symmetry

MDPI

Article

Shadowed Sets-Based Linguistic Term Modeling and Its Application in Multi-Attribute Decision-Making

Huidong Wang [1,*], Shifan He [1], Xiaohong Pan [1] and Chengdong Li [2]

[1] School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan 250014, China; shifanhe0828@163.com (S.H.); xiaohongpan0303@126.com (X.P.)
[2] School of Information and Electrical Engineering, Shandong Jianzhu University, Jinan 250101, China; chengdong.li@foxmail.com
* Correspondence: huidong.wang@ia.ac.cn

Received: 20 October 2018; Accepted: 16 November 2018; Published: 2 December 2018

Abstract: For many multi-attribute decision-making (MADM) problems, linguistic variables are more convenient for people to express the attribute values. In this paper, a novel shadowed set-based method is proposed to deal with linguistic terms, where the linguistic term sets are symmetrical both in meaning and form. Firstly, to effectively express the linguistic variables, we develop a data-driven method to construct the shadowed set model for the linguistic terms. Secondly, the Pythagorean shadowed set is defined, and some theorems are subsequently explored. Thirdly, we propose the score function of the Pythagorean shadowed number and develop a new MADM method on the basis of the Pythagorean shadowed set. Finally, a case study of the supplier selection problem is provided to illustrate the effectiveness of the proposed method, and the superiority of our method is demonstrated by comparison analysis.

Keywords: Pythagorean fuzzy linguistic set; shadowed set; Pythagorean shadowed set; multi-attribute decision-making

1. Introduction

Multi-attribute decision-making (MADM) aims to select the best alternative solution(s) from multiple alternatives and has been widely used in various fields [1–3]. In MADM problems, linguistic terms are a convenient and natural way to describe evaluation information. For example, the decision-makers (DMs) can use linguistic terms such as 'Extremely low', 'Very low', 'Low', 'Fair', 'High', 'Very high', and 'Extremely high' to estimate service quality, product performance, and so forth. Therefore, MADM problems based on linguistic terms have received increasing attention. In [4], Aggarwal proposed a new aggregation operator for linguistic terms, and the effectiveness of the operator was illustrated by a case study on the supplier selection problem. Jin [5] developed two group decision-making methods to handle MADM problems under linguistic set environment, and comparative analysis with other methods was performed to demonstrate the validity and merits of the two methods. Yu [6] proposed an extended TODIM method with unbalanced hesitant fuzzy linguistic term sets for MADM problems. For linguistic decision-making problems, Pei [7] developed a new decision-making method by integrating the fuzzy linguistic multiset and TOPSIS methods, and two practical examples were utilized to verify the feasibility of the proposed approach. For the venture capital problem under a linguistic environment, Cheng [8] proposed an interaction approach. However, the methods mentioned above directly replace the linguistic variables with linguistic subscript in the decision-making process, which may cause distortion of information. To better express linguistic variables, the linguistic 2-tuple [9,10] and linguistic scale function [11,12] were introduced to deal with linguistic sets. Nevertheless, the linguistic 2-tuple and linguistic scale function methods still use linguistic subscript to express language variables in nature. Besides, it is difficult to explain the rationality in theory by simply replacing linguistic word with its linguistic subscript. Furthermore, people may have diverse opinions on identical words, but linguistic subscript can only depict a single meaning for one person, which may lead to information distortion.

As distinct from the linguistic 2-tuple and scale function, shadowed sets [13,14] can effectively construct linguistic terms using a data-driven method, and have recently been attracting more and more attention [15,16]. The membership value of the shadowed set is not a precise number, and its distribution is composed of three different zones: the core zone, shadowed zone, and exclusion zone. The core zone and the exclusion zone take the values of 1 and 0, which means all the elements of both zones are fully compatible with or completely excluded from the linguistic word described by a shadowed set. The shadowed zone is an entire unit interval perceived as a zone of uncertainty, which means we are not sure whether the shadowed zone elements represent the linguistic word described by a shadowed set.

In addition, in many situations, experts may hesitate as to what attribute values should be given by them, due to the increasing complexity. Consequently, Atanassov [17] proposed the intuitionistic fuzzy set (IFS) to express uncertainty, which involves not only membership degree but also non-membership degree. However, the limitation of IFS is that the sum of membership degree and non-membership degree must be no more than 1, which makes it difficult to sufficiently express the ideas of the DMs. Therefore, Yager [18] defined the Pythagorean fuzzy set (PFS), which can effectively express the certainty and uncertainty of experts. Recently, PFS has been introduced to deal with MADM problems [19–22]. Zhang and Xu [19] proposed the operation rules of PFS, and extended the TOPSIS method to PFS. By combining PFS with the hesitant fuzzy set (HFS), a new fuzzy set was defined by Liang and Xu [20], named the hesitant Pythagorean fuzzy set (HPFS), an extended TOPSIS method with HPFS was subsequently proposed. Zhang [21] extended PFS to the interval-valued case, and explored the basic operation rules of the Pythagorean fuzzy set (IVPFS). In addition, a Pythagorean fuzzy QUALIFLEX method was developed by integrating closeness index, and its effectiveness was demonstrated through a hierarchical MADM problem. Combining PFSs with linguistic variables, the definition of Pythagorean fuzzy linguistic set (PFLS) was proposed by Peng and Yang [22] and the operation rules of PFLS was defined, subsequently.

Inspired by the idea of the shadowed set and PFS, we propose a new approach to solve MADM problems under linguistic set environment. Firstly, we define Pythagorean shadowed set and explore some theorems of the shadowed set. Secondly, a score function of the Pythagorean shadowed number is defined and the detailed decision-making procedures-based upon the score function is proposed. Finally, a case study of supplier selection is adopted to verify the feasibility of the proposed approach.

The organization of this paper is as follows. Section 2 presents the preliminaries of the Pythagorean fuzzy set and shadowed set. In Section 3, the shadowed set model of seven-level language term is obtained by a data-driven method. A new score function of Pythagorean shadowed number is introduced in Section 4. Section 5 mainly addresses a new MADM method based on Pythagorean shadowed set. The effectiveness of the proposed approach is demonstrated through a supplier selection problem in Section 6, and comparative analysis is made with the other existing methods. Finally, some conclusions are drawn in Section 7.

2. Preliminaries

2.1. Pythagorean Fuzzy Set (PFS)

Definition 1 ([18,23])**.** *Suppose X is a fixed set. A PFS takes the form of:*

$$P = \{\langle\, x,\ P(u_P(x), v_P(x))\rangle | x \in X\}$$

where $v_p(x)$: $X \to [0,1]$ and $u_p(x)$: $X \to [0,1]$ represent the non-membership function and membership function of $x \in X$, respectively, $u_p^2(x) + v_p^2(x) \le 1$. In addition, $\pi_{p(x)} = \sqrt{1 - u_p^2(x) - v_p^2(x)}$ denotes the hesitation degree of $x \in X$.

For the sake of simplicity, Zhang and Xu [19] named $P(u_P(x), v_P(x))$ the Pythagorean fuzzy number (PFN), expressed by $\beta = P(u_\beta, v_\beta)$, where $u_p(x), v_p(x) \in [0,1]$, $\pi_{p(x)} = \sqrt{1 - u_p^2(x) - v_p^2(x)}$ and $u_p^2(x) + v_p^2(x) \le 1$.

Definition 2 ([9]). *Assume $S = \{s_i | i = 0, \cdots, t, t \in R\}$ is a linguistic term set, s_i is the linguistic evaluation value, t is the granularity of S. Take the seven-level linguistic term as an example: S = {s_0 = Extremely low, s_1 = Very low, s_2 = Low, s_3 = Fair, s_4 = High, s_5 = Very high, s_6 = Extremely high}.*

S must satisfy the following two properties:

(1) There is a negation operator: $neg(s_i) = s_{t-i}$;
(2) If $i < j$ then $S_i < S_j$;

Definition 3 ([22]). *Based on the definition of linguistic term set and PFS, the Pythagorean fuzzy linguistic set (PFLS) takes the form of $D = \left\{ \left\langle s_{\tau(x)}, u_p(x), v_p(x) \right\rangle \middle| x \in X \right\}$, and the Pythagorean fuzzy linguistic number (PFLN) is denoted as $\left\langle s_{\tau(x)}, u_p(x), v_p(x) \right\rangle$, where $s_{\tau(x)}$ is the linguistic evaluation value.*

When the attribute values are represented in the form of linguistic terms in MADM problems, the linguistic variable cannot be directly calculated. Therefore, Xu used the subscript of the linguistic term [24] for computation, Wang put forward a linguistic scale function [11] to convert linguistic terms into crisp numbers, and Herrera converted linguistic terms into fuzzy numbers [25]. However, all those methods still use linguistic subscript to express language variables in nature. To express the fuzziness and uncertainty of linguistic terms, we introduce shadowed set method to cope with linguistic term, and further put forward a new Pythagorean shadowed set.

2.2. Shadowed Set

Definition 4 ([13,14]). *A shadowed set S is a set-valued mapping as follows:*

$$S : U \to \{0, [0,1], 1\}$$

where U is a given universe of discourse.

The core of the shadowed set S is the area where the mapping values of the elements are equal to 1.

$$core(S) = \{x \in U | S(x) = 1\}$$

The elements of U whose mapping values are unit intervals in S compose the shadowed zone of the shadowed set and are expressed as follows,

$$CU(S) = \{x \in U | S(x) = [0,1]\}$$

The elements of U whose mapping values are equal to 0 will be excluded from the shadowed set S.

Definition 5. *$A = [a, b, c, d]$ is called a shadowed number (SN), where a, b are the lower and upper bound of the left-shoulder shadowed part, and c, d are the lower and upper bound of the right-shoulder shadowed part. Figure 1 shows an illustration of shadowed number.*

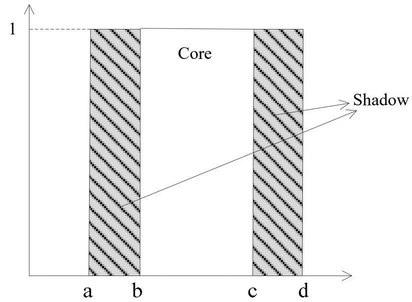

Figure 1. Shadowed number.

2.3. Pythagorean Shadowed Set (PSS)

In this section, we will define the Pythagorean shadowed set and give some properties for it.

Definition 6. *Suppose X is a fixed set, a Pythagorean shadowed set T over X takes the form of*

$$T = \{\langle\, A,\ P(u_P(x), v_P(x))\rangle | x \in X\}$$

where $A = [a, b, c, d]$, a, b are the lower and upper bound of the left-shoulder shadowed part, respectively, and c, d are the lower and upper bound of the right-shoulder shadowed part, respectively. Function $u_p(x)$: $X \to [0, 1]$ and $v_p(x)$: $X \to [0, 1]$ denote the membership function and non-membership function, respectively. ${u^2}_p(x) + {v^2}_p(x) \le 1$, and $\pi_{p(x)} = \sqrt{1 - u_p^2(x) - \nu_p^2(x)}$ denotes the hesitate degree of $x \in X$.

Definition 7. *A Pythagorean shadowed number (PSN) takes the form of:*

$$V = \langle\, A,\ P(u_P(x), v_P(x))\rangle$$

where a, b are the lower and upper bound of the left-shoulder shadowed part, c, d are the lower and upper bound of the right-shoulder shadowed part, and $u_p(x)$: $X \to [0, 1]$ and $v_p(x)$: $X \to [0, 1]$ represent membership function and non-membership function, respectively.

Let $V_1 = \langle\, A_1,\ P(u_P(x_1), v_P(x_1))\rangle$ and $V_2 = \langle\, A_2,\ P(u_P(x_2), v_P(x_2))\rangle$ be two PSNs, where $A_1 = [a_1, b_1, c_1, d_1]$ and $A_2 = [a_2, b_2, c_2, d_2]$, then the operation rules are as follows:

(1) $V_1 + V_2 = \left\langle\ [a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2],\ P\left(\sqrt{(u_p(x_1))^2 + (u_p(x_2))^2 - (u_p(x_1))^2 (u_p(x_2))^2}, v_p(x_1) v_p(x_2)\right)\right\rangle$

(2) $V_1 \times V_2 = \left\langle\ [a_1 \times a_2, b_1 \times b_2, c_1 \times c_2, d_1 \times d_2],\ P\left(u_p(x_1) u_p(x_2), \sqrt{(\nu_p(x_1))^2 + (\nu_p(x_2))^2 - (\nu_p(x_1))^2 (\nu_p(x_2))^2}\right)\right\rangle$

(3) $\lambda V_1 = \left\langle\ [\lambda a_1, \lambda b_1, \lambda c_1, \lambda d_1],\ P\left(\sqrt{1 - \left(1 - (u_p(x_1))^2\right)^\lambda}, (\nu_p(x_1))^\lambda\right)\right\rangle, \lambda \ge 0$

(4) ${V_1}^\lambda = \left\langle\ [a_1^\lambda, b_1^\lambda, c_1^\lambda, d_1^\lambda],\ P\left((u_p(x_1))^\lambda, \sqrt{1 - \left(1 - (\nu_p(x_1))^2\right)^\lambda}\right)\right\rangle, \lambda \ge 0$

Theorem 1. *For any two PSNs $V_1 = \langle\, A_1,\ P(u_P(x_1), v_P(x_1))\rangle$ and $V_2 = \langle\, A_2,\ P(u_P(x_2), v_P(x_2))\rangle$, where$A_1 = [a_1, b_1, c_1, d_1]$ and $A_2 = [a_2, b_2, c_2, d_2]$, the calculation rules satisfy the following properties:*

(1) $V_1 + V_2 = V_2 + V_1$

(2) $V_1 \times V_2 = V_2 \times V_1$

(3) $\lambda(V_1 + V_2) = \lambda V_1 + \lambda V_2, \lambda \geq 0$

(4) $V_1^{\lambda_1+\lambda_2} = V_1^{\lambda_1} + V_2^{\lambda_2}, \lambda_1, \lambda_2 \geq 0$

(5) $\lambda_1 V_1 + \lambda_2 V_1 = (\lambda_1 + \lambda_2)V_1, \lambda_1, \lambda_2 \geq 0$

(6) $V_1^{\lambda} \times V_2^{\lambda} = (V_1 \times V_2)^{\lambda}, \lambda \geq 0$

3. Shadowed Set Model of Linguistic Terms

We collect the interval data for each language word in the form of the seven-level linguistic term listed in Definition 2 and use the collected interval data to construct the shadowed set models for the seven-level linguistic term. The interval data are obtained by means of questionnaire survey. The main framework of our questionnaire is designed to get a proper interval value for each word of the seven-level linguistic term from those respondents according to their experience, habits and common sense. It is necessary for the filled numbers to be accurate to the first decimal place.

We handed out our questionnaires via leaflets, emails and online survey websites to people in different fields, especially to those with a bachelor degree or above. In the end, we got 1205 valid questionnaires, and the questionnaire data were processed by the following interval data preprocessing method to obtain the shadowed number of the seven-level linguistic term.

3.1. Interval Data Preprocessing

Wu and Liu [26,27] proposed an efficient method to preprocess interval data, and we preprocessed the n interval endpoint data $[a_k, b_k](k = 1, 2, \ldots, n)$ based on this method, as follows:

Step 1: Bad data processing. This aims to remove unreasonable results from the surveyed people, whose answers were beyond the range of the universe of discourse U. If the interval endpoints satisfy the following conditions, the interval data are acceptable. Otherwise, they will be rejected.

$$\begin{cases} 0 \leq a_k \leq 10 \\ 0 \leq b_k \leq 10 \quad , k = 1, 2, \ldots, n \\ b_k \geq a_k \end{cases}$$

By this step, some data will be abandoned, and $n^* < n$ interval data will be preserved.

Step 2: Outlier Processing. By using the Box and Whisker test [28], the data that are extremely large or small, i.e., outliers, can be eliminated. Outlier tests can be applied to process the endpoints of interval data and the lengths of interval data $L_k = b_k - a_k$, respectively. Consequently, only the interval endpoints and lengths satisfying the following conditions are kept:

$$\begin{cases} a_k \in [Q_a(0.25) - 1.5IQR_a, Q_a(0.75) + 1.5IQR_a] \\ b_k \in [Q_b(0.25) - 1.5IQR_b, Q_b(0.75) + 1.5IQR_b] \quad , k = 1, 2, \ldots, n^* \\ L_k \in [Q_L(0.25) - 1.5IQR_L, Q_L(0.75) + 1.5IQR_L] \end{cases}$$

where Q_a and IQR_a are respectively the quartile and interquartile ranges of the left endpoints, Q_b and IQR_b are respectively the quartile and interquartile ranges of the right endpoints, Q_L and IQR_L are respectively the quartile and interquartile ranges of the interval data's length. $Q(0.25)$ and $Q(0.75)$ are the first and third quartiles, which include 25% and 75% of the data, respectively. In addition, the interquartile range IQR is the difference between $Q(0.25)$ and $Q(0.75)$; that is to say, IQR contains 50% of the data between $Q(0.25)$ and $Q(0.75)$. The points that are more than $1.5IQR$ below the first quartile or more than $1.5IQR$ above the third quartile are regarded as outliers.

After this step, $m^* < n^*$ interval data will remain.

Then, the following statistics of the m^* interval data are calculated: m_l and σ_l are mean values and standard deviations of the m^* left endpoints, respectively. Similarly, m_r and σ_r represent the mean values and standard deviations of the m^* right endpoints. m_L and σ_L denote the mean values and standard deviations of the lengths of the m^* interval data.

Step 3: Tolerance limit processing. If the remaining intervals satisfy the following conditions, then they will be accepted; otherwise, they will be rejected.

$$\begin{cases} a_k \in [m_l - \eta\sigma_l, m_l + \eta\sigma_l] \\ b_k \in [m_r - \eta\sigma_r, m_r + \eta\sigma_r] \\ L_k \in [m_L - \eta\sigma_L, m_L + \eta\sigma_L] \end{cases}, k = 1, 2, \ldots, m^*$$

where η is the tolerance factor, which represents that we can assure the given limits at least include the proportion $1 - \alpha$ of the measurements with $100 \cdot (1-\gamma)\%$ confidence level. The value of tolerance factor can be obtained from Table 1 [29].

Table 1. Tolerance factor η for several collected data.

m^*	$1-\gamma = 0.95$		$1-\gamma = 0.99$	
	$1-\alpha$		$1-\alpha$	
	0.90	**0.95**	**0.90**	**0.95**
10	2.839	3.379	3.582	4.265
15	2.480	2.954	2.945	3.507
20	2.310	2.752	2.659	3.168
30	2.140	2.549	2.358	2.841
50	1.996	2.379	2.162	2.576
100	1.874	2.233	1.977	2.355
1000	1.709	2.036	1.736	2.718
∞	1.645	1.960	1.645	1.960

After the processing of Step 3, $m^{**} < m^*(1 \leq m^{**} \leq n)$ interval data will be left, and the following statistical characteristics of the m^{**} data will be computed: m_l, σ_l, m_r, σ_r, m_L and σ_L of the left (right) endpoints the m^{**} interval data.

Step 4: Reasonable-interval processing. If the intervals satisfy the following conditions, they will be kept; otherwise, they will be rejected.

$$2m_l - \phi^* \leq a_k < \phi^* < b_k \leq 2m_r - \phi^*$$

where

$$\phi^* = \frac{(m_r\sigma_l^2 - m_l\sigma_r^2) \pm \sigma_l\sigma_r\left[(m_l - m_r)^2 + 2(\sigma_l^2 - \sigma_r^2)\ln(\sigma_l/\sigma_r)\right]^{1/2}}{\sigma_l^2 - \sigma_r^2}$$

After this step, there will be m interval data.

In a word, there will be m interval data after the four processing steps above, which is not greater than the n interval data at the beginning, as shown in Figure 2.

$$n \xrightarrow{\text{Bad data}} n^* \xrightarrow{\text{Outliers}} m^* \xrightarrow[\text{limit}]{\text{Tolerance}} m^{**} \xrightarrow[\text{interval}]{\text{Reasonable}} m$$

Figure 2. The process of data preprocessing.

3.2. Shadowed Set Model of Seven-Level Language Terms

After data preprocessing, the distribution of the remaining interval data is obtained as shown in Figure 3. The intervals of the left-end points and right-end points can reflect the linguistic word's uncertainties from different surveyed persons. Therefore, it is necessary to determine the representative intervals for the left-end points and right-end points to express the uncertainties. As shown in Figure 3, the core area can be determined even if the surveyed people cannot give accurate representative intervals. As a result, the core of the shadowed set is the core area and the uncertain bound of the shadowed set is the representative intervals.

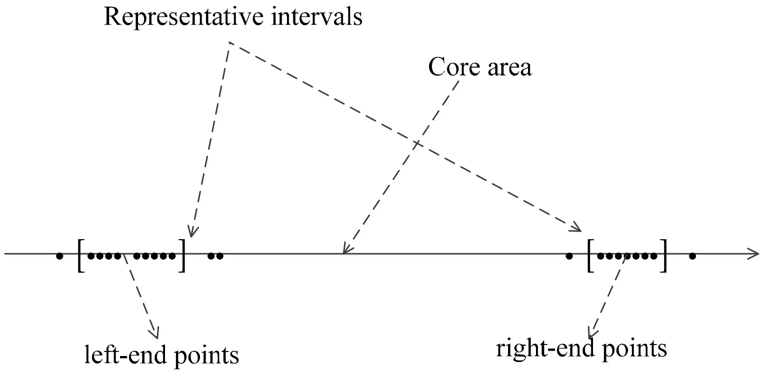

Figure 3. Distribution of the remaining interval data.

Next, we will estimate the representative intervals by the tolerance limit method via the following steps.

Step 1: Calculate the mean m_l and standard deviation σ_l of the remaining left-end points

$$m_l = \frac{\sum\limits_{k=1}^{m} \hat{l}_k}{m} \tag{1}$$

$$\sigma_l = \sqrt{\frac{\sum\limits_{k=1}^{m} \left(\hat{l}_k - m_l\right)^2}{m}} \tag{2}$$

where $\hat{l}_k$ denotes the left-end point of each remaining interval, m is the number of remaining intervals.

Step 2: Determine the representative interval. Let $[L_l, L_r]$ and $[R_l, R_r]$ be the representative intervals of the left-end points and right-end points, respectively.

$$L_l = m_l - \eta * \sigma_l \tag{3}$$

$$L_r = m_l + \eta * \sigma_l \tag{4}$$

where η is the tolerance factor in Table 1.

Then, the representative interval for the right-end points is calculated in the same way.

The parameters γ and α are set to 0.05 and 0.1 in this paper, respectively, and we can obtain a tolerance factor η of 1.709 from Table 1. Take the seven-level language terms as an example: based on the results above, the shadowed set models for seven-level language terms can be constructed as shown in Figure 4.

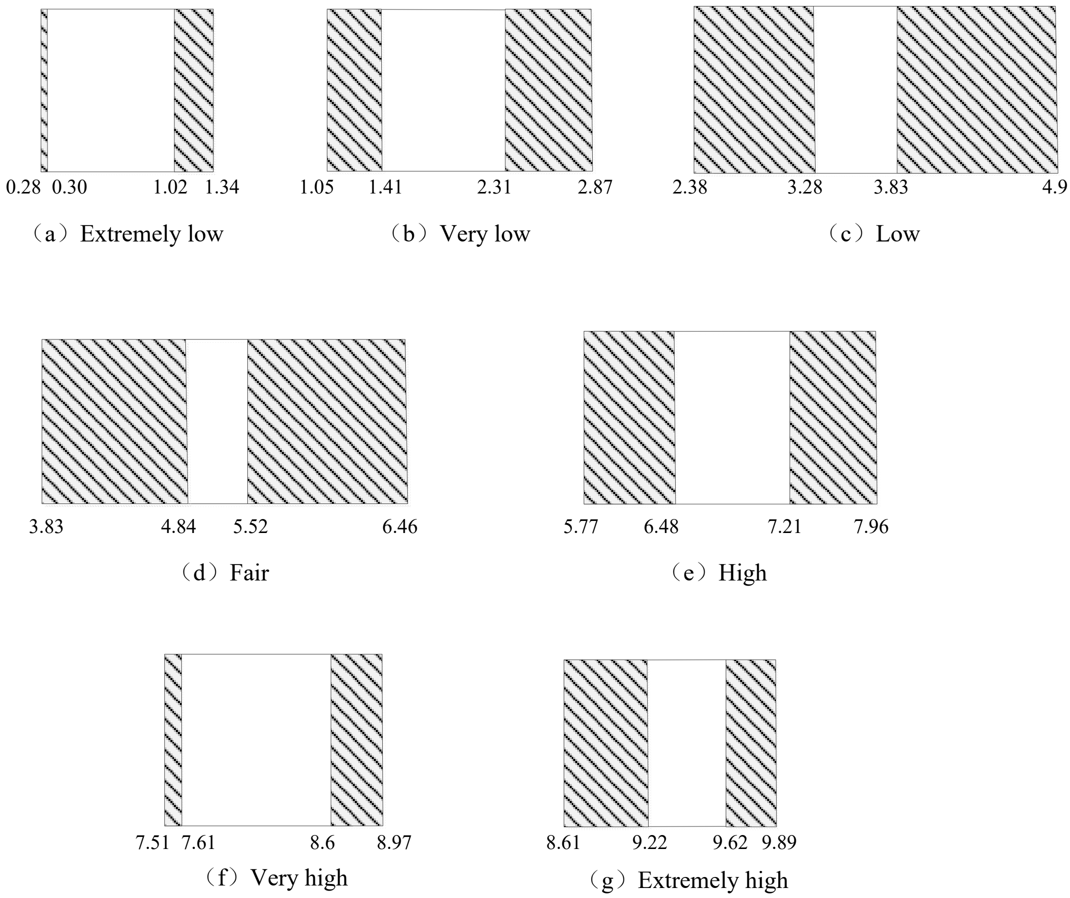

Figure 4. The shadowed set models for seven-level language terms.

4. The Score Function of Pythagorean Shadowed Number

Based on the concepts of shadowed number and Pythagorean shadowed number in Section 2, we will further present the score functions of shadowed number and Pythagorean shadowed number, respectively. Numerical examples will also be given to illustrate the specific calculation process of the two score functions.

According to the central limit theorem, the attribute value r_{ij} given by the decision-maker is stable and tends to be the most likely attribute value at a certain point, so it is believed that r_{ij} obeys the normal distribution within the fuzzy interval. From the tolerance limit method in Section 3.2, we can obtain the distribution of attribute value in the shadowed set $S = \{A_i|U\}$, $A_i = [a_i, b_i, c_i, d_i]$, as shown in Figure 5.

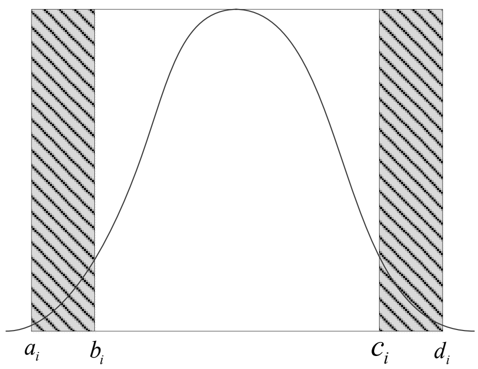

Figure 5. Normal distribution of attribute value.

Definition 8. *The score function of shadowed number A is defined as follows:*

$$\text{score}(A) = a + \int_a^b f(x)\,dx + c - b + \int_c^d f(x)\,dx + d \tag{5}$$

where $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-(x-u)^2/(2\cdot\sigma^2)}$.

According to the 3σ principle of normal distribution:

$$p(r \in [u - 3\sigma, u + 3\sigma]) = 0.9974,\ p(r \in [a, d]) = 0.9974$$

Then $u = \frac{a+d}{2}, \sigma = \frac{d-a}{6}$.

Example 1. *The score function value of shadowed set A_0 for 'High' in Figure 3 can be calculated as follows:*

$$u = a + d = 5.77 + 7.96 = 13.73, \sigma = \frac{d-a}{6} = \frac{7.96 - 5.77}{6} = 0.37$$

$f(x) = 1.08e^{-(x-6.87)^2/0.27}$

Then, we can gain the figure of shadowed number 'High' as shown in Figure 6.

$$\text{score}(A_0) = 5.77 + \int_{5.77}^{6.48} f(x)\,dx + 7.21 - 6.48 + \int_{7.21}^{7.96} f(x)\,dx + 7.96 = 17.9$$

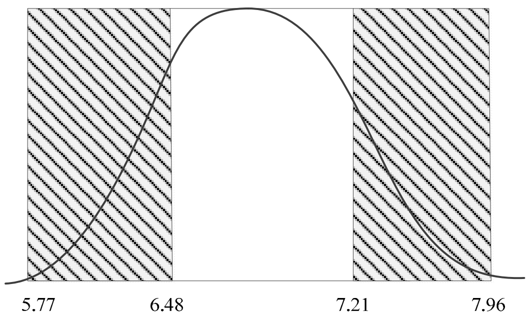

Figure 6. The shadowed number of 'High' and its normal distribution.

In the same way, we can get the score function of shadowed sets for the other six language terms in Figure 3.

Definition 9. *The score function of a Pythagorean shadowed number V is denoted as:*

$$\text{score}(V) = \left(a + \int_a^b f(x)\,dx + c - b + \int_c^d f(x)\,dx + d\right) \cdot (u/v) \tag{6}$$

Example 2. *For a Pythagorean shadowed number $V_1 = \{[5.77, 6.48, 7.21, 7.96], P(0.7, 0.4)\}$, the score function value is:*

$$score(V_1) = \left(5.77 + \int_{5.77}^{6.48} f(x)dx + 7.21 - 6.48 + \int_{7.21}^{7.96} f(x)dx + 7.96\right) \cdot (0.7/0.4) = 31.33$$

where $f(x) = 1.08e^{-(x-6.87)^2/0.27}$.

5. MADM Method Based on the Pythagorean Shadowed Set

With the concept of PSS in mind, we can put forward a novel MADM approach under Pythagorean fuzzy linguistic term circumstances. The diagram of the proposed method is shown in Figure 7. Firstly, present a description of the MADM problem under the Pythagorean linguistic fuzzy circumstances. Secondly, transform the PFLS into PSS through a data-driven method. Thirdly, determine the ranking order of all alternatives so as to obtain the best choice(s) by means of the score function of PSNs and OWA operator. The whole decision-making process is carried out in the following steps.

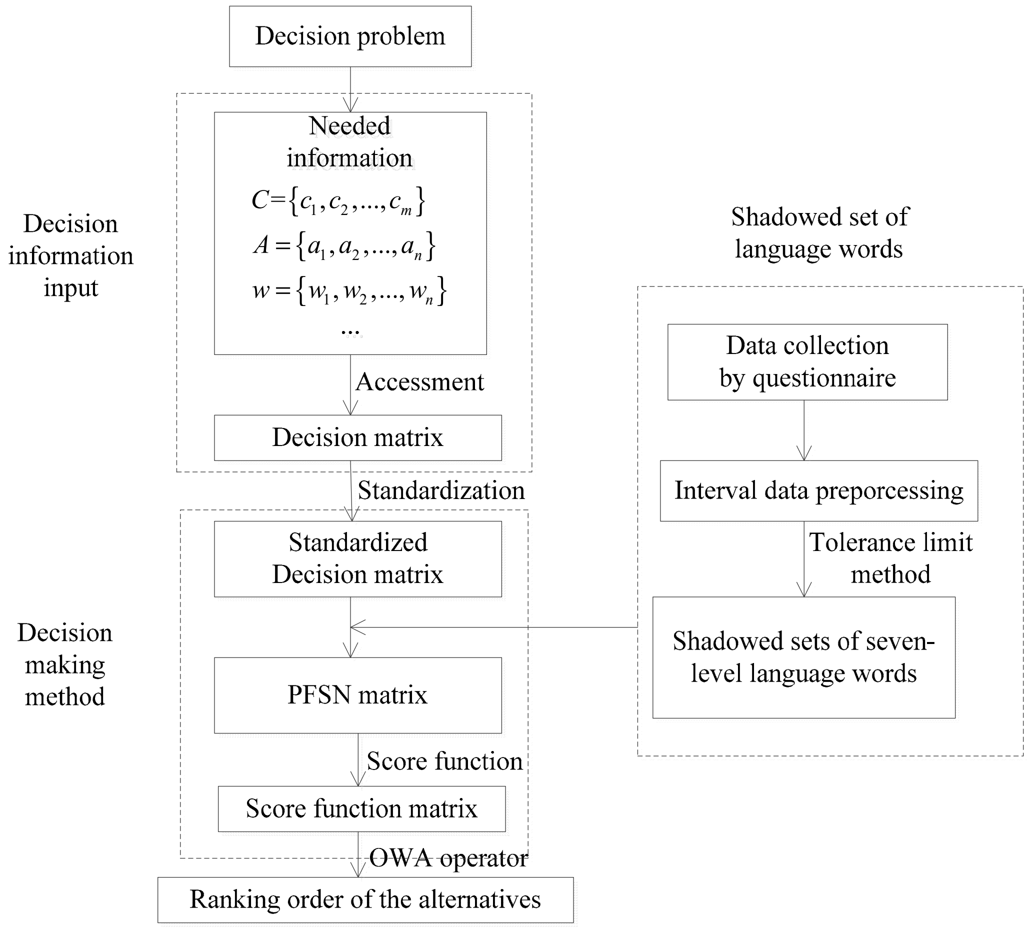

Figure 7. Diagram of the proposed method.

Step 1: Standardized decision matrix. For PFLVs $P_{ij} = \left\langle s_{\tau_{ij}}, P\left(u_p(x_{ij}), v_p(x_{ij})\right)\right\rangle$

For beneficial attributes, $\overline{P_{ij}} = P_{ij} = \left\langle s_{\tau_{ij}}, P\left(u_p(x_{ij}), v_p(x_{ij})\right)\right\rangle$

For cost attributes, $\overline{P_{ij}} = (P_{ij})^{-1} = \left\langle s_{(\tau_{ij})^{-1}}, P\left(v_p(x_{ij}), u_p(x_{ij})\right)\right\rangle$ where $(\tau_{ij})^{-1} = l + 1 - \tau_{ij}$ and l is the number of language term.

Step 2: Collect the data by questionnaire and get the shadowed set of language terms by processing the data. Transform Pythagorean fuzzy linguistic numbers into PSNs using Figure 4.

Step 3: Transform the PFSN decision matrix into score function matrix based on Equation (6).

Step 4: By OWA operator, the attribute values r_{ij} of each alternative a_i are aggregated to obtain the comprehensive attribute values z_i.

$$z_i = OWA_w(r_{i1}, r_{i2}, \ldots, r_{im}) = \sum_{j=1}^{m} w_i r_{ij},\ i = 1, 2, \ldots, n$$

where $w = (w_1, w_2, \ldots, w_m)$ is the criterion weight vector, n is the number of alternatives, m is the number of attribute.

Step 5: Determine the order of all the alternatives in the light of the comprehensive attribute values z_i.

6. Numerical Study

The proposed algorithm will be demonstrated by solving the problem of how to select the most suitable supplier for a company under various evaluation factors. At the same time, comparisons with the linguistic term subscript method and the linguistic scale function method are performed to show the advantages of our approach.

6.1. Supplier Selection Problem

A car company needs to choose appropriate supplier of spare parts. A total of five alternative suppliers are denoted as a_1, a_2, a_3, a_4, a_5. After synthetical consideration, four main factors are taken into account: c_1 Supply capacity, c_2 Delivery timeliness, c_3 Service quality, c_4 Scientific research ability. The criterion weight vector is $w = (0.3, 0.2, 0.4, 0.1)$. Language evaluation of the four attributes adopts the form of seven-level linguistic term, $S = \{s_0, s_1, s_2, s_3, s_4, s_5, s_6\}$ = {Extremely low, Very low, Low, Fair, High, Very high, and Extremely high}. The decision matrix given by experts is shown in Table 2:

Table 2. Decision matrix.

Alternatives	Attributes			
	c_1	c_2	c_3	c_4
a_1	$\langle s_4, P(0.7, 0.4)\rangle$	$\langle s_5, P(0.5, 0.6)\rangle$	$\langle s_2, P(0.7, 0.3)\rangle$	$\langle s_3, P(0.8, 0.4)\rangle$
a_2	$\langle s_5, P(0.6, 0.4)\rangle$	$\langle s_3, P(0.7, 0.4)\rangle$	$\langle s_3, P(0.6, 0.4)\rangle$	$\langle s_4, P(0.7, 0.5)\rangle$
a_3	$\langle s_6, P(0.6, 0.5)\rangle$	$\langle s_3, P(0.8, 0.3)\rangle$	$\langle s_5, P(0.6, 0.5)\rangle$	$\langle s_2, P(0.6, 0.4)\rangle$
a_4	$\langle s_3, P(0.7, 0.3)\rangle$	$\langle s_4, P(0.6, 0.5)\rangle$	$\langle s_3, P(0.7, 0.4)\rangle$	$\langle s_6, P(0.7, 0.6)\rangle$
a_5	$\langle s_4, P(0.7, 0.4)\rangle$	$\langle s_5, P(0.6, 0.5)\rangle$	$\langle s_4, P(0.7, 0.4)\rangle$	$\langle s_3, P(0.8, 0.4)\rangle$

Step 1: c_1, c_2, c_3, c_4 are beneficial attributes. Therefore, the standardized decision matrix is the same with Table 2.

Step 2: Transform PFLNs into Pythagorean shadowed numbers using Figure 4, and the result is shown in Table 3.

Table 3. Decision matrix with PFSN.

Alternatives	Attributes	
	c_1	c_2
a_1	$\langle[5.77, 6.48, 7.21, 7.96], P(0.7, 0.4)\rangle$	$\langle[7.51, 7.61, 8.60, 8.97], P(0.5, 0.6)\rangle$
a_2	$\langle[7.51, 7.61, 8.60, 8.97], P(0.6, 0.4)\rangle$	$\langle[3.83, 4.84, 5.52, 6.46], P(0.7, 0.4)\rangle$
a_3	$\langle[8.61, 9.22, 9.62, 9.89], P(0.6, 0.5)\rangle$	$\langle[3.83, 4.84, 5.52, 6.46], P(0.8, 0.3)\rangle$
a_4	$\langle[3.83, 4.84, 5.52, 6.46], P(0.7, 0.3)\rangle$	$\langle[5.77, 6.48, 7.21, 7.96], P(0.6, 0.5)\rangle$
a_5	$\langle[5.77, 6.48, 7.21, 7.96], P(0.7, 0.4)\rangle$	$\langle[7.51, 7.61, 8.60, 8.97], P(0.6, 0.5)\rangle$
Alternatives	**Attributes**	
	c_3	c_4
a_1	$\langle[2.38, 3.28, 3.83, 4.90], P(0.7, 0.3)\rangle$	$\langle[3.83, 4.84, 5.52, 6.46], P(0.8, 0.4)\rangle$
a_2	$\langle[3.83, 4.84, 5.52, 6.46], P(0.6, 0.4)\rangle$	$\langle[5.77, 6.48, 7.21, 7.96], P(0.7, 0.5)\rangle$
a_3	$\langle[7.51, 7.61, 8.60, 8.97], P(0.6, 0.5)\rangle$	$\langle[2.38, 3.28, 3.83, 4.90], P(0.6, 0.4)\rangle$
a_4	$\langle[3.83, 4.84, 5.52, 6.46], P(0.7, 0.4)\rangle$	$\langle[8.61, 9.22, 9.62, 9.89], P(0.7, 0.6)\rangle$
a_5	$\langle[5.77, 6.48, 7.21, 7.96], P(0.7, 0.4)\rangle$	$\langle[3.83, 4.84, 5.52, 6.46], P(0.8, 0.4)\rangle$

Step 3: Transform the PFSNs decision matrix into score function matrix (shown in Table 4) based on Equation (6).

Table 4. Score function matrix.

Alternatives	Attributes			
	c_1	c_2	c_3	c_4
a_1	31.33	18.39	25.29	26.76
a_2	33.11	23.42	20.07	25.06
a_3	40.97	35.68	26.49	16.26
a_4	31.22	21.48	23.42	39.83
a_5	31.33	26.49	31.33	26.76

Step 4: By OWA operator, the attribute values r_{ij} of each alternative a_i are aggregated to obtain the comprehensive attribute values z_i.

$z_1 = 25.87, z_2 = 25.15, z_3 = 31.65, z_4 = 27.01, z_5 = 29.91$**Step 5:** Rank the alternatives and obtain the best alternative(s) according to the comprehensive attribute values z_i in the Step 4.

$z_3 > z_5 > z_4 > z_1 > z_2$, that means, $a_3 \succ a_5 \succ a_4 \succ a_1 \succ a_2$.

And the alternative a_3 is the best choice of the supplier option problem.

6.2. Comparison Analysis

To verify the superiority of our method, comparations will be made between our approach and the other two approaches, i.e., the linguistic term subscript method [22] and the linguistic scale function method [11,19].

In [22], the score function of $p = \left\langle s_{\tau(x)}, u_A(x), v_A(x) \right\rangle$ is:

$$score(p) = \frac{\tau(x)}{t+1} * \left(\mu_\beta^2 - \nu_\beta^2\right) \tag{7}$$

where $\tau(x)$ is the subscript of the linguistic term, and t is the number of linguistic terms.

We can obtain the comprehensive attribute values z_i based on Equation (7) and the OWA operator.

$z_1 = 0.094, z_2 = 0.104, z_3 = 0.098, z_4 = 0.115, z_5 = 0.147$, and $z_5 > z_4 > z_2 > z_3 > z_1$.

Therefore, the alternative a_5 is the best choice.

In [19], the score function of PFN $\beta = P(u_\beta, v_\beta)$ is:

$$score(\beta) = \mu_\beta^2 - \nu_\beta^2 \tag{8}$$

In [11], the improved linguistic scale function is calculated as follows:

$$f(s_i) = \theta_i = \begin{cases} \frac{m^\alpha - (m-i)^\alpha}{2m^\alpha} & (i = 0, 1, 2, \ldots, m) \\ \frac{m^\beta + (i-m)^\beta}{2m^\beta} & (i = m+1, m+2m \ldots, t) \end{cases} \tag{9}$$

where $\alpha, \beta \in (0, 1]$, $m = \frac{t}{2}$, and t is the number of linguistic terms.

According to the improved linguistic scale Function (8) and score Function (9), we can obtain the score function of $p = \left\langle s_{\tau(x)}, u_A(x), v_A(x) \right\rangle$ as:

$$score(p) = f(s_i) * \left(\mu_\beta^2 - \nu_\beta^2\right) \tag{10}$$

Let $\alpha = \beta = 0.5$. We can obtain the comprehensive attribute values z_i based on Equation (10) and the OWA operator.

$z_1 = 0.07, z_2 = 0.15, z_3 = 0.2, z_4 = 0.14, z_5 = 0.16$, and $z_3 > z_5 > z_2 > z_4 > z_1$.

Therefore, the alternative a_3 is the best choice.

From Table 5, it can be observed that the ranking result obtained via our algorithm is different from the other two methods. By using the linguistic term subscript method, the ranking order is $a_5 \succ a_4 \succ a_2 \succ a_3 \succ a_1$, which is totally different form the results of our method and the language scale function method. The reason is that replacing linguistic words simply with linguistic subscript leads to distortion of information. In fact, the linguistic subscript cannot effectively reflect original decision information. Compared with the linguistic term subscript approach, the linguistic scale function method seems more reasonable for describing the linguistic term information with a so-called language scale function. However, the language scale function still replaces linguistic words with numbers in nature, and information loss or information distortion is still inevitable. On the other hand, different people may have different viewpoints on the same word, but the linguistic subscript and linguistic scale function can only express a single meaning for a word. Compared with the other two methods, we utilize a data-driven method to construct the shadowed set models for the linguistic terms, which cannot only maintain the original decision information as far as possible, but also take different views into account for a single word.

Table 5. Comparison analysis results.

Method	Order of Alternatives
Our method	$a_3 \succ a_5 \succ a_4 \succ a_1 \succ a_2$
Linguistic term subscript method [22]	$a_5 \succ a_4 \succ a_2 \succ a_3 \succ a_1$
Language scale function method [11,19]	$a_3 \succ a_5 \succ a_2 \succ a_4 \succ a_1$

7. Conclusions

A novel method for MADM problems under a linguistic term environment was proposed, combining shadowed sets and Pythagorean fuzzy sets. We defined Pythagorean shadowed numbers and subsequently described their operation rules and basic properties. Based on the operation rules, the score function of Pythagorean shadowed numbers was deduced, and a numerical example was provided to illustrate the computing process. Bearing the above results in mind, we proposed a new MADM approach to deal with linguistic terms. A supplier selection example was used to demonstrate the feasibility of our method. Compared with the linguistic term subscript method and the linguistic scale function method, a data-driven method was adopted to construct the shadowed set models for linguistic terms, which can avoid information loss or information distortion to a great extent. The comparative analysis shows that our method can provide more reasonable and accurate decision-making results by depicting linguistic terms in a more precise manner.

In future research, the proposed method can be extended to other types of shadowed sets, for example, left-shoulder, right-shoulder, non-cored, etc. Additionally, applications in other fields are also worth exploring with our approach.

Author Contributions: Methodology, H.W.; Writing—Original Draft Preparation, S.H.; Formal Analysis, C.L.; Investigation, X.P.; Writing—Review & Editing, H.W.

Funding: NSFC (No. 61402260, 61473176), and partly by the Taishan Scholar Project of Shandong Province.

Acknowledgments: The authors would like to thank the National Natural Science Foundation of China (No. 61402260, 61473176) and the Taishan Scholar Project of Shandong Province for the financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, S.; Zhao, H.R. Fuzzy best-worst multi-criteria decision-making method and its applications. *Knowl.-Based Syst.* **2017**, *121*, 23–31. [CrossRef]
2. Wu, T.; Liu, X.W.; Liu, F. An interval type-2 fuzzy TOPSIS model for large scale group decision making problems with social network information. *Inf. Sci.* **2018**, *432*, 392–410.

3. Li, C.D.; Ding, Z.X.; Yi, J.Q.; Lv, Y.; Zhang, G. Deep belief network based hybrid model for building energy consumption prediction. *Energies* **2018**, *11*, 242. [CrossRef]
4. Manish, A. Adaptive linguistic weighted aggregation operators for multi-criteria decision making. *Appl. Soft Comput.* **2017**, *58*, 690–699.
5. Jin, F.F.; Ni, Z.W.; Pei, L.D.; Chen, H.; Tao, Z.; Zhu, X.; Ni, L. Approaches to group decision making with linguistic preference relations based on multiplicative consistency. *Comput. Ind. Eng.* **2017**, *114*, 69–79. [CrossRef]
6. Yu, W.Y.; Zhang, Z.; Zhong, Q.Y.; Sun, L.L. Extended TODIM for multi-criteria group decision making based on unbalanced hesitant fuzzy linguistic term sets. *Comput. Ind. Eng.* **2017**, *114*, 316–328. [CrossRef]
7. Pei, Z.; Liu, J.; Hao, F.; Zhou, B. FLM-TOPSIS: The fuzzy linguistic multiset TOPSIS method and its application in linguistic decision making. *Inf. Fusion* **2019**, *45*, 266–281.
8. Cheng, X.; Gu, J.; Xu, Z.S. Venture capital group decision-making with interaction under probabilistic linguistic environment. *Knowl.-Based Syst.* **2018**, *140*, 82–91. [CrossRef]
9. Liu, H.B.; Jiang, L.; Martínez, L. A dynamic multi-criteria decision making model with bipolar linguistic term sets. *Exp. Syst. Appl.* **2018**, *95*, 104–112.
10. Wang, H.; Xu, Z.S.; Zeng, X.J. Hesitant fuzzy linguistic term sets for linguistic decision making: Current developments, issues and challenges. *Inf. Fusion* **2018**, *43*, 1–12.
11. Wang, J.Q.; Wu, J.T.; Wang, J.; Zhang, H.-Y.; Chen, X.-H. Interval-valued hesitant fuzzy linguistic sets and their applications in multi-criteria decision-making problems. *Inf. Sci.* **2014**, *288*, 55–72. [CrossRef]
12. Wang, H.D.; Pan, X.H.; He, S.F. A new interval type-2 fuzzy VIKOR method for multi-attribute decision making. *Int. J. Fuzzy Syst.* **2018**. [CrossRef]
13. Pedrycz, W. Shadowed sets: Representing and processing fuzzy sets. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **1998**, *28*, 103–109. [CrossRef] [PubMed]
14. Pedrycz, W. From fuzzy sets to shadowed sets: Interpretation and computing. *Int. J. Intell. Syst.* **2009**, *24*, 48–61. [CrossRef]
15. Yao, Y.Y.; Wang, S.; Deng, X.F. Construction shadowed sets and three-way approximations of fuzzy sets. *Inf. Sci.* **2017**, *412–413*, 132–153. [CrossRef]
16. Li, C.D.; Yi, J.Q.; Wang, H.K.; Zhang, G.; Li, J. Interval data driven construction of shadowed sets with application to linguistic words modelling. *Inf. Sci.* **2018**. [CrossRef]
17. Atanassov, K.T. Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* **1986**, *20*, 87–96. [CrossRef]
18. Yager, L.L. Pythagorean membership grades in multicriteria decision making. *IEEE Trans. Fuzzy Syst.* **2014**, *22*, 958–965. [CrossRef]
19. Zhang, X.L.; Xu, Z.S. Extension of TOPSIS to Multiple Criteria Decision Making with Pythagorean Fuzzy Sets. *Int. J. Intell. Syst.* **2014**, *29*, 1061–1078. [CrossRef]
20. Liang, D.; Xu, Z.S. The new extension of TOPSIS method for multiple criteria decision making with hesitant Pythagorean fuzzy sets. *Appl. Soft Comput.* **2017**, *60*, 167–179. [CrossRef]
21. Zhang, X.L. Multicriteria Pythagorean fuzzy decision analysis: A hierarchical QUALIFLEX approach with the closeness index-based ranking methods. *Inf. Sci.* **2016**, *330*, 104–124. [CrossRef]
22. Peng, J.D.; Yang, Y. multi-attribute group decision making method based on Pythagorean fuzzy linguistic set. *Comput. Eng. Appl.* **2016**, *52*, 50–54. (In Chinese)
23. Yager, L.L. Pythagorean Membership Grades, Complex Numbers, and Decision Making. *Int. J. Intell. Syst.* **2013**, *28*, 436–452. [CrossRef]
24. Xu, Z.S. A method based on linguistic aggregation operators for group decision making with linguistic preference relations. *Inf. Sci.* **2004**, *166*, 19–30. [CrossRef]
25. Herrera, F.; Herrera-Viedma, E. Linguistic decision analysis: Steps for solving decision problems under linguistic information. *Fuzzy Sets Syst.* **2000**, *115*, 67–82. [CrossRef]
26. Liu, F.L.; Mendel, J.M. Encoding words into interval type-2 fuzzy sets using an interval approach. *IEEE Trans. Fuzzy Syst.* **2008**, *16*, 1503–1521.
27. Wu, D.R.; Mendel, J.M.; Coupland, S. Enhanced interval approach for encoding words into interval type-2 fuzzy sets and its convergence analysis. *IEEE Trans. Fuzzy* **2012**, *20*, 499–513.

28. Walpole, R.W.; Myers, R.H.; Myers, A.L.; Ye, K. *Probability & Statistics for Engineers and Scientists*, 8th ed.; Prentice-Hall: Upper Saddleback River, NJ, USA, 2007.
29. Walpole, R.E.; Myers, R.H.; Ye, K. *Probability and Statistics for Engineers and Scientists*; Pearson: London, UK, 2014.