

An Improved Heuristic Attribute Reduction Algorithm Based on Information Entropy in Rough Set

Yang Su-Min^{1,2,*}, Meng Jie¹, Zhang zheng-Bao² and Xie Zhi-Ying²

¹State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, Luoyang 471003, Henan, China; ²Department Of Information Engineering, Shijiazhuang Mechanical Engineering College, Shijiazhuang 050003, China

Abstract: At present, in rough set theory there are two kinds of heuristic attribute reduction algorithms, one is based on discernibility matrix, the other is based on mutual information. But if these algorithms are applied to the non-core information system, there will be much problems, such as too much calculation, excessive reduction, or insufficient reduction. So we propose an improved heuristic attribute reduction algorithm on the basis of rough set theory, in which the attribute importance is dependent on two factors, one is increment of mutual information, the other is information entropy. And we set the attribute with both the largest attribute importance and mutual information among all attributes as the core attribute, by which we solve the problem that causes the computational complexity increasing because of selecting the initial attribute randomly. By the proposed algorithm we can not only improve the efficiency of attribute reduction, but decrease the number of attribute reduction. The validity of the proposed algorithm is verified by two ways of the theoretic analysis and the simulation experiments.

Keywords: Attribute reduction, mutual information, non-core information system, rough set.

1. INTRODUCTION

Rough set theory was first proposed by Poland mathematician Z. Pawlak in 1982, which is one important mathematic tool to describe imperfection and uncertainty. With it we can analyze and dispose incomplete, inconsistent, imprecise, or other deficient information in order to find out the latent knowledge rules. Thus, it is widely applied to do with the information under uncertain circumstance, such as machine learning, decision analysis, process control, pattern recognition and data mining etc. The essence of the rough set theory lies in the measurement to both the attribute importance and attribute reduction. The most significant difference compared with other uncertainty and imprecise theory is that it can objectively describe or deal with uncertainty problem without providing any priori information of the processing data, such as the probability distribution of the statistics, membership function of the fuzzy theory.

Attribute reduction takes a crucial role in analyzing the attribute importance of different attributes in given data, whose purpose is to delete the unnecessary attribute under the condition of ensuring unchanged classification ability of information system. However, for an information system, Wong S. K. M and Ziarko.W [1] have proved that the attribute reduction is a non-deterministic polynomial problem, so most of researchers made use of the heuristic reduction

algorithm so as to obtain the optimal or sub-optimal attribute reduction. But the researchers find that effective attribute reduction can be obtained if we can establish the relationship between knowledge and information, and they studied attribute reduction with the point of view of information entropy. Articles [2-7] present some improved attribute reduction algorithm based on discernibility matrices, which have lower computational complexity and storing capacity. On the basis of conditional information entropy, articles [8, 9] study the computation of a core and attribute reduction in distributed environment. The article [10] analyzed the relationship between attribute reduction and conditional information quantity and gave one new conditional information quantity which cut down the number of attributes and time complexity. The article [11] presented a new reduced definition which integrates the complete and incomplete information systems into the corresponding reduction algorithm. Articles [12-15] studied the incomplete information systems. The article [16] proposed one knowledge reduction algorithm on the basis of the mutual information between the conditional attributes and decision attributes; The article [17] proposed one attribute reduction algorithm based on mutual information gain. Articles [18-22] proposed one attribute reduction algorithm in rough set based on mutual information, which could apply heuristic method to reduce the search space, and could shorten the search time as far as possible, and could finally get an optimal or approximate optimal solution. But the attribute reduction algorithm based on mutual information is the bottom-up approach, whose starting point is from the relative core attribute of decision table, then the most important attributes selected from the other attributes are added to the relative core, and the computing processing is

ended when mutual information of the reduction sets is equal with that of the conditional attribute set. In actual information system, there will be a lack of core attribute. When the core attribute are empty, we must compute mutual information when we add one attribute into reduction attributes every time, the computational complexity increases significantly. Therefore we propose one improved attribute reduction algorithm with using the information entropy and mutual information increment, and set the attribute with the largest attribute importance and mutual information among all attributes as the core attribute, by the means we solve the problem that causes the computational complexity increasing greatly because of selecting the initial attribute randomly, and attribute reduction velocity could be faster than some other algorithms, and the number of attribute reduction is relatively small. The simulation experiments verify that the proposed algorithm can not only ensure the efficiency and accuracy of attribute reduction, but also can guarantee the quality of reduction attribute.

2. BASIC CONCEPTS OF ROUGH SET THEORY

Definition I: $S = (U, A, V, f)$ is set to one information system. Among them, $U = \{U_1, U_2, \dots, U_n\}$ is called as one domain space which is a non- empty finite set. $A = \{a_1, a_2, \dots, a_m\}$ is called as the attribute set, and it is one non-empty finite attribute set too, $V = \cup_{a \in A} V_a$, V_a is attribute's value domain set, f is an information mapping function, which is defined as $f : U \times A \rightarrow V$. When $\forall a \in A$ and $\forall x \in U$, $f(x, a)$ has unique value in the range of V_a . On the other side, $S = (U, A, V, f)$ is called as one decision knowledge system, where $A = C \cup D, C \cap D = \phi$, among them C is one condition attribute set, which is consist of $c_1(x), c_2(x), \dots, c_n(x)$, and D is one decision attribute set, which is consist of $d_1(x), d_2(x), \dots, d_n(x)$.

Definition II: For the given knowledge information system $S = (U, A, V, f)$, when there is $B \subseteq A$ and $B \neq \phi$, the in-discernable relationship between X and Y is defined as following:

$$IND(B) = \left\{ (x, y) \in U \times U : \left(\begin{array}{l} f(x, q) = f(y, q), \\ \forall q \in B \end{array} \right) \right\} \quad (1)$$

Definition III: For the given knowledge representation system $S = (U, A, V, f)$,

$f_a(x_i)$ is the attribute value of a for record x_i , c_{ij} is the i th row and j th column factor of discernibility matrix, the discernibility matrix is defined as follows:

$$c_{ij} = \begin{cases} a | (\alpha \in A) \wedge (f_\alpha(x_i) \neq f_\alpha(x_j)) & x_i \neq x_j \\ \phi & x_i = x_j \end{cases} \quad (2)$$

where $i = 1, 2, 3, \dots, |U|$ and $j = 1, 2, 3, \dots, |A|$.

Definition IV: For the given knowledge information system $S = (U, A, V, f)$,

when $P \subseteq A$ and $X \subseteq U, x \in U$, the positive domain P is defined as following:

$$POS_P(X) = \cup_{X \subseteq U / P} \frac{R(X)}{P}, \text{ where } \underline{R(X)} \text{ is lower approximatant.}$$

Definition V: U is one domain set, X and Y is two equivalent relation of domain U , $U / ind(X) = \{x_1, x_2, \dots, x_m\}$, $U / ind(Y) = \{y_1, y_2, \dots, y_n\}$, then the probability distribution of X and Y with regard to U is defined as follows:

$$\begin{aligned} [X : p] &= \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ p(x_1) & p(x_2) & \dots & p(x_m) \end{bmatrix}, \\ [Y : p] &= \begin{bmatrix} y_1 & y_2 & \dots & y_n \\ p(y_1) & p(y_2) & \dots & p(y_n) \end{bmatrix} \end{aligned} \quad (3)$$

Among them,

$$\begin{aligned} p(x_i) &= \frac{|x_i|}{U}, i = 1, 2, \dots, m; \\ p(y_j) &= \frac{|y_j|}{U}, j = 1, 2, \dots, n; \end{aligned}$$

the symbol $|E|$ is indicated as the sample number whose decision value is V_a .

Definition VI: In accordance with the information theory, the information entropy of knowledge X is defined

$$\text{as: } H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i), \text{ the conditional entropy } H(Y|X) \text{ about the knowledge } X \text{ relative to } Y \text{ is:}$$

$$H(Y|X) = - \sum_{i=1}^n p(x_i) \sum_{j=1}^m p(y_j|x_i) \log p(y_j|x_i) \quad (4)$$

The mutual information $I(X; Y)$ of the knowledge X relative to Y is:

$$I(X; Y) = H(Y) - H(Y|X) \quad (5)$$

Definition VII: U is one domain set, X and Y is two equivalent relation of domain U , if $Ind(X) = Ind(Y)$, then $H(X) = H(Y)$

Definition VIII: The independent necessary and sufficient conditions of equivalent relation P of domain U is that there is $H(R|P - \{R\}) > 0$ for any $R \in P$.

Table 1. Non-core information system.

| U | c_1 | c_2 | c_3 | c_4 | c_5 | c_6 | c_7 | D |
|----------|-------|-------|-------|-------|-------|-------|-------|-----|
| S_1 | 2 | 2 | 3 | 1 | 1 | 4 | 1 | 1 |
| S_2 | 1 | 1 | 4 | 2 | 2 | 2 | 2 | 1 |
| S_3 | 2 | 4 | 2 | 2 | 2 | 4 | 2 | 2 |
| S_4 | 2 | 1 | 4 | 3 | 2 | 2 | 3 | 1 |
| S_5 | 4 | 4 | 1 | 3 | 2 | 4 | 3 | 3 |
| S_6 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 2 |
| S_7 | 1 | 1 | 4 | 3 | 2 | 1 | 3 | 1 |
| S_8 | 2 | 2 | 2 | 2 | 2 | 2 | 4 | 2 |
| S_9 | 3 | 3 | 2 | 4 | 4 | 3 | 4 | 1 |
| S_{10} | 3 | 3 | 2 | 2 | 2 | 2 | 4 | 2 |

3. DEFICIENCY ANALYSIS OF THE EXISTING METHODS

At present, there are two typical attribute reduction algorithms in the rough theory, one is discernibility matrix, the other is mutual information theory, but the above two methods are on the basis of the core attribute, if we apply these algorithms on the information system without core attribute can cause the result of excessive or deficient problem in the number of attribute reduction.

3.1. Deficiency of Attribute Reduction on the Basis of Discernibility Matrix

The core attribute is essential for the attribute reduction on the basis of discernibility matrix, when we obtain all possible attribute combinations including core attribute, we must judge whether the following two conditions are met:

- (1) $\forall c_{ij} \in M_{n \times n}$, when $c_{ij} \neq \phi$, there is $A \cap c_{ij} \neq \phi$;
- (2) A is independent.

where M , c_{ij} is the discernibility matrix and matrix element, A is the attribute reduction of sets. It is very complicated and difficult when we judge the above two conditions, so many researchers improved on it. For example, Hu [3] improved the discernibility matrix, and gave the following conclusions: when and only when c_{ij} is a single, the attribute belonged to the core attribute. But Zhao [5] indicated that Hu's algorithm has some deficiency that the attribute reduction set does not contain the core attribute in some condition, and he proposed an efficient attribute reduction algorithm based on discernibility matrix and the classical greedy strategy.

But if we apply the above two kinds of algorithms on the information system listed by Table 1, there are some problems. From Table 1, we can see that the information system has 7 condition attributes $C = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7\}$, 10 experts give those evaluation results $U = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10}\}$, decision attributes D is one set which is consist of 1, 2 and 3, the value of set C is set as $\{1,2,3,4\}$. By the Hu's algorithm, attribute reduction is obtained $\{c_6, c_7\}$, the division by attribute reduction sets of decision is $U / \{c_6, c_7\} = \{\{S_1\}, \{S_2\}, \{S_3\}, \{S_4\}, \{S_5\}, \{S_6\}, \{S_7\}, \{S_8, S_{10}\}, \{S_9\}\}$, the positive domain of attribute reduction sets for decision set is $POS_{\{c_6, c_7\}}(D) = \{S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}\} = POS_C(D)$. By the Zhao's algorithm, attribute reduction is obtained $\{c_3, c_6\}$ the division by attribute reduction sets of decision is $U / \{c_3, c_6\} = \{\{S_1\}, \{S_2, S_4\}, \{S_3\}, \{S_5\}, \{S_6\}, \{S_7\}, \{S_8, S_{10}\}, \{S_9\}\}$, the positive domain of attribute reduction sets for decision set is $POS_{\{c_3, c_6\}}(D) = \{S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}\} = POS_C(D)$. From the result, we can see the above two algorithms can guarantee that the division of attribute reduction for decision table is equivalent to the division of all condition attributes for the decision table, but the mutual information of the above two algorithms is $I_{hu}(R; D) = 3.1219$, $I_{zhao}(R; D) = 2.9219$ respectively, but $I(C; D) = 3.3219$. That is to say, the above two algorithms are not optimal reduction set, so the attribute reduction quality is not good.

3.2. Deficiency of Attribute Reduction on the Basis of Mutual Information

In the process of decision, our aim is to select those condition attributes which affect the decision largely. We consider the mutual information between condition attribute and decision attribute. The article[18] propose the method that obtained the attribute importance by looking over the increment of mutual information when one attribute was added. It is defined as follows:

$$\begin{aligned} SGF(a, R, Q) &= I(Q; R \cup \{a\}) - I(Q; R) \\ &= H(Q|R) - H(Q|R \cup \{a\}) \end{aligned} \quad (6)$$

Based on the above formula (6), the chosen attributes are that there is more quantity in the domain, but from the information theory, it is to select the one which is chaotic, but the selected attributes are not maybe useful for the decision.

In view of the above problems, the article [20] has made the improvement on the importance of attributes, which is defined as follows:

$$\begin{aligned} SGF(a, R, Q) &= (I(Q|R \cup \{a\}) - I(Q|R)) / H(a) \\ &= (H(Q|R) - H(Q|R \cup \{a\})) / H(a) \end{aligned} \quad (7)$$

The improved method not only considers the increment of mutual information after adding the attribute, but also considers its own information entropy. When the mutual information increment is equal, the smaller $H(a)$ is, the higher attribute importance degree is. But when there is not core attribute, the time complexity is $O(|C||U|^2 + O(|U|^3))$ in the worst case, where $|C|$ is the number of condition attributes, $|U|$ is the samples number of decision domain. But according to the attribute reduction algorithm proposed by [20], the attribute reduction $\{c_1, c_2, c_3, c_4, c_5, c_6, c_7\}$ is obtained without considering core attribute, the number of attribute reduction is unchanged.

4. IMPROVED ATTRIBUTE REDUCTION ALGORITHM

We can see from the analysis in section second, when there is not core attribute, formula (7) is becoming into formula (8):

$$\begin{aligned} SGF(a, R, D) &= (I(D) - I(D|a)) / H(a) \\ &= I(a, D) / H(a) \end{aligned} \quad (8)$$

In order to solve the problem without core attribute, we improve the formula (7), firstly the attribute importance of each condition attribute is calculated by formula (8), and we set the attribute with the maximal attribute important degree as the core attribute

the attribute important degree formula

becomes as follows:

$$\begin{aligned} SGF(a, R^*, Q) &= (I(Q|R^* \cup \{a\}) - I(Q|R^*)) / H(Q|a) \\ &= (H(Q|R^*) - H(Q|R^* \cup \{a\})) / H(a) \end{aligned} \quad (9)$$

On the basis of the formula (9), the author proposes one improved attribute reduction algorithm in this paper. Firstly the attribute importance of each condition attribute is calculated by formula (9), and the attribute with the maximal attribute importance is selected as the core attribute, We evaluate the attribute important degree based on the increment of mutual information and information entropy, and add the attribute that both the increment of mutual information and attribute important degree are the biggest into the reduction set. The specific algorithm description is as follows:

The Input: A compatible system decision table $S = (U, A, V, f)$, where $A = C \cup D, C \cap D = \phi$, among sequence C is one condition attribute set, and D is one decision attribute set, U is the domain set, f is an information function.

The Output: One attribute reduction sets;

(1) The mutual information $I(C; D)$ is calculated between condition attribute C and decision attribute set D ;

(2) All the attribute importance is calculated by the formula (9), and set the attribute with maximal attribute importance degree for the core attribute R^* ;

(3) Let $R = R^*$, the above proceed is performed on the attribute set $R' = C - R, C' = C - R$ as follows:

① For each attribute $c_i \in C$, we calculate $(I(Q|R \cup \{c_i\}) - I(Q|R)) / H(a)$, and select the one that has maximum value C_i , if there is the same value for multiple attributes, we choose one which comes the earliest, then $R = R \cup \{c_i\}$, $C' = C - R$.

② Then we judge whether $I(C; D)$ and $I(R; D)$ is equal, if they are the same, then the next step goes to (4), otherwise goes to ①.

(4) R is a reduction result, and we output it.

5. SIMULATION EXPERIMENT

In order to verify the validity of the proposed algorithm, we validate our algorithm by one non-core information system whose sample is shown in Table 1, the result of the indiscernible relationship sets for attribute D is as follows:

$$\begin{aligned} IND(D) &= \{\{s_1, s_2, s_4, s_7, s_9\}, \{s_3, s_6, \\ & s_8, s_{10}\}, \{s_5\}\} \end{aligned}$$

The below is the detailed process according to the algorithm of section 4:

(1) According to formula (5), the mutual information

Table 2. The attribute mutual information and the attribute important degree.

| | c_1 | c_2 | c_3 | c_4 | c_5 | c_6 | c_7 |
|------------------------------|------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Mutual information increment | 1.07548875 | 1.295461844 | 1.321928095 | 0.797416845 | 0.342609804 | 0.770950594 | 0.770950594 |
| Attribute important degree | 0.6107 | 0.6573 | 0.7159 | 0.4631 | 0.2525 | 0.4175 | 0.4175 |

Table 3. The attribute mutual information and the attribute important degree.

| | $\{c_1, c_3\}$ | $\{c_2, c_3\}$ | $\{c_4, c_3\}$ | $\{c_5, c_3\}$ | $\{c_6, c_3\}$ | $\{c_7, c_3\}$ |
|------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Mutual information increment | 2.3219 | 1.9142 | 2.6464 | 1.7066 | 2.9219 | 2.2879 |
| Attribute important degree | 0.9207 | 0.9054 | 1 | 1 | 1 | 0.8645 |

Table 4. The attribute mutual information and the attribute important degree.

| | $\{c_1, c_3, c_6\}$ | $\{c_2, c_3, c_6\}$ | $\{c_4, c_3, c_6\}$ | $\{c_5, c_3, c_6\}$ | $\{c_7, c_3, c_6\}$ |
|------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Mutual information increment | 3.3219 | 3.0219 | 3.1219 | 2.9219 | 3.1219 |
| Attribute important degree | 1 | 0.968 | 1 | 1 | 1 |

$I(C; D)$ is calculated between condition attribute C and decision attribute set D ;

$$I(C; D) = H(D) - H(D|C) = 3.3219;$$

- (2) According to the formula (9), we calculate the attribute important degree of each attribute in C sets, the results is shown in the Table 2, from the table we can see that the attribute important degree of c_3 is the maximal, so we set c_3 for core attribute, i.e. $R^* = \{c_3\}$;
- (3) We Set $R = R^*$, and perform the following operations on the attribute $R' = C - R, C' = C - R$;
- (4) Then we calculate the attribute important degree of the $\{c_i, c_3\}$ sets, the results is shown in Table 3, it can be seen from the Table 3, c_4, c_5 and c_6 have the same attribute important degree, but c_6 has the maximal mutual information increment, so we add c_6 into the attributes reduction sets;
- (5) But $I(R; D) = 2.9219$, it is not equal to $I(C; D)$;
- (6) On the basis of the algorithm procedure of section 3, we must add other attribute into the reduction attribute sets, then we calculate the attribute important degree of the $\{c_i, c_3, c_6\}$, the results is shown in Table 4. It can be seen from the Table 4, $\{c_1, c_3, c_6\}$ has the maximal mutual information increment, so we add c_1 into the attributes reduction sets. Then $R = R \cup \{c_1\}, C' = C - R$;

- (7) We can see that $I(R; D) = 3.3219$, which is the same as $I(C; D)$, so we can terminate this algorithm. $R = \{c_1, c_3, c_6\}$ is one reduction attribute for the non-core information system.

But according to the attribute reduction algorithm proposed by [20], the reduction set is the same as original condition set that is consequence without the precondition of non-core attribute. From the simulation experiment results, we can see that the proposed algorithm can reduce the number of attribute reduction 57%, and decrease attribute reduction time about 42% while maintaining the information system classification ability unchanged. The division by attribute reduction sets of decision is $U / \{c_1, c_3, c_6\} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}, \{10\}\}$, the positive domain of attribute reduction sets for decision set is $POS_{\{c_1, c_3, c_6\}}(D) = \{1,2,3,4,5,6,7,8,9,10\} = POS_C(D)$, and $I(R; D) = I(C, D)$. The experiment results validate the proposed algorithm has good attribute reduction quality and high efficiency for non-core information system.

CONCLUSION

According to the characteristics of non-core information system, we propose an improved heuristic algorithm of attribute reduction based on rough set theory in this paper. Firstly the attribute with the maximal attribute important degree is assumed to be the initial core attribute. Secondly while adding new attribute at one time, we select the one with both the maximal mutual information and the largest attribute important degree, so we can ensure that the added attribute is surely the most influence on decision attribute. By the proposed algorithm we can overcome these problems,

such as too much computational complexity, excessive reduction, or insufficient reduction. The simulation experiment results not only verify the correctness and validity of the proposed algorithm, but also prove its validity in the aspect of attribute reduction quality and attribute efficiency.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This research is supported by State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System under Grants No. CE-MEE2015K0304B.

REFERENCES

- [1] S. K. M. Wong, and W. Ziarko, "On optional decision rules in decision tables", *Bulletin of Polish Academy of Sciences*, vol. 33, pp. 693-696, 1985.
- [2] A. Skowron, and C. Rauszer, "The discernibility matrices and functions in information systems", In: *Slowinski R(Eds.): Intelligent Decision Support-Handbook of Applications and Advances of the Rough Sets Theory*, Kluwer Academic Publishers, London, pp. 331-362, 1992.
- [3] X. H. Hu, and N. Cercone, "Learning in relational databases: A rough set approach", *Computational Intelligence: An International Journal*, vol. 11, pp. 323-338, 1995.
- [4] M. Yang, and Z. Sun, "Improvement of Discernibility Matrix r the Computation of a core", *Journal of Fudan university (Natural science)*, vol. 43, pp. 865-868, 2004.
- [5] Z. Zhao, "An Improvement Method of A Attribute Reduction", *Science & technology*, vol. 32, pp. 179-181, 2006.
- [6] J. Cheng, J. Zhu, and F. Zhang, "An updated algorithm for attribute reduction based on discernibility matrix", *Journal of Hunan University (Natural Sciences)*, vol. 36, pp. 86-88, 2009.
- [7] Y. Jiang, X. Wang, and Z. Ye, "Attribute reduction algorithm of rough sets based on discernibility Matrix," *Journal of System Simulation*, vol. 20, pp. 3718-3721, 2008.
- [8] Z. Chen, and Q. H. Zhang, "Study on attribute reduction algorithm based on fuzzy rough set", Huaqiao University College of Computer Science and Technology, 2006.
- [9] M. Yang, and P. Yang, "Approximate reduction based on conditional information entropy over vertically partitioned multi-decision table", *Control and Decision*, vol. 23, pp. 1103-1107, 2008.
- [10] W. Zhang, and W. X. Zhangyan, "An attribute reduction algorithm combining probability heuristic", *Information And Knowledge Granularity, Computer Applications and Software*, vol. 30, pp. 43-46, 2013.
- [11] J. Qian, F. Ye, X. P. Meng, and D. Liu, "Attribute reduction algorithm based on new conditional information quantity", *Systems Engineering and Electronics*, vol. 12, pp. 2154-2157, 2007.
- [12] S. Teng, S. Zhou, J. Sun, and Z. Li, "Attribute reduction algorithm based on conditional entropy under incomplete information system", *Journal of National University of Defense Technology*, vol. 32, pp. 90-94, 2010.
- [13] J. Y. Liang, and Z. B. Xu, "The algorithm on knowledge reduction in incomplete information systems", *international Journal of Uncertainty, Fuzziness and Knowledge based Systems*, vol. 10, pp. 95-103, 2002.
- [14] L. Yee, J.M. Ma, W.X. Zhang and T. J. Li, "Dependence-space-based attribute reduction in inconsistent decision information systems", *International Journal of Approximate Reasoning*, vol. 3, pp. 623-630, 2008.
- [15] X. B. Pei, and Y. Z. Wang, "Novel approaches to knowledge reduction in inconsistent decision systems", In: *Natural Language Processing and knowledge engineering. IEEE NLP-KE'05 Proceedings of 2005 IEEE International Conference*, pp. 746-749, 2005.
- [16] D. Miao, and S. Fan, "The calculation of knowledge granulation and its application", *System engineering theory & practice*, vol. 1, pp. 48-56, 2002.
- [17] P. Jia, J. Dai, Y. Pan, and M. Zhu, "Novel algorithm for attribute reduction based on mutual-information gain ratio," *Journal of Zhe Jiang University (Engineering Science)*, vol. 40, pp. 1041-1044, 2006.
- [18] G. K. Wang, "Rough theory and knowledge acquisition", Xi'an jiao tong university press, 2001.
- [19] H. Wang, B. Yao, and H. Hu, "The method of ascertaining weight based on rough sets theory", *Computer Engineering And Application*, vol. 36, pp. 20-21, 2006.
- [20] Z. Tan, Z. Xu, and S. Wang, "An improved ascertaining weight based on rough sets theory", *Computer Engineering And Application*, vol. 48, pp. 115-118, 2012.
- [21] Y. H. Yanyan, "Knowledge reduction algorithm based on mutual information", *Tsinghua Science And Technology*, vol. 47, pp. 1903-1906, 2007.
- [22] X. Bao, and C. Liu, "A new method of ascertaining attribute weight based on rough sets theory", *Journal of Management*, vol. 6, pp. 729-732, 2009.

Received: June 10, 2015

Revised: July 29, 2015

Accepted: August 15, 2015

© Su-Min et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.