

A Rough Set Multi-Knowledge Extraction Algorithm and Its Formal Concept Analysis

Zhengqiong Zhu¹, Hui Li¹, Guangyao Dai¹, Ajith Abraham^{1,2,3}, Wanqing Yang⁴

¹School of Information, Dalian Maritime University, Dalian 116026, China

²Machine Intelligence Research Labs, Seattle, WA, USA

³IT4Innovations, VSB-Technical University of Ostrava, Czech Republic

⁴State Grid Dalian Electric Power Supply Company, Dalian, China

zhengqiong_zhu@126.com, lihui@dlmu.edu.cn, ajith.abraham@ieee.org

Abstract—Rough set theory provides an effective method to reduce attributes and extract knowledge. This paper represents a rough set multi-knowledge extraction algorithm and its formal concept analysis. The proposed algorithm can obtain multi-reducts by using rough set in decision table. The formal concept analysis is used to obtain rules from the main values of the attributes influencing the decision making and these rules build a multi-knowledge. Experimental results show that the proposed multi-knowledge extraction algorithm is efficient.

Keywords-Multi-Knowledge, Rough Set, Formal Concept Analysis, Rule

I. INTRODUCTION

Rough set theory [1], [2] is a mathematical tool for attribute reduction and has been widely applied to many fields, for instance, machine learning [3], knowledge discovery [4], [5], artificial intelligence [6], etc. The aim of attribute reduction is to remove redundant knowledge on premise of keeping the decision-making ability unchanged.

The traditional method for attribute reduction is to find a good reduct [7], [8], and a single knowledge [9] is represented by a good reduct. However, when a single knowledge classifies new objects, it may cause errors. Reducts of decision table are usually not unique in real world. So the concept of multi-knowledge [10], [11] from multi-reducts is proposed.

Formal concept analysis (FCA) [12], [13] is a powerful tool for a rule extraction. FCA can be used for conjunctive description of concepts. In FCA, extension of concept is represented as the set of all objects from this concept and intension is represented as the set of common attributes from the objects. The concept lattice [14] is the core data structure of FCA. Each node of concept lattice is a concept, which is composed of extension and intension. A concept lattice is very beneficial for obtaining the important rules from multi-knowledge. Combining rough set theory and FCA [15], [16] can get more valuable knowledge. In this paper, we present a new algorithm to extract multi-knowledge. We obtain the multi-reducts using the rough set and the important rules using the FCA. The rest of the paper is organized as follows. Section II recalls the basics of rough set and FCA,

while Section III shows the details of the multi-knowledge extraction algorithm. Experiment results and analysis are provided in Section IV and the conclusion is discussed in Section V.

II. METHODOLOGY

The rough set theory is used to find multi-reducts and FCA is used to obtain more important rules to build multi-knowledge.

A. Formal Concept Analysis (FCA)

In FCA, data is presented by formal concept, which is defined as follows:

A formal context $K = (G, M, I)$, where G is the set of all objects; M is the set of all attributes; $(g, m) \in I$ presents that all of objects in $g \in G$ have these attributes in $m \in M$.

In the formal context, two mappings f about the common attributes of $\forall O \subseteq G$ and g about the common objects of $\forall Q \subseteq M$ can be defined as follows:

$$\begin{aligned} f(O) &= \{m \in M \mid \forall g \in O, gIm\} \\ g(Q) &= \{g \in G \mid \forall m \in Q, gIm\} \end{aligned} \quad (1)$$

A formal concept of the formal context is a pair $(O \subseteq G, Q \subseteq M)$, where $f(O) = Q, g(Q) = O$. O is called a extension of (O, Q) , and Q is called a intension of (O, Q) . We represents a extension as $Extent(F)$ and a intension as $Intent(F)$ for formal concept F .

The relation between subconcept and superconcept plays a outstanding role in FCA. The relation of formal concepts F_1 and F_2 means if $F_1 \leq F_2$, $Extent(F_1) \subseteq Extent(F_2)$ and $Intent(F_2) \subseteq Intent(F_1)$.

$F_1 \prec F_2$ represents the condition is not existence, which is $\exists F$ and $F_1 < F < F_2$ and F_2 is called upper-neighbor noted $\omega(F_1)$ of F_1 .

$G = \{1, \dots, i, \dots, n\}$ and $O_1, O_2 \subseteq G$. The sequence of elements in O_1 and O_2 are ascending according to the sequence of elements in G . This kind of sequence for G and its subsets is called a lexicographical order.

$O_1 < O_2$ in the is lexicographical order defined as follows:

$$O_1 \cap \{1, 2, \dots, i-1\} = O_2 \cap \{1, 2, \dots, i-1\} \quad (2)$$

where i represents the first different element from left to right between O_1 and O_2 .

The minimum extent of O_1 with a lexicographical order is $O_1 \oplus i$ as follows

$$O_1 \oplus i = g(f((O_1 \cap \{1, 2, \dots, i-1\}) \cup \{i\}))$$

$$i = \arg \max_{i \in G-O_1} (O_1 < O_1 \oplus i) \quad (3)$$

The concept lattice is composed of all the concepts. A node of the concept lattice is a formal concept and describes the relationship between objects and attributes. A formal concept can be form by a Hasse [17].

B. Multi-Knowledge Extraction Based on Rough Set

A decision table is denoted by $T = (U, C, D, V, f)$, where U is the universe of discourse. $C \cap D = \emptyset$, where C is the condition attributes and D is the decision attributes. The value of D is $(d_1, d_2, \dots, d_{|D|})$. $V = \bigcup_{c \in C \cup D} V_c$ (V_c is the value set of the attribute c). f is the total decision function such that $f(x, c) \in V_c$ for every $c \in C \cup D, x \in U$.

The D -positive region on equivalence classes U/E is defined ad follows:

$$POS_E(D) = \bigcup_{X \in U/E \wedge \forall x, y \in X \Rightarrow f(x, D) = f(y, D)} X \quad (4)$$

where $U/E = \{E_1, E_2, \dots, E_m\}$ represents the partition of the condition attributes ($E \subseteq C$) for U .

If $POS_E(D) = POS_C(D)$ and $POS_{E-\{e\}}(D) \neq POS_C(D)$ for $\forall e \in E$, E is a reduct of decision table. The set including many reducts like E is called multi-reducts MR .

In order to improve the efficiency of attribute reduction algorithm, the process of attribute reduction can be based on a simplified decision table. A simplified decision table is defined as $S' = (U', C, D, V, f)$, where $U' = \{u'_1, u'_2, \dots, u'_n\}$ is the set of these first elements in the set $U/C = \{[u'_1]_C, [u'_2]_C, \dots, [u'_n]_C\}$, and if the value of decision attribute in $[u'_i]_C$ is the same, u'_i belongs to U'_{pos} ; if not, u'_i belongs to U'_{neg} .

The importance of attribute is defined as follows:

$$sig_E(e) = |U'_{E \cup \{e\}} - U'_E|$$

$$U'_E = \left\{ \bigcup_{X \subseteq U'_{pos} \wedge |X/D|=1} X \right\} \cup \left\{ \bigcup_{X \subseteq U'_{neg}} X \right\} \quad (5)$$

where $X \in U'/E, E \subseteq C$ and $\forall e \in (C - E)$. The greater the value of sig is, the more important the attribute is.

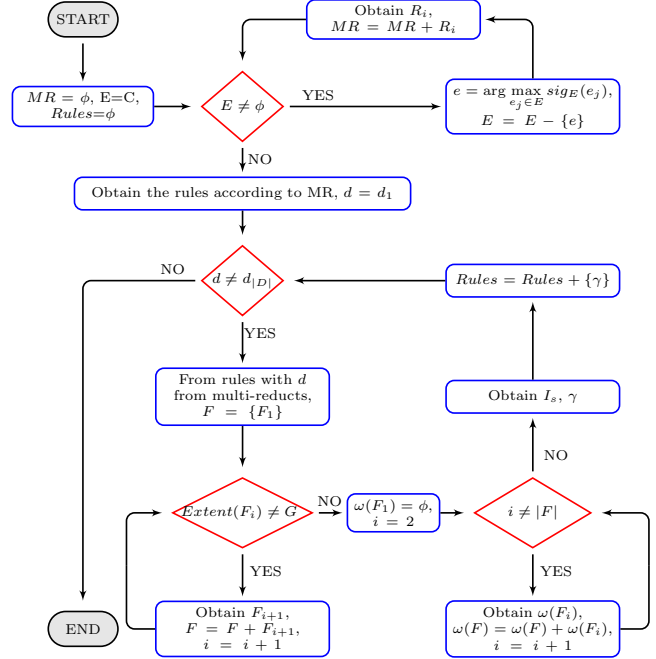


Figure 1. The flow chart of the algorithm

III. MULTI-KNOWLEDGE EXTRACTION ALGORITHM

Firstly, a good reduct is achieved in a decision table T . Then multi-reducts can be obtained by changing the sequences of condition attributes C . U/C is got by using a quick partition based on radix sort [18]. The process for getting a reduct is repeated $|C|$ times. Finally, the rules can be extracted by these multi-reducts.

The rules whose decision value is the same can form a context. The procedure is proposed to obtain the extensions and intensions of all the concepts of a context. Firstly, initialize a extension $O = g(f(\emptyset))$, $O \subseteq G$. G and its subsets are sorted by the lexicographical order. The elements in $G - O$ are selected from maximum to minimum until $O < O \oplus l$ is the first to be satisfied. And the next extension and the next intension of O are respectively $O \oplus l$ and $f(O \oplus l)$. Finally, the other extensions and intensions are calculated in accordance with the above steps until a extension is G .

The relation among concepts is obtained by upper-neighbors. Then some chief characteristics influencing the decision can be found by getting implication sets and implication relation. Finally, multi-knowledge is formed, which have many important rules from these chief characteristics. The pseudo-code for a rough set multi-knowledge extraction algorithm and its formal concept analysis (MEMFCA) is illustrated in Algorithm 1. The flow chart is illustrated in Figure 1.

IV. EXPERIMENTAL RESULTS OF QUALITATIVE_BANKRUPTCY

Our algorithm was implemented in the C language. The computation environment was an Intel® Core™ i5-3337U CPU @1.80 GHz processor with 4G memory.

Algorithm 1 MEMFCA Algorithm

Input:

A decision table $S = (U, C, D, V, f)$.

Output:

The important rules set *Rules*.

```

1:  $MR = \emptyset, E = C, Rules = \emptyset$ ;
2: for  $i = 1$  to  $|C|$  do
3:    $e = \arg \max_{e_j \in E} sig_E(e_j)$  using Equ. (5);
4:    $E = E - \{e\}, R_i = \{e\}$ ;
5:   while  $U' \neq \emptyset$  do
6:     Calculate  $U/R_i$  and obtain  $U', U'_{pos}, U'_{neg}$ ;
7:      $r = \arg \max_{r_t \in C-R_i} sig_{R_i}(r_t)$ ;
8:     Compute  $U'_{R_i}$  using Equ. (5);
9:      $R_i = R_i \cup \{r\}, U = U' - U'_{R_i}$ ;
10:  end while
11:  If  $R_i \notin MR, MR = MR + R_i$ ;
12: end for
13: Obtain the rules according to the multi-reducts  $MR$ ;
14: for  $d = d_1$  to  $d_{|D|}$  do
15:   Form rules with  $d$  from multi-reducts
16:   Initialize all concepts set  $F = \{F_1\}$  and  $i = 1$ , where
    $Extent(F_1) = g(f(\emptyset))$  and  $Intent(F_1) = f(\emptyset)$ ;
17:   while  $Extent(F_i) \neq G$  do
18:      $O = Extent(F_i)$ ;
19:      $l = \arg \max_{l \in G-O} (O < O \oplus l)$ ;
20:      $Extent(F_{i+1}) = O \oplus l,$ 
      $Intent(F_{i+1}) = f(O \oplus l)$ ;
21:      $F = F + F_{i+1}, i = i + 1$ ;
22:   end while
23:    $F'_i = F_{|F|-i+1}$ ;
24:    $F' = \{F'_1, F'_2, \dots, F'_{|F|}\}, \omega(F'_1) = \emptyset$ ;
25:   for  $i = 2$  to  $|F|$  do
26:     for  $j = 1$  to  $i$  do
27:       If  $F'_i \prec F'_j, \omega(F'_i) = \omega(F'_i) \cup \{j\}$ ;
28:     end for
29:   end for
30:   Obtain implication sets  $I_s$  from concept lattices of  $F'$ ;
31:   Compute the frequency of  $\forall m \in M$  according to  $I_s$ ;
32:   Obtain the important rules  $\gamma$  according to the frequency of  $m$ ;
33:    $Rules = Rules + \{\gamma\}$ ;
34: end for

```

A. Results

The experiment data is Qualitative_Bankruptcy data from UCI. Attribute information: (P =Positive, A =Average, N =Negative, B =Bankruptcy, NB =Non-Bankruptcy).

- Industrial Risk (Ir): $\{P, A, N\}$
- Management Risk (Mr): $\{P, A, N\}$
- Financial Flexibility (Ff): $\{P, A, N\}$
- Credibility (Cr): $\{P, A, N\}$
- Competitiveness (Co): $\{P, A, N\}$
- Operating Risk (Or): $\{P, A, N\}$

The condition attributes above are denoted by $a_1, a_2, a_3, a_4, a_5, a_6$ in turn. Class d is the decision attribute which includes B and NB in the decision table.

Firstly, multi-reducts (MR) can be obtained by using the Algorithm 1 as follows:

$$MR = \{\{a_5, a_3, a_1\}, \{a_2, a_5, a_3\}, \{a_4, a_5, a_3\}, \{a_6, a_5, a_1\}\}$$

The rules with NB can be got by MR and presented in Table I.

Table I
RULES WITH DECISION NB FROM MULTI-REDUCTS

No.	a_1	a_2	a_3	a_4	a_5	a_6	d
1	A	*	N	*	A	*	NB
2	*	*	P	*	A	*	NB
3	*	P	*	*	A	*	NB
4	*	A	A	*	A	*	NB
5	*	N	A	*	A	*	NB
6	*	A	P	*	A	*	NB
7	*	N	P	*	A	*	NB
8	*	*	N	P	A	*	NB
9	*	*	P	A	A	*	NB
10	*	*	P	P	A	*	NB
11	A	*	*	*	A	A	NB
12	N	*	*	*	A	N	NB
13	P	*	*	*	A	N	NB
14	*	*	*	*	A	P	NB
15	*	*	*	*	P	*	NB
16	*	*	A	*	A	*	NB

In Table II, we present the concepts and upper-neighbors in Table I. a_{ij} expresses the j th value of the attribute a_i [19]. For example, a_{32} represents the value of the third attribute (FF) is 2 (A). The Hasse of Table II is shown in Figure 2.

The implication set can be achieved by using a general method [20] in FCA and is shown in Table III. The implication relation in implication set is investigated further and the frequency of the implication is given in Table IV.

We can find some chief characteristics influencing the decision NB from the frequency of implication in Table

Table II
ALL THE CONCEPTS IN FORMAL CONTEXT OF RULES WITH NB

No.	extension	instension	upper-neighbor
1	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16	\emptyset	\emptyset
2	1,2,3,4,5,6,7,8,9,10,11,12,13,14,16	a_{52}	1
3	1,8	a_{12}, a_{52}	2
4	1,11	a_{12}, a_{52}	2
5	1	a_{12}, a_{33}, a_{52}	3,4
6	2,6,7,9,10	a_{31}, a_{52}	2
7	3	a_{21}, a_{52}	2
8	4,5,16	a_{32}, a_{52}	2
9	4,6	a_{22}, a_{52}	2
10	4	a_{22}, a_{32}, a_{52}	8,9
11	5,7	a_{23}, a_{52}	2
12	5	a_{23}, a_{32}, a_{52}	8,11
13	6	a_{22}, a_{31}, a_{52}	6,9
14	7	a_{23}, a_{31}, a_{52}	6,11
15	8,10	a_{41}, a_{52}	2
16	8	a_{33}, a_{41}, a_{52}	15,3
17	9	a_{31}, a_{42}, a_{52}	6
18	10	a_{31}, a_{41}, a_{52}	15,6
19	11	a_{12}, a_{52}, a_{62}	4
20	12,13	a_{52}, a_{63}	2
21	12	a_{13}, a_{52}, a_{63}	20
22	13	a_{11}, a_{52}, a_{63}	20
23	14	a_{52}, a_{61}	2
24	15	a_{51}	1
25	\emptyset	$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}, a_{41}, a_{41}, a_{51}, a_{52}, a_{61}, a_{62}, a_{63}$	5,7,10,12,13,14,16,17,18,19,21,22,23,24

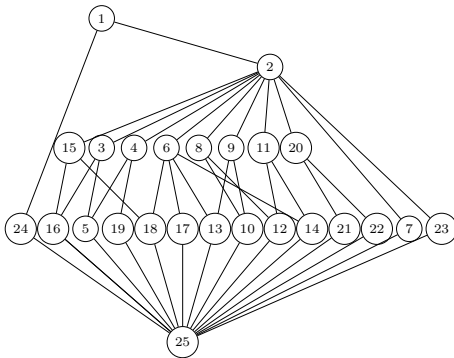


Figure 2. The Hasse of Concept Lattice for Context of Rules with NB

Table III
IMPLICATION SET FOR CONTEXT OF RULES WITH NB

Implication Sets		
[1] $a_{33} \Rightarrow a_{52}$	[6] $a_{22} \Rightarrow a_{52}$	[11] $a_{63} \Rightarrow a_{52}$
[2] $a_{12} \Rightarrow a_{52}$	[7] $a_{23} \Rightarrow a_{52}$	[12] $a_{13} \Rightarrow a_{52}, a_{63}$
[3] $a_{31} \Rightarrow a_{52}$	[8] $a_{41} \Rightarrow a_{52}$	[13] $a_{11} \Rightarrow a_{52}, a_{63}$
[4] $a_{21} \Rightarrow a_{52}$	[9] $a_{42} \Rightarrow a_{31}, a_{52}$	[14] $a_{61} \Rightarrow a_{52}$
[5] $a_{32} \Rightarrow a_{52}$	[10] $a_{62} \Rightarrow a_{12}, a_{52}$	

IV. The most important characteristic is a_{52} which means the value of condition attribute Co is A . The next important characteristic is that the value of Or is N .

Table IV
IMPLICATION RELATION FOR CONTEXT OF RULES WITH NB

Superconcept	Subconcept	Frequency
a_{12}	a_{62}	1
a_{31}	a_{42}	1
a_{52}	$a_{33}; a_{12}; a_{31}; a_{21}; a_{32}; a_{22}; a_{23}; a_{41}; a_{42}; a_{62}; a_{63}; a_{13}; a_{11}; a_{61}$	14
a_{63}	$a_{13}; a_{11}$	2

B. Discussion

The multi-knowledge is composed of some rules extracted according to Table I with IV. The frequency of a_{52} and a_{63} is more than a threshold which is set 1 in Table IV.

These rules including both a_{52} and a_{63} may be important and these important rules are extracted as follows:

- 1) If $a_1 = A, a_3 = N, a_5 = A$, then $d = NB$;
- 2) If $a_3 = P, a_5 = A$, then $d = NB$;
- 3) If $a_2 = A, a_3 = P, a_5 = A$, then $d = NB$;
- 4) If $a_1 = N, a_3 = P, a_5 = A$, then $d = NB$;
- 5) If $a_3 = P, a_4 = A, a_5 = A$, then $d = NB$;
- 6) If $a_3 = P, a_4 = P, a_5 = A$, then $d = NB$;
- 7) If $a_1 = N, a_5 = A, a_6 = N$, then $d = NB$;
- 8) If $a_1 = P, a_5 = A, a_6 = N$, then $d = NB$;

For example, Rule 1) represents if Ir is 'Average', Ff is 'Negative' and the Co is 'Average', then the decision is 'Non-Bankruptcy'.

These rules including either a_{52} or a_{63} may be ordinary and they are obtained as follows:

- 1) If $a_2 = P$, $a_5 = A$, then $d = NB$;
- 2) If $a_2 = A$, $a_3 = A$, $a_5 = A$, then $d = NB$;
- 3) If $a_2 = N$, $a_3 = N$, $a_5 = A$, then $d = NB$;
- 4) If $a_3 = N$, $a_4 = P$, $a_5 = A$, then $d = NB$;
- 5) If $a_1 = A$, $a_5 = A$, $a_6 = A$, then $d = NB$;
- 6) If $a_5 = A$, $a_6 = P$, then $d = NB$;
- 7) If $a_3 = A$, $a_5 = A$, then $d = NB$;

The multi-knowledge using MEMFCA algorithm is formed eight important rules from Qualitative-Bankruptcy data and seven ordinary rules.

V. CONCLUSIONS

In this paper, the aim is to extract multi-reducts for building a multi-knowledge. The multi-knowledge is composed of more rules in order to adapt to changing conditions. The two theories, rough set and FCA, are respectively used to extract multi-reducts and obtain rules. These rules can form a multi-knowledge. The multi-knowledge using FCA is based on the chief characteristics influencing the decision for reducing complexity of decision. The experiment results represent multi-knowledge with more rules and the number of rules can be according to realistic environments.

ACKNOWLEDGMENT

The authors would like to thank Chao Yang and Shiyuan Che for their scientific collaboration in this research work. This work is supported partly by the Fundamental Research Funds for the State Grid Liaoning Electric Power Supply Co. Ltd. (Grant No.2014YF-29), the National Natural Science Foundation of China (Grant No. 61173035, 61472058), the Program for New Century Excellent Talents in University (Grant No. NCET-11-0861), and the Fundamental Research Funds for the Central Universities.

REFERENCES

- [1] Z. Pawlak, "Rough sets," *International Journal of Computer & Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.
- [2] A. Kusiak, "Rough set theory: a data mining tool for semiconductor manufacturing," *IEEE Transactions on Electronics Packaging Manufacturing*, vol. 24, no. 1, pp. 44–50, 2001.
- [3] Y. Jin, *Multi-objective machine learning*. Springer, 2006, vol. 16.
- [4] Y. Leung, M. M. Fischer, W. Wu, and J. Mi, "A rough set approach for the discovery of classification rules in interval-valued information systems," *International Journal of Approximate Reasoning*, vol. 47, no. 2, pp. 233–246, 2008.
- [5] J. Stepaniuk, *Rough-Granular Computing in Knowledge Discovery and Data Mining*. Springer, 2008, vol. 152.
- [6] C. Ramos, J. C. Augusto, and D. Shapiro, "Ambient intelligence: the next step for artificial intelligence," *Intelligent Systems, IEEE*, vol. 23, no. 2, pp. 15–18, 2008.
- [7] X. Jia, W. Liao, Z. Tang, and L. Shang, "Minimum cost attribute reduction in decision-theoretic rough set models," *Information Sciences*, vol. 219, pp. 151–167, 2013.
- [8] T. Fan, C. Liau, and D. Liu, "A relational perspective of attribute reduction in rough set-based data analysis," *European Journal of Operational Research*, vol. 213, no. 1, pp. 270–278, 2011.
- [9] Q. Wu and D. Bell, "Multi-knowledge extraction and application," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*. Springer, 2003, pp. 274–278.
- [10] Q. Wu, D. Bell, and M. McGinnity, "Multi-knowledge for decision making," *Knowledge and Information Systems*, vol. 7, pp. 246–266, 2005.
- [11] C. Yang, H. Liu, Y. Sun, and A. Abraham, "Multi-knowledge extraction from violent crime datasets using swarm rough algorithm," in *Proceedings of 12th International Conference on Hybrid Intelligent Systems*. IEEE, 2012, pp. 560–565.
- [12] J. Poelmans, P. Elzinga, S. Viaene, and G. Dedene, "Formal concept analysis in knowledge discovery: a survey," in *Conceptual Structures: From Information to Intelligence*. Springer, 2010, pp. 139–153.
- [13] P. Tonella and M. Ceccato, "Aspect mining through the formal concept analysis of execution traces," in *Proceedings of 11th Working Conference on Reverse Engineering*. IEEE, 2004, pp. 112–121.
- [14] J. Li, C. Mei, J. Wang, and X. Zhang, "Rule-preserved object compression in formal decision contexts using concept lattices," *Knowledge-Based Systems*, 2014.
- [15] J. Shyng, H. Shieh, and G. Tzeng, "An integration method combining rough set theory with formal concept analysis for personal investment portfolios," *Knowledge-Based Systems*, vol. 23, no. 6, pp. 586–597, 2010.
- [16] X. Kang, D. Li, S. Wang, and K. Qu, "Rough set model based on formal concept analysis," *Information Sciences*, vol. 222, pp. 611–625, 2013.
- [17] J. Baixeries, L. Szathmary, P. Valtchev, and R. Godin, "Yet a faster algorithm for building the hasse diagram of a concept lattice," in *Formal Concept Analysis*. Springer, 2009, pp. 162–177.
- [18] Z. Xu, Z. Liu, B. Yang, and W. Song, "Quick attribute reduction algorithm with complexity of $\max(o(|c||u|), o(|c|^2|u/c|))$," *Chinese Journal of Computers*, vol. 29, no. 3, pp. 391–399, 2006.
- [19] B. Ganter, R. Wille, and C. Franzke, *Formal concept analysis: mathematical foundations*. Springer-Verlag New York, Inc., 1997.
- [20] Z. Wang, K. Hu, X. Hu, and D. Zhang, "General and incremental algorithms of rule extraction based on concept lattice," *Chinese Journal of Computers*, vol. 22, no. 1, pp. 66–70, 1999.