A Parallel Attribute Reduction Algorithm based on Affinity Propagation Clustering

Hong Zhu
School of Computer Science and Technology, China University of Mining and Technology
School of Medical Information, Xuzhou Medical College, Xuzhou, China, 221000
Email: zhuhongwin@126.com

Shifei Ding, Xinzheng Xu and Li Xu
School of Computer Science and Technology, China University of Mining and Technology,
Email: dingsf@cumt.edu.cn

***Abstract*—As information technology is developing rapidly, massive and high dimensional data sets have appeared in abundance. The existing attribute reduction methods are encountering bottleneck problem of timeliness and spatiality. AP(Affinity Propagation) is an efficient and fast clustering algorithm for large dataset compared with the existing clustering algorithms. This paper discusses attribute clustering method in order to reduce attributes and provides a kind of parallel attribute reduction algorithm based on Affinity Propagation （APPAR） clustering. The attribute set is clustered into several subsets by Affinity Propagation algorithm first, and then the reductions of these subsets are proposed concurrently in order to get attribute reduction set of the whole data set. The whole algorithm has been improved in the two sides so as to largely increase the algorithm's speed. Experimental results show that the APPAR method is outperforming traditional attribute reduction algorithm for huge and high dimensional dataset processing.**

***Index Terms*—attribute clustering, attribute reduction, parallel computing**

Ⅰ. INTRODUCTION

The rapid development of Internet and various information system results in massive, high dimensional complex data which are often incomplete, reliability, inaccuracy and inconsistency. For high dimensional data, the redundant attributes of samplers will not only increase the complexity of the calculation, but also affect the accuracy of final result.

Rough set has a wide range of applications in pre-processing of massive high-dimensional complex data[1]. Attribute reduction is one of the core issues of rough set theory. The biggest characteristic of attribute reduction algorithm based on the rough set is to keep the same classification ability. As information technology is developing rapidly, massive and high dimensional data sets have appeared in abundance. The existing attribute reduction methods are encountering bottleneck problem of timeliness and spatiality. Although any experts made unremitting efforts, the reduction of time and space complexity is still current focus of research in this field[2-10]. People turn to reduce dimension through attribute clustering.

Attribute clustering is to cluster attribute set into several subsets according the distance between every two attributes. So, after clustering, the attributes those separating capacity is similarity are divided into the same cluster. These clusters are the subsets of original attribute set. Representative attributes are produced from each subset, and other attributes are reduced. Yet it's easy to loss information contained in original attribute set.

According to high dimension data attribute reduction, this paper provides a parallel attribute reduction algorithm based on AP clustering. The method clusters attribute set into several subsets first, and then the reductions of these subsets are proposed concurrently in order to get attribute reduction set of the whole attribute set. Algorithm can adjust the number and size of subsets automatically through the change of parameters.

The remainder of this paper is organized as follows. The basic theory and methods of attribute clustering are presented in Section II. Section III introduces a novel clustering method--Affinity propagation clustering algorithm. A parallel attribute reduction algorithm based on Affinity Propagation clustering is presented in Section IV. Some simulation experimental evaluations are discussed to show the performance of the developed methods in Section V. The paper ends with conclusion in Section VI.

Manuscript received May 10, 2012; revised June 1, 2012; accepted July 1, 2012.

Corresponding author: Ding Shifei

doi:10.4304/jcp.8.4.990-997

II. ATTRIBUTE CLUSTERING

Data clustering is to group a set of data (without a predefined class attribute), based on the conceptual clustering principle: maximizing the intraclass similarity and minimizing the interclass similarity[11]. Clustering is one of the important research contents in the field of pattern recognition, image processing, data mining, machine learning and so on. It plays a vital role in aspect of identifying data's intrinsic structure. The study of cluster analysis is always the hot focus because of its importance, multiple application fields and cross-cutting features with other research direction.

As an unsupervised machine learning method, cluster analysis has been widely used in natural and social science. It classifies some objects into several clusters, making the differences of the objects in distinct classes as large as possible while in the same as small as possible. The most essential point—"clustering" is found among samples. The analysis results not only reveal the intrinsic differences and connections between the data, but also provide an important basis for the further data mining. The typical clustering algorithms are listed: the method based on hierarchy (gathering, splitting), division (K-means, K-center), density (DBSCAN, OPTICS, DENCLUE), grid (STING, Wave Cluster, CLIQUE), model (statistical methods, neural network methods), constraint and so on.

According to the difference of classification objects, clustering analysis is divided into sample clustering and attributes clustering. The former is called Q clustering and the later is called R clustering. But most of the work is focused on sample clustering in data mining. But attribute clustering has very important applications in many fields such as data preprocessing, association rules mining and so on. The essence of attribute clustering is knowledge reduction. Knowledge reduction is to remove not important or redundant knowledge under the condition of keeping the decision-making ability in knowledge base. Minimum reduction (contain minimum attribute reduction) is expected.

As information technology is developing rapidly, massive and high dimensional data sets have appeared in abundance. The existing attribute reduction methods are encountering bottleneck problem of timeliness and spatiality. Although any experts made unremitting efforts, the reduction of time and space complexity is still current focus of research in this field. People turn to reduce dimension through attribute clustering.

The aim of attribute clustering is to cluster attributes into several subsets according the similarity (such as distance) between attributes. The similarity between attributes in different subsets is rather less and in the same subsets is comparatively large. So the distinguish ability of attributes are similar in each group. The distance between clusters is the distance between attributes which represent their clusters. And then from every attribute subsets, we select representative attributes which have the same distinguish ability as their subsets. The representative attributes of each attribute subsets consist of attribute reduction set.

Attribute clustering has three key problems:

1) Select attribute similarity function: there are many methods suitable for attribute clustering, such as distance method, related coefficient method, angle cosine method and so on.
2) Select clustering algorithm: all data clustering algorithms are applicable to attribute clustering theoretically as long as the similarity function is reasonable.
3) Select representative attributes from each attribute subset: there are many methods for selecting representative attributes such as clustering center, entropy of information, weighted attributes method and so on.

III. AP CLUSTERING ALGORITHM

A. The Disadvantage of K-means Algorithm

The k-means algorithm is an evolutionary algorithm that gains its name from its method of operation. The algorithm clusters observations into k groups, where k is provided as an input parameter. It then assigns each observation to clusters based upon the observation's proximity to the mean of the cluster. The cluster's mean is then recomputed and the process begins again. Here's how the algorithm works:

Step 1: Choosing the number of clusters k

Step 2:The algorithm arbitrarily selects k points as the initial cluster centers ("means").

Step 3: Each point in the dataset is assigned to the closed cluster, based upon the Euclidean distance(or other distance) between each point and each cluster center.

Step 4:Each cluster center is recomputed as the average of the points in that cluster.

Steps 3 and 4 repeat until the clusters converge. Convergence may be defined differently depending upon the implementation, but it normally means that either no observations change clusters when steps 3 and 4 are repeated or that the changes do not make a material difference in the definition of the clusters.

K-means algorithm which has a certain self-adaptive and can achieve dynamic clustering is a classical algorithm for cluster analysis, but one of the main disadvantages to k-means is the fact that you must specify the number of clusters as an input to the algorithm. As designed, the algorithm is not capable of determining the appropriate number of clusters and depends upon the user to identify this in advance.

B. Affinity Propagation Clustering

Affinity propagation clustering (AP) is a novel message passing algorithm and first be proposed by Frey and Dueck in Science[12]. Different from algorithms like k-centers clustering, affinity

propagation doesn't fix the cluster number. In contrast, it considers all data points as candidate exemplars by simultaneously to avoid an unlucky initializations. Affinity propagation takes as input measures of similarity between pairs of data points. Real-valued messages are exchanged between data points until a high-quality set of exemplars and corresponding clusters gradually emerges. Affinity propagation has been used to cluster images of faces, detect genes in microarray data, identify representative sentences in this manuscript, and identify cities that are efficiently accessed by airline travel. Affinity propagation found clusters with much lower error than other methods, and it did so in less than one-hundredth the amount of time[13-15].

In AP algorithm, the similarities s (i, j) = $-|| x_i - x_j ||^2$ between any two data points x_i and x_j are stored in N*N matrix. s (i, j) and s (j, i) can take different values, this is different from K means algorithm. Before clustering, AP takes as input a real number s(k,k) for each data point k so that data points with larger values of s(k,k) are more likely to be chosen as exemplars. These values are referred to as "preferences". If a priori, all data points are equally suitable as exemplars, the preferences should be set to a common value. This value can be varied to produce different numbers of clusters.

There are two kinds of message exchanged between data points. The "responsibility" r(i,k), sent from data point i to candidate exemplar point k, reflects the accumulated evidence for how well-suited point k is to serve as the exemplar for point i, taking into account other potential exemplars for point i. The "availability" a(i,k), sent from candidate exemplar point k to point i, reflects the accumulated evidence for how appropriate it would be for point i to choose point k as its exemplar, taking into account the support from other points that point k should be an exemplar(Fig.1).

$$a(i,k) \leftarrow \min\left\{0, r(k,k) + \sum_{i's.t.i' \notin \{i,k\}} \max\{0, r(i',k)\}\right\} \quad (2)$$

$$a(k,k) \leftarrow \sum_{i's.t.i' \neq k} \max(0, r(i',k)) \quad (3)$$

At any point during affinity propagation, availabilities and responsibilities can be combined to identify exemplars. For point i, the value of k that maximizes a(i,k) + r(i,k) either identifies point i as an exemplar if k = i, or identifies the data point that is the exemplar for point i. The message-passing procedure may be terminated after a fixed number of iterations, after changes in the messages fall below a threshold, or after the local decisions stay constant for some number of iterations.

When updating the messages, it is important that they be damped to avoid numerical oscillations that arise in some circumstances. Each message is set to λ times its value from the previous iteration(Rold or Aold) plus (1 - λ) times its prescribed updated value.

$$R = (1-\lambda) * R + \lambda * Rold \quad (4)$$

$$A = (1-\lambda) * A + \lambda * Aold \quad (5)$$

The damping factor l is between 0 and 1. We often uses a default damping factor of 0.5 in our experiments. And each iteration of affinity propagation consisted of (i) updating all responsibilities given the availabilities, (ii) updating all availabilities given the responsibilities, and (iii) combining availabilities and responsibilities to monitor the exemplar decisions and terminate the algorithm when these decisions did not change for a fixed iterations.

IV. A Parallel Attribute Reduction Algorithm Based On AP Clustering

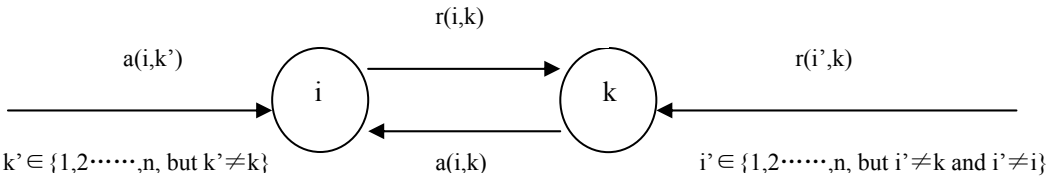

Figure 1. Message passing procedure between data points

A. Attribute Reduction

The responsibilities and availabilities are computed using the rules:

$$r(i,k) \leftarrow s(i,k) - \max_{k's.t.k' \neq k} \{a(i,k') + s(i,k')\} \quad (1)$$

Rough set theory[16]is firstly proposed by Z. Pawlak in 1982. It is a mathematical tool which can deal with imprecise, inconsistent, incomplete information and knowledge quantitatively. It has been applied in such fields as machine learning, data mining, intelligent data analysis and control algorithm acquiring, etc. Knowledge reduction is one of the most important contributions of rough set theory to machine learning, pattern recognition and data mining. The most notable difference with other theories which deal with uncertain and imprecise problem is: it does not need to provide any prior

information except the necessary data sets, so the description and processing for uncertainty problem is more objective. Rough set theory has become an important intelligent information processing technology by its unique approach and innovative idea. Its main idea is: at the premise of keeping the classification ability unchanged, a problem of decision-making or classification rules can be gotten through the simplification of knowledge. Among these existing knowledge reduction methods, one group method focuses on the indiscernibility relation in a universe that captures the equivalence of objects, while the other group considers the discernibility relation that explores the differences of objects. For indiscernibility relation, one can employ it to induce a partition of the universe and thereby to construct positive regions whose objects can be undoubtedly classified into a certain class with respect to the selected attributes. Thus, knowledge reduction algorithms based on positive regions have been proposed[17-19]. For discernibility relation, we have knowledge reduction algorithms based on a discernibility matrix and information entropy. Reduction methods based on discernibility matrix[20] have high cost of storage with space complexity O(m*n2) for a large decision table with n objects and m conditional attributes. The problem of finding a minimal reduction of a given information system was proven to be a NP-hard problem. As information technology is developing rapidly, massive and high dimensional data sets have appeared in abundance. The existing attribute reduction methods are encountering bottleneck problem of timeliness and spatiality. Although any experts made unremitting efforts, the reduction of time and space complexity is still current focus of research in this field. People turn to reduce dimension through attribute clustering.

B. Attribute Reduction based on AP Clustering

Attribute reduction algorithm based on AP clustering selects AP clustering algorithm to cluster attribute set of high dimension mass data sets. Besides the reason mentioned in A of III, the time complexity of K-means clustering is important. Time complexity of K-means is n * K * m, including sample n, K for category number and m for sample dimension. This time complexity is quite appreciable If we adopt the method of exhaustion to search for the optimal method. AP clustering method considers all data points as candidate exemplars by simultaneously to avoid unlucky initializations. Besides it can deal with the large-scale data, and can take as input general nonmetric similarities. Although its time complexity is $n^2 \log n$, it could search for the optimal method without exhaustion method. So it exhibits fast execution speed with low error rate.

After AP clustering, attribute set is divided into several subsets. In each subset, the distinguish ability of attributes are similar. Then APPAR algorithm reduces each attribute subsets concurrently and form a reduction set. At last, APPAR reduces the reduction set and obtains the final attribute reduction set of the whole data set. This conclusion is based on Proposition 1.

Definition 1. Let S=(U,A,V,f) be an information system, for attribute set $X \subseteq A$, we definite the classification of U（U/X）: for two objects u, $v \in U$ are of the same class, if and only if for every $a \in X$, a(u)=a(v).

Proposition 1. Let S=(U,A,V,f) be an information system, let A1,A2,……,An be any subset of A, let P1 be the reduction set of A1, P2 be the reduction set of A2……, Pn be the reduction set of An, and let P be the reduction set of {P1,P2,……,Pn}, then P is the reduction set of A.

Proof. From the definition 1, for attribute set $A1 \subseteq A$, the classification of U（U/A1）: for two objects u, $v \in U$ are of the same class, if and only if for every $a \in A1$, a(u)=a(v). From the hypothesis, P1 is the reduction set of A1, so P1 has the same ability of classification as A1, thus the classification of U（U/P1）: for two objects u, $v \in U$ are of the same class, if and only if for every $a \in A1$, a(u)=a(v).

For attribute set $A2 \subseteq A$, the classification of U（U/A2）: for two objects u, $v \in U$ are of the same class, if and only if for every $b \in A2$, b(u)=b(v). From the hypothesis, P2 is the reduction set of A2, so P2 has the same ability of classification as A2, thus the classification of U（U/P2）: for two objects u, $v \in U$ are of the same class, if and only if for every $b \in A2$, b(u)=b(v).

For attribute set $A1 \subseteq A$ and $A2 \subseteq A$, the classification of U（U/A1A2）: for two objects u, $v \in U$ are of the same class, if and only if for every $a \in A1$, $b \in A2$, a(u)=a(v) and b(u)=b(v) are simultaneously true.

From the hypothesis of definition 1, if P is the reduction of P1 and P2, the classification of U（U/P）: for two objects u, $v \in U$ are of the same class, if and only if for every $a \in A1$, $b \in A2$, a(u)=a(v) and b(u)=b(v) are simultaneously true.

So, P is the reduction of attribute subset A1 and A2.

And so on, if P is the reduction set of {P1,P2,……,Pn}, then P is the reduction set of the whole attribute set A.

C. Parallel Attribute Reduction Algorithm based on Affinity Propagation （APPAR）

Massively parallel computer (MPC) has been trying to pursuit the goal of high-performance. With the entrance and mature of Single Chip Multiprocessors to the main stream markets, parallel programming is particularly important. But at present, supporting

parallel computing programming model relatively lags behind and has no corresponding standard. To a certain extent, this leads to the result that parallel programming ideas is still far from the mainstream program designer. So, it is unrealistic to rely on the compiler to complete serial code to parallel code transformation without changing programming habit. In order to implement high performance, programmers should be devoted to the development of parallel degree of applied problems. The basic strategy is to refine calculation granularity[21-23].

Degree of parallelism is number of processes executed concurrently. Parallel granularity is computing load performed by operations between two parallelism or interaction. Degree of parallelism and parallel granularity are reciprocal. Increasing parallel granularity will reduce degree of parallelism. Fine — grain parallelism is emphasized in the design of massively parallel computers to acquire higher performance. Fine granularity parallel can fully mining potential parallelism of application problems, and increase the degree of parallelism of algorithm. Along with the calculation granularity refined, total communication traffic will be increased, and communication cost either. Fine granularity has enough parallelism, so it can hide communication delays through overlapping communication and computing technology, to decrease negative performance impact caused by the increasement of communication cost brought by fine granularity.

Multithreading calculation is a commonly used method to achieve fine granularity parallel. As long as the problem itself has enough parallelism, the multiple light load processes in the multi-threading calculation can ensure calculation has enough fine granularity, so as to improve the actual degree of parallelism of the computation and make full use rich hardware resources of MPC. Multi-threaded calculation can cover communication delays through threads of switching. Multi-thread switching cost be solved through special multithreading processor in massively parallel system.

The aim of attribute clustering is to obtain clearly results which have practical significance, strong resolution and representative. After AP clustering, APPAR algorithm reduces each attribute subsets concurrently in order to improve the efficiency of the algorithm.

Bernstein criteria about two programs p1 and p2 which can be executed in parallel is that: P1 input variables set and P2 output variables set do not intersect and vice versa. Their output variables set also not intersect. For APPAR algorithm, each attribute subset is input set and they do not intersect and vice versa. Their output sets do not intersect. So APPAR algorithm can be executed concurrently on each attribute set using multi-thread. This method could refine calculation granularity and improve parallel degree so as to implement high performance.

A parallel attribute reduction algorithm based on AP clustering has four steps. It is described as the following:

Algorithm 1: A parallel attribute reduction algorithm based on AP clustering

Input: Data set

Output: The reduction set of attributes

Step 1: Calculate similarity matrix

（a）Data normalization

Data normalization can make each attribute value be united in a common numerical characteristics range.

$$X = \frac{X' - \overline{X}'}{C} \tag{6}$$

In Eq. (6), X′ are original data, $\overline{X}'$ are the average of the original data, C is the variance of the original data.

The normalized data can be compressed into [0,1] by using extreme value standardization formula in Eq. (7).

$$X = \frac{X' - X'_{\min}}{X'_{\max} - X'_{\min}} \tag{7}$$

(b)Calculate similarity matrix elements

Calculate similarity matrix elements in order to get the similarity relation matrix S.

Step 2: Attribute clustering by AP algorithm

(a) Initialization

s(k,k) are assigned the same value and the value is also assigned to parameter P; assign initial values to r(i,k) and a(i,k),and store in matrixes R and A; assign initial value toλ.

(b）Iteration:

Calculate R:

1) calculate R

calculate r(i,k) (k=k’)

2) R=(1-λ)*R+λ*Rold

Calculate A:

1) calculate A

2) calculate a(k,k)

3) A=(1-λ)*A+λ*Aold

Judge whether the algorithm meets the following conditions, if one of them is satisfied, the iteration may be terminated.

- Exceed the maximum iterating times
- the change of Information falls below a given threshold
- the selected clustering center remains stable

(c) Output attributes clustering results

Step 3: Select representative attributes in parallel

In order to balance load, each attribute subsets should be reduced concurrently following the method

below：

（a）Calculate discernibility matrix M= { Mst } and core C

(b) Calculate frequency of each attribute $F(a_i)$

(c) P_i= C

$Q=\{Mst \cap P_i \neq \emptyset\}$

$M=M-Q$

$B=A-P_i$

$F(a_q)=Max\{F(a_i)\}$

$Pi=Pi \cup \{ a_q \}$

Repeat the process above, until M= Ø

Pi is the reduction of one attribute subset.

(d) Calculate the reduction set of {P1,P2,……Pn} to form attribute set P

Step 4: Output P, and it is the reduction set of the whole attribute set

V. EXPERIMENTS

Experiment used four famous data sets in UCI data set for the test. Glass Identification data set has 214 objects which are divided into float and non float including 10 condition attributes and a decision attribute. Except incomplete objects, Mushroom data set has 5644 objects, including 22 condition attributes and a decision attribute. Table 1 shows the characteristics of the four data sets:

TABLE Ⅰ

CHARACTERISTICS OF FOUR DATA SETS IN UCI

Data set	Number of samples	Number of attributes	Number of class
Iris	150	4	3
Glass Identification	214	10	2
Ionosphere	351	34	2
Mushroom	5644	22	2

The results of attribute reduction using original method and APPAR are compared as follows(table 2 and table 3):

TABLE Ⅱ

THE COMPARISONS OF REDUCTION RESULTS BETWEEN AR ALGORITHM AND APPAR ALGORITHM

Data set	Reduction results	
	AR	APPAR
Iris	{a2,a3,a4}	{a2,a3,a4}
Glass Identification	{ a1,a3,a5, a6,a7}	{ a1,a3,a5, a6,a7}
Ionosphere	{a14,a16,a28}	{a14,a16,a28}
Mushroom	{a2,a3,a4,a6, a8,a10,a13, a14,a15,a16, a21,a22}	{a2,a3,a4,a6, a8,a10,a13, a14,a15,a16, a21,a22}

TABLE Ⅲ.

THE COMPARISONS OF RUNTIME BETWEEN AR ALGORITHM AND APPAR ALGORITHM

Data set	Runtime （s）	
	AR	APPAR
Iris	0.6875	1.0157
Glass Identification	8.0756	7.0312
Ionosphere	5.2116E+002	3.1516E+002
Mushroom	7.2308E+003	2.9723E+003

For Iris data set, APPAR and AR algorithm have the same result {a2,a3,a4} , but APPAR algorithm is lower than original AR. Because the variation of attributes sets is small after clustering and attributes clustering waste a lot of time. But for Mushroom data set, for the same result {a2,a3,a4,a6,a8,a10, a13,a14,a15,a16,a21,a22}, APPAD algorithm has obvious advantages. After AP clustering, attributes of Mushroom are clustered into three subsets： $\{a_1,a_2,a_4,a_6,a_7,a_{11},a_{13},a_{20},a_{22}\}$, $\{a_3,a_5,a_{14},a_{15},a_{16},a_{21}\}$, $\{a_8,a_{19}\}$. The reductions of these three subsets are proposed concurrently. At last, the reduction results is {a2,a3,a4,a6,a8,a10,a13, a14,a15,a16,a21,a22}. But APPAR algorithm is faster than original AR. From the test of Glass Identification and Ionosphere data sets, we can draw the same conclusion. So, for large data sets reduction, APPAR algorithm is superiority.

VI. CONCLUSION

Attribute reduction is one of the core issues of rough set theory. The biggest characteristic of attribute reduction algorithm based on the rough set is to keep the same classification ability. As information technology is developing rapidly, massive and high dimensional data sets have appeared in abundance. The existing attribute reduction methods are encountering bottleneck problem of timeliness and spatiality.

This paper provides a kind of parallel attribute reduction algorithm based on AP clustering. The attribute

set is clustered into several subsets by AP algorithm first, and then the reductions of these subsets are proposed concurrently in order to get attribute reduction set of the whole data set. Because subsets after AP clustering is smaller than original attribute set, the existing attribute reduction algorithm is efficient. APPAR combines the fast and effective advantage of AP and the advantage of existing attribute reduction algorithm. It provides a new thought and method for high dimensions and massive data sets. The experimental results show that for large data sets reduction, APPAR algorithm is superiority.

Although the algorithm shows its effectiveness, there are still some problems we need to further study. We will develop the APPAR algorithm further in order to improve the effectiveness in finding subspace.

VII. ACKNOWLEDGEMENTS

This work is supported in part by a grant from the Basic Research Program (Natural Science Foundation) of Jiangsu Province of China (No.BK2009093), the National Nature Science Foundation of China (No.60975039), the Opening Foundation of Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (No.IIP2010-1), and the Scientific Innovation Research of College Graduate in Jiangsu Province (No.CXZZ11_0296).

REFERENCE

[1] Wang Guoyin, Yao Yiyu. A survey on rough set theory and applications. Chinese Journal of Computers, 2009, 32(7): 1229-1246

[2] Yang Ming. An incremental updating algorithm for attribute reduction based on improved discernibility matrix. Chinese Journal of Computers, 2007,30(5):815-822(in Chinese)

[3] Wang Guoyin, Yu Hong, Yang Dachun. Decision table reduction based on conditional information entropy. Chinese Journal of Computers, 2002, 25(7): 759-766

[4] Xu Zhangyan , Liu Zuopeng, Yang Bingru, Song Wei. A quick attribute reduction algorithm with complexity of max { O(| C| | U|), O (| C | ^{2} | U/ C |) }. Chinese Journal of Computers, 2006, 29(3) : 391-399

[5] Liu Shaohui, Sheng Qiujian, Wu Bin, Shi Zhongzhi, Hu Fei. Research on efficient algorithms for Rough set methods. Chinese Journal of Computers, 2003, 26(5): 524-529

[6] Ye Dongyi, Chen Zhaojiong. A new discernibility matrix and the computation of a core. Acta Electronica Sinica, 2002, 30(7) : 1086-1088

[7] Wang Guoyin. The computation method of core attribute in decision table. Chinese Journal of Computers, 2003, 26(5) :611- 615

[8] Liu Qing, Liu Shaohui, Zhengfei. Rough logic and its applications in data reduction. Journal of Software, 2001, 12 (3): 415-419

[9] Hu Li-hua, Ding Shi-fei, Ding Hao. Research on heuristic attributes reduction algorithm of rough sets. Computer Engineering and Design,2011,32(4):1438-1441

[10] Yang Chuanjian, Ge Hao, Wang Zhisheng. Overview of attribute reduction based on rough set. Application Research of Computers, 2012,29(1):16-20

[11] Jiawei Han, Micheline Kamber. Data Mining: Concepts and Techniques (Second Edition). Massachusetts: Morgan Kaufmann Publishers, 2006

[12] B.J. Frey, D. Dueck, Clustering by passing messages between data points. Science, 2007, 315(5814): 972-976

[13] Dueck D, Frey B J, Jojic N, et al. Constructing treatment portfolios using affinity propagation[C]. Proceedings of 12th Annual International Conference, RECOMB 2008. Singapore. 3.30-4.2,2008: 360-371.

[14] Dueck, D, Frey, BJ, "Non-metric affinity propagation for unsupervised image categorization. Proceedings of the IEEE International Conference on Computer Vision (ICCV), October 2007.

[15] Yu Xiao, Jian Yu. Semi-supervised clustering based on affinity propagation algorithm. Journal of software, 2008,19(11): 2803-2813.

[16] PAWLAK Z, “Rough sets,” International Journal of Information and Computer Science, 1982,11(5):341-356

[17] K. Y. Hu, Y. C. Lu, and C. Y. Shi, Feature ranking in rough sets, AI Communications, 2003, 16: 41-50

[18] X. H. Hu, and N. Cercone, Learning in relational database: A rough set approach, International Journal of Computational Intelligence, 1995, 11: 323-338

[19] Z. Pawlak, A. Skowron, Rough sets: Some extensions. Information Sciences, 2007, 177: 28-40

[20] A. Skowron, C. Rauszer, The discernibility functions matrics and functions in information systems, In Slowinski, R. ed.: Intelligent Decision Support - Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publisher, Dordrecht 1992: 331-362

[21] Hong Gongbing. Fine-grain parallelism and multithreaded computing[J]. Computer Research and Development, 1996, 33(6):473-480

[22] Xia Fei, Dou Yong , Xu Jiaqing et al. Fine grained parallel zuker algorithm accelerator with storage optimization on FPGA[J]. Computer Research and Development, 2011,48(4):709-719

[23] Yu Lei, Liu Zhiyong. Study on Fine-grained Synchronization in Many-Core Architecture[C]// 10th ACIS International Conference on Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing, Washington: IEEE, 2009:524-529

Hong Zhu, born in 1970, received the M.A’s degree in applied computer Technology from China University of Mining and Technology. She is an associate professor at Xuzhou Medical College. Since 2009, she has been a Ph.D. degree candidate in applied computer Technology from the China University of Mining and Technology. Her research interests includes granule computing, clustering, parallel computing et al.

Email: zhuhongwin@126.com

Shifei Ding 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, 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 80 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 distributed intelligence and knowledge engineering, CAAI, professional committee of machine learning, CAAI, and professional committee of rough set and soft computing, CAAI. He 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.
Email: dingsf@cumt.edu.cn; dingshifei@sina.com

Xinzheng Xu, born in 1980. He has been a Ph.D. degree candidate in applied computer Technology from the China University of Mining and Technology. His research interests include intelligent information processing and granular computing et al.
Email: xxzheng@cumt.edu.cn

Li Xu, born in 1986. She has been a M.A's degree candidate in applied computer Technology from the China University of Mining and Technology. Her research interests include intelligent information processing and granular computing et al.
Email: xl412@126.com