DOI: 10.1051/matecconf/201710002023

Selecting Visualization Alternative Based on Uncertain Theory

Ping Huang[1,*], Zhiying Lv[1,2] and Liwei Zheng[3]

[1]College of Management, Chengdu University of Information Technology, Chengdu 610103,China
[2]College of Mathematics, University of Electronic Science and Technology of China, Chengdu 610103,China
[3]College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610103,China
Corresponding Email: 229059014@qq.com

Abstract. Multiple attribute decision making (MADM) is an efficient way to solve complex systems, and has wide application. This research develops MCDM model based on uncertain theory, used for selecting a suitable visualization alternative for tourism. First, in order to achieve desirable decision making, a new concept is proposed, which is called the best and the worst reference uncertain linguistic variable as a datum uncertain linguistic variable. At the same time, a new method for ranking uncertain linguistic variable is also presented. Second, based on the preference order relation of attributes given by the experts, a new score function is introduced to get the weight vector of attributes. Finally, the evaluation system of tourism big data visualization alternatives is constructed and the order of those alternatives is acquired by the decision method.

Keywords. MADM, score function, reference uncertain linguistic variable, tourism big data visualization.

1 Introduction

Big data [1] has been risen as national strategic resources. In recent years, after many developed countries such as Europe, the United States, Japan and South Korea took large data as a national level strategy, China have brought the construction of large data into the national strategic choice. Large datawas written in the government work report in 2014. In January 2015, the National Tourism Administration in the "Guidance on the promotion of the development of intellectual tourism," said: "By 2020, large data mining and intelligent marketing capabilities will be improved significantly, mobile e-commerce, tourism data system analysis, artificial intelligence technology will be applied in the tourism industry more widely. And the main task of building tourism data has been clear many times in the views.

Tourism big data visualization [2] decision-making is an emerging field of the internet age. In the internet context, tourism big data has a large scale and has a high complexity. It is a huge challenge to search, analysis and understand the large number of unstructured data and semi structured data. The value of the rich information behind the hidden is reflected only through the collection, analysis, interpretation and expression. And visualization is the most effective way for people to understand easily the value of complex information data. Therefore, it is significant to study the decision making of tourism big data visualization. It can not only promote the fusion and innovation of data mining, analysis techniques and methods, computer graphics technology and decision theory and methods, but also have a transformative influence on thinking and methods of government departments, tourism enterprises and tourists. Therefore, it can provide a more rapid, effective and scientific decision-making protection.

In the process of decision, there are some difficulties for experts to express their preference degrees with crisp numerical values. So, it is another possible way to use linguistic labels [3], which represents qualitative aspects values. We will consider a finite and totally ordered label set $S = \{s_0, s_1, \cdots, s_{U-1}\}$ in the usual sense and with odd cardinality, where each label s_i represents a possible value for a linguistic real variable. The number of linguistic terms in the set S is called the cardinality of S. In the symbolic computation process, the discrete linguistic set S is extended to a continuous interval $\overline{S} = \{s_\alpha : \alpha \in [0,t]\}$. $\overline{S}$ is called an extended linguistic term set associated with S[4]. Let $s \in \overline{S}$, $I(s)$ is denoted as the position index of s and called the gradation of s in $\overline{S}$. For example, $I(s_\alpha) = \alpha$. The 2-tuple fuzzy linguistic representation model represents the linguistic information by means of a 2-tuple (s, α), where s is a linguistic label and the numerical value and α represents the value of the symbolic translation [4, 5].

In this paper, the focus is on the extension of discrete support model for MADM [6, 7] in which the experts express their opinions by means of uncertain linguistic setting in stead of precise numerical values. In order to get the weight vector, a new method by the definition of score function of attribute and the preference order of attribute presented by different experts is presented. Moreover， a new ranking method is presented to rank uncertain linguistic variable, which is based on the best and the worst reference uncertain linguistic variable.

2 Preliminaries

In this section, the basicconcepts and their extensions of the this paper are briefly introduced.

Definition 1 Let $\tilde{s} = [s_{a^-}, s_{a^+}]$, where $s_{a^-}, s_{a^+} \in \overline{S}$, s_{a^-} and s_{a^+} are lower and upper limits, respectively. Then $\tilde{s}$ is called an uncertain linguistic variable.

Definition 2 Let $\{\tilde{s}_{a_1}, \tilde{s}_{a_2}, \cdots, \tilde{s}_{a_n}\}$ be a set of uncertain linguistic variables, where $\tilde{s}_{a_i} = [s_{a_i^-}, s_{a_i^+}] \in \tilde{S}$.The associated weight vector is $\omega = (\omega_1, \omega_2, \cdots, \omega_n)$, where $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$. The uncertain linguistic weighted average (ULWA)operator comes from the extended linguistic representation model is defined as:

$$ULWA(\tilde{s}_{a_1}, \tilde{s}_{a_2}, \cdots, \tilde{s}_{a_n}) = \tilde{s}_a = [s_{a^-}, s_{a^+}] \tag{1}$$

where

$$\begin{aligned} a &= [a^-, a^+] \\ &= [\sum_{i=1}^{n} \omega_i I(s_{a_i^-}), \sum_{i=1}^{n} \omega_i I(s_{a_i^+})]. \\ &= [\sum_{i=1}^{n} \omega_i a_i^-, \sum_{i=1}^{n} \omega_i a_i^+] \end{aligned}$$

The deviation measures have been discussed between two uncertain linguistic terms. The following deviation measure for two uncertain linguistic variables is first introduced.

Definition 3 Let $\tilde{s}_{a_1} = [s_{a_1^-}, s_{a_1^+}]$ and $\tilde{s}_{a_2} = [s_{a_2^-}, s_{a_2^+}]$ be two uncertain linguistic variables. The deviation measure between $\tilde{s}_{a_1}$ and $\tilde{s}_{a_2}$ is defined as follows:

$$d(\widetilde{s}_1,\widetilde{s}_2)=\frac{1}{2U}(|I(s_{a_1^-})-I(s_{a_2^-})|^2+|I(s_{a_1^+})-I(s_{a_2^+})|^2) \\ =\frac{1}{2U}(|a_1^- - a_2^-|^2+|a_1^+ - a_2^+|^2) \quad (2)$$

where U is the number of linguistic terms in the set S .

Clearly, $0\le d(\widetilde{s}_1,\widetilde{s}_2)\le 1$; $d(\widetilde{s}_1,\widetilde{s}_2)=0$ if and only if $\widetilde{s}_1=\widetilde{s}_2$.

Definition 4 Suppose $s_{a_i}=[s_{a_i^-},s_{a_i^+}], i\in I$ is a group of linguistic interval valuables, the best and the worst referenceuncertain linguistic valuable as a datum uncertainlinguistic variables are $s_p=[s_{p^-},s_{p^+}], i\in I$ and $s_q=[s_{q^-},s_{q^+}], i\in I$, where

$$s_{p^+}=\max_{i\in I}\{s_{a_i^+}\},$$

$$s_{p^-}=\max_{i\in I}\{\{s_{a_i^-},s_{a_i^+}\}\setminus s_{p^+}\},$$

$$s_{q^-}=\min_{i\in I}\{s_{a_i^-}\},$$

$$s_{q^+}=\min_{i\in I}\{\{s_{a_i^-},s_{a_i^+}\}\setminus s_{q^-}\}.$$

Definition 5 Suppose $s_{a_i}=[s_{a_i^-},s_{a_i^+}], i\in I$ is a group of uncertain linguistic variables, the best and the worst referenceuncertain linguistic valuable as a datum uncertainlinguistic variables are $s_p=[s_{p^-},s_{p^+}], i\in I$ and $s_q=[s_{q^-},s_{q^+}], i\in I$, then the relative correlation coefficient of s_{a_i} is defined as:

$$\alpha_i=\frac{d(s_{a_i},s_q)-d(s_{a_i},s_p)}{d(s_p,s_q)} \quad (3)$$

Obviously,

(1) $-1\le\alpha_i\le 1,\quad i\in I.$

(2) If $\alpha_i>\alpha_j, \alpha_j>\alpha_k$, then $\alpha_i>\alpha_k, i,j,k\in I.$

For $\alpha_i>\alpha_j, \alpha_j>\alpha_k, i,j,k\in I$, then

$$\frac{d(s_{a_i},s_q)-d(s_{a_i},s_p)}{d(s_p,s_q)} \\ >\frac{d(s_{a_j},s_q)-d(s_{a_j},s_p)}{d(s_p,s_q)} \\ >\frac{d(s_{a_k},s_q)-d(s_{a_k},s_p)}{d(s_p,s_q)}$$

We get $\alpha_i > \alpha_k$, so can rank uncertain linguistic variables based on their relative correlation coefficient. **Theorem 1** Suppose $s_{a_i} = [s_{a_i^-}, s_{a_i^+}], i \in I$ is a group of linguistic interval valuable, the best and the worst referenceuncertain linguistic valuable as a datum uncertain linguistic variables are $s_p = [s_{p^-}, s_{p^+}], i \in I$ and

$$s_q = [s_{q^-}, s_{q^+}], i \in I \text{, then } s_p \ge s_{a_i} \ge s_q .$$

Proof. For $\forall k \in I$, according to the definition of s_p, we get $s_{a_k^+} \le s_{p^+}$. So if $s_{a_k^+} \le s_{p^+}$, suppose $s_{a_l^+} \le s_{p^+}$, $l \in I$, then $s_{p^-} = \max_{\substack{i \in I \\ i \ne l}} \{\{s_{a_i^-}, s_{a_i^+}\} \setminus s_{p^+}\}$, so $s_{a_k^-} \le s_{p^-}$.

If $s_{a_k^+} = s_{p^+}$,then $s_{p^-} = \max_{\substack{i \in I \\ i \ne l}} \{\{s_{a_i^-}, s_{a_i^+}\} \setminus s_{a_k^+}\}$. Obviously, $s_{a_k^-} \le s_{p^-}$ .Overall, $s_{a_k} \le s_p$.

so, for $\forall i \in I, s_{a_i} \le s_p$. In the same way, $s_{a_i} \ge s_q$.

Definition 6 Suppose $C = \{c_1, c_2, \cdots, c_m\}$ is the set of attributes, the score function of attribute $c_i \in C$ is defined as:

$$s_i = \sum_{\substack{i,j=1 \\ i \ne j}}^{m} r_{ij} ,$$

where

$$r_{ij} = \begin{cases} 1 & c_i \succ c_j \\ 0.5 & c_i \approx c_j \quad c_i, c_j \in C \\ 0 & c_i \prec c_j \end{cases} . \tag{4}$$

3 The solution approach

Let $A = \{A_1, A_2, \cdots, A_n\}$ be the alternative set and $C = \{c_1, c_2, \cdots, c_m\}$ be the set of all attributes. Assume $\omega = (\omega_1, \omega_2, \cdots, \omega_m)^T$ is the weight vector of attributes, such that $\sum_{j=1}^{m} \omega_j = 1$, $\omega_j \ge 0$ and ω_j denotes the weight of attribute c_j. Let $E = \{e_1, e_2, \cdots, e_t\}$ be the set of experts who give the preference of attributes and $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_t)$ be the weight vector ofthe experts, where $\lambda_k \in (0,1)$, $\sum_{k=1}^{t} \lambda_k = 1$. Suppose $\widetilde{S} = (s_{a_{ij}})_{n \times m}$ is the uncertain linguistic decision matrix given by the experts, where $s_{a_{ij}} = [s_{a_{ij}^-}, s_{a_{ij}^+}]$ represents the performance of the alternative A_i with respect to the attribute c_j.

Now, we will give the process of decision based on the mentioned method before.

Step 1. Based on Eq .(4), compute the weight vector of attributes as:

$$\varpi_j = \frac{\sum_{k=1}^{t} \lambda_k s_j^k}{\sum_{j=1}^{m} \sum_{k=1}^{t} \lambda_k s_j^k} \quad (j = 1, 2, \cdots, m) \tag{5}$$

where s_j^k denote the score of c_j given by expert $e_k, k = 1, 2, \cdots, t$.

Step 2. Based on the decision matrix and Eq.(1), for every alternative A_i, compute the corresponding weighted decision value s_{a_i} as follows:

$$s_{a_i} = ULWA(s_{a_{i1}}, s_{a_{i2}}, \cdots, s_{a_{im}}) = [s_{a_i^-}, s_{a_i^+}] \quad (6)$$

Where

$$[a_i^-, a_i^+] = [\sum_{j=1}^{m} \omega_j I(s_{a_{ij}^-}), \sum_{j=1}^{m} \omega_j I(s_{a_{ij}^+})] = [\sum_{j=1}^{m} \omega_j a_{ij}^-, \sum_{j=1}^{m} \omega_j a_{ij}^+]$$

Step 3. For the set $\{s_{a_1}, s_{a_2}, \cdots, s_{a_n}\}$, decide the the best and the worst referenceuncertain linguistic valuable as a datum linguistic interval $s_p = [s_{p^-}, s_{p^+}]$ and $s_q = [s_{q^-}, s_{q^+}]$,based on definition 4,that is

$$s_{p^+} = \max_{i=1,2,\cdots n} \{s_{a_i^+}\},$$

$$s_{p^-} = \max_{i=1,2,\cdots n} \{\{s_{a_i^-}, s_{a_i^+}\} \setminus s_{p^+}\},$$

$$s_{q^-} = \min_{i=1,2,\cdots,n} \{s_{a_i^-}\},$$

$$s_{q^+} = \min_{i=1,2,\cdots,n} \{\{s_{a_i^-}, s_{a_i^+}\} \setminus s_{q^-}\}. \quad (7)$$

Step 4. Then, compute the relative correlation coefficient α_i of $s_{a_i}, (i = 1,2,\cdots,n)$ by Eq.(3) , that is

$$\alpha_i = \frac{d(s_{a_i}, s_q) - d(s_{a_i}, s_p)}{d(s_p, s_q)} \quad (8)$$

then the bigger the α_i , the better the alternative A_i , so the overall ranking is acquired.

4 The establishment of the evaluation

At present, the tourism big data visualization has gradually been widely used in industry. The characteristics of tourism big data visualization alternative can be describes by a group of factor attributes, and experts can evaluate it with these attributes. Based on the National Tourism Administration website and the latest research results, experts summarized the evaluation rules and related factors generally considered in the process of evaluating the tourism big data visualization alternative, then construct the evaluation system, which will be presented as follows:

Let $A = \{A_1, A_2, \cdots, A_5\}$ be the set of alternatives, where A_1 is statistic graph tool, A_2 is information graph tool, A_3 is map tool, A_4 is vectored graph tool, A_5 is statistic graph measured tool.

Let $C = \{c_1, c_2, c_3, c_4\}$ be the set of attributes, where c_1 is for the government, c_2 is for the enterprise, c_3 is for the tourism practitioners, c_4 is for the broad masses of the visitor.

Suppose there are four experts who give the preference order of attributes as follows:

$$c_4 \succ c_1 \approx c_3 \succ c_2$$

$$c_2 \succ c_1 \succ c_3 \succ c_4$$

$$c_4 \succ c_2 \succ c_1 \succ c_3$$

$$c_1 \succ c_2 \succ c_3 \succ c_4$$

After many years 'research analysis by many experts, we can construct the decision matrix as follows:

$$\widetilde{S}=\begin{pmatrix}[s_7,s_9] & [s_7,s_9] & [s_5,s_7] & [s_5,s_6]\\ [s_4,s_6] & [s_3,s_4] & [s_6,s_8] & [s_7,s_9]\\ [s_5,s_7] & [s_5,s_6] & [s_5,s_7] & [s_7,s_9]\\ [s_4,s_6] & [s_7,s_8] & [s_3,s_4] & [s_3,s_5]\\ [s_5,s_6] & [s_5,s_7] & [s_7,s_8] & [s_6,s_7]\end{pmatrix}$$

In the following, the proposed method is applied to solve this problem and the computational procedure is summarized in the following.

Step1.Using Eq.(5), compute the weight vector of attribute:

$$\omega_1=0.333,$$
$$\omega_2=0.308,$$
$$\omega_3=0.112,$$
$$\omega_4=0.247$$

Step 2. Compute the weighted decision value by Eq.(6):

$$s_{a_1}=[(s_6,0.282),(s_8,0.035)]\,,\; s_{a_2}=[(s_5,-0.343),(s_6,0.349)]\,,$$

$$s_{a_3}=[(s_5,0.494),(s_7,0.186)]\,,$$

$$s_{a_4}=[(s_5,-0.435),(s_6,0.145)]\,,$$

$$s_{a_5}=[(s_5,0.471),(s_7,-0.221)]$$

Step 3. Decide the best and the worst referenceuncertain linguistic valuables by Eq.(7):

$$s_p=[(s_7,0.186),(s_8,0.035)]\,,$$
$$s_q=[(s_5,-0.435),(s_5,-0.343)]$$

Step4.Compute the relative correlation coefficient α_i of $z_i,(i=1,2,\cdots,n)$ by Eq.(8):

$$\alpha_1=0.785,$$
$$\alpha_2=-0.348,$$
$$\alpha_3=0.201,$$
$$\alpha_4=-0.435,$$

Then, we can rank all the alternatives:

$$A_1\succ A_3\succ A_5\succ A_2\succ A_4$$

and thus the most desirable alternative is A_1.

5 Conclusion

There are many types of visualization alternatives faced by traveler, so it is hard to see which one is better. In order to deal with this problem, visualization alternative evaluations should be done to help people to decide on the better option for tourism. In this paper, a new multiple attribute decision method based on uncertain theory is given to evaluate visualization alternatives. Compared to traditional way of assessment, the new selection model proposed by this article is less liable to fault and more feasible to practice. This method can not only be used to solve visualization alternatives selection problem, but it can also be used to deal with select problem of many similar issues. In the practice application, people can not only adopt the most desirable alternative, but also combine several desirable alternative.In future research, the decision model presented in this paper will be extended according to the frequency of the application in different scenarios and group decision [3,8].

Acknowledgment

The National Natural Science Foundation Project of China (No. 61203285;No.61673285; No.11671284)

The Province Department of Science Soft Science Project in Sichuan (2016ZR0095)

The high level research team of the major projects division of Sichuan province(Sichuan letter [2015] no.17-5)
The Project of Chengdu University of InformationTechnology (CRFKYTZ201644)

References

1. F.F. Luo, The systematic review about the theory of big data and the development of practice, Business, 11(2016):219.
2. Sébastien Liment, Millian Poquet, Sophie Robert, A frameworkfor efficient dynamic interactive scientific visualization, Procedia Computer Science. **29** (2014):692-702.
3. Z.S. Xu, Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment, Information Sciences, 168(2004): 171-184.
4. D.F. Li, Multiple attribute decision making method using extended linguistic variables, Int J Uncertain Fuzziness Known Based System17(2009): 793-806.
5. Herrera F, Marteinez L, A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multiple experts decision making. IEEE Transactions on SystemsManand cybernetics, 31(2001):227-234.
6. Z. Y. Lv, A fuzzy multiple attribute decision making methodbased on possibility degree, Journal of Intelligent & Fuzzy Systems. 31 (2016) 787–794.
7. D.F. Li, Multiple attribute decision making method using extended linguistic variables, Int J Uncertain Fuzziness Known Based System,17(2009):793-806.
8. J.P. Xu, Z.B. Wu, A discrete consensus support for multiple attribute group decision making, Knowledge-Based systems 24(2011): 1196-1202..