

Regional Economic Evaluation Based on a Novel Fuzzy Clustering Method in Big data

Liu Zhaohui¹, Li Liyan² and Xu Wenjing³

¹*Enrollment and employment, Hebei Normal University of Science & Technology
Qinhuangdao, Hebei, P.R.China,066004*

²*Department of Economics, Hebei Qinhuangdao Institute of Technology
Qinhuangdao, Hebei 066100*

³*College of Finance, Hebei Normal University of Science & Technology
Qinhuangdao, Hebei, P.R.China, 066004*

¹ zjcliu@126.com, ² 583074965@qq.com, ³ liying7688@126.com

Abstract

Cloud computing, Internet of things, social networks and other new services make human society's data type and size increase at unprecedented rates. In order to analyze and make use of the huge data resources, the effective data analysis technique needed to be used. Thus, an optimized fuzzy clustering method GAPSO-FCM is proposed in this paper. The GA and PSO algorithm is adopted in order to overcome FCM be sensitive to initial value and noise problems and easy to fall into local minimum value. The combination of PSO and GA algorithm can partly solve the problem that is easy to appear premature phenomenon to PSO algorithm. It improves the global search ability by combining GAPSO algorithm and FCM algorithm. At last, the algorithm is used to analyze the comprehensive economic strength of cities in the Yangtze River Delta.

Key words: Regional economy, Evaluation, FCM, GA, PSO

1. Introduction

The rise of cloud computing, Internet of things technology, data is increasing and accumulating in the hitherto unknown speed. In scientific research, computer simulation, Internet applications, e-commerce and other fields, the data quantity presents the fast growth tendency. In order to discover knowledge from data, guide people's decision, data need to be analyzed deeply. People not only need the data to understand what is happening now, but also need to predict what will happen by using data, so as to make active preparations in action. For example, by predicting GDP in advance, people can make future policy and planning.

Regional economy is the production complex which is produced by the social economic activities and social economic relations of economic development in a certain area. It reflects the objective laws of economic development in different areas and the relation between the intension and extension. The spatial distributions of cities in Yangtze River Delta Area have become increasingly important. The development orientation of the central cities and sub-regional central cities is deeply concerned. From the comparison of the internal area of Yangtze River Delta, economic development of cities is imbalance since they are very different from each other. In order to have a more specific and clear understanding of the level of regional economic development in the Yangtze River Delta, the fuzzy clustering method is used to analyze the comprehensive economic strength of 16 cities. Thus it can be understood

clearly about the economic development between cities.

In many of the clustering algorithms, the theory of fuzzy C-means (FCM) algorithm is relatively mature. It is deduced based on hard c means (HCM) [1]. The introduction of parameter m by Bezdek [2] makes the FCM objective function extended to the infinite family so the FCM algorithm becomes a general algorithm in this field. Bezdek [2] also discussed the problems of convergence of FCM, and proved it converge to an extreme value. In recent years, the FCM algorithm has become the key areas to further research. Then there are lots of researches applied this method in many fields [3-6]. Richard *et al.*, [7] studied the effects of attribute weight of each data in FCM algorithm and achieved higher accuracy of clustering by changing the distribution of data. Fan [8] proposed a single point approximation weighted FCM algorithm. It is by means of probability statistics and sample attribute weights of the original data to adjust the data for the uniform distribution and it achieved good results. Li *et al.*, [9] proposed a new fuzzy clustering algorithm based on feature weight according to different effects on each dimension vectors in pattern classification. Xing and Hua [10] proposed an adaptive expert model based on FCM in order to solve the problem of uneven distribution of samples. Li and Yu [11] studied on a fuzzy clustering algorithm for noise sensitive and less robust data problem. The traditional FCM algorithm is mainly to solve the clustering problem between data points in space. It is only suitable for the convex type data, but not applicable to non-convex data. Schikuta and Erhart[12] pointed that the Bang clustering system had improved in processing the data structure, but when increasing the data dimension, the effect is not ideal. The merits of FCM algorithm on the effect of data clustering depend on the initial values of parameters, such as the number of cluster C , the initial cluster center, index weight m and so on. Chiu [13] partly solve the problem of clustering number C and the initial cluster center selection, proposed an adaptive decision method of clustering number C and the initial clustering center. In order to determine the optimal clustering number C , the scholars construct many different validity criterion functions based on the indicators of compactness and separation. At present, the effective criterion function is mainly divided into two categories: 1. only related with fuzzy membership grade, such as v_{PE} and v_{PC} from Bezdek [2]. 2. Related with fuzzy membership grade and the classification data set itself, such as v_{XB} from Xie [14], v_{WSJ} from Sun [15]. But it did not solve the problem that FCM algorithm is easy to fall into local minima and sensitive to the initial defects, poor stability and accuracy by the introduction of effective criterion function. Thus, some scholars adapted intelligent algorithm with FCM algorithm to solve the shortcomings of FCM. Zhu *et al.*, [16] used iterative Self organizing Data Analysis Techniques Algorithm (ISODATA) and genetic algorithm (GA) to optimize the FCM algorithm and the convergence of the algorithm is proved. Although it can automatically search the optimal cluster number, the class center vector is not optimized. Thus, the optimal results become instability. Tang *et al.*, [17] used Particle Swarm Optimization (PSO) into FCM to solve the problem of FCM which is much more sensitive to the initialization and easier to fall into local optimization. Chen et al [18] adapted simulated annealing (SA) to optimize the FCM algorithm and the problem of sensitivity with initial clustering center is solved by the method.

In this paper, an optimized fuzzy clustering method GAPSO-FCM is proposed in order to overcome FCM be sensitive to initial value and noise problems and easy to fall into local minimum value. The overall structure of the study is as follows: In Section 2, the standard FCM algorithm is introduced. In Section 3, we introduce the Genetic algorithm and PSO algorithm. The concrete steps of the GAPSO-FCM algorithm are presented. In Section 4, the hybrid optimization algorithm is used to analyze the comprehensive economic strength of cities in the Yangtze River Delta. At last, a comprehensive analysis of the result of the

experiment is made in Section 5.

2. Big Data and Data Mining

The rapid growth of information brought demands for information analysis tool more and higher. People want to be able to automatically obtain the knowledge of potential significance from large amounts of data. The Data analysis system is also changing. Fig 1 shows the analysis platform of Facebook.

Big data itself is a relatively abstract concept. Literally, it indicates that the data size is large. There is not a universally accepted definition for big data. Different definitions are from the characteristics of data and using these features to describe and summarize its definition. The representative view thinks big data should have 3 characteristics: volume, variety and velocity.

Data mining, as an aspect of knowledge discovery, is to discover the implicit, effective, useful, comprehensible pattern from a large number of unordered data. Then find the useful knowledge and provide decision support for users. Among them, the clustering analysis is one of the main methods of data mining. It can be used as an individual tool to find the data distribution of some further information. It can also be used as a preprocessing step analysis of other data mining algorithms. In big data background, how to use the rapid processing of data clustering algorithm and solve the practical problems will become a hot issue.

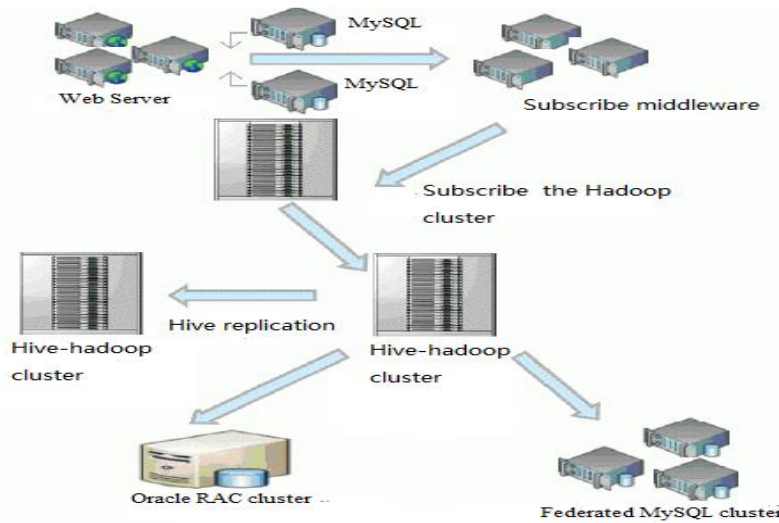


Figure 1. Architecture of the Analysis Platform in Facebook

3. Standard FCM Algorithm

Let $X = \{x_i, i = 1, 2, \dots, n, x_i \in R^s\}$ be the set of n metadata. FCM method is to divide X into c subsets and $V = \{v_i, i = 1, 2, \dots, c\}$ are c clustering centers. The minimized objective function which is defined as

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \tag{1}$$

Where u_{ij} is the membership value, m is the weighting parameter. d_{ij} is the distance between two points that is shown bellow

$$d_{ij} = \|x_j - v_i\| = \sqrt{\sum_{k=1}^s (x_{jk} - x_{ik})^2} \quad (2)$$

Where $i = 1, 2, \dots, c, j = 1, 2, \dots, n$.

And v_i is the centroid of the i th cluster that is

$$v_i = \{v_{i1}, v_{i2}, \dots, v_{im}\}, \quad i = 1, 2, \dots, c$$

The cluster centers are calculated by the following equation:

$$v_{ik} = \frac{\sum_{j=1}^n u_{ij}^m x_{jk}}{\sum_{j=1}^n u_{ij}^m} \quad (3)$$

Where $i = 1, 2, \dots, c, j = 1, 2, \dots, n$.

The membership functions is updated by the following equation

$$u_{ij}(s+1) = \left(\sum_{k=1}^c \left(\frac{d_{ik}}{d_{kj}} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (4)$$

In addition the sum of u_{ij} should meet the following condition

$$\sum_{i=1}^c u_{ij} = 1 \quad (1 \leq j \leq n) ; 0 \leq u_{ij} \leq 1, \forall i, j ; \quad (5)$$

Finally, the termination condition is determined by:

$$\|U^{(s+1)} - U^{(s)}\| < \varepsilon \quad (6)$$

where ε is error level for the termination of iteration which varies between 0 and 1. And the procedure of FCM algorithm is described in Figure 2.

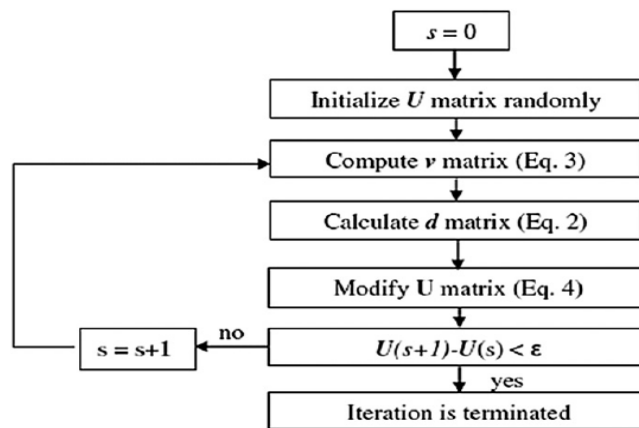


Figure 2. The Procedure of FCM Algorithm

4. Optimized FCM Algorithm

4.1. Principle of GA Algorithm and PSO Algorithm

The Genetic algorithm (GA) is widely used in many fields [19-22]. The optimum solution of Genetic algorithm (GA) is based on an iterative search from individuals. The algorithm first initializes population and randomly generates chromosome set in certain scale. Then calculate the fitness value of each individual. Do copy operation based on selection operator. Crossover operation is performed on the basis of crossover operator and mutation operation is according to the mutation operator. Determine whether the termination condition is satisfied. If it meets the condition, then output optimal solution. Otherwise repeat the above steps except initialization operation. The flow chart of GA algorithm is shown in Figure 3.

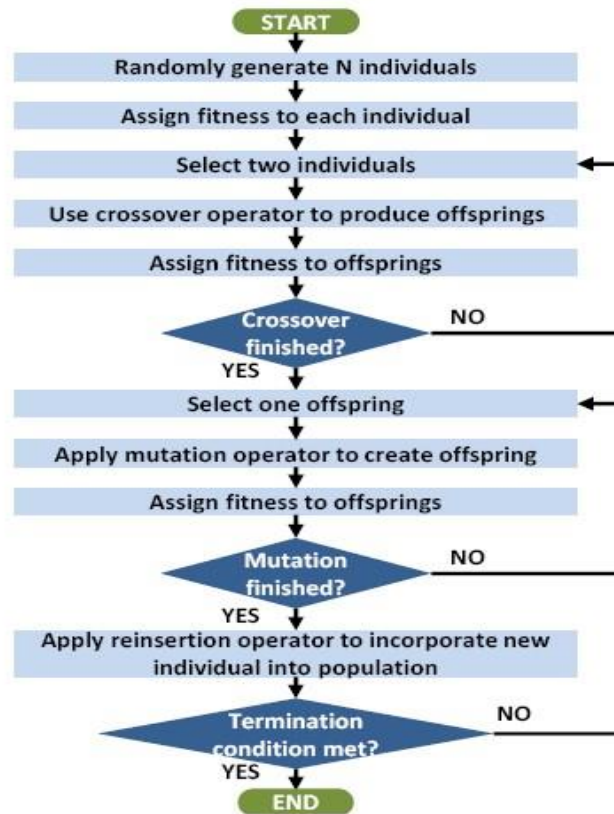


Figure 3. Flow Chart of GA Algorithm

In 1995, Eberhart and Kennedy [23] proposed the Particle Swarm Optimization (PSO) algorithm. The PSO also has been applied in a wide variety of applications [24-27]. The algorithm can be described as follows: A group of m particles fly at certain speed in search space. Each particle considers the history best point of itself and the history best point of the other particles within population in the search. Then the particle changes its position on this basis.

The position of l th particle is

$$x_l = (x_{l1}, x_{l2}, \dots, x_{ln})$$

The speed of l th particle is

$$v_i = (v_{i1}, v_{i2}, \dots, v_{is})$$

The history best point of lth particle is

$$p_i = (p_{i1}, p_{i2}, \dots, p_{is})$$

The history best point of the group is

$$p_g = (p_{g1}, p_{g2}, \dots, p_{gs})$$

The updating formula of particle about speed and position is

$$v_{ij}^{(t+1)} = w v_{ij}^{(t)} + c_1 \delta_1 (p_{ij}^{(t)} - v_{ij}^{(t)}) + c_2 \delta_2 (p_{gj}^{(t)} - v_{ij}^{(t)}) \quad (7)$$

$$x_{ij}^{(t+1)} = x_{ij}^{(t)} + v_{ij}^{(t+1)} \quad (8)$$

Where w is inertia weight. c_1 and c_2 are learning factors. δ_1 and δ_2 are random numbers subjects to the uniform distribution in the interval of $[0, 1]$.

Generally, the particle should have good exploration ability at the beginning of the flight and have good development ability later. Thus, a time varying weight can be achieved to the target. Let

$$w_i = w_{max} - \frac{w_{max} - w_{min}}{Iter_{max}} * i^n \quad (9)$$

Where $w_i \in [w_{min}, w_{max}]$, $Iter_{max}$ is the maximum number of iterations.

In order to let the particle has better self-learning ability and smaller social learning ability at the beginning and has smaller self-learning ability and better social learning ability later, the learning factors can be described like

$$c_1 = (c_{1b} - c_{1a}) \frac{Iter}{Iter_{max}} + c_{1a} \quad (10)$$

$$c_2 = (c_{2b} - c_{2a}) \frac{Iter}{Iter_{max}} + c_{2a} \quad (11)$$

Where $c_{1a}, c_{1b}, c_{2a}, c_{2b}$ are initial value and final value of c_1, c_2 , $Iter$ is current number of iterations.

4.2. Process of GAPSO-FCM

The GA algorithm is used to find the optimal number of clusters. And use the validity criterion function as its fitness function. The PSO algorithm is adopted to optimize the central vector of clusters. Here the valid criterion function V_{WSJ} [15] is

$$V_{WSJ}(U, V, c) = Scat(c) + \frac{Sep(c)}{Sep(c_{max})} \quad (12)$$

Where $Scat(c) = \frac{\frac{1}{c} \sum_{i=1}^c \|\sigma(V_i)\|}{\|\sigma(X)\|} \quad (13)$

$$Sep(c) = \frac{(\max_{i \neq j} \|V_i - V_j\|)^2}{(\min_{i \neq j} \|V_i - V_j\|)^2} \sum_{i=1}^c (\sum_{j=1}^c \|V_i - V_j\|^2)^{-1} \quad (14)$$

$$\sigma(X) = \{\sigma(X)^1, \sigma(X)^2, \dots, \sigma(X)^s\}^T \quad (15)$$

$$\sigma(X)^j = \frac{1}{n} \sum_{j=1}^n x_j^p - \bar{x}^p \quad (16)$$

$$\sigma(V_i)^p = \frac{1}{n} \sum_{j=1}^n u_{ij} (x_j^p - V_i^p)^2 \quad (17)$$

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j \quad (18)$$

The flow chart of the GAPSO-FCM algorithm is shown in Figure 3. The algorithm process is as follows:

Step 1: Randomly generate the binary population $G_k = \{G_1^k, G_2^k, \dots, G_n^k\}$ which length is $l+2$ and size is n . From the beginning of $k=0$, the top l digits is for binary code, the $(l+1)$ th digit is decimal code corresponding to the binary code. The $(l+2)$ th digit is fitness function value of the individual. For example, assume the binary code of G_i^k is (001010), then the chromosome length is 8. The coding structure is shown in Table 1.

Table 1. Coding Structure of Chromosome

0	0	1	0	1	0	11	$f(v)$
---	---	---	---	---	---	----	--------

Step 2: Calculate $c_i^k (i = 1, 2, \dots, n)$ of the individual $G_i^k (i = 1, 2, \dots, n)$. That is turning the binary code to decimal code.

Step 3: Code the clustering center of v_i^k . Randomly generate the population which has m particles. $x = (x_1, x_2, \dots, x_m)^T$. The position of l th particle is $x_l = (x_{l1}, x_{l2}, \dots, x_{lc_l})$, $l = 1, 2, \dots, m$. The component of x_l represents the clustering center. Initialize the position x_l and speed v_l of particles. The position of each particle is a $c_i^k \times n$ matrix, where c_i^k is the numbers of clustering center. n is the number of samples.

Step 4: Calculate the membership matrix according to Eq.(4)

$$U^k(0) = (u_{ij}^k(0))_{c_i^k \times n}$$

Step 5: If

$$\max_{ij} |u_{ij}^k(t+1) - u_{ij}^k(t)| < \varepsilon$$

Then go to step 8, otherwise go to step 6.

Step 6: Set the fitness function of PSO algorithm be

$$f(x_i) = \frac{1}{J_m(U, V) + 1} \quad (19)$$

Then calculate the fitness value of each particle. For each particle, compare its fitness value with the individual best position fitness value $f(p_i)$. If $f(x_i) > f(p_i)$, use the current location to update individual historical best position. In addition, compare the individual best

position fitness value $f(p_i)$ with the group best position fitness value $f(p_g)$. If

$f(p_i) > f(p_g)$, use the individual historical best position to update the group best position.

Step 7: Update the position and speed of particles, then go to Step 4.

Step 8: Calculate the fitness function $G(U^*, V^*, c_i) = f(v)$, $f(v)$ is the function about the valid criterion function

Step 9: Let $G(U^*, V^*, c_i) = \max(\min)G(U^*, V^*, c_i)$, c_i^* is the optimal individual of current population. Keep the optimal individual of current population, and then do the selection operation, crossover operation and mutation operation. A new population $G_{k+1} = \{G_1^{k+1}, G_2^{k+1}, \dots, G_n^{k+1}\}$ will be get from above operation. Then turn back to Step2. When the algorithm reaches the maximum number of iterations, the algorithm terminates.

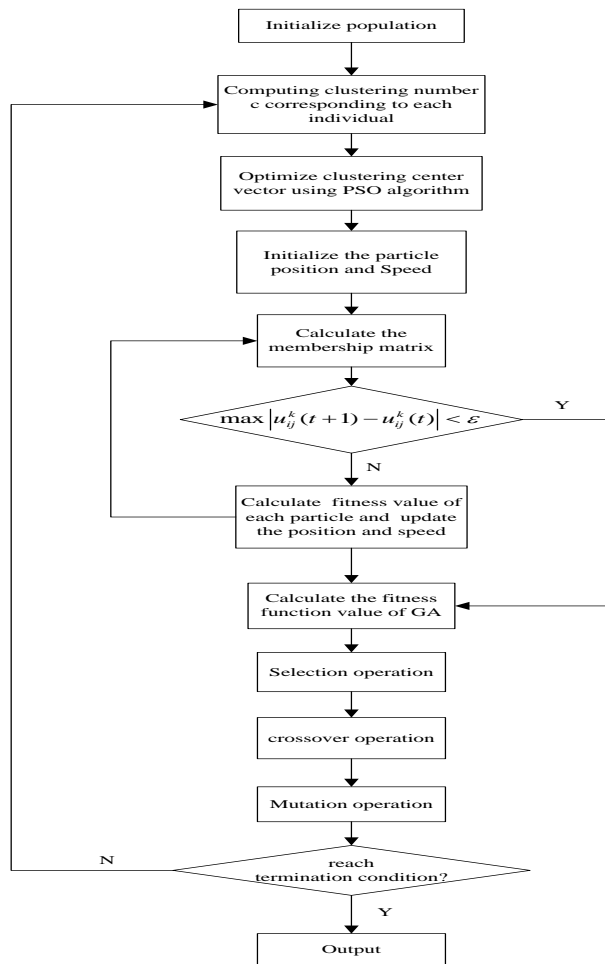


Figure 4. Flow Chart of the GAPSO-FCM Algorithm

5. Model Construct and Experiment

The GAPSO-FCM algorithm is adopted to estimate the economic development level of regional economy in Yangtze River Delta Area. 16 cities are considered in Yangtze River

Delta Area of china in this study. There are Shanghai, Nanjing, Hangzhou, Ningbo, Suzhou, Wuxi, Zhenjiang, Changzhou, Nantong, Yangzhou, Taizhou (Jiangsu), Shaoxin, Huzhou, Jiaxing, Zhoushan and Taizhou (Zhejiang).

In order to eliminate the influence of the unit, the data standardization method should be used. The method is shown in Eq. (20)

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (20)$$

In addition, a scientific and complete index system needs to be established in order to better evaluate the city development ability. The evaluation index system of economic development level is given in Table 2.

Table 2. Evaluation Index System

Variable	First level index	Second level index	unit
X1	City development level factor	Total retail sales of social consumer goods	Million yuan
X2		Total wages of workers in cities and towns	Million yuan
X3	industrialization level factor	Gross industrial production per capita	Yuan
X4		Total profits of industrial enterprises	Million yuan
X5		Main business income of industrial enterprises	Million yuan
X6		Average balance net value of fixed assets	Million yuan
X7	Regional industrial structure factor	GDP contribution rate of second industry	%
X8		GDP contribution rate of third industry	%
X9		Proportion of non-agricultural industries	%
X10	Potential economic growth factor	Regional revenue accounted for the proportion of GDP	%
X11		Balance of deposits in financial institutions at the end of the year	Million yuan
X12	Total economy level factor	GDP	Million yuan
X13		Total fiscal revenue	Million yuan

The evaluation index system includes five first index such as city development level factor, industrialization level factor, regional industrial structure factor, potential economic growth factor and total economy level factor. It also has thirteen second level index, such as total retail sales of social consumer goods, main business income of industrial enterprises, GDP and so on.

Then we construct the GAPSO-FCM model through these indexes. The result of the algorithm is shown in Table 3 and Table 4. The Figure 5 shows the fitness curve of GAPSO-FCM algorithm.

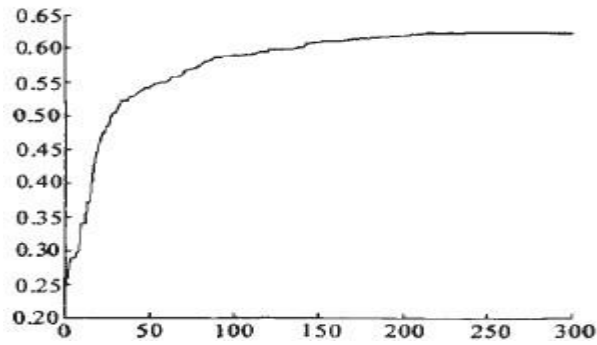


Figure 5. Fitness Curve of GAPSO-FCM Algorithm

Table 3. The Result of GAPSO-FCM Algorithm

Cases	City	category
1	Shanghai	1
2	Nanjing	3
3	Hangzhou	3
4	Ningbo	3
5	Suzhou	2
6	Wuxi	3
7	Zhenjiang	4
8	Changzhou	4
9	Nantong	4
10	Yangzhou	4
11	Taizhou(Jiangsu)	4
12	Shaoxin	4
13	Huzhou	4
14	Jiaxing	4
15	Zhoushan	4
16	Taizhou(Zhejiang)	4

Table 4. City Classification

category	cities
1	Shanghai
2	Suzhou
3	Hangzhou, Nanjing, Wuxi, Ningbo Zhenjiang, Changzhou, Nantong, Yangzhou,
4	Taizhou(Jiangsu), Shaoxin, Huzhou, Jiaxing, Zhoushan and Taizhou(Zhejiang)

6. Result Analysis and Conclusion

According to the results of GAPSO-FCM algorithm, the Yangtze River Delta Area is divided into four regions through comprehensive analysis and comparison. Then the law of economic development of 16 cities in Yangtze River Delta Area is reflected by analysis the different cities from different categories.

The first category, Shanghai, is the fastest growing economic region in the Yangtze River Delta Area. Shanghai is the most comprehensive strength of the city. The GDP etc. which reflect the city's economic development level is much higher than the other cities. And the social development, economic development and the level of opening up are comprehensive optimal. Undoubtedly Shanghai is the most suited to the region's economic center. Overall, Shanghai should maintain and strengthen the core status of the region and the vitality of economic development. At the same time, it needs to pay attention to prevent the emergence of hollow phenomenon which is often appearing in the western cities and maintain the prosperity of city.

The second category, Suzhou, the overall strength of the Suzhou is after Shanghai. Suzhou is the second circle of development in Yangtze River Delta. The economic development of Suzhou is relatively slow according to Shanghai. But it is also in the upstream level. Suzhou as the fulcrum of Shanghai and Wuxi will have strong development and play an irreplaceable role in the Yangtze River Delta Area. Suzhou should maintain the advantages in the future development, and speed up the urban sustainable development, improve the development level of city.

The third category includes Hangzhou, Nanjing, Wuxi, Ningbo. The economic development of these four cities is quite equal. They all have their distinctive aspects and are the second circle of development in Yangtze River Delta. These cities should work together to promote the overall development of the region. Hangzhou is the pivot of Shanghai and Zhejiang province. It is also the capital city of Zhejiang province. The comprehensive economic strength of Wuxi and Ningbo is equal to Hangzhou. They all have a strong economic strength and play an irreplaceable role. Nanjing is the pivot of Shanghai and Suzhou, Subei. It also connects with the vast hinterland, such as Anhui province. Nanjing is the capital city and has excellent development environment. It plays a leading role in the regional economy. These four cities should continue to strengthen the regional communication function and economic function to lead the region developing fast.

The fourth category includes Zhenjiang, Changzhou, Nantong, Yangzhou, Taizhou(Jiangsu), Shaoxin, Huzhou, Jiaxing, Zhoushan and Taizhou(Zhejiang). The scale of the city in this category is small. They should continue to strengthen communication with other cities and accelerate their economic development in the macro environment. At the same time, these cities should pay more attention to their respective characteristics. In addition, district and county economy is one of the key factors to the local economy. The radiation effect of central cities was not obvious in the cities of this category. So it can be considered to establish some core cities to drive the development of the new part of the city so as to better realize regional integration.

References

- [1] R. O. Duda and P. E. Hart, "Pattern classification and scene analysis", New York: John Wiley & Sons: 230-235, (1973).
- [2] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms", New York: Plenum Press, (1981).
- [3] K. L. Wu and M. S. Yang, "Alternative c-means clustering algorithms", Pattern Recognition, vol. 35, (2002), pp. 2267-2278.

- [4] G. C. Karmakar and L. S. Dooley, "A generic fuzzy rule based image segmentation algorithm", *Pattern Recognition Letters*, vol. 23, no. 10, (2002), pp. 1215–1227.
- [5] S. R. Kannan, S. Ramathilagam, R. Devi and A. Sathya, "Robust kernel FCM in segmentation of breast medical images", *Expert Systems with Applications*, vol. 38, no. 4, (2011), pp. 4382–4389.
- [6] M. Mehrabi, M. Sharifpur and J. P. Meyer, "Application of the FCM-based neuro-fuzzy inference system and genetic algorithm-polynomial neural network approaches to modelling the thermal conductivity of alumina–water nanofluids", *International Communications in Heat and Mass Transfer*, vol. 39, no. 7, (2012), pp. 971–977.
- [7] R. Nock and F. Nielsen, "On Weighting Clustering", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, (2006), pp. 1223–1235.
- [8] F. Jianchao, H. Min and W. Jun, "Single Point Iterative Weighted Fuzzy C-Means Clustering Algorithm for Remote Sensing Image Segmentation", *Pattern Recognition*, vol. 42, (2009), pp. 2527–2540.
- [9] L. Jie, G. Xinbo and J. Licheng, "A new feature weighted fuzzy clustering algorithm", *Acta Electronica Sinica*, vol. 34, no. 1, (2006), pp. 89–92.
- [10] X. HongJie and H. Baogang, "An adaptive fuzzy C-mean clustering-based mixture of experts model for unlabeled data classification", *Nero Computing*, vol. 71, (2008), pp. 1008–1021.
- [11] L. Cuixia and Y. Jian, "Study on the classification of a kind of fuzzy clustering algorithm", *Journal of Beijing Jiao Tong University*, vol. 29, no. 2, (2005), pp. 17–21.
- [12] E. Schikuta and M. Erhart, "The BANG-clustering system: Grid-based data analysis", *Lecture Notes in Computer Science*, vol. 1280, (1997), pp. 513–524.
- [13] S. Chiu, "A cluster estimation method with extension to fuzzy model identification", *Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence*, vol. 2, (1994), pp. 1240–1245.
- [14] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 8, (1991), pp. 841–847.
- [15] H. J. Sun, S. G. Wang and Q. S. Jiang, "FCM-based model selection algorithms for determining the number of clusters", *Pattern Recognition*, vol. 37, no. 7, (2004), pp. 2027–2037.
- [16] K. J. Zhu, S. H. Su and J. L. Li, "Optimal Number of Clusters and the Best Partition in Fuzzy C-mean", *Systems Engineering-Theory & Practice*, vol. 25, no. 3, (2005), pp. 52–61.
- [17] X. L. Tang, L. Zhuang, Y. G. Li and C. X. Cao, "Intrusion Detection Based on Particle Swarm Optimization and Fuzzy c-means Clustering", *Computer Engineering*, vol. 4, no. 4, (2008), pp. 13–15.
- [18] Y. J. Chen, X. Y. Wang, D. Z. Lin and Y. M. Liu, "Application of Fuzzy C-Means Clustering Based on Simulated Annealing to Stock Market", *Journal of East China University of Science And Technology(Natural Science Edition)*, vol. 31, no. 4, (2005), pp. 538–541.
- [19] H. Malekzhehtab and A. A. Golafshani, "Damage Detection in an Offshore Jacket Platform using Genetic Algorithm based Finite Element Model Updating with Noisy Modal Data", *Procedia Engineering*, vol. 54, (2013), pp. 480–490.
- [20] G. Kanagaraj, S. G. Ponnambalam and N. Jawahar, "A hybrid cuckoo search and genetic algorithm for reliability–redundancy allocation problems", *Computers & Industrial Engineering*, vol. 66, no. 4, (2013), pp. 1115–1124.
- [21] B. Stojanovic, M. Milivojevic, M. Ivanovic, N. Milivojevic and D. Divac, "Adaptive system for dam behavior modeling based on linear regression and genetic algorithms", *Advances in Engineering Software*, vol. 65, (2013), pp. 182–190.
- [22] N. Maehara and Y. Shimoda, "Application of the genetic algorithm and downhill simplex methods (Nelder–Mead methods) in the search for the optimum chiller configuration", *Applied Thermal Engineering*, vol. 61, no. 2–3, (2013), pp. 433–442.
- [23] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory", *Micro Machine and Human Science, Proceedings of the Sixth International Symposium*, (1995), pp. 39–43.
- [24] M. A. Mohandes, "Modeling global solar radiation using Particle Swarm Optimization (PSO)", *Solar Energy*, vol. 86, no. 11, (2012), pp. 3137–3145.
- [25] M. I. Menhas, L. Wang, M. R. Fei and H. Pan, "Comparative performance analysis of various binary coded PSO algorithms in multivariable PID controller design", *Expert Systems with Applications*, vol. 39, no. 4, (2012), pp. 4390–4401.
- [26] S. A. Taher, A. Karimian and M. Hasani, "A new method for optimal location and sizing of capacitors in distorted distribution networks using PSO algorithm", *Simulation Modelling Practice and Theory*, vol. 19, no. 2, (2011), pp. 662–672.
- [27] I. Ziari and A. Jalilian, "Optimal placement and sizing of multiple APLCs using a modified discrete PSO", *International Journal of Electrical Power & Energy Systems*, vol. 43, no. 1, (2012), pp. 630–639.