Consensus reaching process.

Use the HPDCM to manage opinion differences. Table 4 shows the consensus iterations.

In conclusion, after three consensus iterations (including one hard adjustment and two soft adjustments), the group consensus index satisfies the consensus threshold. Let $t^* = 3$, and the final opinions of the clusters can be obtained as $G_k^3 (k = 1, 2, \ldots, 6)$.

Process 4 Selection process.

Step 4.1 Aggregate all the clusters' opinions to obtain the group opinion via Eq. (3):

$$R_c = \begin{pmatrix}
\left\{\begin{array}{l} s_{-3}(0.1715), s_{-2}(0.148), s_{-1}(0.2116), s_0(0.0937), \\ s_1(0.168), s_2(0.1334), s_3(0.0739) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.0461), s_{-2}(0.2101), s_{-1}(0.2349), s_0(0.1238), \\ s_1(0.2604), s_2(0.0547), s_3(0.0701) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.086), s_{-2}(0.2399), s_{-1}(0.1439), s_0(0.269), \\ s_1(0.1584), s_2(0.1028), s_3(0) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0.0406), s_{-2}(0.0515), s_{-1}(0.1458), s_0(0.2173), \\ s_1(0.1082), s_2(0.2195), s_3(0.2172) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0.0504), s_{-1}(0.1659), s_0(0.2464), \\ s_1(0.2903), s_2(0.2094), s_3(0.0375) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.0226), s_{-2}(0.1518), s_{-1}(0.1864), s_0(0.307), \\ s_1(0.1746), s_2(0.0961), s_3(0.0615) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0.045), s_{-2}(0.0311), s_{-1}(0.3031), s_0(0.1539), \\ s_1(0.2198), s_2(0.0711), s_3(0.1761) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.0926), s_{-2}(0.0839), s_{-1}(0.187), s_0(0.3122), \\ s_1(0.1895), s_2(0.1129), s_3(0.0219) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.0057), s_{-2}(0.062), s_{-1}(0.2705), s_0(0.2153), \\ s_1(0.2141), s_2(0.1555), s_3(0.0768) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0.0216), s_{-2}(0.1011), s_{-1}(0.157), s_0(0.2611), \\ s_1(0.0519), s_2(0.4072), s_3(0) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.196), s_{-2}(0.0667), s_{-1}(0.3346), s_0(0.0992), \\ s_1(0.1843), s_2(0.0692), s_3(0.0501) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.0242), s_{-2}(0.1133), s_{-1}(0.1819), s_0(0.2911), \\ s_1(0.178), s_2(0.1897), s_3(0.022) \end{array}\right\} \\
\left\{\begin{array}{l} s_{-3}(0.0856), s_{-2}(0.1248), s_{-1}(0.1129), s_0(0.1323), \\ s_1(0.2006), s_2(0.2018), s_3(0.142) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0.1194), s_{-2}(0.1515), s_{-1}(0.1622), s_0(0.1663), \\ s_1(0.1912), s_2(0.1152), s_3(0.0943) \end{array}\right\} & \left\{\begin{array}{l} s_{-3}(0), s_{-2}(0.1183), s_{-1}(0.1965), s_0(0.2072), \\ s_1(0.1236), s_2(0.2806), s_3(0.0738) \end{array}\right\}
\end{pmatrix}$$

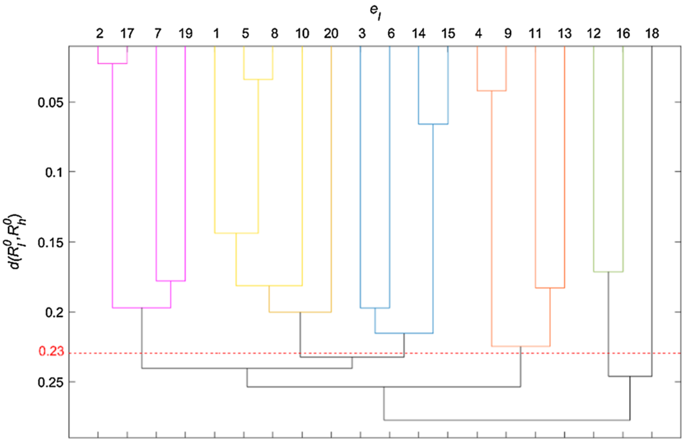

Fig. 5 Visual clustering processes using Algorithm 1

Table 3 Clustering results by using Algorithm 1

c_k	n_k	e_l	θ_k^0	c_k	n_k	e_l	θ_k^0
c_1	4	e_2, e_{17}, e_7, e_{19}	0.1967	c_2	5	e_1, e_5, e_8, e_{10}, e_{20}	0.197
c_3	4	e_3, e_6, e_{14}, e_{15}	0.1933	c_4	4	e_4, e_9, e_{11}, e_{13}	0.1936
c_5	2	e_{12}, e_{16}	0.1097	c_6	1	e_{18}	0.1097

Table 4 Consensus iterations using the HPDCM

t	GCI before CRP	Identified opinion	Hard/soft adjustment	Given adjustment coefficient (matrix)	Punishment coefficient (matrix)	GCI after CRP
1	0.8305	G_6^0	Hard	0.5	0.47	0.8576
2	0.8576	G_5^1	Soft	$\begin{pmatrix} 0.2 & 0.2 & 0.1 \\ 0.2 & 0 & 0.2 \\ 0.1 & 0.3 & 0.1 \\ 0.2 & 0.1 & 0.1 \\ 0.1 & 0.3 & 0.2 \end{pmatrix}$	$\begin{pmatrix} 0.3282 & 0.2036 & 0 \\ 0.466 & 0 & 0.5939 \\ 0.1966 & 0 & 0.1609 \\ 0.3662 & 0.4083 & 0 \\ 0.1684 & 0.7025 & 0.3788 \end{pmatrix}$	0.8741
3	0.8741	G_4^2	Soft	$\begin{pmatrix} 0.2 & 0.5 & 0 \\ 0.2 & 0.2 & 0 \\ 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0.1 \\ 0.2 & 0.2 & 0.4 \end{pmatrix}$	$\begin{pmatrix} 0.2027 & 0.5443 & 0 \\ 0.3314 & 0 & 0 \\ 0.3739 & 0.0929 & 0.2729 \\ 0 & 0 & 0.4173 \\ 0.2449 & 0.222 & 0.37 \end{pmatrix}$	0.8829

Step 4.2 Calculate the overall criterion values $Z(x_i)(i = 1, 2, \ldots, m)$ based on Eq. (3), as follows:

$$Z(x_1) = \{s_{-3}(0.1051), s_{-2}(0.182), s_{-1}(0.1858), s_0(0.148), s_1(0.181), s_2(0.0976), s_3(0.0499)\},$$
$$Z(x_2) = \{s_{-3}(0.0229), s_{-2}(0.0793), s_{-1}(0.1552), s_0(0.2322), s_1(0.1726), s_2(0.1697), s_3(0.1138)\},$$
$$Z(x_3) = \{s_{-3}(0.0469), s_{-2}(0.0552), s_{-1}(0.2376), s_0(0.2046), s_1(0.1949), s_2(0.1052), s_3(0.0978)\},$$
$$Z(x_4) = \{s_{-3}(0.0737), s_{-2}(0.0916), s_{-1}(0.2031), s_0(0.2068), s_1(0.1244), s_2(0.2267), s_3(0.0215)\},$$
$$Z(x_5) = \{s_{-3}(0.688), s_{-2}(0.1253), s_{-1}(0.1452), s_0(0.1561), s_1(0.1651), s_2(0.1871), s_3(0.1037)\}$$

Step 4.3 Based on the score function of PLTSs proposed by Pang et al. (2016), we can calculate the scores of the overall criterion values:
$E(Z(x_1)) = s_{-0.3392}$, $E(Z(x_2)) = s_{0.4709}$, $E(Z(x_3)) = s_{0.21}$, $E(Z(x_4)) = s_{0.035}$, $E(Z(x_5)) = s_{0.2482}$
Therefore, the ranking is $x_2 \succ x_5 \succ x_3 \succ x_4 \succ x_1$, and the best solution is x_2. That is, the DMs are more likely to include the supplier in Ho Chi Minh City on the supplier list.

6 Comparative Analysis and Discussion

We discuss some additional important issues about the HPDCM in Sect. 6.1, including a comparison of hard adjustment and soft adjustment and the determination of relevant parameters. Our proposal is compared with other linguistic models in Sect. 6.2. Section 6.3 presents the managerial implications involved in the practical application of the proposed model. Note that the data used in the comparative analysis and discussion originate from the case study in Sect. 5.

6.1 Further Discussion of the Hierarchical Punishment-Driven Consensus Model

This section analyzes two important issues regarding the HPDCM: the comparison between hard adjustment and soft adjustment and the calculation of the threshold for the attainment rate of the consensus index.

6.1.1 Hard Adjustment and Soft Adjustment

To eliminate DMs' subjective attitudes towards adjustment, we adopt the punishment coefficient in implementing the CRP. Figure 6 shows the consensus results obtained by hard adjustment and soft adjustment. First, regardless of which type of adjustment strategy is utilized, the group consensus index continues to rise as the number of consensus iterations increases. This indicates that the two adjustment strategies are effective in promoting consensus improvement. However, different adjustment strategies lead to different degrees of opinion distortion. For example, as shown in Fig. 6, the distortion degree is 0.5464 when using hard adjustment after the first iteration, while the soft adjustment results

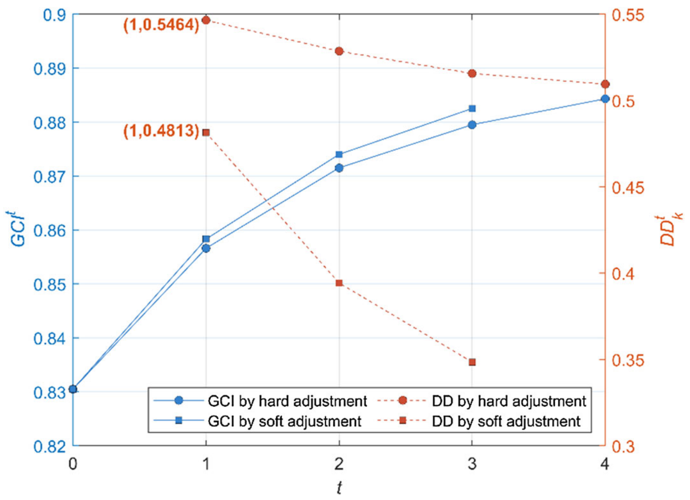

Fig. 6 Consensus results obtained by hard adjustment and soft adjustment. *Note*: GCI and DD represent the group consensus index and the distortion degree, respectively

in only a 0.4813 distortion degree. This is because adopting a hard adjustment strategy causes some matrix elements that should not have been adjusted to be adjusted.

Although, as noted above, soft adjustment performs better than hard adjustment in terms of opinion distortion, this does not mean that soft adjustment can completely replace hard adjustment. Hard adjustment requires only one adjustment coefficient, which is easy to obtain, especially when no punishment coefficient is used as a reference. The hard adjustment may lead to undue distortion, because even if a certain matrix element already satisfies the consensus threshold, it will still be adjusted. Soft adjustment refines the adjustment range to the matrix element level, which helps reduce unnecessary distortion. However, for the DMs, too much time is spent providing adjustment coefficients according to the different consensus indexes of matrix elements. We conclude that hard adjustment is better suited to situations where (1) the decision process has just begun, (2) the identified matrix is low-dimensional, or (3) the decision is time-critical. When the group consensus index has increased significantly and most elements in the identified opinion have already met the consensus threshold, soft adjustment takes precedence.

6.1.2 Determination of the Relevant Parameters

In this section, we discuss how to assign the following two parameters: the threshold of the attainment rate of the consensus index $\overline{ARoCI}$ and the power of the punishment coefficient σ.

The attainment rate of the consensus index plays a decisive role in whether to adopt hard adjustment or soft adjustment. A high threshold can delay the use of soft adjustment strategy, which may increase opinion distortion. A low threshold will lead to premature use of soft adjustment strategy; this may make it difficult for the cluster to provide the adjustment coefficient at the matrix element level because there may be multiple elements in the identified opinion that need to be adjusted. We suggest that determining the threshold should follow two principles:

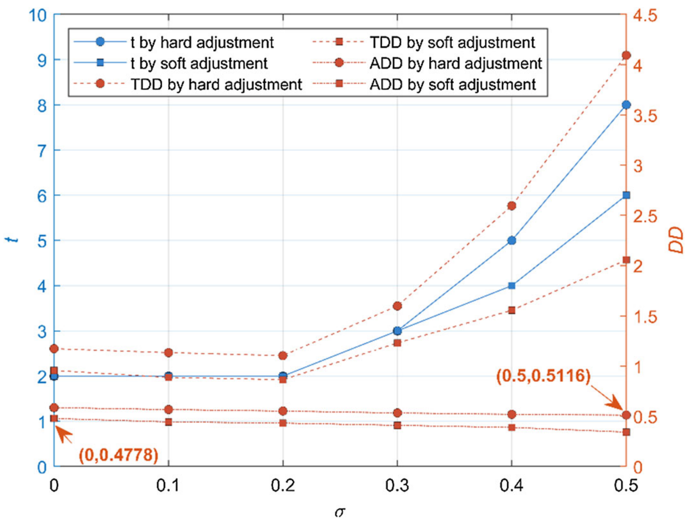

Fig. 7 Consensus results when σ increases from 0 to 0.5. *Note*: TDD represents the total distortion degree, which is calculated by $\sum_t DD_k^t$. ADD is the average distortion degree, which is calculated by $1/t\sum_t DD_k^t$

1. Computer software is used to simulate the CRP so as to provide reference for practical operation.
2. There are two factors to consider: the number of consensus iterations and the degree of opinion distortion.

Assume that the following requirements are met: the maximum number of iterations is 5 and the opinion distortion is no more than 0.5 in each iteration. We implement punishment-driven consensus iterations using MATLAB software. Figure 7 shows the simulation results of the number of iterations and the distortion degree when σ increases from 0 to 0.5. When $a = 0.5$, the number of iterations reaches 8 by using the hard adjustment strategy (or 6 by using the soft adjustment strategy). According to the function characteristic of punishment coefficient, the setting $a > 0.5$ will lead to a greater number of iterations. Therefore, Fig. 7 only depicts the simulation results under the condition of $\sigma \in [0, 0.5]$. We find that as σ increases, the number of iterations and the total distortion degree increase, but from the overall trend, the average distortion degree decreases. The average distortion degree is defined as the total distortion degree divided by the number of iterations. Determining the parameter σ includes the following steps.

Step 1 Observe the simulation of the number of iterations and determine the membership interval of σ

If adopting the hard adjustment strategy, then we can set $\sigma \in [0, 0.4]$; If adopting the soft adjustment strategy, then we can set $\sigma \in [0, 0.5)$. By taking the intersection of the above two intervals, the membership interval is temporarily determined as $\sigma \in [0, 0.4]$.

Step 2 Modify the temporary interval based on the simulation of opinion distortion

From Fig. 7, the average distortion will always be greater than 0.5 if the hard adjustment strategy is adopted, while the opposite result will be obtained if the soft adjustment strategy is adopted. Therefore, σ should be assigned a larger value so that the average

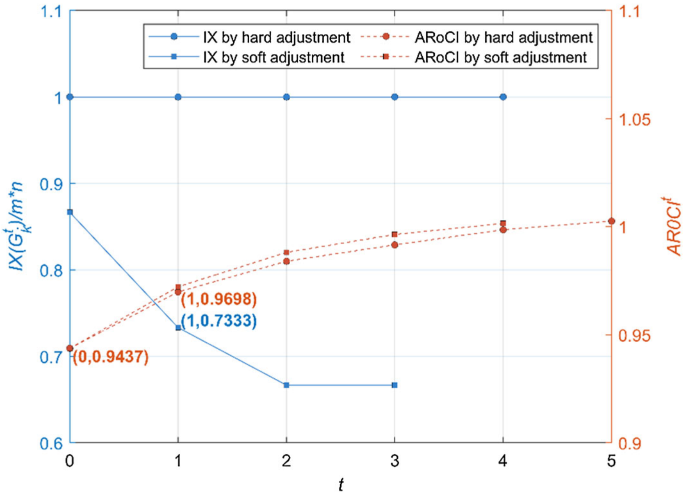

Fig. 8 Changes in some parameters in the CRP. *Note*: Set $\sigma = 0.4$. *IX* represents the proportion of adjusted matrix elements in the identified opinion. *ARoCI* is the attainment rate of the consensus index

distortion is as small as possible. Based on the obtained membership interval, we can set $\sigma = 0.4$. In this case, the soft adjustment strategy must be adopted at least once.

Without loss of generality, we assume that the hard adjustment strategy is allowed to be used once. In fact, this case can only happen at the beginning of the CRP. Figure 8 shows the changes of some parameters in the CRP when setting $\sigma = 0.4$. From Fig. 8, the following results can be obtained $ARoCI^0 = 0.9437$, $ARoCI^1 = 0.9698$, and $IX\left(G^1_k\right)\Big/m * n = 0.7333$. To ensure that the hard adjustment strategy is adopted in the first iteration, we set $\overline{ARoCI} > ARoCI^0$. In the second iteration the soft adjustment strategy can be adopted when setting $\overline{ARoCI} \leq ARoCI^1 = 0.9698$ and $\overline{IX} \geq 0.7333$. In conclusion, we have that $\overline{ARoCI} \in (0.9437, 0.9698]$ and $\overline{IX} \geq 0.7333$. Specifically, the parameters can be set as $\overline{ARoCI} = 0.95$ and $\overline{IX} = 1$.

6.2 Comparison to Other Linguistic Models

This section presents a comparison of our proposal with Pang et al.'s (2016) research and HFLTSs in the decision-making process.

An HFLTS represents hesitation among multiple linguistic terms but cannot reflect the importance of these linguistic terms. In the kind of case we are considering, if the decision information is expressed by HFLTSs, this can be interpreted as indicating that all possible linguistic terms in an HFLTS have the same importance degree. Accordingly, an HFLTS can be written in the form of a PLTS, in which all the linguistic terms have the same possibility. For instance, if the opinion $v^0_{1,13} = \{s_{-1}(0.1), s_2(0.9)\}$ is represented using an HFLTS, it is changed to $\{s_{-1}(0.5), s_2(0.5)\}$.

On the other hand, since the concept of PLTSs was proposed by Pang et al. (2016), research on aggregation and distance measurement for PLTSs has attracted much attention. In this section, we compare our proposal with Pang et al.'s (2016) model in terms of the

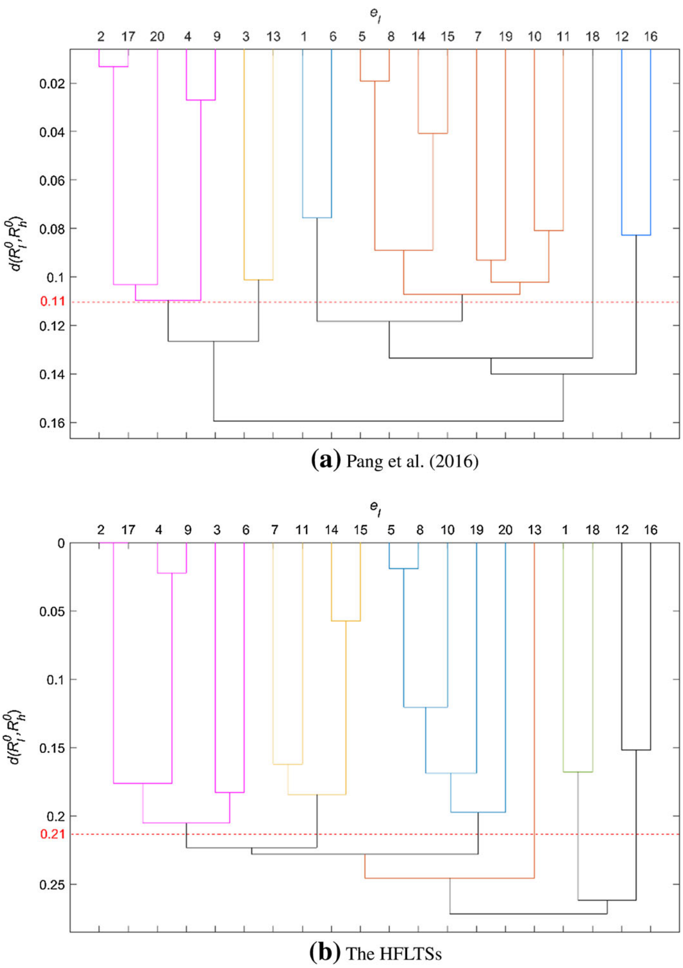

Fig. 9 Visualization of the clustering process when taking different approaches to decision information

effect that processing PLTSs has on the decision results. In general, there are two models for addressing PLTSs: (1) Probability values are incorporated into the linguistic terms in the process of information aggregation (see Pang et al. (2016)). In this model, the

Table 5 Clustering results of the first three clusters formed by using different linguistic models

Linguistic model	First three clusters formed	Distances between the DMs in each cluster
Pang et al. (2016)	$\{e_2, e_{17}\}$, $\{e_5, e_8\}$, $\{e_4, e_9\}$	0.0229, 0.0341, 0.0421
The HFLTSs	$\{e_2, e_{17}\}$, $\{e_5, e_8\}$, $\{e_4, e_9\}$	0, 0.0191, 0.0222
Our proposal	$\{e_2, e_{17}\}$, $\{e_4, e_9\}$, $\{e_5, e_8\}$	0.0132, 0.0191, 0.027

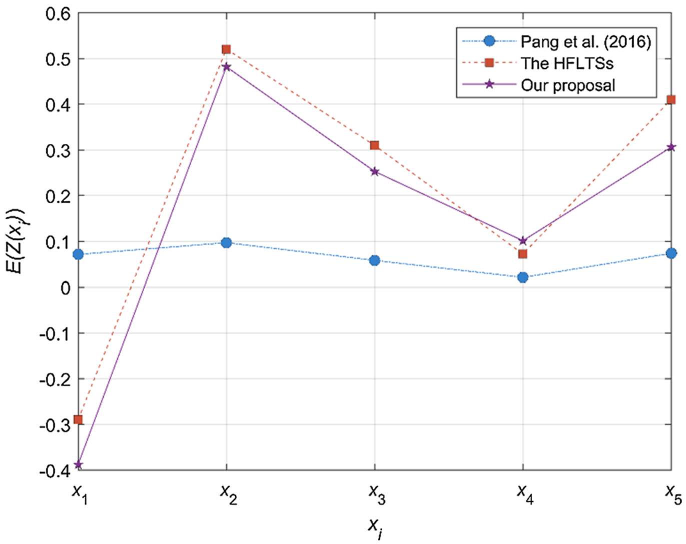

Fig. 10 Initial rankings of the alternatives obtained by using different linguistic models

probabilities of linguistic terms are not reflected in the result, and the original discrete linguistic terms are transformed into virtual ones. (2) Similar to our proposal, before entering the selection process, the information aggregation only involves operations on probability values corresponding to the same linguistic terms and does not merge probability values and linguistic terms.

To ensure comparability, we set the same number of clusters, i.e., $K = 6$. From the visual clustering processes shown in Figs. 5 and 9, the most significant differences are the order in which new clusters are formed and the distances between DMs. Table 5 depicts the results of the first three clusters formed by using different linguistic models. Our proposal differs from the other two models in forming the second and third clusters. This is due to the distance measure. As $d\left(R_5^0, R_8^0\right) > d\left(R_4^0, R_9^0\right)$ in our proposal, the second cluster is composed of e_4, e_9; however, in the other two models, since $d\left(R_5^0, R_8^0\right) < d\left(R_4^0, R_9^0\right)$, DMs e_5, e_8 form the second cluster. It is important to note that if HFLTSs are used, the distance between R_2^0 and R_{17}^2 is 0, which is different from the result obtained by using our proposal or Pang et al.'s (2016) model. This is due to the fact that HFLTSs do not reflect the weights of the linguistic terms.

Owing to different clustering results, the quantitative comparison of the CRP seems unfair. Nonetheless, we consider that it is difficult for DMs to understand the meaning of virtual linguistic terms, as opposed to the discrete linguistic terms with corresponding probabilities. Additionally, the alternative rankings prior to the CRP are depicted in Fig. 10. All three linguistic models yield the same optimal solution (i.e., x_2) and the same ranking (i.e., $x_2 \succ x_5 \succ x_3 \succ x_4 \succ x_1$). However, there are large differences in the scores of the alternatives, although this does not affect the selection of the best alternative.

Above all, our proposal overcomes the shortcomings of the other two models in processing linguistic information. Unlike HFLTSs, which represent hesitations regarding multiple linguistic terms, our proposal based on PLTSs considers the importance of the linguistic terms, which is more in line with actual circumstances. On the other hand, using Pang et al.'s (2016) model to handle PLTSs leads to the probabilities of linguistic terms not being reflected in the result. In this way, the unique feature of PLTSs compared with ordinary linguistic term sets is lost. Furthermore, according to the statement "the ordinal discrete linguistic terms are used to evaluate alternatives, while the virtual linguistic terms can only appear in the operation and ranking" (Xu 2009), virtual linguistic terms should be avoided in the CRP. Therefore, our proposal insists on preserving the original discrete linguistic term structure of PLTSs in the processes of clustering, aggregating and consensus reaching, and only after beginning the selection process are virtual linguistic terms introduced to calculate the scores of PLTSs.

6.3 Managerial Implications

The proposed HPDCM can be applied to the practical case of supplier selection in the manufacturing and service industries, and the following important issues are raised:

1. Supplier selection is a strategic and complex decision with risks that is closely related to the survival and sustainable development of enterprises. It is therefore necessary to measure the group consensus index and make any useful adjustments to ensure that the result is accepted by the majority. This arrangement can reduce the risks and losses of decision-making errors, but more importantly, only when a result is approved by the majority can it obtain the maximum practical support from the majority of departments.
2. The increase in the consensus index is due to some DMs adjusting their own views, whether voluntarily or involuntarily. In this way, the adjusted opinion may be somewhat distorted from the original opinion. Therefore, it is important to measure the degree of opinion distortion and use it as an indicator to evaluate the performance of the CRP. In business practice, some DMs insist on their opinions and are unwilling to compromise for the sake of consensus. Even if the adjustments are forced by the punishment-driven consensus model, the reasonableness of the adjusted opinion needs to be examined. The distortion degree is an important index to test rationality. If the distortion of a DM's opinion is high in an iteration (close to 1), the DM's original opinion is likely to deviate significantly from the majority opinion.

7 Conclusions

In this paper, a hierarchical punishment-driven consensus model for probabilistic linguistic large-group decision-making was developed, and an application of the model to global supplier selection was presented. The obtained results show that the new model can overcome the scalability challenge and soften the human supervision of the CRP. The major contributions of this study are as follows:

1 The PL-LGDM was defined, and its characteristics were detailed. To improve the performance of PLTSs in the CRP, the rules governing their normalization and operations were redefined.
2 A hierarchical punishment-driven consensus model for PL-LGDM was developed, including hard adjustment and soft adjustment, which aimed to reduce the distortion of opinions with the goal of achieving a consensus within a specified number of iterations. Before the CRP begins, the large group was classified into several subgroups by hierarchical clustering, and the opinions of clusters were generated. This operation can overcome the scalability challenge when the number of DMs involved is sufficiently large. When the necessary parameters are set, the model can automatically provide guidance for the adjustment of opinions, thereby reducing the human supervision cost of the CRP.
3 The model was applied to a case study of global supplier selection, and the application implications were discussed.

It is worth noting that there are always social relationships between DMs (Ding et al. 2019; Wu et al. 2018). The influence of social relationships on the CRP should be considered. As further research, we will study the theoretical framework of consensus building in LGDM problems in the context of social networks.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Number 71901151); the Major Project for National Natural Science Foundation of China (Grant Numbers 71991461, 91846301); the Natural Science Foundation of SZU (Grant Number 2019025); and the Special Fund Project of Scientific and Technological Innovation Cultivation for Guangdong University Students in 2019 (Grant Number pdjh2019b0025).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Awasthi A, Govindan K, Gold S (2018) Multi-tier sustainable global supplier selection using a fuzzy AHP-VIKOR based approach. Int J Prod Econ 195:106–117
Bai CZ, Zhang R, Shen S, Huang CF, Fan X (2018) Interval-valued probabilistic linguistic term sets in multi-criteria group decision making. Int J Intell Syst 33(6):1301–1321
Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy c-means clustering algorithm. Comput Geosci 10(2–3):191–203
Cabrerizo FJ, Morente-Molinera JA, Pedrycz W, Taghavi A, Herrera-Viedma E (2018) Granulating linguistic information in decision making under consensus and consistency. Expert Syst Appl 99:83–92
Ding RX, Wang XQ, 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. Inf Fusion 50:251–272
Dong Y, Zhuang Y, Chen K, Tai X (2006) A hierarchical clustering algorithm based on fuzzy graph connectedness. Fuzzy Sets Syst 157(13):1760–1774

Du ZJ, Yu SM, Xu XH (2020) Managing noncooperative behaviors in large-scale group decision-making: integration of independent and supervised consensus-reaching models. Inf Sci. https://doi.org/10.1016/j.ins.2020.03.100
Gou X, Xu Z, Herrera F (2018) Consensus reaching process for large-scale group decision making with double hierarchy hesitant fuzzy linguistic preference relations. Knowl Based Syst 157:20–33
Herrera F, Herrera-Viedma E, Verdegay JL (1996) A linguistic decision process in group decision making. Group Decis Negot 5(2):165–176
Herrera F, Alonso S, Chiclana F, Herrera-Viedma E (2009) Computing with words in decision making: foundations, trends and prospects. Fuzzy Optim Decis Mak 8(4):337–364
Johnson S (1967) Hierarchical clustering schemes. Psychometrika 32(3):241–254
Kacprzyk J, Fedrizzi M (1988) A 'soft'measure of consensus in the setting of partial (fuzzy) preferences. Eur J Oper Res 34(3):316–325
Kim B, Park KS, Jung SY, Park SH (2018) Offshoring and outsourcing in a global supply chain: impact of the arm's length regulation on transfer pricing. Eur J Oper Res 266(1):88–98
Labella Á, Liu Y, Rodríguez RM, Martínez L (2018) Analyzing the performance of classical consensus models in large scale group decision making: a comparative study. Appl Soft Comput 67:677–690
Liu P, He L, Yu X (2016) Generalized hybrid aggregation operators based on the 2-dimension uncertain linguistic information for multiple attribute group decision making. Group Decis Negot 25(1):103–126
Massanet S, Riera JV, Torrens J, Herrera-Viedma E (2014) A new linguistic computational model based on discrete fuzzy numbers for computing with words. Inf Sci 258:277–290
Merigó JM, Gil-Lafuente AM (2013) Induced 2-tuple linguistic generalized aggregation operators and their application in decision-making. Inf Sci 236:1–16
Morente-Molinera JA, Kou G, Pang C, Cabrerizo FJ, Herrera-Viedma E (2019) An automatic procedure to create fuzzy ontologies from users' opinions using sentiment analysis procedures and multi-granular fuzzy linguistic modelling methods. Inf Sci 476:222–238
Palomares I, Martínez L, Herrera F (2014) A consensus model to detect and manage noncooperative behaviors in large-scale group decision making. IEEE Trans Fuzzy Syst 22(3):516–530
Pang Q, Wang H, Xu Z (2016) Probabilistic linguistic term sets in multi–attribute group decision making. Inf Sci 369:128–143
Park JH, Gwak MG, Kwun YC (2011) Uncertain linguistic harmonic mean operators and their applications to multiple attribute group decision making. Computing 93(1):47–64
Quesada FJ, Palomares I, Martínez L (2015) Managing experts behavior in large-scale consensus reaching processes with uninorm aggregation operators. Appl Soft Comput 35:873–887
Reefke H, Sundaram D (2018) Sustainable supply chain management: decision models for transformation and maturity. Decis Support Syst 113:56–72
Rodríguez RM, Martínez L, Herrera F (2012) Hesitant fuzzy linguistic term sets for decision making. IEEE Trans Fuzzy Syst 20(1):109–119
Rodríguez RM, Labella Á, De Tré G, Martínez L (2018) A large scale consensus reaching process managing group hesitation. Knowl Based Syst 159:86–97
Saint S, Lawson JR (1994) Rules for reaching consensus: a modern approach to decision making. Jossey-Bass, San Francisco
Song Y, Li G (2019) A large-scale group decision-making with incomplete multi-granular probabilistic linguistic term sets and its application in sustainable supplier selection. J Oper Res Soc 70(5):827–841
Viswanadham N, Samvedi A (2013) Supplier selection based on supply chain ecosystem, performance and risk criteria. Int J Prod Res 51(21):6484–6498
Wu X, Liao H (2019) A consensus-based probabilistic linguistic gained and lost dominance score method. Eur J Oper Res 272(3):1017–1027
Wu Z, Xu J (2015) Possibility distribution-based approach for MAGDM with hesitant fuzzy linguistic information. IEEE Trans Cybern 46(3):694–705
Wu Z, Xu J (2016) Managing consistency and consensus in group decision making with hesitant fuzzy linguistic preference relations. Omega 65:28–40
Wu Z, Xu J (2018) A consensus model for large-scale group decision making with hesitant fuzzy information and changeable clusters. Inf Fusion 41:217–231
Wu T, Liu X, Liu F (2018) An interval type-2 fuzzy TOPSIS model for large scale group decision making problems with social network information. Inf Sci 432:392–410
Xu ZS (2005) Deviation measures of linguistic preference relations in group decision making. Omega 33:249–254
Xu ZS (2009) An interactive approach to multiple attribute group decision making with multigranular uncertain linguistic information. Group Decis Negot 18(2):119–145

Xu XH, Du ZJ, Chen XH (2015) Consensus model for multi-criteria large-group emergency decision making considering non-cooperative behaviors and minority opinions. Decis Support Syst 79:150–160
Xu XH, Du ZJ, Chen XH, Cai CG (2019) Confidence consensus-based model for large-scale group decision making: a novel approach to managing non-cooperative behaviors. Inf Sci 477:410–427
Yu SM, Wang J, Wang JQ, Li L (2018) A multi-criteria decision-making model for hotel selection with linguistic distribution assessments. Appl Soft Comput 67:741–755
Yücenur GN, Vayvay Özalp, Demire NÇ (2011) Supplier selection problem in global supply chains by AHP and ANP approaches under fuzzy environment. Int J Adv Manuf Technol 56(5–8):823–833
Zhang Y, Xu Z, Wang H, Liao H (2016) Consistency-based risk assessment with probabilistic linguistic preference relation. Appl Soft Comput 49:817–833
Zhang Y, Xu Z, Liao H (2017) A consensus process for group decision making with probabilistic linguistic preference relations. Inf Sci 414:260–275
Zuheros C, Li CC, Cabrerizo FJ, Dong YC, Herrera-Viedma E, Herrera F (2018) Computing with words: revisiting the qualitative scale. Int J Uncertain Fuzz Knowl Based Syst 26(Suppl. 2):127–143

Affiliations

Sumin Yu[1] · **Zhijiao Du**[2] · **Xuanhua Xu**[3]

[1] Institute of Big Data Intelligent Management and Decision, College of Management, Shenzhen University, Shenzhen 518060, China

[2] Business School, Sun Yat-Sen University, Guangzhou 510275, China

[3] School of Business, Central South University, Changsha 410083, China