Attribute Granulation Based on Attribute Discernibility and AP Algorithm

Hong Zhu
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China 221116
School of Medical Information, Xuzhou Medical College, Xuzhou, China 221000
Email: zhuhongwin@126.com

Shifei Ding
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China 221116
Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China 100190
Email: dingsf@cumt.edu.cn

Han Zhao
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China 221116

Lina Bao
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China 221116

***Abstract*—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. The existing attribute reduction methods are encountering bottleneck problem of timeliness and spatiality. In order to looking for a relatively coarse attributes granularity of problem solving, this paper proposes an efficient attribute granulation method to remove redundancy attribute. The method calculates the similarity of attributes according attribute discernibility first, and then clusters attributes into several group through affinity propagation clustering algorithm. At last, representative attributes are produced through some algorithms to form a coarser attribute granularity. Experimental results show that the attribute granulation method based on affinity propagation clustering algorithm(AGAP) method is a more efficient algorithm than traditional attribute reduction algorithm(AR).**

***Index Terms*—attribute granulation, attribute dependability, AP clustering, parallel computing**

I. INTRODUCTION

High dimensional data is a phenomenon in real-world data mining applications[1]. The trend we see with data for the past decade is towards more observations and high dimensions [2]. 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. But these attributes are not important equally. What people think of above all is dimension reduction. That is selecting some important attributes to compose an attribute set through certain algorithm. This attribute set can be the subset of the original attribute set, or be a new attribute set derived from the original attribute set through certain algorithm. Thus we can solve problems on the coarser attribute granularity.

Rough set has a wide range of applications in pre-processing of massive high-dimensional complex data. 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[3-8]. A number of attempts have been made to improve attribute reduction algorithm[9-11]. Some 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. We can solve problems on the coarser attribute granularity.

Manuscript received 2012; revised2012; accepted 2012.

Corresponding author: Shifei Ding

doi:10.4304/jsw.8.4.834-841

This paper puts forth attribute granule and related concepts , and then provides an attribute granulation method based on AP clustering algorithm according to attribute dependability. The method calculates the attribute dependability first, and then clusters the attributes into several groups according to attribute dependability. So, representative attributes are produced through some algorithms to form a coarser attribute granularity. Algorithm can adjust the number and size of attribute granule automatically through the change of parameters.

II. The Definition And Division Of Attribute Granule

Definition 1.Given a information system IS = (U, A), A is the attribute set, A1 , A2 , ……An are subsets of A, A1 , A2 , ……An are called attribute granules of universe U.

Definition 2.A1 , A2 , ……An are attribute granules of universe U, the size of attribute granule is defined as $d(A_i) = Card(A_i) = |A_i|$.

In other words, the size of attribute granule is the number of attributes contained in attribute granule.

Supposed R is the sum of the equivalence relations [12], and $R_1, R_2 \in R$, for any two elements of the whole researching objects: x, y, if exits $xR_1y \Rightarrow xR_2y$, we often say R_1 is more detailed than R_2. That is, given two equivalence relations corresponding to two different partitions, if one division set is included in the other, it shows that the latter set is bigger than the former one. The former subdivides the latter, denoting: $R_2 < R_1$. According to this principle, we can get a sequence of equivalence relations: $R_1 < ... < R_{n-1} < R_n$. R_1 is the "biggest"(fuzzy) relationship, R_n is the "smallest"(detailed)relationship, so a n-level tree will be obtained. All the leaf nodes constitute the domain, representing the smallest division, bottom-up each layer is a partition of the domain.

Definition3. A_1 , A_2 , ……A_n are attribute granules of universe U, the corresponding equivalence relations R_1、R_2.....R_n is a sequence of equivalence relations, and $R_1 < ... < R_{n-1} < R_n$, the coarser degree of granularity is defined as : A_1 is the most coarse-grained attribute granule, An is the most fine-grained attribute granule .

In a general way, the bigger attribute granule contains more attributes and the corresponding equivalence relation is finer. So the attribute granule is finer and vice versa.

We can get a sequence of attribute granule from coarse to fine by the way of adding attribute into attribute set. So that we can get a granularity world from coarse to fine. This is illustrated in figure 1.

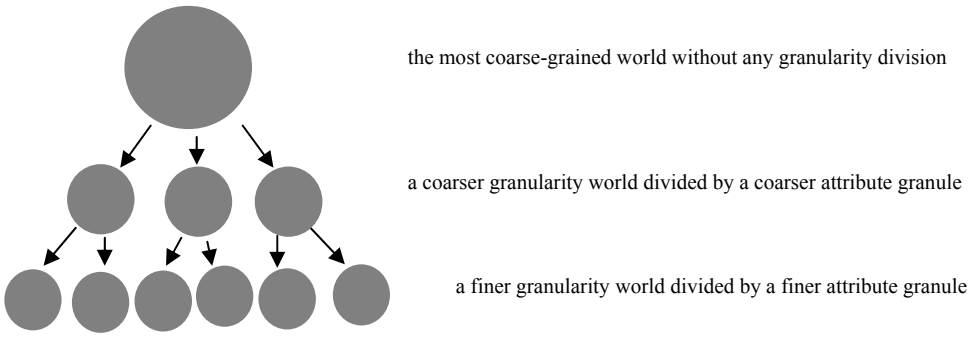

Figure 1. A sequence of attribute granule from coarse to fine

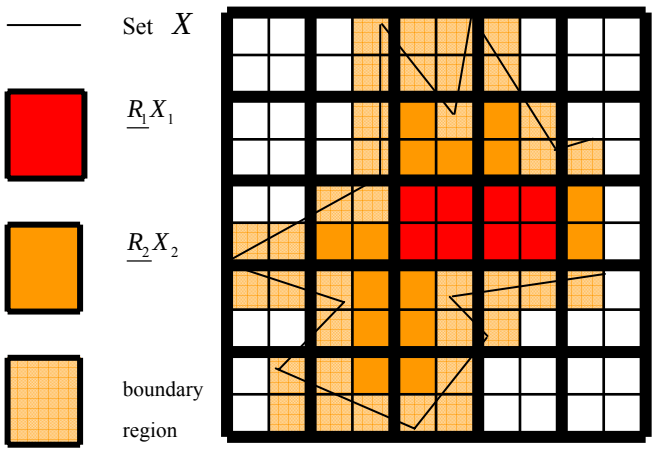

Figure 2. The effect of attribute set division on boundary region and positive regions.

Figure 2 illustrate the effect of attribute set division from coarse to fine on boundary region and positive regions.

Attribute granule is a partition of attribute set from different perspectives and different levels. Attribute granules are obtained in the way of granulating the original attribute set. The most coarse-grained attribute granule is each attribute, and the

fine-grained attribute granule is the whole attribute set.

Attribute granule is a kind of the information granules. There are several methods to granulating attribute set: the division based on equivalence relation, the division based on fuzzy set, the division based on information entropy, the division based on concept lattice, the division based on clustering algorithm, and so on.

The methods of attribute granulation are divided into two classes: attribute reduction and feature dimension. The former can reduce knowledge under maintained the same classification ability of decision table; the latter does not emphasize the same classification ability.

III. AP Clustering Algorithm

Affinity propagation clustering (AP) is a novel message passing algorithm and first be proposed by Frey and Dueck in Science[13]. 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[14-16].

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. The responsibilities and availabilities are computed using the rules:

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

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

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

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.

$$R = (1-\lambda) * R + \lambda * Rold \quad (4)$$
$$A = (1-\lambda) * A + \lambda * Aold \quad (5)$$

IV. Attribute Granulation Based On Attribute Discernibility And AP Algorithm(AGAP)

A. 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 [17]. 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 (the application multi-domains as well as) cross-cutting features with other research direction.

According to the difference of classification objects, clustering analysis is divided into sample clustering and attribute 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.

B. Clustering based Attribute Granulation

Attribute discernibility ability is the base to evaluate the similarity of every two attributes of AGAD algorithm.

Definition 4. For a knowledge base K= (U,S) , ∀P,Q∈IND(K) ,

$$r_P(Q) = k = \frac{|\ pos_P(Q)|}{|U|} = \frac{|\bigcup_{X \in U/Q} \underline{P}(X)|}{|U|}$$

$r_P(Q)$ is defined as the dependency of attribute Q to attribute P.

AGAP algorithm selects $r_P(Q)$ as the distinguish ability measurement method of attributes. Because this measure is easy to understand and calculation, and is suitable for various applications.

(1) Decision table

If Q is a decision attribute D and P is a condition attribute C, $r_C(D)$ means the dependency of attribute D to attribute C. Condition attributes have similar discernibility ability if they have similar $r_C(D)$.

It's important to realize that $r_C(D)$ reflect discernibility ability of single attribute, and it does not reflect attribute significance. The value of $r_C(D)$ could not indicate the degree of its discernibility ability. We can cluster attributes into several groups according their $r_C(D)$, and then, select representative attributes from each group to form an attribute set. In this way, a coarser attribute granule is received. Table 1 is a decision table.

TABLE I

A DECISION TABLE

U	a	b	c	d
1	2	2	0	1
2	1	2	0	0
3	1	2	0	1
4	0	0	0	0
5	1	0	1	0
6	2	0	1	1

The dependency of decision attribute to the condition attribute is calculated as below:

$\gamma_a(d) = 3/6 \quad \gamma_b(d) = 0$

$\gamma_c(d) = 0$

The dependency of decision attribute d to condition attribute a, b, c are 0.5, 0, 0. This means attribute b has the similar discernibility ability to c. So, the original attribute set is divided into two clusters: $\{a\},\{b,c\}$ if we use some clustering algorithm. For the second set $\{b,c\}$, representative attribute is produced after using some algorithm based on attribute significance, information entropy or other measures. Here we choose a simple method to select representative attribute. That is selecting any one of it to represent the attribute set. The result of attribute granulation is $\{a,b\}$ or $\{a,c\}$.

This method is supported by attribute reduction based on discernibility matrix. The discernibility matrix (table 2)could be obtained on the basis of table 1:

TABLE II

THE DISCERNIBILITY MATRIX BASED ON TABLE 1

	1	2	3	4	5	6
1						
2	a					
3	a					
4	a,b	a,b	a,b			
5	a,b,c	b,c	b,c			
6		a,b,c	a,b,c	a,c	a	

The discernibility function is :

$$f = a(a \vee b)(a \vee b \vee c)(a \vee b)(a \vee b \vee c)$$
$$(a \vee b)(b \vee c)(a \vee c)a$$
$$= ab \vee ac$$

So, the attribute reduction set is $\{a,b\}$ or $\{a,c\}$, and algorithm is verified.

(2) Information system

In information system, attribute clustering is performed according to attribute relative dependency. And, at sometime, we select attribute relative dependency to calculate discernibility ability can obtain a more precise answer.

The method has four steps:

a) Calculate attribute relative dependency relation matrix based on attribute relative dependency
b) Calculate the distance between every two attributes based on attribute relative dependency relation matrix
c) Cluster attribute set into several subsets
d) Select Representative attribute

Table 3 is an information system.

TABLEIII
AN INFORMATION SYSTEM

U	a	b	c	d
1	0	1	2	0
2	1	2	0	2
3	1	0	1	0
4	2	1	0	1
5	1	1	0	2

Attribute relative dependency are list below:

$\gamma_b(a) = 3/5$ $\gamma_c(a) = 0$
$\gamma_d(a) = 3/5$ $\gamma_a(b) = 2/5$
$\gamma_c(b) = 0$ $\gamma_d(b) = 1/5$
$\gamma_a(c) = 2/5$ $\gamma_b(c) = 2/5$
$\gamma_d(c) = 5/5$ $\gamma_a(d) = 4/5$
$\gamma_b(d) = 2/5$ $\gamma_c(d) = 4/5$

So, attribute relative dependency relation matrix is shown as table 4:

TABLEIV
ATTRIBUTE RELATIVE DEPENDENCY RELATION MATRIX

	a	b	c	d
a	1	3/5	0	3/5
b	2/5	1	0	1/5
c	2/5	2/5	1	1
d	4/5	2/5	4/5	1

According to this attribute relative dependency relation matrix, we could obtain the distance between attributes:

$$|ab| = \sqrt{(3/5)^2 + (2/5)^2 + (2/5)^2} = \sqrt{17}/5$$

$$|ac| = \sqrt{(1/5)^2 + (1/5)^2 + 1 + (2/5)^2} = \sqrt{31}/5$$

$$|ad| = \sqrt{(1/5)^2 + (1/5)^2 + (4/5)^2 + (2/5)^2} = \sqrt{22}/5$$

$$|bc| = \sqrt{(3/5)^2 + 1 + (4/5)^2} = \sqrt{50}/5$$

$$|bd| = \sqrt{(2/5)^2 + (3/5)^2 + (4/5)^2 + (4/5)^2} = \sqrt{45}/5$$

$$|cd| = \sqrt{(2/5)^2 + (1/5)^2} = \sqrt{5}/5$$

So, according clustering method, the original attribute set is divided into two clusters {a, c, d} and {b}. Here we select representative attribute by the method based on information entropy.

The information entropy of attribute a, c, d are -0.412, -0.412, and -0.46. Attribute d is selected to represent the attribute set $\{a,c,d\}$. The result of attribute granulation is $\{b,d\}$. It's the coarser attribute granule. The attribute reduction set is $\{a,b\}$ or $\{b,d\}$ through using discernibility matrix. So the algorithm is verified.

C. A Parallel Attribute Granulation Algorithm based on Attribute Discernibility and AP Clustering

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 [18-22].

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 AGAP algorithm, attribute subsets obtained through AP clustering method are input sets and they do not intersect and vice versa. Their output sets do not intersect either. So selecting representative attribute 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.

We select AP algorithm because it does not need initialing clustering centers and the number of clusters.

A parallel attribute granulation algorithm based on AP clustering has the following three steps:

（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

AGAP calculate similarity matrix elements according B of section Ⅳin order to get the similarity relation matrix S. In general, we often use attribute relative dependency as criterion of attribute discernibility ability. Experimental result shows that this method can improve accuracy of algorithm.

(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 to λ.

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

(3）Select representative attributes in parallel

Selecting representative attribute can be executed concurrently on each attribute set using attribute importance degrees , information entropy , or other method. Thus, a coarser attribute granularity is produced.

V. EXPERIMENTS

Firstly, Iris data set is used to verify our method. It has 150 objects, including 4condition attributes and a decision attribute. The dependency of attribute D to conditional attribute a_1, a_2, a_3, a_4 are0.2133, 0.1267, 0.7800, 0.6533. Using AP clustering method, they are divided into two clusters: {a_1, a_2}, { a_3 } , { a_4}. And after selecting representative attributes, a coarser attribute granularity is produced as {a_2, a_3, a_4}. But the AGAP is more efficient than traditional attribute reduction method. This can be seen in table Ⅵ.

Experiment used other three 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 5642 objects, including 22 condition attributes and a decision attribute. Table Ⅴ shows the characteristics of the four data sets:

TABLE Ⅴ.

CHARACTERISTICS OF FOUR DATA SETS IN UCI

Data set	Number of samples	Number of attributes
Iris	150	4
Glass Identification	213	10
Ionosphere	351	34
Mushroom	5642	22

The results of attribute granulation using attribute reduction(AR) and AGAP are compared as follows(table Ⅵ):

TABLE Ⅵ
THE COMPARISONS BETWEEN AR ALGORITHM AND AGAP ALGORITHM

Data set	Attribute granulation results		Runtime （s）	
	AR	AGAP	AR	AGAP
Iris	$\{a_2,a_3,a_4\}$	$\{a_2,a_3,a_4\}$	0.6875	0.2832
Glass Identification	$\{a_1,a_3,a_5,a_6,a_7\}$	$\{a_1,a_3,a_6,a_7\}$	8.0756	2.0312
Ionosphere	$\{a_{14},a_{16},a_{28}\}$	$\{a_1,a_4,a_{14},a_{16},a_{28}\}$	5.2116E+002	5.9639
Mushroom	$\{a_2,a_3,a_4,a_6,a_8,a_{10},a_{13},a_{14},a_{15},a_{16},a_{21},a_{22}\}$	$\{a_2,a_3,a_4,a_5,a_6,a_8,a_{10},a_{13},a_{14},a_{15},a_{21}\}$	7.2308E+003	139.3169

For Iris data set, AGAP and AR algorithm have the same result $\{a_2,a_3,a_4\}$, but traditional algorithm is lower than AGAP. For Glass Identification data set, the result of AGAP lost attribute a_5. For Ionosphere data set, the result of AGAP has some extra attributes a_1,a_4.. For Mushroom data set, the result of AGAP lost attribute a_{16} , a_{22} and has one more arribute a_5. But AGAP algorithm has obvious advantages in efficiency.

Experiments show that AGAP algorithm could not reach the accuracy of traditional attribute reduction method, but the efficiency of AGAP is much higher than that of traditional attribute reduction method. AGAP method is outperforming traditional attribute reduction algorithm for huge and high dimensional dataset processing when precision request of attribute granulation is not rigid.

VI. CONCLUSION

This paper provides a method of attribute granulation-Attribute granulation based on attribute discernibility and AP algorithm. The method divided the original attribute set into several clusters through AP algorithm according attribute discernibility. And then representative attributes are produced through some algorithms to form a coarser attribute granularity. Experiments show that the efficiency of AGAP is much higher than that of traditional attribute reduction method. But sometimes AGAP algorithm could not reach the accuracy of traditional attribute reduction method.

ACKNOWLEDGMENT

This work is supported by the National Key Basic Research Program of China (No. 2013CB329502), the National Natural Science Foundation of China (No.41074003), the Opening Foundation of the Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (IIP2010-1), and the Scientific Innovation Research of College Graduate in Jiangsu Province (No.CXZZ11_0296).

REFERENCES

[1] Liping Jing, Michael K. Ng, and Joshua Zhexue Huang, An entropy weighting k-Means algorithm for subspace clustering of high-dimensional sparse data, Transactions on knowledge and data engineering, 2007, 19(8): 1026-1041
[2] D. Donoho, High-dimensional data analysis: the curses and blessings of dimensionality, American Mathematical Society-Mathematical Challenges of the 21st Century, Los Angeles, CA, USA, 2000.
[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] Shifei Ding, Hong Zhu,Weikuan Jia,Chunyang Su. A survey on feature extraction for pattern recognition.Artificial Intelligence Review,2012, 37(3):169-180
[9] Yandong Zhai, Chunguang Zhou, Yingjuan Sun, et al. Approach of Rule Extracting Based on Attribute Significance and Decision Classification. Journal of Software, 2012,7(2):514-520
[10]Taorong Qiu, Yuyuan Li, Xiaoming Bai , et al. The Difference Degree of Condition Attributes and Its Application in the Reduction of Attributes. Journal of Software,2012,7(5): 1087-1091.
[11]Lin Sun, Jiucheng Xu, Zhan'ao Xue, et al. Decision Degree-based Decision Tree Technology for Rule Extraction. Journal of Software, 2012,7(7):1769-1779
[12] Chen Jie, Zhang Yingchun, Zhang Ling, et al, "Analysis and application of clustering based on information granularity," Journal of Image and Graphics, Vol.12, No.1, pp.87-91, 2007.
[13] B.J. Frey, D. Dueck, Clustering by passing messages between data points. Science, 2007, 315(5814): 972-976
[14] 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.
[15] Dueck, D, Frey, BJ, "Non-metric affinity propagation for unsupervised image categorization. Proceedings of the IEEE International Conference on Computer Vision (ICCV), October 2007.
[16] Yu Xiao, Jian Yu. Semi-supervised clustering based on affinity propagation algorithm. Journal of software, 19(11): 2803-2813, 2008.
[17] Jiawei Han, Micheline Kamber. Data Mining: Concepts and Techniques (Second Edition). Massachusetts: Morgan Kaufmann Publishers, 2006
[18] Intel Asia Pacific R & D Co., Ltd., Parallel Technology Co., Ltd. Beijing. The release of multi-core potential: Guide to the Intel Parallel Studio parallel development[M]. Beijing: Tsinghua University Press, 2010.
[19] Zhu Hong, Ding Shifei, Xu Xinzheng. An AP clustering algorithm of fine-grain parallelism based on improved attribute reduction. Computer Research and Development, 2012, 49(12):2638-2644.
[20]Hong Gongbing. Fine-grain parallelism and multithreaded computing[J]. Computer Research and Development, 1996, 33(6):473-480
[21] 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
[22] 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 is an associate professor at Xuzhou Medical College. Since 2009, she has been a Ph.D. candidate in computer science and technology from the China University of Mining and Technology, and her supervisor is Prof. Shifei Ding. She received her M.Sc. degree in computer science and technology from China University of Mining and Technology in 2007. Her research interests includes granule computing, clustering, parallel computing et al.
Email: zhuhongwin@126.com

Shifei Ding is a professor and Ph.D. supervisor at China University of Mining and Technology. 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. He received his B.Sc., M.Sc. degree in mathematics and computer science in 1987, 1998 respectively from Qufu Normal University, received Ph.D. degree in computer science from Shandong University of Science and Technology in 2004, and received post Ph.D. degree in computer science from Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, and Chinese Academy of Sciences in 2006.
In addition, 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 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) et al.
Email: dingsf@cumt.edu.cn; dingshifei@sina.com

Han Zhao is currently a graduate student now studying in School of Computer Science and Technology, China University of Mining and Technology, and his supervisor is Prof. Shifei Ding. He received her B.Sc. degree in computer science from China University of Mining and Technology in 2012. His research interests include cloud computing, feature selection, pattern recognition, machine learning et al.

Lina Bao is currently a graduate student now studying in School of Computer Science and Technology, China University of Mining and Technology, and her supervisor is Prof. Shifei Ding. She received her B.Sc. degree in computer science from China University of Mining and Technology in 2012. His research interests include cloud computing, feature selection, pattern recognition, machine learning et al.