

An Improvement of Grey Integrated Clustering Method

Tianhui Wang, Xiaolu Li, Bingjun Li*

College of Information and Management Science, Henan Agricultural University, Zhengzhou, China

Email: tianhui_wang2011@163.com, *zzbjun@163.com

How to cite this paper: Wang, T.H., Li, X.L. and Li, B.J. (2017) An Improvement of Grey Integrated Clustering Method. *Open Journal of Applied Sciences*, 7, 140-146.
<https://doi.org/10.4236/ojapps.2017.74012>

Received: February 18, 2017

Accepted: April 25, 2017

Published: April 30, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In the grey clustering assessment, the grey integrated clustering method can solve the problem that there are not distinguished differences of grey clustering coefficient. This paper proposes an improved grey integrated clustering method based on the existing problem that there are some deficiencies in the division of the value scope of the integrated cluster coefficients, and proves the effectiveness of the improved method through the empirical analysis.

Keywords

Grey Cluster, Cluster Coefficient, Distinguished Difference

1. Introduction

Grey clustering assessment and analysis is an important part of grey system theory, which has been one of the most important techniques in the theory discussion and practical application. On the basis of the grey variable weight clustering method established by Professor Deng Julong [1]; Professor Liu Sifeng proposed a grey fixed weight clustering evaluation method [2] [3]; Grey optimal cluster model was established by Xiao Xinping [4]; Xu Xiuli discussed the improvement measures of grey clustering analysis [5]; The new grey clustering evaluation model based on the triangle whitenization weight function was also constructed by Professor Liu Sifeng [6]; Mi Chuanmin put forward the entropy weight method to determine the weight of each index in grey cluster [7]; Zhang Qishan discussed the measurement of characteristics of a grey cluster result [8], the problem of grey cluster based on the hazy set is studied by Zhang [9]; Grey interval clustering decision method was proposed by Luo Dang [10]. The above scholars studied from different aspects of grey clustering method and evaluation analysis. However, the above-mentioned various grey clustering analysis methods are based on the maximum principle of the clustering coefficient vector

component to determine the grey clustering object's grey classifications [11]. But in practical application, there are not distinguished differences of grey cluster coefficient. Dang Yaoguo put forward a new method of grey integrated clustering, which attempts to solve the problem of grey clustering that there are not distinguished differences of grey cluster coefficient [12]. In the integrated clustering method, the clustering object is divided into s classifications, and the value scope of the integrated clustering coefficient is divided into s disjoint equal length interval. In the grey clustering assessment, the grey integrated clustering method can solve the problem that there are not distinguished differences of grey cluster coefficient. However, there are some deficiencies in the division of the value scope of the integrated cluster coefficients [13]. This paper puts forward an improved method based on the existing problem, which is more accurate in distinguishing the grey classifications of grey clustering object, and the empirical analysis proves that the improved grey integrated clustering method is more valuable [14].

2. Integrated Grey Clustering Method and Its Improvement

Definition 1 Assume that there are n clustering objects, m clustering indexes, s different grey classifications, observation value of clustering object i on the clustering index j is x_{ij} ($i=1,2,\dots,j=1,2,\dots,m$) and the whitenization weight function of the i -th clustering object of clustering index j about k grey classification is f_j^k ($j=1,2,\dots,m;k=1,2,\dots,s$), the value of whitenization weight function of the i -th clustering object of clustering index j on k grey classification is $f_{ij}^k = f_j^k(x_{ij})$, the whitenization weight function matrix about k grey classifications is $F_k = (f_{ij}^k)_{n \times m}$ ($i=1,2,\dots,n;j=1,2,\dots,m$). Determine the grey clustering weight η_{ij} ($j=1,2,\dots,m$) of each index. Generally, when there is no enough information, the weight can be set as $\eta_{ij} = \frac{1}{m}$ [15].

$$\sigma_i^k = \sum_{j=1}^m f_{ij}^k \eta_{ij} \quad (1)$$

$$\sigma_i = (\sigma_i^1, \sigma_i^2, \dots, \sigma_i^n) \quad (i=1,2,\dots,n) \quad (2)$$

where σ_i^k is clustering coefficient of clustering object i belongs to k grey classification, σ_i is called the clustering coefficient vector of the clustering object i .

Definition 2 Order $\delta_i^k = \frac{\sigma_i^k}{\sum_{k=1}^s \sigma_i^k}$, where δ_i^k is represents normalized clustering coefficient of the clustering object i belongs to k grey classification. Call Equation (3) as normalized clustering coefficient vector of clustering object i [16].

$$\delta_i = (\delta_i^1, \delta_i^2, \dots, \delta_i^s) \quad (i=1,2,\dots,n) \quad (3)$$

Definition 3 Assume that there are n clustering objects s different grey classifications, order $\varepsilon = (1, 2, \dots, s-1, s)^\top$, Equation (4) is called the integrated clustering coefficient of clustering object i .

$$\omega_i = \delta_i \cdot \varepsilon = \sum_{k=1}^s k \cdot \delta_i^k \quad (i=1, 2, \dots, n) \quad (4)$$

Definition 4 Because of the integrated clustering coefficient $1 \leq \omega_i \leq s$, the clustering object is divided into s classifications, and the range of the integrated clustering coefficient is divided into s disjoint equal length interval, as shown in Equation (5).

$$\begin{aligned} & [1, 1+(s-1)/s], [1+(s-1)/s, 1+2(s-1)/s], \dots, \\ & [1+(k-1)(s-1)/s, 1+k(s-1)/s], \dots, [s-(s-1)/s, s] \end{aligned} \quad (5)$$

When there is no significant difference in the clustering coefficient of the i cluster and the integrated clustering coefficient of the object i is $\omega_i \in [1+(k-1)(s-1)/s, 1+k(s-1)/s]$, we call the object i belongs to the k grey classification.

The accurate degree of integrated grey clustering evaluation coefficient is very important, and the grey degree of clustering coefficient is too big or too small may directly lead to the clustering objects belonging to different grey classification, in the definition 3 $\varepsilon = (1, 2, \dots, s-1, s)^T$, integrated clustering coefficients $\omega_i = \delta_i \cdot \varepsilon = \sum_{k=1}^s k \cdot \delta_i^k$ ($i=1, 2, \dots, n$) of numerical and nature of the clustering coefficient exist error, this paper gives a new method, which the integrated clustering coefficients of accuracy is greatly improved, such as the definition 5 [17].

Definition 5 Assume that there are n clustering object, s different grey classification, then $\varepsilon = (1, \sqrt{2}, \dots, \sqrt{s-1}, \sqrt{s})^T$, Equation (6) is called integrated clustering coefficient of the clustering object i .

$$\omega_i = \delta_i \cdot \varepsilon = \sum_{k=1}^s k \cdot \delta_i^k \quad (i=1, 2, \dots, n) \quad (6)$$

Definition 6 The integrated clustering coefficient $1 \leq \omega_i \leq \sqrt{s}$, and the range of the integrated clustering coefficient is divided into s disjoint equal length interval, as shown in Equation (7).

$$\begin{aligned} & [1, 1+(\sqrt{s}-1)/s], [1+(\sqrt{s}-1)/s, 1+2 \cdot (\sqrt{s}-1)/s], \dots, \\ & [1+(k-1) \cdot (\sqrt{s}-1)/s, 1+k \cdot (\sqrt{s}-1)/s], \dots, [\sqrt{s}-(\sqrt{s}-1)/s, \sqrt{s}] \end{aligned} \quad (7)$$

When there is no significant difference in the clustering coefficient of the clustering object i , If the integrated clustering coefficient of the object i is $\omega_i \in [1+(k-1) \cdot (\sqrt{s}-1)/s, 1+k \cdot (\sqrt{s}-1)/s]$, the object i belongs to the k grey classification [18].

This paper regards integrated clustering coefficients of ω_i and integrated clustering coefficients range median z_i (such as:

$z_i = 1 + (2k-1)(\max \omega_i - 1)/2s$) of similar degree L_i as the object i belong to k grey class of precision degree. Order $L_i = (1 - |\omega_i - z_i| / \max \omega_i) \cdot 100\%$. In other words, the distance of integrated clustering coefficients ω_i and integrated clustering coefficients the center value z_i of the interval is closer, the difference is smaller, L_i is bigger, indicating the degree of i objects belong to k grey classification is stronger, this integrated clustering coefficients of numerical accuracy is

higher; conversely, L_i is smaller, indicating that the integrated clustering coefficients of numerical accuracy is poorer, the object i belongs to k grey classification is relatively weaker.

3. Empirical Analysis

The selection of mining method depends on coal mine geological structure and mining technological conditions. There is a wide variation between technical level of the selected mining method and the economic benefits, as the difference geological structure of coal mine or a coal mine in different lots. In order to improve the economic benefit of coal mining, which enterprises could select the appropriate mining method to achieve the economic benefit of coal enterprises best. How to make an objective and integrated evaluation of different mining methods. This paper selects original integrated clustering method and the improved integrated clustering method on different mining methods for analysis and evaluation. A coal mine uses four kinds of methods for coal mining: fully mechanized mining, top grade general mining, general mining and blasting mining, which are regard as clustering objects, then this paper respectively takes the working face per unit area yield, mining effect, the equipment investment and the mining cost as the clustering index. According to the good, medium and poor three class classifications, the different indexes data of four different kinds of mining methods are shown in **Table 1** [19].

Through the investigation of the experts, the whitenization weight function $f_j^k(\cdot)$ ($j = 1, 2, 3, 4; k = 1, 2, 3$) of the j index about the k grey classification are [20]:

$$f_1^1 [2.16, 3.24, -, -], f_1^2 [1.08, 2.16, -, 3.24], f_1^3 [-, -, 1.08, 2.16], f_2^1 [9.6, 14.40, -, -], f_2^2 [4.8, 9.6, -, 14.4], f_2^3 [-, -, 4.8, 9.6], f_3^1 [390, 780, -, -], f_3^2 [390, 780, -, 1170], f_3^3 [-, -, 780, 1170], f_4^1 [-, -, 6.5, 13], f_4^2 [6.5, 13, -, 19.5], f_4^3 [13, 19.5, -, -].$$

According to the Delphi method, the weights of work surface area per unit area yield, mining efficiency, equipment investment and the cost are respectively as follows: $\eta_1 = 0.4547, \eta_2 = 0.2631, \eta_3 = 0.1411, \eta_4 = 0.1411$. Using the software to calculate the results of integrated clustering method and improved integrated clustering method, evaluating the influence degree of economic benefit of different coal mining method in this coal mine, then analyzing and comparing the

Table 1. The observation data of four different coal mining methods in coal mine.

| Clustering object | Per unit area yield (Million tons/month) | Mining efficiency (Tons per work) | equipment investment (Million yuan) | Mining cost (Yuan/ton) |
|--------------------------|---|--------------------------------------|---|---------------------------|
| Fully mechanized mining | 4.34 | 16.37 | 2046 | 10.2 |
| Top grade general mining | 1.76 | 10.83 | 1096 | 18.67 |
| General mining | 1.08 | 6.32 | 523 | 13.72 |
| Blasting mining | 1.44 | 4.81 | 250 | 9.43 |

accuracy of the integrated clustering coefficients calculated by the two methods, namely the grey degree of i to k . Specific results are shown in **Tables 2-4**.

It can be seen from **Table 2** and **Table 3** that the integrated clustering method and the improved integrated clustering method are less influence and the clustering results have no big changes in the economic benefit which brought by different coal mining method in macroscopic word. However, it can be seen from **Table 4** that the accuracy of clustering coefficients of clustering object elevate from fully mechanized mining: 91.9%, top grade Mining: 89.6%, general mining: 94.5%, blasting mining: 90.3% to fully mechanized mining: 95.2%, top

Table 2. The improved integrated clustering method for the evaluation of different mining methods in coal mine.

| Clustering object | Clustering coefficient | | | integrated clustering coefficient | | | ω_i | Improved clustering results |
|--------------------------|------------------------|------------|------------|-----------------------------------|------------|------------|------------|-----------------------------|
| | good | preferably | difference | good | preferably | difference | | |
| Fully mechanized mining | 0.801 | 0.080 | 0 | 0.909 | 0.091 | 0 | 1.038 | good |
| Top grade general mining | 0.067 | 0.548 | 0.438 | 0.064 | 0.520 | 0.416 | 1.520 | poor |
| General mining | 0.093 | 0.374 | 0.650 | 0.083 | 0.335 | 0.582 | 1.565 | poor |
| Blasting mining | 0.219 | 0.440 | 0.613 | 0.17 | 0.343 | 0.480 | 1.487 | medium |

Table 3. Integrated clustering method for evaluation and analysis of different mining methods in coal mine.

| Clustering object | Clustering coefficient | | | integrated clustering coefficient | | | ω_i | Improved clustering results |
|--------------------------|------------------------|------------|------------|-----------------------------------|------------|------------|------------|-----------------------------|
| | good | preferably | difference | good | preferably | difference | | |
| Fully mechanized mining | 0.801 | 0.080 | 0 | 0.909 | 0.091 | 0 | 1.091 | good |
| Top grade general mining | 0.067 | 0.548 | 0.438 | 0.064 | 0.520 | 0.416 | 2.352 | poor |
| General mining | 0.093 | 0.374 | 0.650 | 0.083 | 0.335 | 0.582 | 2.499 | poor |
| Blasting mining | 0.219 | 0.440 | 0.613 | 0.17 | 0.343 | 0.480 | 2.297 | medium |

Table 4. The accuracy of the integrated clustering coefficient.

| Clustering object | The accuracy of the integrated clustering coefficient $\omega_i L_i$ | The accuracy of the improved integrated clustering coefficient $\omega_i L_i$ |
|--------------------------|--|---|
| Fully mechanized mining | 91.9% | 95.2% |
| Top grade general mining | 89.6% | 94.8% |
| General mining | 94.5% | 97.4% |
| Blasting mining | 90.3% | 93.0% |

grade general Mining: 94.8%, general mining: 97.4%, blasting mining: 93.0%. The accuracy of improved integrated clustering coefficient is higher than the original integrated clustering coefficients. That is to say, the error of the improved method is smaller, the accuracy has been greatly improved, and also shows that the object i belongs to the k grey class. The improved method is obviously better than the original integrated clustering method.

Evaluation of the coal mine has not changed in the two methods, however, integrated clustering coefficients of improved integrated clustering method numerical error is smaller, and the degree of membership is more bright and strong [21]. Obviously, the improved clustering results are more in line with the actual, more real and effective, and provides more objective, science, truth, effective evaluation method for coal mining enterprises to take what kinds of mining method to improve economic benefits.

4. Conclusion

This paper makes some improvement about the original integrated clustering coefficient, further analyzes and improves the value interval of integrated clustering coefficient. The original integrated clustering coefficient $1 \leq \omega_i \leq s$ is divided into s disjoint equal length interval, as shown in Equation (5). If the integrated clustering coefficient $\omega_i \in [1 + (k-1)(s-1)/s, 1 + k(s-1)/s)$ of the object i , then the object i belongs to k grey classification. The improved integrated clustering coefficient $\omega_i (1 \leq \omega_i \leq \sqrt{s})$ is divided into s disjoint equal length interval, as shown in Equation (7). If the integrated clustering coefficient of object i $\omega_i \in [1 + (k-1) \cdot (\sqrt{s}-1)/s, 1 + k \cdot (\sqrt{s}-1)/s)$ is k , then the object i belongs to k grey classification. The improved clustering result is more objective and reasonable, and the clustering coefficient is more accurate with small error, belonging to more distinct interval which conforms to objective facts. Finally, this paper takes four different coal-mining methods of a certain mine as examples, according to the influence of the different coal mining method of economic benefit of coal enterprises income, analyzes the four different kinds of mining methods category with improved integrated clustering method, shows the improved integrated clustering method more applicable and has better effectiveness.

Acknowledgements

The authors are grateful to anonymous referees for their helpful and constructive comments on this paper. This work is supported by the soft-science foundation of Henan Province (172400410015) and the Philosophy and Social Science Foundation of Henan Province (2016BJJ022).

References

- [1] Deng, J.L. (2002) Grey System Theory. Huazhong University of Science and Technology Press, Wuhan, 300-309.
- [2] Liu, S.F., Guo, T.B. and Dang, Y.G. (2011) Grey System Theory and Its Application. 5th Edition, Science Press, Beijing, 113-118.

- [3] Liu, S.F. (1993) Grey Clustering Analysis of Fixed Weights: A New Method of Grey System. Agriculture Press, Beijing.
- [4] Xiao, X.P. and Xiao, W. (1997) Grey Optimal Theory Model of Classification and Its Application. *Operations Research and Management Science*, **6**, 21-26.
- [5] Xu, X.L. and Luo, J. (2002) An Improved Grey Clustering Method. *Journal of Xiamen University*, **1**, 26-29.
- [6] Dang, Y.G., Liu, S.F. and Zhai, Z.J. (2004) Study on Grey Entropy Weight Clustering Decision Making. *Statistics and Decision*, **10**, 4-6.
- [7] Mi, C.M., Liu, S.F. and Dang, Y.G. (2006) Research on Grey Entropy Weight Clustering Decision Method. *Systems Engineering and Electronics*, **12**, 1823-1825.
- [8] Zhang, Q.S. (2002) The Measure of Grey Characteristics of Grey Clustering Result. *Chinese Journal of Management Science*, **2**, 54-56.
- [9] Zhang, Q.S. and Deng, J.L. (1995) On Grey Clustering in Grey Hazy Set. *The Journal of Grey System*, **4**, 377-390.
- [10] Luo, D. (2007) The Method for Grey Interval Clustering Decision. *Journal of Zhengzhou University: Nature Science Edition*, **1**, 119-124.
- [11] Wang, H.Z. (2014) The Redesign and Step Improvement of the Value Interval of Grey Integrated Cluster Evaluation. *Statistics and Decision*, **22**, 7-10.
- [12] Dang, Y.G., Liu, S.F. and Liu, B. (2005) Study on the Integrated Grey Clustering Method under the Clustering Coefficient No-Distinguished Difference. *China Journal of Management Science*, **4**, 69-73.
- [13] Dong, F.Y., Liu, J.J. and Liu, B. (2010) Improved Grey Integrated Clustering Method and Its Application in the Evaluation to Rural Economic Development of Henan Province. *System Sciences and Comprehensive Studies in Agriculture*, **4**, 479-483.
- [14] Li, Y.M., Luo, I.M. and Lu, F.H. (2007) An Improved Method of Grey Clustering Evaluation. *Statistics and Decision*, **1**, 20-21.
- [15] Zhang, W.Q. (1997) Discussion on the Application of Weight in Clustering Analysis. *Information and Statistics Forum*, **2**, 44-49.
- [16] Wang, H.Z., Qiang, F.J. and Zhu, F.Y. (2014) Reconstruction of Grey Clustering Decision Process and Grey Adjustment Coefficient. *Statistics and Decision*, **14**, 8-11.
- [17] Xiao, X.P. (1996) Improvement and Application of Grey Clustering Analysis. *Systems Engineering*, **2**, 71-74.
- [18] Pei, L.L., Chen, W.M. and Shen, C.G. (2012) Study on Optimization of Grey Clustering Evaluation Model. *Journal of Inner Mongolia Normal University (Natural Science Edition)*, **5**, 462-466+473.
- [19] Liu, J.P., Ji, C.S. and Li, H. (2001) Evaluation of Coal-Mining Methods with Weighted Grey Clustering Analysis. *Journal of China Coal Society*, **5**, 493-495.
- [20] Tang, Q.H. (1996) Discussion on the Method of Grey Class Whitening Function of Grey Clustering. *Sichuan Ordnance Journal*, **4**, 4-7.
- [21] Chen, X.H. and Liu, R. (2006) Improved Clustering Algorithm and Its Application in Complex Large Group Decision Making. *Systems Engineering and Electronics*, **11**, 1695-1699.

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact ojapps@scirp.org