Rough Control Rule Mining Model Based on Decision Interval Concept Lattice and Its Application

Ailing Sun, Chunying Zhang(✉), Liya Wang, and Zhijiang Wang(✉)

College of Science, North China University of Science and Technology, Tangshan, Hebei, China
salangelbaby@163.com, {zchunying,wangzj}@ncst.edu.cn, wang_liya@126.com

Abstract. Fusing the structure feature of interval concept lattice and the actual needs of rough control rules, we have constructed the decision interval concept lattice, further more, we also have built a rules mining model of rough control based on decision interval concept lattice, in order to achieve the optimality between rough control mining cost and control efficiency. Firstly, we have preprocessed the collected original data, so that we can transform it into Boolean formal context form, and then we have constructed the decision interval concept lattice in rough control; secondly, we have established the control rules mining algorithm based on decision interval concept lattice. By analyzing and judging redundant rules, we have formed the rough control association rule base in end. Analysis shows that under the premise of improving the reliability of rules, we have achieved the rough control optimization goal between cost and efficiency. Finally, the model of reservoir scheduling has verified its feasibility and efficiency.

Keywords: Rough control · Decision interval concept lattice · Attribute discretization · Decision rule mining

1 Introduction

Rough control is a new type method of intelligent rules control [1], rules extraction is the crucial link in intelligent control. The precision of rules directly affect the efficiency and accuracy.

Many researchers have adopted a lot of means to extract decision rules and a variety of algorithms have been developed. Dong et al. [2] presented the rough rule mining algorithm based on the theory of variable precision rough set. Huang [3] proposed the method of attribute reduction and rule extraction under the decision background. Rough control has achieved some success in industrial control applications [4, 5], however, when the variables in the control system is continuous, there are some problems, such as the establishment of a decision table, discretization of continuous variables, the consistency of rules and so on. Rough control always limited application, for its reason, there are some problems of rules in large number and low efficiency. In this paper, we put forward the decision interval concept lattice based on interval

W. Chen et al. (Eds.): BDTA 2015, CCIS 590, pp. 83–92, 2016.
DOI: 10.1007/978-981-10-0457-5_9

concept lattice [6]. Based on the feature that concept lattice extension must meet a certain amount of the intension property, and its intension is determined by conditional attributes and decision attributes, so we adopt the theory of decision interval concept lattice to mine the decision rules, which become more decisive than traditional rules. The method we put forward not only reduce the mining cost, but also improve the control efficiency.

In this paper, we have designed the decision rule mining model of rough control based on decision interval concept lattice. Firstly we construct the decision interval concept lattice, then mine the decision interval rule based on the decision interval concept lattice. The model we constructed has achieved the optimum between rules mining cost, efficiency and reliability. The rationality of model analysis is presented, further more its feasibility and efficiency are proved by an example.

2 Decision Interval Concept Lattice

2.1 Basic Concepts

Definition 2.1. Let $(U, C \times D, R)$ be a decision context. $RL(U, C \times D, R)$ is the decision interval concept lattice constructed by $(U, C \times D, R)$, and (M, N, Y) is a decision interval concept based on RL. Where C is the set of conditional attributes, D is the set of decision attributes. In the interval $[\alpha, \beta](0 \leq \alpha \leq \beta \leq 1)$, α upper bound extension M^{α} and β lower bound extension M^{β} are defined respectively by Eqs. (1) and (2)

$$M^{\alpha} = \{x | x \in M, |f(x) \cap Y| / |Y| \geq \alpha, 0 \leq \alpha \leq 1\} \tag{1}$$

$$M^{\beta} = \{x | x \in M, |f(x) \cap Y| / |Y| \geq \beta, 0 \leq \alpha \leq \beta \leq 1\} \tag{2}$$

Where Y is the intension of concept, among them, $Y = C' \cup D'$, $C' \subseteq C$, $D' \subseteq D$. If $Y \neq \phi$, then $D' \neq \phi$, C' is the set of child conditional attributes of C, and D' is the set of child decision attributes of D. $|Y|$ is the number of elements in Y, namely cardinal number. M^{α} is the set of objects that may be covered by at least $\alpha \times |Y|$ attributes in Y; M^{β} is the set of objects that may be covered by at least $\beta \times |Y|$ attributes in Y.

Definition 2.2. Let $(U, C \times D, R)$ be a decision context. The ternary ordered pair $(M^{\alpha}, M^{\beta}, Y)$ is called decision interval concept, where Y is intension, it contains conditional intension and decision intension, namely decision concept description; M^{α} is α upper bound extension and M^{β} is β lower bound extension.

Definition 2.3. Let $L_{\alpha}^{\beta}(U, C \times D, R)$ be the set of decision interval concepts getting from $(U, C \times D, R)$ in $[\alpha, \beta]$. If $(M_1^{\alpha}, M_1^{\beta}, Y_1) \leq (M_2^{\alpha}, M_2^{\beta}, Y_2) \Leftrightarrow C_1 \supseteq C_2, D_1 \supseteq D_2$, then "$\leq$" is called partial order relation on $L_{\alpha}^{\beta}(U, C \times D, R)$.

Definition 2.4. Let $L_{\alpha}^{\beta}(U, C \times D, R)$ be the set of decision interval concepts getting from $(U, C \times D, R)$ in $[\alpha, \beta]$. If all of the concepts in $L_{\alpha}^{\beta}(U, C \times D, R)$ meet the partial order relation "$\leq$", then $L_{\alpha}^{\beta}(U, C \times D, R)$ is called decision interval concept lattice.

Definition 2.5. Let $G_1 = \left(M_1^{\alpha}, M_1^{\beta}, Y_1\right)$ and $G_2 = \left(M_2^{\alpha}, M_2^{\beta}, Y_2\right)$ be two nodes in decision interval concept lattice, and they have the relationship $G_1 \leq G_2 \Leftrightarrow C_1 \supseteq C_2$. If there is no G_3, which meet $G_1 \leq G_3 \leq G_2$, then G_2 is a parent node (immediate predecessor) of G_1 and G_1 is a child node (immediate successor) of G_2.

2.2 Decision Interval Rule

Definition 2.6. Let $(U, C \times D, R)$ be a decision context. D is the set of decision attributes, U is the set of rule objects, and $C \times D$ is the set of rule projects. R describes the relationship between U and $C \times D$. For $A \subseteq C$ and $B \subseteq D$, then $A_{\alpha}^{\beta} \Rightarrow B$ is a decision interval rule getting from $(U, C \times D, R)$ in $[\alpha, \beta]$.

Definition 2.7. For the decision interval rule $A_{\alpha}^{\beta} \Rightarrow B$, if RO_{α} is the set of objects in B that meets the degree of α, RO_{β} is the set of objects in B that meets the degree of β, then the set of objects in A that meets the degree of $[\alpha, \beta]$ also meets the possible degree in B, which is defined the roughness of interval rules by (3)

$$\gamma = \rho(RO_{\beta}/RO_{\alpha}) = \left|RO_{\beta}\right|/\left|RO_{\alpha}\right| \tag{3}$$

$0 \leq \gamma \leq 1$. For decision interval rule, the lower roughness, the more accurate.

Definition 2.8. For the decision interval rule $r : A_{\alpha}^{\beta} \Rightarrow B$, If $\forall p \in A$, the weight of p based on B is defined by (4)

$$\delta(p, r) = |g(p)|/|\cup g(y)|_{\alpha}, (y \in A) \tag{4}$$

$|\cup g(y)|_{\alpha}$ is the number of objects in A that meets the degree of α.

2.3 Construction Algorithm for Decision Interval Concept Lattice

Algorithm: $(DICLCA)$ *Construction Algorithm for Decision Interval Concept Lattice*
INPUT: Decision Context $(U, C \times D, R)$
OUTPUT: Decision Interval Concept Lattice L_{α}^{β}
(1) Suppose α, β, determine the intension of decision interval concept, then get the initial node set G.
The intension of decision interval concept is determined by conditional attributes and decision attributes. If conditional attributes are $A = \{a_1, a_2, \ldots a_m\}$, $B = \{b_1, b_2, \ldots b_n\}$ and so on, and decision attributes is $D = \{d_1, d_2, \ldots d_l\}$, then for decision attribute d_2, the set of intension is $\left\{a_i b_j \ldots d_2\right\}$, $i = 0, 1 \ldots m, j = 0, 1 \ldots n$, $m, n \ldots$ can be 0 in a different time. $|\cdot|$ is the number of elements, then the number of intension is $\left|a_i b_j \ldots\right|$, $i = 0, 1 \ldots m$, $j = 0, 1 \ldots n$, $m, n \ldots$ can be 0 in a different time.

(2) Get the upper bound extension M_i^{α} and lower bound extension M_i^{β}.
(3) Construct the lattice. For the initial node set G, determine the layer and the parent-child relation according to the relationship between precursor and successor.

2.4 Mining Algorithm for Decision Interval Rule

Algorithm: $(DIRMA)$ *Mining Algorithm for Decision Interval Rule*
INPUT: Decision Interval Concept Lattice $L_{\alpha}^{\beta}(U, C \times D, R)$, parameters α, β
OUTPUT: Decision Interval Rule
(1) For parameters α, β, use the breadth-first traversal method to get the set of decision concept node, namely *Dcset*. The object "x" of every concept node in *Dcset* must meet the requirement of conditional attributes A in $[\alpha, \beta]$.
Let $L_{\alpha}^{\beta}(U, C \times D, R)$ be the set of decision interval concept lattice, and $A \subseteq C$, if $\exists y$, meets $y \in f(x)$, $y \in A$, and $\alpha \leq |y|/|A| \leq \beta$ $(0 \leq \alpha \leq \beta \leq 1)$, then "$x$" is called that it meets the conditional attributes A at the degree in $[\alpha, \beta]$.
(2) For each concept node in *Dcset*, mining rules $r : A_{\alpha}^{\beta} \Rightarrow B$, to form the set of decision Rules, namely *Diset*.
Let every intension of concept node contains conditional attributes and decision attributes, so one node can extract one rule, the former in rule is conditional attributes and the latter is decision attributes. Repeat the above steps, the set of decision Rules *Diset* be formed.
(3) For every rule in *Diset*, calculating its roughness and attribute contribution respectively, and then judging the accuracy and reliability based on its size of roughness and attribute contribution., so that we can extract the more optimal rules. That is, removing undesirable rules, the set of final decision rule be formed, namely *Disset*.

3 Mining Model of Decisions Interval Rule in Rough Control

3.1 Model Design

The simplest decision rules in rule base is used to realize the goal of rough control, which actually queries the same or similar condition attributes of decision rules in the rule base, to be applied to rough control. Using the mining algorithm of decision interval concept lattice can get a set of decision interval rules in rough control, the mining model is shown in Fig. 1.

The most prominent advantages of rough control rules mining model based on decision interval concept lattice is that it guarantees the reliability of the rules, at the same time, realizes the optimality between the scale of rule base and the mining cost. The optimality depends on the interval parameter Settings, because of interval parameter determines the structure of constructed decision interval concept lattice, and then affects the number of interval association rules and its accuracy. So adopting the decision rule extraction method to extract rules based on decision interval concept

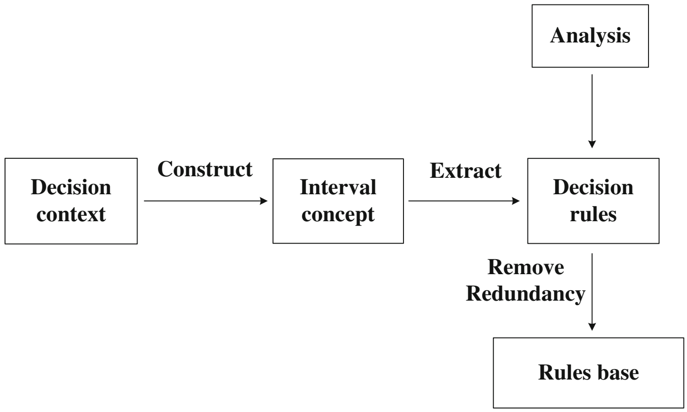

Fig. 1. Decision rule acquisition process

lattice, we can realize the optimization between mining cost and efficiency. Specific steps are as follows:

Step 1, Collect and preprocess the original data so as to get a decision table.

(1) According to the practical background in rough control, the observation and control quantity be considered as condition attributes and decision attributes respectively. Record the control strategy adopted by dispatch staff on representational state in rough control so as to form the original decision table;
(2) In most cases the data in industrial control is continuous. Using the method of decision interval concept lattice to mining control rules, we can discretize the continuous variables. With the aid of the background knowledge in industrial processes, we discrete continuous data, and mark each discrete interval with numbers;
(3) Using the method of rough set theory to reduce and combine data, after that we get the processed original decision table.

Step 2, Build the Boolean form of decision context table.

(1) According to the attributes condition showed by original decision table, which had got in above steps, we mark attributes and construct the set of attributes. The marking numbers are subscript of set elements. For example, an attribute is marked by A, the appeared marking number in the table are $2, 3, 4$, then the set of attribute is $A = \{a_2, a_3, a_4\}$;
(2) According to the corresponding attribute set, we build the Boolean form of decision context table.

Step 3, Suppose α, β, use the algorithm of Sect. 2.3 to construct the decision interval concept lattice, which is matched by practical background in rough control.

Step 4, Use the algorithm of Sect. 2.4 to extract decision interval rules in rough control.

Step 5, Calculate the roughness and attributes contribution. Taking various factors on roughness, attributes contribution, and actual cost into consideration. Removing the undesirable rules, ultimately we get the rough control rule base.

In the process of constructing the decision interval concept lattice, the intension is determined by condition attributes and decision attributes. Therefore, the extracted rules mean that for some different rules, taking same measures, we can get all results we expected. For this reason, it improves the control efficiency.

3.2 Model Analysis

The decision rule acquisition methods of decision interval concept lattice is divided into two parts: construct the decision interval concept lattice that meet the practical meaning in rough control and mining rough control decision rules. Among them, let the original data table be transformed into the Boolean context table is the key step. So we should discrete those continuous attributes. Such as a condition attribute marked by A, discrete it, it will become $a_1, a_2, \ldots, a_k$. Each object has attributes at most one. Therefore, the *DICLCA* is different from the reference [7], and it can form $|a_i b_j \ldots|$ decision interval concepts.

The decision interval concept extension contains a certain amount or proportion of objects set in intension, so the extracted rules become more targeted. Removing undesirable rules, consequently the reliability of the rules and control efficiency have been improved. In the process of traditional rough control rules extraction, in some respects, the establishment of decision table, the discretization of continuous, and the consistency of decision rules, its time complexity becomes more larger, and mining cost highly. The model guarantees the reliability of the rules, at the same time, improves the mining precision of rough control rule and control efficiency.

4 Case Study

Reservoir is a complex system. In this example, under the premise of the function is mainly electricity, the first hydropower station is a large hydropower station that power generation and flood play important roles simultaneously. An upstream hydropower station, that is, the second hydropower station have great influence on the first hydropower station. Mining decision rules of reservoir dispatching, specific steps are as follows:

(1) Collect and preprocess the original data

Analysis specific conditions of the first hydropower station. Considering observation about the main factors influencing the hydropower station scheduling as condition attribute set, the control quantity as decision attribute set. Discrete those continuous attribute. The specific process is as follows:

Condition attributes: a– water discharge by the second hydropower station
b– the natural runoff condition

Decision attributes: d– daily average electricity by hydropower station

(a) The range of discretized water discharge by the second hydropower station (m^3/s)

$$1 - [100 - 200) \quad 2 - [200 - 300) \quad 3 - [300 - 400) \quad 4 - [400 - 500)$$

(b) The range of discretized natural runoff condition

$$1{-}\text{dry} \quad 2{-}\text{moderate} \quad 3{-}\text{flood}$$

(d) The range of discretized daily average electricity (kWh/d)

$$1 - [250 - 300) \quad 2 - [300 - 350) \quad 3 - [350 - 400) \quad 4 - [400 - 450)$$

Record the measure that is taken by scheduling persons from January to June, merge the same decision, finally the original data table is formed. As is shown in Table 1.

Table 1. Original data table

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

(2) Construct the decision formal context table of reservoir scheduling. According to the data shown in Table 1, mark discretized attribute with numbers and build attribute set. It can get $A = \{a_2, a_3, a_4\}$ $B = \{b_1, b_2, b_3\}$ $D = \{d_3, d_3, d_4\}$. The decision form table is shown in Table 2.

Table 2. Decision form table

U	a_1	a_2	a_3	b_1	b_2	b_3	d_2	d_3
1	0	0	1	1	0	0	1	0
2	0	1	0	1	0	0	1	0
3	0	1	0	0	1	0	1	0
4	0	0	1	0	1	0	0	1
5	0	1	0	0	0	1	0	1
6	1	0	0	0	0	1	0	1

(3) Construct the decision interval concept lattice in reservoir scheduling.

For the decision context in Table 2, supposing α, β, and then get the concept intension, the upper bound M_i^{α} and lower bound extension M_i^{β}. It is concluded that

Table 3. Decision interval concept

Concept	M^{α}	M^{β}	Intension	Concept	M^{α}	M^{β}	Intension
F_1	ϕ	ϕ	ϕ	F_{17}	$\{456\}$	$\{6\}$	a_1d_3
F_2	$\{1236\}$	ϕ	a_1d_2	F_{18}	$\{23456\}$	$\{5\}$	a_2d_3
F_3	$\{1235\}$	$\{2\}$	a_2d_2	F_{19}	$\{1456\}$	$\{4\}$	a_3d_3
F_4	$\{6789\}$	$\{1\}$	a_3d_2	F_{20}	$\{12456\}$	ϕ	b_1d_3
F_5	$\{123\}$	$\{12\}$	b_1d_2	F_{21}	$\{3456\}$	$\{4\}$	b_2d_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
F_{13}	$\{235\}$	ϕ	$a_2b_3d_2$	F_{29}	$\{14\}$	ϕ	$a_3b_1d_3$
F_{14}	$\{12\}$	$\{1\}$	$a_3b_1d_2$	F_{30}	$\{4\}$	$\{4\}$	$a_3b_2d_3$
F_{15}	$\{134\}$	ϕ	$a_3b_2d_2$	F_{31}	$\{456\}$	ϕ	$a_3b_3d_3$
F_{16}	$\{1\}$	ϕ	$a_3b_3d_2$	F_0	ϕ	ϕ	Ω

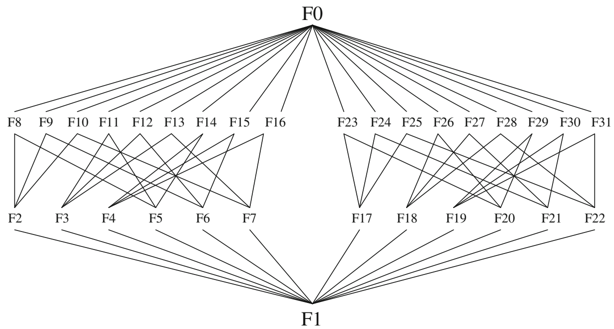

Fig. 2. Decision interval concept lattice based on reservoir scheduling

decision interval concept is shown in Table 3. The decision interval concept lattice is shown in Fig. 2

(4) Extract control rules in reservoir scheduling.

For parameters α, β, use the breadth-first traversal method to get the set of decision concept node. Because of the feature that cardinal number of condition attribute must meet $\alpha \leq |y|/|A| \leq \beta$, so we can get the set of decision concept node $Dcset = \{F_3\, F_4\, F_5\, F_6\, F_{11}\, F_{12}\, F_{14}\, F_{17}\, F_{18}\, F_{19}\, F_{21}\, F_{22}\, F_{25}\, F_{28}\, F_{30}\}$. Calculate its roughness and attribute contribution, and all the association rules of the roughness and attribute contribution as is shown in Table 4.

(5) The construction and optimization of rule base in reservoir scheduling

According to the roughness and contribution, we remove those undesirable rules. For example, comparing the rules $a_2 \Rightarrow d_3$ with $a_3 \Rightarrow d_3$, under the premise that they have the same attribute contribution, but the former roughness is lower than the latter, thus, the latter is a desirable rule. Repeat steps, the final set of rules *Disset* can be

Table 4. Decision interval rules

Rules	Roughness	Attribute contribution	Rules	Roughness	Attribute contribution
$a_2 \Rightarrow d_2$	25 %	$(a_2, \delta) = 100\,\%$	$a_2 \Rightarrow d_3$	20 %	$(a_2, \delta) = 100\,\%$
$a_2b_2 \Rightarrow d_2$	25 %	$(a_3, \delta) = 100\,\%$	$a_3 \Rightarrow d_3$	25 %	$(a_3, \delta) = 100\,\%$
$b_1 \Rightarrow d_2$	67 %	$(b_1, \delta) = 100\,\%$	$b_2 \Rightarrow d_3$	25 %	$(b_2, \delta) = 100\,\%$
$b_2 \Rightarrow d_2$	67 %	$(b_2, \delta) = 100\,\%$	$b_3 \Rightarrow d_3$	34 %	$(b_3, \delta) = 100\,\%$
$a_2b_1 \Rightarrow d_2$	34 %	$(a_2, \delta) = 75\,\%$	$a_1b_3 \Rightarrow d_3$	50 %	$(a_1, \delta) = 50\,\%$
		$(b_1, \delta) = 50\,\%$			$(b_3, \delta) = 100\,\%$
$a_2b_2 \Rightarrow d_2$	50 %	$(a_2, \delta) = 75\,\%$	$a_2b_2 \Rightarrow d_2$	50 %	$(a_2, \delta) = 75\,\%$
		$(b_2, \delta) = 50\,\%$			$(b_2, \delta) = 50\,\%$
$a_3b_1 \Rightarrow d_2$	50 %	$(a_3, \delta) = 67\,\%$	$a_3b_2 \Rightarrow d_3$	100 %	$(a_3, \delta) = 67\,\%$
		$(b_1, \delta) = 67\,\%$			$(b_2, \delta) = 67\,\%$
$a_1 \Rightarrow d_3$	34 %	$(a_1, \delta) = 100\,\%$			

obtained. Under the practical cost, remove redundant rules secondly. The final reservoir scheduling rule base in rough control is

$$\{a_2 \Rightarrow d_2, b_1 \Rightarrow d_2, a_2b_1 \Rightarrow d_2, a_3b_1 \Rightarrow d_2, a_1 \Rightarrow d_3, b_2 \Rightarrow d_3, a_1b_3 \Rightarrow d_3, a_2b_3 \Rightarrow d_3, a_3b_2 \Rightarrow d_3\}$$

5 Conclusions

In this paper, we put forward the rule mining model based on the theory of the decision interval concept lattice in rough control. From the initial data preprocessing to the decision interval concept lattice construction in rough control, and the decision control rules extraction, the model has carried on the detailed design. Model analysis summarizes the two main parts of the model, constructing decision interval concept lattice and decision rule mining. Compared with the traditional methods in rough control, model highlight the reliability and effectiveness. It is verified by a case that the model improves its feasibility.

Acknowledgements. This work is partially supported by the National Natural Science Foundation of China (Grant No. 61370168, 61472340), Conditional Construction Project of Hebei Province Technology Hall (Grant No. 14960112D). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

1. Aixian, P., Yun, G.: Problems in rough control. J. Ocean Univ. China **39**(6), 1315–1320 (2009)
2. Dong, W., Wang, J., Gu, S.: The rule acquisition algorithm based on theory of variable precision rough set. Comput. Eng. **14**(1), 73–75 (2007)

3. Huang, J.: Attribute reduction and rule acquisition based on the rough concept lattice. Software **32**(10), 16–23 (2011)
4. Peters, J.F., Skowron, A., Suraj, Z.: An application of rough set methods in control design. Fundamenta Informaticae **43**(1), 269–290 (2011)
5. Wang, H., Rong, Y., Wang, T.: Rough control for hot rolled laminar cooling. **In:** International Conference on Industrial Mechatronics and Automation (2010)
6. Liu, B., Zhang, C.: New concept lattice structure—interval concept lattice. Comput. Sci. **39**(8), 273–277 (2012)
7. Zhang, C., Wang, L.: The incremental generation algorithm of interval concept lattice based on attribute power set. Appl. Res. Comput. **31**(3), 731–734 (2014)