View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Water Science and Engineering, Sep. 2008, Vol. 1, No. 3, 73-82
ISSN 1674-2370, *http://kkb.hhu.edu.cn*, e-mail: *wse@hhu.edu.cn*

An information-based rough set approach to critical engineering factor identification

Jing Ji*[1, 2], Zheng Dongjian[1, 2]

1. College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, P. R. China
2. National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, Hohai University, Nanjing 210098, P. R. China

Abstract: In order to analyze the main critical engineering factors, an information-based rough set approach that considers conditional information entropy as a measurement of information has been developed. An algorithm for continuous attribute discretization based on conditional information entropy and an algorithm for rule extraction considering the supports of rules are proposed. The initial decision system is established by collecting enough monitoring data. Then, the continuous attributes are discretized, and the condition attributes are reduced. Finally, the rules that indicate the action law of the main factors are extracted and the results are explained. By applying this approach to a crack in an arch gravity dam, it can be concluded that the water level and the temperature are the main factors affecting the crack opening, and there is a negative correlation between the crack opening and the temperature. This conclusion corresponds with the observation that cracks in most concrete dams are influenced mainly by water level and temperature, and the influence of temperature is more evident.

Key words: *rough set*; *information entropy*; *critical engineering factor*; *crack opening*

DOI: 10.3882/j.issn.1674-2370.2008.03.008

1 Introduction

The safety of a critical engineering structure is influenced by various factors that cannot be summarized in brief, so it is necessary to develop an effective method to analyze and evaluate the effects of these factors. Cracking is one of the common defects of concrete structures. The existence and extension of cracks weaken the bearing capacity of concrete structures. With the analysis of influential factors of cracking, appropriate measures can be taken to control the formation and extension of cracks. Li (2003), Li and Ye (2006), Yu (2006), and Han (2007) have used the rough set approach to analyze the main factors affecting a crack in a concrete dam and gotten reasonable results. The rough set approach is an effective tool for dealing with imprecise, incompatible and incomplete data, and has been widely used in many fields. However, the conventional rough set approach is based on an algebraic perspective; it does not make full use of the information of the boundary domain. In this study, we used the rough set approach with an information perspective to analyze the main factors affecting crack opening, and improved the discretization and rule extraction algorithms.

This work was supported by the National Natural Science Foundation of China (Grant No. 50539010 and 50579010) and the Ertan Hydropower Development Company Joint Research Project (Grant No. 50539030-1-3).
*Corresponding author (e-mail: *iamj@163.com*)
Received May 21, 2008; accepted Jul. 22, 2008

2 Overview of the rough set approach

The rough set approach (Pawlak 1982, 1985) was introduced by Pawlak in 1982. It has been proven to be an effective tool for dealing with imprecise, incompatible and incomplete data. Its main function is to extract decisions or classification rules from a decision system through the simplification of knowledge.

The subjects to which rough set approach is applied can be presented in the form of a decision table. Let *S* be a decision table: $S = (U, A, V, f)$, where *U* is a certain set called a universe; *A* is a set of attributes usually divided into subsets *C* and *D*, which denote the set of condition attributes and the set of decision attributes, respectively; *V* is the set of value domains of attributes; and *f* is an information function that assigns values to attributes. Some basic concepts and algorithms can be found in related references (Pawlak 1982; Pawlak 1985; Wang 2001; Li 2003; Li and Ye 2006; Yu 2006; Han 2007).

The algorithm presented in the references (Li 2003; Li and Ye 2006; Yu 2006; Han 2007) is based on an algebraic perspective. This algorithm determines whether an attribute can be removed according to whether the positive domain (the set of objects that can be classified precisely) changes during attribute reduction. Because it does not consider whether the distribution of objects of the boundary domain (the set of objects that cannot be classified precisely) changes, it does not make full use of the information of the boundary domain. In this paper, the rough set approach was developed to include the information perspective, using the conditional information entropy to measure the uncertainty of the rough set.

3 Attribute reduction based on conditional information entropy

From the information perspective, the uncertainty of the rough set can be measured by the conditional information entropy. Let $X = \{X_1, X_2, \cdots, X_n\}$ and $Y = \{Y_1, Y_2, \cdots, Y_m\}$, and consider *X* and *Y* the partitions of *U* according to condition attributes *C* and decision attributes *D*, respectively. The conditional information entropy of decision attributes *D* under the condition attributes *C* is defined as

$$H(D \mid C) = -\sum_{i=1}^{n} \left\{ P(X_i) \sum_{j=1}^{m} P(Y_j \mid X_i) \lg \left[P(Y_j \mid X_i) \right] \right\} \tag{1}$$

where $P(X_i) = \dfrac{|X_i|}{|U|}$ is the probability distribution of partition *X*, $P(Y_j \mid X_i)$ is the probability distribution of partition *Y* under the condition of partition *X*, and $P(Y_j \mid X_i) = \dfrac{|Y_j \cap X_i|}{|X_i|}$ (Wang 2001; Wang et al. 2002; Wang 2003).

The conditional information entropy reflects the uncertainty of the classification of decision attributes by condition attributes. If $H(D \mid C) = 0$, then the classifications derived from condition attributes can be placed into the classifications derived from decision attributes precisely. The larger $H(D \mid C)$ is, the more uncertain the classification of decision attributes by condition attributes.

For a condition attribute $r \in C$, if

$$H(D|C) = H(D|C - \{r\}) \tag{2}$$

then attribute *r* is redundant (Wang 2001; Wang et al. 2002; Wang 2003). Removing *r* from C will not cause the information to vary.

For a subset $C' \subset C$, we say C' is a reduction of the set of condition attributes C with respect to the set of decision attributes D, if and only if $H(D|C) = H(D|C')$, and for any attribute $t \in C'$, $H(D|C') \neq H(D|C' - \{t\})$ (Wang 2001; Wang et al. 2002; Wang 2003).

This is attribute reduction defined from information entropy. Since the entropy of the positive domain is 0, the entropy is generated only by the boundary domain. If neither the positive domain nor the distribution of objects of the boundary domain changes during attribute reduction, then this attribute can be reduced. This rough set approach with an information perspective makes full use of information of the positive domain and the boundary domain.

4 A discretization algorithm based on conditional information entropy

When we deal with a decision table using the rough set approach, the attributes should have discrete values. If an attribute has a continuous value, it must be discretized.

Continuous attributes are discretized by the selection of a set of breakpoints to divide the continuous range, and the minimization of the number of breakpoints while maintaining the decision-making ability. Wang (2001) presented a discretization method based on the significance of attributes. It reduces breakpoints as much as possible while maintaining the degree of inconsistency of the decision table. In this paper, the idea was defined from the information perspective, using the conditional information entropy as a measurement of the uncertainty of the decision table, and a discretization algorithm based on conditional information entropy was proposed. Its main objective is to reduce breakpoints as much as possible, on the premise that the conditional information entropy remains unchanged. If a breakpoint is removed but the conditional information entropy remains unchanged, then the breakpoint is redundant. Otherwise, the breakpoint is significant and cannot be removed. The main calculation steps are as follows:

(1) The condition attributes are ranked from small to large according to their significance. If some attributes are of identical significance, they are ranked according to their number of breakpoints, from that with the most breakpoints to that with the fewest. The significance of a condition attribute is defined as

$$SGF(a, D) = H(D) - H(D|\{a\}) \tag{3}$$

where $SGF(a, D)$ is the significance of the condition attribute a. The larger $SGF(a, D)$ is, the more significant the condition attribute a is to the decision attributes D; and $H(D)$ is the information entropy of the decision attributes D, defined as

$$H(D) = -\sum_{j=1}^{m} P(Y_j)\lg\left[P(Y_j)\right] \tag{4}$$

where $P(Y_j) = \dfrac{|Y_j|}{|U|}$ is the probability distribution of partition Y.

(2) For each continuous condition attribute, the values of the condition attribute are ranked from small to large. Each value is scanned. If two adjacent values v_i and v_{i-1} are identical, the next value is scanned. Otherwise, a breakpoint $c_k = (v_i + v_{i-1})/2$ is set and then the next value is scanned. This generates the initial set of breakpoints.

(3) The significance of each of the initial set of breakpoints is considered. If the smaller of the two adjacent values of the breakpoint has been changed into the larger one and the conditional information entropy of the decision table remains the same, this breakpoint is removed. Otherwise, this breakpoint remains and the changed value is restored.

If all of the breakpoints for a condition attribute can be removed, it means that reducing this attribute will not change the conditional information entropy. Therefore, for the continuous condition attributes, the process of discretization is simultaneously the process of attribute reduction.

5 A rule extraction algorithm considering supports of rules

The process of rule extraction is also the process of the reduction of attribute values. Through the reduction of attribute values, concise rules can be extracted from the decision table. Many scholars have already proposed methods of doing this (Chang et al. 1999; Wang 2001; Yang et al. 2003), such as inductive value reduction, heuristic value reduction, value reduction based on a decision matrix and so on. When there are duplicate records in the decision table, the conventional way is to just remove the duplicate records. In fact, the duplicate records contain information on supports of rules. The support of a rule reflects the proportion of the decision system that the rule occupies. The more times a rule is duplicated, the larger the support and the more common the rule. This study proposes a new rule extraction algorithm considering supports of rules. This new algorithm allows the extracted rules to have as large a support as possible, so that they will be more common.

The support of the rule "if X_i, then Y_j" is defined as

$$\sup(X_i \to Y_j) = \frac{|X_i \cap Y_j|}{|U|} \tag{5}$$

The main calculation steps are as follows:

(1) Combine the duplicate records in the decision table and calculate the supports of rules.

(2) Consider each value in the decision table row by row and column by column. If a

value’s removal causes rules to conflict with one another, then this value cannot be removed. Otherwise, this value can be removed and marked with “?”.

(3) Consider each record in the decision table. If there are values marked with “?” in a record, then consider each of these values. If removing a value will cause rules to conflict with one another, then restore this value. Otherwise, remove this value and record the increase of support of the rule corresponding to this record. Find the value that, after removal, will cause the largest increase of support and mark it with “*”. Repeat this process until all of the values marked with “?” are modified and then move on to the next record (a mark of “*” means that this value has been removed).

(4) Delete the duplicate rows in the decision table after value reduction, and calculate the supports of rules.

6 Main steps of critical engineering factor identification using rough set approach

According to the analysis above, the main steps of critical engineering factor identification using the rough set approach are: first, to collect enough monitoring data, discretize continuous attributes and establish the decision system; second, to reduce condition attributes according to conditional information entropy and extract rules; and finally, to analyze the results in engineering practice. The recommended steps are described in detail below:

(1) Data collection: Sufficient monitoring data are collected and the initial decision system is established. The factors that are being analyzed are defined as decision attributes, and other factors that might affect the decision attributes are defined as condition attributes.

(2) Discretization: In practice, monitoring data such as crack openings, water level and temperature usually have continuous values. These continuous attributes must be discretized before attribute reduction.

(3) Attribute reduction: With the reduction of condition attributes, the redundant attributes are removed and the significant attributes are reserved. In this way, the factors that are significant can be found.

(4) Rule extraction: With the reduction of attribute values, rules can be extracted from the decision table. These rules indicate how the condition attributes affect the decision attributes.

(5) Analysis of results: The results of the rough set approach are analyzed and the main factors are identified and explained.

7 Case study

The rough set approach was applied to crack analysis of a concrete gravity arch dam with a crest elevation of 126.3 m, a maximum dam height of 76.3 m, a dam crest arc length of 419 m, and a body divided into 28 dam sections from left to right. The crack was at an elevation of

105 m on the downstream side of the dam. The measurement point of the crack opening from which the monitoring data were collected was on dam section 18.

(1) Data collection: Water level, temperature, rainfall, uplift pressure and crack opening of the latest five years from 2002 were monitored and collected (Table 1).

Table 1 Original data

Date	Water level (m)	Tempera-ture (℃)	Rain-fall (mm)	Uplift pressure (kPa)	Crack opening (mm)	Date	Water level (m)	Tempera-ture (℃)	Rain-fall (mm)	Uplift pressure (kPa)	Crack opening (mm)
2002-04-29	112.01	21.1	30	592.5	2.83	2005-02-14	107.76	5.9	18	579.1	3.93
2002-05-13	113.04	19.6	33	592.5	2.83	2005-03-14	107.67	4.7	0	584.1	3.73
2002-07-08	114.56	24.8	0	592.9	2.79	2005-04-11	106.12	11.1	11	577.4	2.97
2002-08-05	114.76	30.0	0	592.9	2.79	2005-05-09	107.43	20.1	0	578.4	2.81
2002-09-02	113.73	27.0	0	592.9	2.77	2005-06-06	107.39	24.6	0	579.3	2.75
2002-09-30	112.57	20.0	0	592.4	2.85	2005-07-04	105.85	31.8	0	579.2	2.75
2002-10-14	112.20	21.8	0	592.1	2.87	2005-08-01	107.32	31.4	0	578.5	2.79
2002-11-11	109.85	16.1	0	591.7	3.18	2005-08-29	107.02	26.2	0	577.3	2.67
2002-12-09	109.84	−0.3	0	590.7	3.67	2005-09-26	107.72	23.8	0	577.3	2.73
2003-01-06	111.23	1.1	0	590.4	4.04	2005-10-24	107.01	13.6	0	578.3	3.11
2003-02-03	110.78	4.9	0	590.7	3.79	2005-11-21	105.38	9.9	0	578.0	3.37
2003-03-03	111.53	7.5	2	590.6	3.39	2005-12-26	103.21	5.9	0	576.5	4.11
2003-03-31	113.20	20.1	36	591.0	2.89	2006-03-13	108.68	3.9	0	583.7	3.57
2003-04-28	112.81	18.8	9	591.3	2.81	2006-04-10	108.47	18.0	0	581.7	2.87
2003-05-26	110.70	21.3	0	593.2	2.81	2006-05-08	108.73	21.6	15	582.7	2.87
2003-06-23	107.35	26.3	3	592.9	2.71	2006-07-03	108.74	32.2	0	585.2	2.75
2003-07-21	110.81	29.1	0	592.1	2.75	2006-07-31	108.90	30.5	0	583.4	2.77
2003-08-18	108.73	25.3	0	572.5	2.77	2006-08-28	107.57	29.6	0	581.2	2.78
2003-10-13	106.25	12.4	0	571.8	3.04	2006-09-25	106.91	21.3	0	580.1	3.03
2003-11-10	104.81	7.2	1	572.8	3.27	2006-10-23	106.66	17.6	0	578.6	3.05
2003-12-08	104.48	6.5	20	572.5	3.62	2006-11-20	106.38	13.4	0	579.5	3.27
2004-01-26	105.07	1.1	0	575.8	4.07	2006-12-18	106.41	2.1	0	580.3	3.97
2004-02-16	104.98	11.5	0	575.0	3.56	2007-01-15	106.68	3.6	4	581.2	4.13
2004-03-15	105.15	12.5	0	575.4	3.14	2007-02-26	107.26	9.5	0	584.2	3.51
2004-05-10	106.68	21.8	0	575.4	2.69	2007-03-26	110.25	15.8	8	588.0	3.09
2004-07-05	113.50	28.5	0	580.3	2.73	2007-04-23	109.49	15.6	4	583.6	2.87
2004-08-30	111.46	24.8	0	579.1	2.79	2007-05-21	109.34	25.3	0	583.4	2.77
2004-10-25	107.45	16.8	0	578.5	3.06	2007-06-18	108.78	24.6	0	583.2	2.77
2004-11-22	105.89	11.1	0	576.4	3.37	2007-07-02	109.10	26.0	34	584.0	2.77
2005-01-17	105.06	2.1	0	571.8	4.12						

(2) Discretization: The water level, temperature, rainfall and uplift pressure were considered condition attributes, denoted by c_1, c_2, c_3 and c_4, respectively. The crack opening was considered a decision attribute, denoted by d.

The statistical method was used to discretize measured values of the crack opening, and breakpoints were set according to a specified probability: $p_k = p(x \leqslant x_k)$ (x_k is a breakpoint). In theory, the division of the decision attributes influences the division of the condition attributes; the more meticulously the decision attributes are divided, the more meticulously the condition attributes should be divided to maintain the explanatory capability of the decision attributes. In this example, a trial method was used to find a group of probability values to divide the decision attribute (crack opening) that minimizes the number of breakpoints of condition attributes. Groups of two probability values were calculated from $\{0.1, 0.2\}$ to $\{0.8, 0.9\}$, with an incremental probability increase of 0.1. The results are presented in Table 2.

Table 2 Trial results

Group of probability values	Number of breakpoints of condition attributes	Group of probability values	Number of breakpoints of condition attributes	Group of probability values	Number of breakpoints of condition attributes
{0.1, 0.2}	12	{0.2, 0.7}	24	{0.4, 0.8}	23
{0.1, 0.3}	9	{0.2, 0.8}	9	{0.4, 0.9}	18
{0.1, 0.4}	16	{0.2, 0.9}	12	{0.5, 0.6}	17
{0.1, 0.5}	17	{0.3, 0.4}	9	{0.5, 0.7}	23
{0.1, 0.6}	23	{0.3, 0.5}	11	{0.5, 0.8}	21
{0.1, 0.7}	27	{0.3, 0.6}	17	{0.5, 0.9}	20
{0.1, 0.8}	10	{0.3, 0.7}	21	{0.6, 0.7}	24
{0.1, 0.9}	19	{0.3, 0.8}	4	{0.6, 0.8}	23
{0.2, 0.3}	8	{0.3, 0.9}	7	{0.6, 0.9}	22
{0.2, 0.4}	13	{0.4, 0.5}	13	{0.7, 0.8}	22
{0.2, 0.5}	15	{0.4, 0.6}	19	{0.7, 0.9}	24
{0.2, 0.6}	21	{0.4, 0.7}	25	{0.8, 0.9}	6

The group of probability values $\{0.3, 0.8\}$ is the only one that reduces the number of breakpoints of condition attributes to the minimum of the four (Table 2). The probability values $\{0.3, 0.8\}$ were chosen to divide the decision attribute and the breakpoints of decision attribute $\{2.79, 3.565\}$ were obtained. That is to say, 30% of the values of decision attribute were less than 2.79 mm, 50% of the values of decision attribute were between 2.79 mm and 3.565 mm, and 20% of the values of decision attribute were greater than 3.565 mm. After the discretization of decision attribute, we used the discretization method based on conditional information entropy to discretize continuous condition attributes, and calculated the significance of each of the condition attributes: $SGF(c_1, d) = 0.435$, $SGF(c_2, d) = 0.445$, $SGF(c_3, d) = 0.103$, and $SGF(c_4, d) = 0.374$. Continuous condition attributes were discretized in the order c_3, c_4, c_1, and c_2, according to their significances. After the discretization, all the breakpoints of c_3 and c_4 were removed. This means that c_3 and c_4 can be removed without causing the variation of information. The sets of breakpoints of c_1 and c_2 were {112.105} and {6.85, 21.675, 24.7}, respectively. The discretization partitions are presented in Table 3. The decision table after discretization is presented in Table 4.

Table 3 Discretization partitions

Value of attributes after discretization	Attributes		
	c_1	c_2	d
0	≤112.105 m	≤6.85 ℃	≤2.79 mm
1	>112.105 m	6.85-21.675 ℃	2.79-3.565 mm
2		21.675-24.7 ℃	>3.565 mm
3		>24.7 ℃	

Table 4 Decision table after discretization

Samples	c_1	c_2	d	Samples	c_1	c_2	d
1	0	1	1	7	1	2	1
2	1	1	1	8	0	1	1
3	1	3	0	9	0	0	2
4	1	3	0	10	0	0	2
5	1	3	0	…	…	…	…
6	1	1	1				

(3) Attribute reduction: The attribute reduction algorithm based on conditional information entropy was used to remove redundant condition attributes. Wang et al. (2002) proposed two algorithms, CEBARKCC and CEBARKNC. Wang (2003) proposed an algorithm to calculate the core set of attributes. In this example, because all the condition attributes were continuous, all the redundant condition attributes were removed after the process of discretization. Here, c_3 and c_4 were removed, and c_1 and c_2 were reserved. The algorithm proposed by Wang (2003) was used to calculate the core set of the decision table after discretization (Table 4), and the result was $\{c_1, c_2\}$. This verified that the process of discretization is also the process of attribute reduction based on conditional information entropy.

(4) Rule extraction: The original rules, after the combination of duplicate records and the calculation of supports, are presented in Table 5. The new rule extraction algorithm proposed in this paper was used to extract rules that have large supports. The extracted rules are presented in Table 6.

Table 5 Original rules

c_1	c_2	d	Supports
0	0	2	0.203
0	1	1	0.373
0	2	0	0.068
0	3	0	0.203
1	1	1	0.068
1	2	1	0.017
1	3	0	0.068

Table 6 Extracted rules

c_1	c_2	d	Supports
*	0	2	0.203
*	1	1	0.441
0	2	0	0.068
*	3	0	0.271
1	2	1	0.017

(5) Analysis of results: The condition attributes c_1 and c_2 were reserved. This means that the crack opening is mainly affected by the temperature and the water level. The support of the second of the five extracted rules was 0.441, meaning that the second rule covered 44.1% of the samples of the decision table. The third rule and the fifth rule covered small percentages of the samples of the decision table. The first, the second and the forth rules covered 91.6% of the samples of the decision table, and decisions could be made only by attribute c_2. This means that the temperature is the most significant factor affecting the crack opening. The crack opening tends to be larger at lower temperatures. This conclusion is consistent with the actual variation law of crack opening. According to the analysis results of Li (2003), cracks in most concrete dams are influenced by

water level and temperature, and the influence of temperature is more evident.

It could be pointed out that, in practice, some factors are correlated with each other, including water level and uplift pressure. This means that some attributes have overlapping information from the information perspective, but since they are not perfectly correlated with each other, they also have their individual, unique information. A condition attribute's information can be divided into two parts: overlapping information and unique information. If a condition attribute is significant to the decision attributes, not only its overlapping information but also its unique information should have a significant effect on the decision attribute. If this attribute is removed, the conditional information entropy of the decision system will change, because the effect of its unique information cannot be replaced by other condition attributes. This means that the condition attribute is significant and should not be removed. Otherwise, if a condition attribute is not significant to the decision attribute, its unique information should have no significant effect on the decision attribute. If this attribute is removed, the conditional information entropy will not change, because even if its overlapping information has an effect on the decision attribute, the effect can be replaced by other condition attributes. This means that the attribute is redundant.

8 Conclusions

(1) In this study, the rough set approach with an information perspective was used to analyze the main critical engineering factors. This approach makes full use of information of the boundary domain.

(2) A discretization algorithm for continuous condition attributes based on conditional information entropy has been proposed. It makes full use of information of the boundary domain when discretizing the continuous attributes. For those continuous condition attributes, the process of discretization is also the process of attribute reduction based on conditional information entropy.

(3) A rule extraction algorithm considering supports of rules has been proposed. This new algorithm allows the extracted rules to have as large a support as possible, so they will be more common.

(4) By applying the rough set approach to the analysis of a crack in an arch gravity dam, it can be concluded that the water level and temperature are the main factors affecting the crack opening. Of the two of them, temperature is the more significant factor. There is a negative correlation between the crack opening and the temperature. This conclusion corresponds with the actual condition of the dam.

References

Chang, L. Y., Wang, G. Y., and Wu, Y. 1999. An approach for attribute reduction and rule generation based on rough set theory. *Journal of Software*, 10(11), 1206-1211. (in Chinese)

Han, S. D. 2007. *Study on the Safety Analysis and Early-warning Method for Dams with Cracks*. M.E. Dissertation. Nanjing: Hohai University. (in Chinese)

Li, X. H. 2003. *Evolution Rule and Abnormality Diagnosis of Crack of Important Hydraulic Concrete Structure*. Ph.D. Dissertation. Nanjing: Hohai University. (in Chinese)

Li, X. H., and Ye, Y. H. 2006. Approach of rough set to extracting cause of crack hydraulic concrete structure. *Journal of Southeast University (Natural Science Edition)*, S36, 145-149. (in Chinese)

Pawlak, Z. 1982. Rough sets. *International Journal of Computer and Information Sciences*, 11(5), 341-356.

Pawlak, Z. 1985. Rough sets and decision tables. *Lecture Notes in Computer Science*, 208, 187-196.

Wang, G. Y. 2001. *Rough Set Theory and Knowledge Acquisition*. Xi'an: Xi'an Jiaotong University Press. (in Chinese)

Wang, G. Y., Yu, H., and Yang, D. C. 2002. Decision table reduction based on conditional information entropy. *Chinese Journal of Computers*, 25(7), 759-766. (in Chinese)

Wang, G. Y. 2003. Calculation methods for core attributes of decision table. *Chinese Journal of Computers*, 26(5), 611-615. (in Chinese)

Yang, Z. F., Guo, J. F., and Chang, F. 2003. A value reduction algorithm based on rough set theory. *Computer Engineering*, 29(9), 96-97. (in Chinese)

Yu, P. 2006. *Study on the Method of Cause and Stabilization of Cracks in Concrete Dams*. M.E. Dissertation. Nanjing: Hohai University. (in Chinese)