Applied Mathematics & Information Sciences
An International Journal

Research of Individual Neural Network Generation and Ensemble Algorithm Based on Quotient Space Granularity Clustering

Li Hui[1], *Ding Shifei*[1,2]

[1] School of Computer Science and Technology, China University of Mining and Technology, Xuzhou Jiangsu 221116 P. R. China
[2] Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190 P. R. China

Received: 1 Nov. 2012; Revised 2 Dec. 2012; Accepted 20 Dec. 2012
Published online: 1 Mar. 2013

Abstract: The aim of this paper is to develop an individual neural network generation and ensemble algorithm based on quotient space granularity clustering. Firstly, we give the characteristics of the quotient space granularity and affinity propagation(AP) clustering. Secondly, we introduce the quotient space concept to the AP clustering analysis, which can find an optimal granularity from all possible granularities. Then using improved AP clustering algorithm to seek optimal results of sample clustering and using different individual neural network to learn different categories of samples so that the degree of difference between networks and the generalization ability of neural network ensemble(NNE) can be improved. Further, according to the degree of correlation between the input data and the sample category to adaptively adjust ensemble weights. The algorithm proposed here is not only a method of generating the individual neural networks, but also can adaptively adjust ensemble weights of individual neural network. Experiments show that our proposed method is validity and correctness.

Keywords: Neural Network Ensemble(NNE), Affinity Propagation(AP) clustering, Quotient Space theory, Granularity Analysis

1. Introduction

Hansen and Salamon proved in 1990: by training multiple neural networks and integrating the results can significantly improve the generalization ability of the neural network system[1]. Because of excellent performance and accuracy significantly improved in the aspect of the generalization ability, neural network ensemble(NNE), proposed since 1990, has been extensively studied by many scholars. The studies of NNE are mainly focused on three aspects: how to generate the individual neural network, and how to integrate the individual neural network, and how to determine the structure of the NNE. Zhou[2] proposed a selective NNE based on genetic algorithm, GASEN, which used genetic algorithm to choose neural network with the large degree of difference, rather than integrated all the individual neural networks. The experiments show that the generalization capability and accuracy of GASEN are higher than ordinary algorithms, such as Boosting and Bagging. Li Kai, et al.[3] and Li Guozheng et al.[4] proposed a selective NNE based on clustering techniques, which clustered network by calculating the degree of difference between the individual neural networks, and then excluded related neural networks to reduce the scale of NNE. Experiments show that this algorithm is feasible. Md. et al.[5] proposed a new ensemble algorithm CNNE, which established a smaller NNE firstly, and then constructed the optimal structure of the network by increasing the hidden layer neurons and the number of individual neural network. Wang[6] improved the algorithm proposed by Zhou[2] by adding a bias, in some extent, which improved the accuracy of the NNE. Pitoyo[7] proposed an adaptive NNE algorithms which learned from incomplete samples, this algorithm improved the accuracy of the neural network by eliminating incorrect samples. Yao[8] proposed a NNE algorithm by using evolutionary computing method to adjust the weights.

* Corresponding author: e-mail: lihui400@163.com

Three aspects of NNE are interdependent, closely linked. Any unreasonable aspect will lead to the performance of NNE decline. Unfortunately, these three aspects are often studied alone, such as, GASEN studied structure of NNE, but not generation algorithm of individual neural network. Bagging and Boosting algorithm studied the individual neural network generation algorithm, but not the structure and the ensemble of NNE.

Based on the above analysis, comprehensively considering the three aspect of NNE, we develop an individual neural network generation and ensemble algorithm based on quotient space granularity clustering by introducing the quotient space concept to the AP clustering analysis, which can find an optimal granularity from all possible granularities to make the clustering criteria of similarity in class and different between classes maximization. Using different individual neural network to learn different categories of samples so that the degree of difference between the networks can be improved. Making the number of individual neural networks equal to the number of sample categories to determine the structure of the NNE. The learning samples of individual neural network and cognitive ability of the different sides of an issue is different, so for the different input data the processing power of the individual neural networks is different. The fixed weights of NNE which may reduce the performance of some of the individual neural network lead to the decline in the performance of the overall NNE. The integer weights can can effectively improve the performance of the neural network[9]. According to the degree of correlation between the input data and the sample category to adaptively adjust ensemble weights can improve the ensemble accuracy of NNE. The algorithm proposed here is not only a method of generating the individual neural networks, but also can adaptively adjust ensemble weights of individual neural network. The following chapters are organized as follows. Chapter 2 introduces the related knowledge, such as Affinity Propagation(AP) clustering, the quotient space granularity and granularity of combination and decomposition. Chapter 3 proposes the improved AP clustering algorithm based on the quotient space granularity analysis by introducing the quotient space concept to the AP clustering analysis. On this basis, we give the architecture, algorithm description and analysis of NNE. In chapter 4, through two sets of experiments, we verify effectiveness of the improved AP clustering algorithm based on the quotient space granularity analysis and the performance of individual neural network generation and ensemble algorithm based on quotient space granularity clustering our proposed. Finally, chapter 5 elaborates the overall conclusions and directions for further research.

2. Related knowledge

2.1. Affinity Propagation(AP) clustering

Affinity Propagation(AP) clustering, as a new clustering technique based on message passing, has been proposed by Frey and Dueck[10,11]. AP clustering simultaneously considers all data points as potential exemplars, and recursively transmits real-valued messages along the edges of the network until a set of good centers is generated[12]. The AP clustering algorithm has been shown very useful for many applications in facility location[13-15], image recognition[16], image segmentation[17,18], text mining[19,20], biomedicine[10,21], key frame extraction of video[22], image Retrieval[23], etc. Unlike previous methods, AP algorithm[24] which not requires data sets with a particular structure is mainly based on the similarity between the N sample points to cluster. These similarities compose of N $\times$ N similarity matrix S, such as: S (i, j) denotes the similarity between sample point i and j. The main diagonal values of matrix S is also known as Preference(P), which is the criteria whether corresponding sample points can become the cluster centers. In general, the higher the value, the greater the possibility of becoming the clustering center. AP algorithm mainly depends on a kind of "information transfer" mechanism to achieve the clustering of the data sets. This message passing mechanism consists mainly of two types of information: Responsibility and Availability. Responsibility denotes the information spread from the sample points to the candidate class center which shows the suitability of the candidate class center corresponding to the sample point as the potential center, and the higher the value, the more likely of the candidate class center becoming the actual class center. Availability denotes the information spread from the candidate class center to the sample points which shows the extent of polymerization of the candidate center to sample points, and the higher the value, the more likely of the sample points belonging to a certain category.

2.2. The quotient space granularity

Academician Zhang Bo and Professor Zhang Ling independently proposed the quotient space theory in the study of problem solving, which uses subset to describe the concept. The concept of different granularity reflects the subset of the different granularity. A cluster of concepts constitute a partition of space: Quotient Space (knowledge base), and different concept clusters constitute the quotient space. As for Granular computing(GrC), it is equivalent to study the relationship and conversion between the subsets of a given knowledge base. This model describes the problem with a triple (X, F, T), where X is the domain; F is the set of attributes; T is a topology on X. In this model, granularity of the domain is equivalent to a given equivalence relation R or a partition, so the quotient set corresponding to R is denoted by [x], and the triple ([X], [F]

[T]) can be called the quotient space corresponding to R. The quotient space theory is to study the relationship, synthesis, decomposition, and reasoning between each quotient space. By the different equivalence relation, we can construct different granular world, which can simplify the problem to make the problem easy to be simplified and analyzed. To introduce the concept of the quotient space into clustering analysis is to find an optimal granularity in all possible granularities.

2.3. Granularity of combination and decomposition

The aim of introduction of granularity analysis theory is to cluster effectively. But if the granularity is too coarse, details of the problem may be hidden, on the contrary, if the granularity is too fine, we can not excavate sample knowledge because each sample is a class in these circumstances. In order to seek a suitable granularity, we can take advantage of the combination and the decomposition method to adjust the granularity.

Definition 1.*assuming R_1 and R_2 are two equivalence relations on the universe X, and R is also a equivalence relation on the universe X, which meets $R_1 < R$ and $R_2 < R$. If no existing R' which meets $R_1 < R'$, $R_2 < R'$, $R < R'$, we call R is the product of R_1 and R_2, denoted as $R = R_1 \otimes R_2$.*

Definition 2.*assuming R_1 and R_2 are two equivalence relations on the universe X, and R is also a equivalence relation on the universe X, which meets $R_1 > R$ and $R_2 > R$. If no existing R' which meets $R_1 > R'$, $R_2 > R'$, $R > R'$, we call R is the sum of R_1 and R_2, denoted as $R = R_1 \oplus R_2$.*

From the above definition, $R_1 \otimes R_2$ is the most coarse upper bound of division R_1 and R_2, $R_1 \oplus R_2$ is the finest lower bound of division R_1 and R_2. When making a granularity analysis for a specific clustering problem, we can gain the quotient space S by equivalence relations division and gain the preliminary conclusion A_1. According to the judgment, if the granularity is partial coarse, we can take a granularity R_1' (A_1 as guideline) and let $R_2 = R_1 \otimes R_1$', and then gain the conclusion A_2 by making the analysis on the R_2 again. The above procedure may be repeated with each repetition, the granularity will be refined gradually. Similarly, the granularity will get coarse gradually. By combining the method of combination and decomposition, we can find a suitable granularity ultimately.

3. The NNE algorithm based on quotient space granularity clustering

3.1. Sample clustering based on improved AP clustering algorithm

In the iterative process, AP algorithm continuously updates the values of Responsibility and Availability of each point until a number of class centers is automatically generated, while the rest of the data points are assigned to the corresponding class group. The steps of AP are described as follows:

Step 1: Initialize the algorithm. Calculate the initial similarity matrix S, and initialize the initial value of the Preference (P).

Step 2: Calculate Responsibility between sample points.

$R(i,k) \leftarrow s(i,k) - max_{j \neq k}(s(i,j) + A(i,j))$,
A(i, j) denotes Availability of j to i.

Step 3: Calculate Availability between sample points.

$$A(i,k) \leftarrow min\{0, R(k,k) + \sum_{j \neq i,k} max(0, R(j,k))\}$$

$$A(k,k) \leftarrow \sum_{j \neq k} max(0, R(j,k))$$

Step 4: Update Responsibility and Availability.

$$R_{i+1}(i,k) = \lambda \cdot R_i(i,k) + (1-\lambda) \cdot R_{i+1}^{old}(i,k)$$

$$A_{i+1}(i,k) = \lambda \cdot A_i(i,k) + (1-\lambda) \cdot A_{i+1}^{old}(i,k), \lambda \in [0.5, 1)$$

$A_{i+1}(k,k) = P(k) - max[A_{i+1}(k,j) + S_{i+1}(k,j)]$,
$j \in \{1, 2, \ldots, N\}, j \neq k$

Where, λ is the convergence coefficient which is used to adjust convergence speed and the stability during the iterative process.

Step 5: Terminate the calculation when the number of iterations exceeds the pre-set maximum or when the number of cluster centers does not change in the iterations, and determine the class centers and the corresponding sample points; Otherwise, return to Step 2.

Step 6: Record the centers and their number according to the results of AP clustering

Clerc Maurice's research shows that using the constriction factor can effectively guarantee that the algorithm converges[25]. The constriction factor can be expressed as:

$$\rho = \frac{2}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|}, \varphi > 4 \qquad (1)$$

In order to accelerate the algorithm fast convergence, updating the formula of Responsibility and Availability for the following changes:

$$R_{i+1}(i,k) = \rho \cdot \lambda \cdot R_i(i,k) + (1-\lambda) \cdot R_{i+1}^{old}(i,k) \qquad (2)$$

$$A_{i+1}(i,k) = \rho \cdot \lambda \cdot A_i(i,k) + (1-\lambda) \cdot A_{i+1}^{old}(i,k) \quad (3)$$

where, $\lambda \in [0.5,1)$, ρ is the constriction factor which can effectively guarantee the algorithm convergence.

3.2. Description of the improved AP clustering algorithm based on the quotient space granularity analysis

Applying quotient space granularity principle to the improved AP clustering, we can design the corresponding clustering algorithm, the algorithm is described as follows:

Step 1: Use the improved AP algorithm to cluster the samples, and calculate the centers and the sample number.
Step 2: Analyze the clustering results by the method of granularity analysis. According to the expertise and the actual situation, if the granularity is partial coarse, return **Step 1** to refine this granularity, otherwise, return **Step 4**;

Step 3: Calculate the similarity of clustering results between any two clusters;
Step 4: Gain the largest similarity sim_{max} by comparing the all similarity values;
Step 5: Give the similarity threshold s. If $sim_{max} <$ s, the clustering is end. Otherwise, combining the largest similarity of the two clusters until the clustering is end;
Step 6: Continue to use the granularity analysis method to analyze the clustering results and determine the appropriate granularity based on the expertise and the actual situation, and get the final clustering results. At this time we mean the similarity threshold s. Then exit the algorithm.

3.3. The description and analysis of NNE algorithm

Based on the above analysis, we introduce the quotient space concept to the AP clustering analysis, which can find an optimal granularity from all possible granularities. Then using improved AP clustering algorithm to seek optimal sample clustering results and using different individual neural network to learn different categories of samples so that the degree of difference between the networks can be improved. Further, according to the degree of correlation between the input data and the sample category to adaptively adjust ensemble weights. The algorithm proposed here is not only a method of generating the individual neural networks, but also can adaptively adjust ensemble weights of individual neural network. Figure1 shows the architecture of NNE our proposed. The NNE algorithm can be described as follows:

Step 1: Use improved AP clustering algorithm based on the quotient space granularity analysis to cluster the samples and record the centers;
Step 2: Use different individual neural network to learn different categories of samples and generate the individual neural network, and randomly initialize ensemble weights;
Step 3: Make Correlation Analysis(CA) on the input data to determine the ensemble weights of individual neural network;
Step 4: Integrate and output results.

Suppose that the sample is divided into m classes, $x_1, x_2, \cdots, x_m$, its characteristics of each center is $c_1, c_2, \cdots, c_m$. Input data $d = \{d_1, d_2, \cdots, d_m\}$. Ensemble weight and the distance between d and each class are described as follows:

$$diff(d,c_j) = \sum_{i=1}^{k} |d_i - c_{ji}|, 0 < j < m \quad (4)$$

$$w_j = \frac{diff(d,c_j)}{\sum diff(d,c_j)}, where, w_j > 0, \sum_{j=1}^{m} w_j = 1 \quad (5)$$

Suppose the output of network j is V^j, the ensemble result is described as follows:

$$\overline{V} = \sum_{j=1}^{m} w_j V^j \quad (6)$$

Suppose that using N individual neural network ensemble to learn sample X, the weights are w_j, output of network j to input X is $V_j(X)$. Output of NNE is denoted as follows(A^j is the degree of difference of network j, $\overline{A}$ is the degree of difference of ENN):

$$A^j = (V^j(X) - \overline{A}(X))^2 \quad (7)$$

$$\overline{A} = \sum_{i=1}^{m} w_i A^i \quad (8)$$

$$\frac{\partial A^j}{\partial V^j} = 2(V^j - \overline{V}(X)(1-w_j)) \quad (9)$$

$$\frac{\partial \overline{A}}{\partial V^j} = \frac{\partial \overline{A}}{\partial A^j}\frac{\partial A^j}{\partial V^j} = 2(V^j - \overline{V}(X)(1-w_j))(1-w_j)w_j \quad (10)$$

Above formula shows that the degree of difference of network is directly related to its weight and the discrepancy between output and ensemble output. For NNE with fixed weights, the degree of difference is only related with output. For adaptive weights, the degree of difference of the different individual network is related with training set and the input data of the network.

Table 1 Parameters of datasets

Dataset	Number of classes	Number of testing sets	Number of training sets	Number of attributions
Iris	3	75	75	4
Wine	3	89	89	13

Table 2 The results of AP clustering for there different values of preference(P)

Dataset	Number of classes	Numbers of clustering centers		
		P=median(S)/2	P=median(S)	P=2*median(S)
Iris	3	10	7	5
Wine	3	15	9	7
Dataset	Number of classes	Time(s)		
		P=median(S)/2	P=median(S)	P=2*median(S)
Iris	3	4.797	5.172	5.391
Wine	3	4.813	4.437	4.422
Dataset	Number of classes	Numbers of iterations		
		P=median(S)/2	P=median(S)	P=2*median(S)
Iris	3	152	158	178
Wine	3	157	140	141

Table 3 The clustering accuracy rate of our method

Method	Clustering accuracy rate(%)	
	Iris	Wine
Our method(P=median(S)/2)	95.82	92.25
Our method(P=median(S))	95.82	92.25
Our method(P=2*median(S))	93.20	93.36
K-means	88.67	86.52
Neighborhood covering(mean of 50 times)	95.58	91.49

Table 4 Parameters of datasets

Dataset	Formula	Variable
Friedman#1	$y = 10sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5$	$x_i \sim U[0,1]$
Friedman#2	$y = \sqrt{x_1^2 + (x_2 x_3 - (\frac{1}{x_2 x_4}))^2}$	$x_1 \sim U[0,100], x_2 \sim U[40\pi, 560\pi], x_3 \sim U[0,1], x_4 \sim U[0,11]$
Friedman#3	$y = tan^{-1}\frac{x_2 x_3 - \frac{1}{x_2 x_4}}{x_1}$	$x_1 \sim U[0,100], x_2 \sim U[40\pi, 560\pi], x_3 \sim U[0,1], x_4 \sim U[0,11]$

Table 5 The generalization error between bagging, boosting and our proposed algorithm

Dataset	Generalization error		
	Bagging	Boosting	Our proposed algorithm
Friedman#1	1.055	1.032	0.876
Friedman#2	0.955	0.956	0.823
Friedman#3	0.982	0.983	0.981

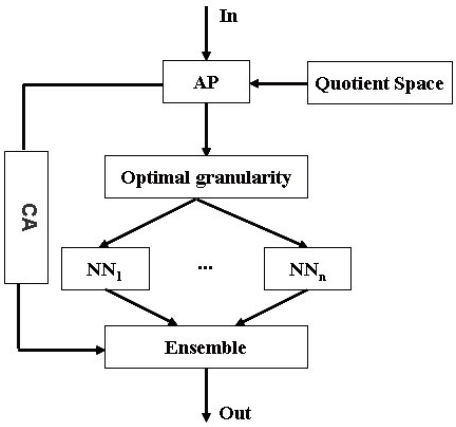

Fig. 1 Architecture of NNE

4. Experimental results

1)In this section we present two datasets from the UCI Machine Learning Repository[26] to verify characteristics and accuracy of the improved AP clustering algorithm based on the quotient space granularity analysis our proposed.

Table 1 shows the sample size, the number of classes, and the number of attributions of Iris and Wine.

As we know, among all the parameters of AP clustering, the most important is the preference (P), which directly influences the number of clusters. Note that P is negative. In our experiments, we select three different values of P to deal with the datasets, including the median of similarity matrix S (denoted as median(S)), its half (median(S)/2), and its double (2*median(S)). The results of clustering using the different P are shown in Table 2. Table3 lists the clustering accuracy rate of our proposed algorithm, k-means and neighborhood covering. Overall, our proposed algorithm has better accuracy.

2)In this section we present three groups of multiple regression datasets to verify the effectiveness of individual neural network generation and ensemble algorithm based on quotient space granularity clustering our proposed. Table 4 shows the parameters of sample datasets.

When generating experimental datasets, x_1 subjects to uniform distribution. According to the proportion of 80% and 20%, the dataset is Randomly divided into independent training dataset and test dataset, and then random noise which follows a normal distribution N (0, 1) is be added to the training dataset. The validation dataset does not add the noise. Table 5 shows the generalization error between bagging, boosting and our proposed algorithm. Table5 shows that generalization error of our proposed algorithm is relatively small.

5. Conclusion

This paper develops an individual neural network generation and ensemble algorithm based on quotient space granularity clustering by introducing the quotient space concept to the AP clustering analysis, which can find an optimal granularity from all possible granularities. Then using this improved AP clustering algorithm to seek optimal sample clustering results and using clustering results to determine the number of individual neural networks. Using different individual neural network to learn different categories of samples so that the degree of difference between the networks and the generalization performance of NNE can be improved. Further, calculating the correlation between the input data and sample class center, according to the degree of correlation between the input data and the sample category to dynamically and adaptively adjust ensemble weights. The algorithm our proposed is not only a method of training the individual neural networks, but also can adaptively adjust ensemble weights of individual neural network. The experiments show that our proposed method is validity and correctness. In the future, we will further improve the NNE algorithm by adding the ideas of multi-granularity, and further improve the efficiency of NNE.

Acknowledgement

This work is supported by the National Key Basic Research Program of China(No.2013CB329502), the National Natural Science Foundation of China(No.41074003), and the Opening Foundation of Key Laboratory of Intelligent Information Processing, Chinese Academy of Sciences (No.IIP2010-1).

References

[1] Hansen L K, Salamon P. Neural network ensemble. IEEE Trans on Pattern Analysis and Machine Intelligence, 1990, **12**(10): 993-1001.

[2] Zhou Zhihua, Wu Jianxin, Tang Wei. Ensembling neural networks: many could be better than all. Artificial Intelligence, 2002, **137**(2):239-263

[3] Li Kai, Huang Houkuan. A selective approach to neural network ensemble based on clustering technology. Journal of Computer Research and Development, 2005, **42**(4): 594-598.

[4] Li Zhengguo, Yang Jie, Kong Ansheng, Chen Nianyi. Clustering algorithm based selective ensemble. Journal of Fudan University (Natural Science), 2004, **43**(5): 689-692.

[5] MdMonirul IslamXin Yao and Kazuyuki MuraseA Constructive Algorithm for Training Cooperative Neural Network EnsemblesIEEE Transactions On Neural Networks200314(4)820 834

[6] Wang Zhengqun, Chen Shifu, Chen Zhaoqian and Xie Junyuan. A parallel learning approach for neural network ensemble.Artificial Intelligence, 2004, 1200-1205.
[7] Pitoyo Hartono and Shuji HashimoloAdaptive Neural Netwrok Ensemble That Learns From Imperfect Supervisor. Proceedings of the 9th Intern ational Conference on Neural Information Processing(ICONIP02). 2002, 5: 2561-2565
[8] Xin YaoManfred Fischert and Gavin Brown Neural Network Ensembles and Their Application to Tr c Flow Prediction in Telecommunications NetworksIEEE TransEvo1Comp20014693- 698
[9] Bao Jian, Zhou Liangjie. Analysis on complexity of neural networks using integer weights. Applied Mathematics & Information Sciences, 2012, **6**(2):317-323.
[10] Frey B. J. , Dueck D. Clustering by passing messages between data points. Science, 2007, **315** (5814): 972-976.
[11] Liu Xiaoyong, Fu Hui. A fast affinity propagation clustering algorithm. Journal of Shandong University (Engineering Science), 2011, **41**(4): 20-28.
[12] Xia Dingyin, Wu Fei, Zhang Xuqing, Zhuang Yueting. Local and global approaches of affinity propagation clustering for large scale data. Journal of Zhejiang University-Science A, 2008, **9**(10): 1373-1381. [doi:10.1631/jzus.A0720058].
[13] Lazic N., Givoni I., Aarabi P., et al. FLoSS: Facility location for subspace segmentation / /Proceedings of 12th International Conference on Computer Vision (ICCV). Kyoto: IEEE Press, 2009: 825-832.
[14] Lazic N., Frey B. J., Aarabi P. Solving the uncapacitated facility location problem using message passing algorithms // Proceedings o f 13th International Conference on Artificial Intelligence and Statistics(AISTATS). Sardinia: Microtome Publishing, 2010:429-436.
[15] Tang Dongming, Zhu Qingxin, Yang Fan, et al. Solving largescale location problem using affinity propagation clustering. Application Research of Computers, 2010, **27**(3): 841-844.
[16] Dueck D., Frey B. J. Nonmetric affinity propagation for unsupervised imag ecategorization / / Proceedings of 11th International Conference on Computer Vision (ICCV). Riode Janeiro: IEEE Press, 2007: 1-8.
[17] Zhang Renyan, Zhao Hongliang, Lu Xiao, et al. Grey image segmentation method based on affinity propagation clustering. Journal of Naval University of Engineering, 2009, **21**(3): 33-37.
[18] Givoni I. E., Frey B. J. Semi-supervised affinity propagation with instance-level constraints / /Proceedings of 12th International Conference on Artificial Intelligence and Statistics (AISTATS). Florida: Microtome Publishing, 2009: 161-168.
[19] Guan Renchu, Pei Zhili, Shi Xiaohu, et al. Weight affinity progagation and its application to text clustering. Journal of Computer Research and Development, 2010, **47**(10): 1733-1740.
[20] Xu Xinzheng, Ding Shifei, et al. Optimizing radial basis function neural network based on rough sets and affinity propagation clustering algorithm. Journal of Zhejiang University-SCIENCE C (Computers & Electronics), 2012, **13**(2): 131-138.
[21] Dueck D., Frey B. J., Jojic N., et al. Constructing treatment portfolios using affinity propagation / / Proceedings of International Conference on Research in Computational Molecular Biology (RECOMB). Singapore: Springer, 2008: 360-371.
[22] Xu Wenzhu, Xu Lihong. Adaptive key-frame extraction based on affinity propagation clustering. Computer Science, 2010(1): 268-270.
[23] Xiang Peisu. New CBIR system based on the affinity propagation clustering algorithm. Journal of Southwest University for Nationalities: Natural Science Edition, 2010, **36**(4): 624-627.
[24] Frey B. J. , Dueck D. Clustering by passing messages between data points. Science, 2007, **315** (5814): 972-976.
[25] Dueck D., Frey B. J., Jojic N., et al. Constructing treatment portfolios using affinity propagation / / Proceedings of International Conference on Research in Computational Molecular Biology (RECOMB). Singapore: Springer, 2008: 360-371.
[26] Li Yaqin, Yang Huizhong. Multi-model modeling method based on affinity propagation clustering and Gaussian processes. Computers and Applied Chemistry, 2010, **27**(1): 51-54.

Li Hui was born in Xuzhou, China, in 1978, received his MS degree from Suzhou University of Science and Technology in 2007. He is a doctoral student of China University of Mining and Technology and his advisor was Professor Ding Shifei. His research interests are in the areas of Artificial Intelligence, Pattern Recognition, Neural networks, Granular Computing et al.

Ding Shifei received his bachelor's degree and master's degree from Qufu Normal University in 1987 and 1998 respectively. He received his Ph.D degree from Shandong University of Science and Technology in 2004. He received postdoctoral degree from Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, and Chinese Academy of Sciences in 2006. And now, he works in China University of Mining and Technology as a professor and Ph.D supervisor. His research interests include intelligent information processing, pattern recognition, machine learning, data mining, and granular computing et al. He has published 3 books, and more than 100 research papers in journals and international conferences. Prof. Ding is a senior member of China computer Federation (CCF), and China Association for Artificial Intelligence (CAAI). He is a member of professional committee of Artificial Intelligence and Pattern Recognition, CCF, professional committee of distributed intelligence and knowledge engineering, CAAI, professional committee of machine learning, CAAI, and professional committee of rough set and soft computing, CAAI. He is an associate Editor-in-Chief, Topic Editor-in-Chief (TEIC) for International Journal of Digital Content Technology and its Applications (JDCTA), acts as an editor for Journal of Convergence Information Technology (JCIT), International Journal of Digital Content Technology and its Applications (JDCTA). Meanwhile, he is a reviewer for Journal of Information Science (JIS), Information Sciences (INS), Computational Statistics and Data Analysis (CSTA), IEEE Transactions on Fuzzy Systems (IEEE TFS), Applied Soft Computing (ASOC), Computational Statistics and Data Analysis (CSDA), International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI) et al.