Study on Updating Algorithm of Attribute Coordinate Evaluation Model

Xiaolin Xu[1(✉)], Guanglin Xu[2], and Jiali Feng[3]

[1] Shanghai Polytechnic University, Shanghai, China
xlxu2001@163.com
[2] Shanghai Lixin University of Commerce, Shanghai, China
glxu@outlook.com
[3] Shanghai Maritime University, Shanghai, China
jlfeng@shmtu.edu.cn

Abstract. Evaluation model based on attribute coordinate has made some achievements in both theoretical research and practical applications. However, if the new evaluation samples are added, the evaluation model needs to be reconstructed rather than the dynamic updating. Almost no progress has been made on how to dynamically update the evaluation model. Thus, this paper puts forward a dynamic updating algorithm based on barycentric coordinates and satisfaction function to effectively solve this problem. The experiment results show the reasonability and effectiveness of this algorithm.

Keywords: Comprehensive evaluation · Attribute coordinate evaluation · Barycentric coordinates · Global satisfaction

1 Introduction

The main difficulty of comprehensive evaluation is the weight determination of each attribute of evaluated objects by evaluators. The characteristic of evaluation method of attribute coordinate is that the evaluation way is very close to the normal thinking pattern of people and the corresponding preference curve can be accurately constructed according to the preference of the evaluator. Hence, this method can not only learn about the experience of experts but also give full play to the advantages of machine learning [1–5]. After more than ten years of research and application, the evaluation method of attribute coordinate has made a certain achievements in many fields [6–14]. But after the new evaluation samples are added, the evaluation model needs to be reconstructed rather than the dynamic update. There has been no progress on the study on how to dynamically update the evaluation model. This paper proposes a kind of update model to solve this problem. The remaining parts of this paper are organized as follows. Section 2 introduces the related work including the introduction to updating strategies and evaluation method of attribute coordinate. Section 3 elaborates the algorithm through the flow chart, which combines the evaluation method with strategy

The work was supported by the Key Disciplines of Computer Science and Technology of Shanghai Polytechnic University (No. XXKZD1604).

D.-S. Huang et al. (Eds.): ICIC 2017, Part III, LNAI 10363, pp. 653–662, 2017.
DOI: 10.1007/978-3-319-63315-2_57

III. Section 4 carries out the experiments and validates the updating algorithm. Section 5 comes to the conclusion.

2 Related Work

2.1 Classification on Updating Strategy

(1) The updating of evaluation model

After the completion of evaluation model, if new evaluation sample data are collected, the evaluation model needs to be correspondingly updated. So the updating process refers to the process changing from the evaluation on old samples to that on both new and old samples. For the comprehensive evaluation, the updating of the evaluation model is an essential process. The purpose and effect of evaluation model update are that the original understanding is amended by the latest information to decrease the subjective or objective inaccuracy caused by the imperfect understanding. Moreover, the latest information could be various and contain other information in the comprehensive evaluation, such as the expert experiences and preferences.

The updating mode of evaluation model of attribute coordinate is shown in Fig. 1. The figure indicates that, in the evaluation depending on the weight obtained from samples, when new samples are added, they are likely to have a certain impact on old samples like changing the weight of samples, which further gives rise to the change of evaluation result. Old samples and new samples need to work together to generate new samples in light of a certain weight calculation method.

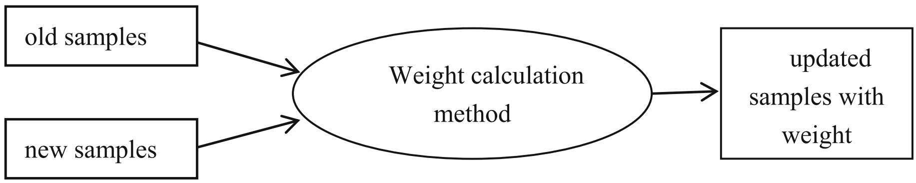

Fig. 1. Updating mode of samples with weight

(2) Updating Strategy

We divide the updating methods of evaluation model into three strategies: Strategy I, Strategy II and Strategy III.

Strategy I

Strategy I refers to the algorithm that after the new data are added to the original data, apply the original evaluation model to update the evaluation results. Its main characteristic is that both old sample data and comprehensive evaluation model are known. Although the updating method is easy, the result may be beyond accuracy, thus the updating is meaningless.

Strategy II

Strategy II refers to the algorithm that all the data join to calculate the weight. Its main characteristic is that the new and old data are thought of as the original data to reconstruct the new evaluation model. Strictly, this kind of updating is the standard to verify the effectiveness of other updating algorithms. However the drawback is it might take enormously long time.

Strategy III

Strategy III refers to the algorithm that the newly collected data are used to supplement and affect the original comprehensive evaluation model. Its core is how to impose the new weight on the latest information.

Obviously, from Strategy I to Strategy III, as the constraint conditions are more and more complex, the updating algorithm will be more and more difficult.

2.2 Introduction to Attribute Coordinate Evaluation

Comprehensive evaluation is to evaluate whether the evaluated objects with multi attributes are good or not. After attribute values with certain unified dimension are given to all the evaluated objects, the optimal evaluated object is the solution A = (10, …,10) that each attribute value is full score (assuming the full score is 10). However, there does not exist the optimal solution in the comprehensive evaluation but generally exists the satisfactory solution. Hence, the satisfactory solution can be obtained instead of the optimal solution in the practical decisions. Thus the comprehensive evaluation could only require the most satisfactory solution meeting some weighted conditions. When evaluators evaluate multi-attribute objects, they often think that some attributes are important, namely, attributes can be given certain weights in the evaluation model. The importance of different attributes would change with the degree to which good evaluated objects or bad evaluated objects belong, namely, the weight of an attribute would dynamically change in the evaluation model.

(1) Solution of attribute barycentric coordinate

It is assumed that $S_T = \{x_i = (x_{i1}, \ldots, x_{im}) \mid \sum_{j=1}^{m} x_{ij} = T\}$ is a hyper plane with the total score T. $x_i = (x_{i1}, \ldots, x_{im})$ are supposed to be independent attribute values. The intersection of S_T and X $(S_T \cap X)$ is a *(n-1)*-dimensional simplex (shown in Fig. 2, e.g. $S_{100} \cap X = \Delta ABC$, and $\Delta A'B'C'$ is a two-dimensional simplex that the total score hyperplane S_T of $10 < T < 10 \times n$ intersects X, $(S_T \cap X) = \Delta A'B'C'$).

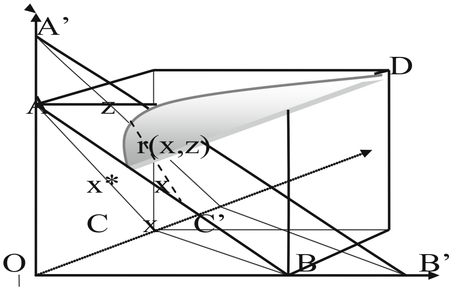

Fig. 2. Two-dimensional simplex from $ST \cap X = \Delta A'B'C'$

Let $\{x_k, k = 1,\ldots,s\} \subseteq ST \cap X$ be the sample solution set and $Z = \{z_1, \cdots, z_n\}$ be the evaluator set. The decision maker z_i selects t sets of satisfactory solutions $\{x_h,\ h = 1,\ldots,t\}$, and each solution is marked as $w_h(x_h)$. Because the solution space X is a convex set with respect to score $w_h(x_h)$, the barycenter $b(\{x_h(z_i)\})$ of $\{x_h(z_i)\}$ can be solved as formula (1) by the weighted average method ($w_h(x_h)$ is taken as the weight).

$$b(\{x_h(z_i)\}) = \left(\frac{\sum_{h=1}^{t} w_h^i x_{1h}}{\sum_{h=1}^{t} w_{1h}^i}, \cdots, \frac{\sum_{h=1}^{t} w_{mh}^i x_{mh}}{\sum_{h=1}^{t} w_{mh}^i} \right) \tag{1}$$

(2) Solution of attribute barycentric curve

Obviously, when we have large enough training sample set $\{x_k\}$ and more enough training times, the solution sets selected by decision makers z_i are enough, the barycenter $b(\{x_h(z_i)\})$ would be close to the local most satisfactory solution x*|T, namely, $\lim_{h\to\infty} b(\{x_h(z)\}) \to x^*|T$. When the number of decision makers Z is more than 1, $\lim_{h\to\infty} b'(\{x_h(Z)\})$ is the local most satisfactory solution of decision makers Z in $S_T \cap X$, namely, the barycenter of decision makers. The set of all local most satisfactory solutions $\{b'(\{x_h(Z)\})|T \in [10, 10 \times m]\}$ can be obtained after T covers the interval $[10, 10 \times m]$. The set would be a line, written as $L(b'(\{x_h(Z)\}))$, and this line is called as the local most satisfactory linear solution of decision makers Z. The line can be obtained by polynomial curve fitting method such as the following polynomial function (2):

$$G(T) = a_0 + a_1 T + a_2 T_2 \tag{2}$$

In this situation, three local most satisfactory solutions of decision makers in S100 ∩ X, ST ∩ X and S10 × m ∩ X are taken as three interpolation points to be substituted into Lagrange interpolation formula (3) to calculate the local most satisfactory solution line.

$$g_i(T) = \frac{(T - x_1^*)(T - x_2^*)}{(x_0^* - x_1^*)(x_0^* - x_2^*)} a_{i0} + \frac{(T - x_0^*)(T - x_2^*)}{(x_1^* - x_0^*)(x_1^* - x_2^*)} a_{i1} + \frac{(T - x_0^*)(T - x_1^*)}{(x_2^* - x_0^*)(x_2^* - x_1^*)} a_{i2} \tag{3}$$

2.3 The Work of the Paper

After the calculation of barycentric point $b'(\{x_h\})$, if the new sample data are collected again, Strategy II is used to substitute both new and old sample data into the formula (3) for recalculation. The advantage of this method is that the calculation results are very accurate, but the disadvantages are that repeated marks on old samples are unavoidable. Besides, the calculation will not be feasible if the old data are lost. However, if Strategy III is used, only the marks on new samples are needed, hence this algorithm is simple and effective. In the paper, we integrate the Strategy III into the evaluation model of attribute coordinate to accomplish the updating of the model in the case of new sample data being added into the current model. The combined algorithm is described as follows.

3 Updating Algorithm of Evaluation Model of Attribute Coordinate

3.1 Algorithm Process

Figure 3 illustrates the process of the algorithm in the way of flow chart.

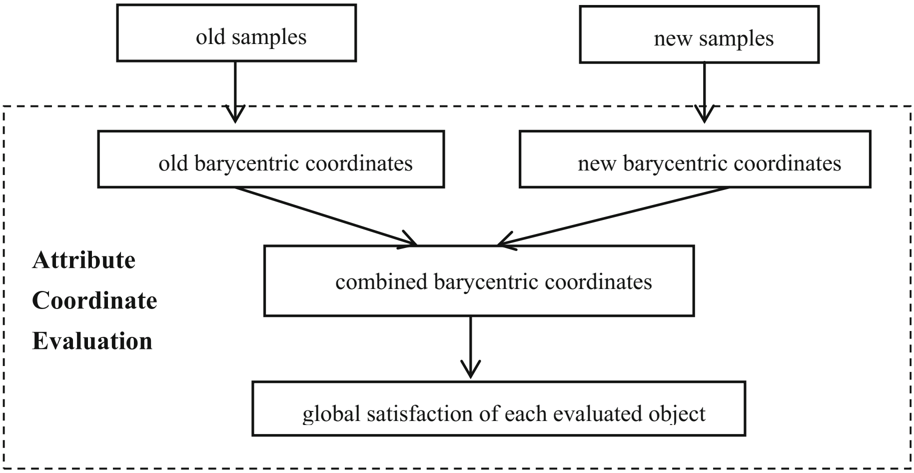

Fig. 3. The process of the algorithm

3.2 Updating of Attribute Barycentric Coordinates

After calculating old barycentric coordinates, if t new samples are added and they are marked as $w_t(\{x_h\})$, respectively, the formula (1) can be updated as formula (4). Where s is the number of old samples, $x_1,\ldots, x_m$ are the old barycentric coordinates.

$$b^u = \left[\frac{\frac{\sum_{h=1}^{t} w_h^i x_{1h}}{\sum_{h=1}^{t} w_{1h}^i} + sx_1}{t+s}, \cdots, \frac{\frac{\sum_{h=1}^{t} w_{mh}^i x_{mh}}{\sum_{h=1}^{t} w_{mh}^i} + sx_m}{t+s} \right] \tag{4}$$

The new barycentric coordinate figured out is $b^{\mathrm{u}}(\{x_h(z_i)\})$, and the barycentric coordinate of old model is $b^u(\{x_h(z_i)\})$. Their norm can be solved according to the formula (5) and be calculated by Euclidean distance, as green segment shown in Fig. 4.

$$e = \|b^{\mathrm{u}}(\{x_h(z_i)\}) - b'(\{x_h(z_i)\})\| \tag{5}$$

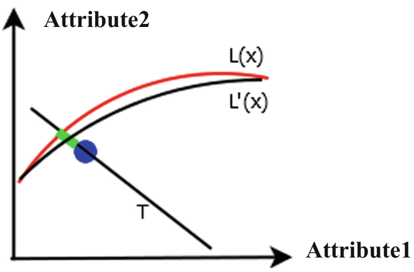

Fig. 4. The old and updated barycentric curve (Color figure online)

Figure 4 also shows the old psychological barycenter curve L(x) and the updated psychological barycentric curve L'(x). T represents a certain total score obtained from Attribute1 and Attribute2.

3.3 Satisfaction Calculation after Updating Barycentric Coordinates

When there exists the new psychological standard point $b^u(\{x_h(z_i)\})|T$ of Z in $S_T \cap X$, taking $b^u(\{x_h(z_i)\})|T$ as the standard, the global satisfaction function (6) is used for the satisfaction evaluation of all the solutions in $S_T \cap X$.

$$sat(x, Z) = \lambda(x, Z) * \exp\left(-\frac{\sum_{j=1}^{m} w_j \left| x_{ij} - b_{ij}^u(\{x_h(Z)) \right|}{\sum_{j=1}^{m} \delta_j (b_{ij}^u(\{x_h(Z)) - \delta_j)} \right) \tag{6}$$

Where $sat(x_{i,}Z)$ is the satisfaction of evaluated object x_i from evaluator Z, whose value is expected to be between 0 and 1. x_{ij} is each attribute value. $\left| x_{ij} - b_{ij}^u(\{x_h(Z)) \right|$ is to measure the difference between each attribute value and the corresponding barycentric value. w_j and δ_j are used as the factor which can be adjusted to make the satisfaction comparable value in the case where the original results are not desirable. It is found from the above learning that all the solutions in $L(b^u(\{x_h(Z)\}))$ are determined by taking the local most satisfactory solution as the standard in different simplexes $S_T \cap X$ as well as calculating the maximal satisfaction by formula (6). The correction coefficient $\lambda(x, Z)$ is described as $\lambda(x, Z) = \left(\frac{\sum_{i=1}^{m} x_{ij}}{\sum_{j=1}^{m} X_j} \right)^S$, Where $\sum_{j=1}^{m} X_j$ is the sum of X_j with each attribute value the full score, $\sum_{ij=1}^{m} x_{ij}$ is the sum of each attribute value x_{ij} of solution

x_i, and $S = \left(\frac{\sum_{j=1}^{m} Xj}{\sum_{i=1}^{m} xij} \right)$. Formula (6) can reach consistent satisfaction $sat(x, Z)$ in the entire solution space, so λ(x,Z) is called the global consistent coefficient.

4 Experiments

The data set comes from the scores of college entrance examination of a certain city in 2015 and contains more than 4,000 objects to be evaluated. Moreover, each object contains four attributes: Chinese, mathematics, English and comprehensive discipline. This simulation adopts the first three attributes in order to get a more intuitive three-dimensional chart.

4.1 Experiments of Updated Barycentric Coordinates

According to evaluation model of attribute coordinates, firstly, a hyper plane with the total score of 370 is selected, on which there are 26 sample points. Then the 26 sample points are marked by five experts. It is assumed that 18 sample points are initial samples and the rest (8 sample points)are the supplementary samples. The above attribute coordinate updating algorithm is used to update the barycentric coordinates of attributes.

(a) Calculate the original barycentric coordinates. According to formula (1), the barycentric coordinates of 18 sample points are figured out, as black points shown in Fig. 5.
(b) Calculate the new barycentric coordinates by the updating method of Strategy II. According to formula (1), the barycentric coordinates of 26 sample points are figured out, as green points shown in Fig. 5.

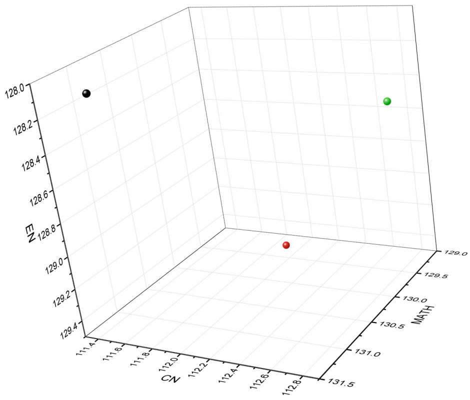

Fig. 5. Comparison among original and updated barycentric coordinates with Strategy II and III (Color figure online)

Table 1. Comparison of original and updated barycentric coordinates

Barycentric coordinates	Chinese	Maths	English
Original barycentric coordinates	111.47	131.11	128.12
Updated coordinates by Strategy II	112.74	129.24	128.51
Updated coordinates by Strategy III	111.99	129.50	129.41

(c) Calculate the new barycentric coordinates by the updating method of Strategy III. The barycentric coordinates integrated with the other 8 sample points are obtained according to formula (4), as red points shown in Fig. 5.

Table 1 shows the results of barycentric coordinates in terms of the above three cases.

As we had stated in Sect. 2, the updated barycentric coordinates by Strategy III provides more accurate result as the standard to verify the effectiveness of other updating algorithms. It can be seen from Table 1 that the updated barycentric coordinates by Strategy III cannot be completely equal to the updated barycentric coordinates by Strategy II but be close to it.

Figure 5 illustrates the position of the three barycentric coordinates in three dimentional space, each point including three attributes indicated with CN (Chinese), MATH (mathematics), and EN (English).

In Fig. 6, the black point is the barycentric coordinates which are figured out from the original samples. Then the satisfaction curve illustrates the degrees of satisfaction of samples according to formula (6), as the red line shown in Fig. 6. The barycentric coordinates of 8 sample points evaluated by experts are shown as the blue point. The new barycentric coordinates updated by formula (4) is shown as the red point. Finally, the updated satisfaction curve is figured out as the yellow line (solid line).

From Fig. 6, the satisfaction curve after adding sample is basically close to the original satisfaction curve (dotted line), which indicates that our updating algorithm is effective. Compared with the Strategy II which needs to calculate the barycentric coordinates for all the samples again, this algorithm greatly avoids the repeated evaluation on old samples and effectively reduces the time of evaluation.

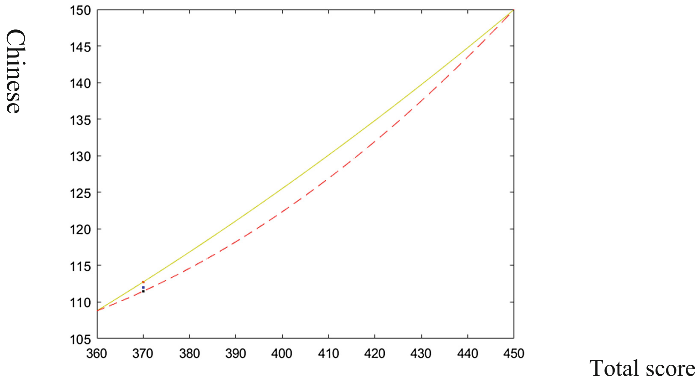

Fig. 6. Comparison of original and updated barycentric curves (Color figure online)

5 Conclusions

This paper studies how to update the existing evaluation model based on the evaluation of attribute coordinate after the new sample data are added and presents the updating algorithms of barycentric coordinates and satisfaction function. The simulation of comparison is carried out according to the real data. It can be clearly seen from the calculation results that the updating strategy of Strategy III is used to figure out the new barycentric coordinates and then to calculate the new satisfaction for all samples by calculating the barycentric coordinates of supplementary samples. This can not only realize the effective update of evaluation system but also greatly decrease the calculation work. In the further work, we would conduct the multi-expert evaluation on the increased samples, integrate the supplementary samples into the old evaluation model, and then observe how they affect the evaluation results.

References

1. Feng, J.L., Zhao, T.A.: KDD model based on conversion of quantity-quality features of attributes. J. Comput. Res. Dev. **37**(9), 1114–1119 (2000)
2. Feng, J.L., Feng, J.J.: Granularity transformation of qualitative criterion, orthogonality of qualitative mapping system, and pattern recognition. In: IEEE International Conference on Granular Computing, pp. 211–216 (2006)
3. Xu, G.L., Xu, X.L.: Study on evaluation model of attribute barycentric coordinates. Int. J. Grid Distrib. Comput. **9**(9), 115–128 (2016)
4. Wu, Q.F., Feng, J.L.: A kind of evaluation and decision model based on analysis and learning of attribute coordinate. J. Nanjing Univ. (Nat. Sci.) **39**(2), 182–188 (2006)
5. Xu, X.L., Feng, J.L.: A quantification method of qualitative indices based on inverse conversion degree functions. In: Enterprise Systems Conference, pp. 261–264 (2014)
6. Duan, X.Y., Yu, S.Q., Xu, G.L., Feng, J.L.: Evaluation on comprehensive competitiveness of inland port based on the method of arrtribute theory. J. Pattern Recogn. Image Proress. **4**, 373–380 (2013)
7. Xu, G.L., Min, S.: Research on multi-agent comprehensive evaluation model based on attribute coordinate. In: IEEE International Conference on Granular Computing (GrC), pp. 556–562 (2012)
8. Xu, X.L., Xu, G.L., Feng, J.L.: A kind of synthetic evaluation method based on the attribute computing network. In: IEEE International Conference on Granular Computing (GrC), pp. 644–647 (2009)
9. Xu, X.L., Xu, G.L.: Research on ranking model based on multi-user attribute comprehensive evaluation method. In: Applied Mechanics & Materials, pp. 644–650 (2014)
10. Xu, X.L., Xu, G.L.: A recommendation ranking model based on credit. In: IEEE International Conference on Granular Computing (GrC), pp. 569–572 (2012)
11. Xu, X.L, Feng, J.L.: Research and implementation of image encryption algorithm based on zigzag transformation and inner product polarization vector. In: IEEE International Conference on Granular Computing, vol. 95, no. 1, pp. 556–561 (2010)
12. Xu, G.L., Feng, J.L., Liu, Y.C.: Pattern recognition method based on the attribute computing network. J. Comput. Res. Dev. (S1) (2008)

13. Xu, G.L., Wang, L.F.: Evaluation of aberrant methylation gene forecasting tumor risk value in attribute theory. J. Basic Sci. Eng. **16**(2) (2008)
14. Feng, J.L., Bi, J.Y.: Lung cancer cells' recognition based on attribute theory. J. Guangxi Normal Univ. (Nat. Sci. Ed.) **29**(3) (2011)