A Novel Algorithm for Classification Rule Discovery based on Concept Granule Structure

Zhao Jian [1,2*] , Leng Kong [3]
[1] School of Computer Technology and Engineering
Changchun Institute of Technology
[2] Mathematics Institute, Jilin University
Changchun 130012, China
[3] Department of Engineering Sciences
NCT, Yong-in, Kyunggi-do 101,
Republic of Korea
zjay_yvette1005@163.com

Journal of Digital
Information Management

ABSTRACT: *This study established concept elements based on granular computing theory and the isomorphic relation between rated scales in formal concept analysis (FCA) and constructed the correlation of the concept elements. A concept granule was constructed by studying the mapping relation between concept elements. The common polymerization and extension forms of the concept granule were given. We studied the condition in which the granular structure in a conceptual system is purified, as well as the formation mechanism and generalized cohesiveness of concept granules. An algorithm for classification rule discovery algorithm based on concept granule structure—the GRD algorithm—was created. According to experimental results, the proposed GRD algorithm has higher classification accuracy, simpler rule set, and better generalization than traditional algorithms for classification rule discovery. The formal representation based on conceptual elements shows that a knowledge representation model that is complete in terms of semantic description can be built.*

Subject Categories and Descriptors
H.2.8 [Database Applications]: Data mining; **I.2.4 [Knowledge Representation Formalisms and Methods**]: Representations (procedural and rule-based)

General Terms: Classification, Rule mining, Formal Concept Analysis

Keywords: Classification, Rule mining, Formal Concept Analysis, Rough Set, Granular Computing

Received: 1 November 2015, Revised: 27 December 2015, Accepted: 30 December 2015

1. Introduction

Big data analysis and mining are the current research hotspots in the field of computers. The algorithm for classification rule discovery can generate classification rule sets, which can supervise classification and display classification results to users in understandable forms (concepts). Thus, the said algorithm has an important role in the field of big data analysis[1]. Many related studies have been reported. The uncertainty problems of sampling resolved with the rule of fuzzy-interval neural network extraction[2]. An optimum combination of rules is found by the genetic algorithm through group evolution, thereby improving the classification accuracy[3,4]. The coverage of rules and generalization of rule sets are enhanced through a pruning strategy[5]. Classification accuracy is improved through a multi-dimensional coverage fuzzy function[6]. Although these studies have made some achievements, obtaining a simple classification rule set with high classification capability remains an urgent problem in the field. This problem is due to the continuous increase of the data size and rules discovered by algorithms. A large classification rule set also has low classification accuracy and is difficult to understand from the perspective of users.

Human knowledge comes from intuitive experiences about the real world. A perfect conceptual system is formed both in content and structure through the aggregation of sensorymaterials into continuous experience, which is the basis of cognition, leaning, and reasoning for human evolution are exceptionally complex and are characterized by the following features: (1) A conceptual system consists of a large amount of concepts and the various complex relationships among them. These concepts are linked with one another; that is, one concept can be associated with others and is therefore rich and relevant. (2) Human beings can learn, reason, and solve problems using known concepts, indicating that the conceptual system of human beings has systemic reasonability and a constitutive property. (3) The process of learning of human beings continues to develop, accompanied with the increasing richness of the conceptual system, which can accept new supplementary concepts and associations with new ones, thereby showing openness. (4) Human beings can understand and resolve problems in different fields using the conceptual system, demonstrating that such system possesses completeness and universality.

However, the existing representation methods and theories of knowledge all have various shortcomings in the study of the structure of the conceptual system. The major problems include failing to represent the richness and relevance of a conceptual system, the lack of attention to the incidence relation among concepts, the incapability to show the constitutive property of a conceptual system, and the failure to reflect the openness and constructive property of the system.

Thus, establishing a general framework is urgent to stimulate the cognition and reasoning mechanism of human beings through the characterization of the relevance and dynamic constructive property of the elements in a conceptual system. Based on the different understanding of the thinking ability of humans, many theories are established to simulate the ways of obtaining knowledge and the ways consciousness handles problems. For such theories as granular computing and others, complex questions are decomposed into simple ones with different granularities to seek solutions.

2. Related Work

Several influential methods and theories have emerged in such areas as knowledge representation and conceptual system construction. Formal concept analysis (FCA)[7,8] which was proposed by Prof. Wille, is a reconstruction of lattice theory. Radim Belohlave presented a preliminary study of the basic level of concepts in the FCA framework, which is an important phenomenon studied in the psychology of concepts[9,10]. Zhang WX et al. [11] established a granular computing model based on concept lattice according to the binary relation between object and attribute. This model is a concept hierarchical structure based on the dependency and causal relationships of the knowledge body in connotation (attribute set) and extension (object set) and is capable of discovering knowledge from different generalized and instantiated knowledge systems. Dubois et al. recently introduced a possibility-theoretic view of FCA and evaluated the set valued counterparts of four set functions—potential or actual and possibility or necessity—that underlie bipolar possibility theory[12,13].

The following problems currently exist in the study of conceptual system models using the theories and methods based on FCA:

1) Semantic notation excessively depends on external knowledge, such as in a fuzzy set[14]; the subordination of one object to a fuzzy concept depends on its subordinate function. Subordinate function is a type of prior knowledge; that is, it does not rely on other objects in the discourse domain and is generally obtained through the use of statistical approaches. Thus, with strong subjectivity, the subordinate function is difficult to use as the basis to represent a self-organizing conceptual system characterized by dynamic constitution.

2) Studies are lacking on the dependency relationship among concept elements in the semantic level and constitutive property of the structure of the conceptual system.

The present study proposes that a conceptual system based on the granular computing paradigm can be used as the frame to represent the conceptual system and investigate such issues as the structure and constructive mechanism of the system and overcome the above limitations.

3. Granular Structure of Conceptual System

From the perspective of granular computing, human beings have the ability of global analysis, that is, to handle problems in a different granular world, as well as to observe and analyze objective phenomena in case of different granularities. “Granularity” means that all contents that are included in the phenomenon can be organized under different levels according to their similarities. A granular structure is generally defined as follows.

Definition 3.1. (Granular structure): A granular structure is described with a triad (E_G, I_G, R), where sets E_G and I_G represent the granular extension and granular connotation, respectively. The binary relation $R \subseteq E_G \times I_G$ is an implication relation that indicates the context of granular structures in special backgrounds.

A granular structure can be used to present the basic form of elements in a conceptual system. In a special background, a granular extension refers to the elements that are contained in a singular granular structure in the discourse domain, a granular connotation refers to the attribute and similarity of elements in the granularstructure, and the implication relation reflects the special relationship between granular extension and granular connotation.

In this study, the granular structure of concept elements is constructed through the isomorphism based on rough set (RS)[15] and FCA theories. Refer to [16,17] for the proving process of Proposition 3.1.

Proposition 3.1. Knowledge base (U, A), U is the discourse domain, $A := \{B_m \mid m \in M\}$. $\mathbb{S}(U,A) := ((U, A, W, I), (S_B \mid B \in A))$, where S_B is the rated scale that can derive the background (U, N, J). Let γ represent the object concept mapping of (U, N, J). If $(u,v) \in IND(P)$ ($(u,v) \in U^2$, $P \subseteq A$), then $\gamma(u) = \gamma(v)$.

As shown in the preceding paragraph, in a knowledge base (U, A), a multi-valued background can be obtained under the effect of a scale operator, and the background that is derived by the multi-valued background is expressed as follows:

$$\left(U, \left\{\left(B, [u]_B\right) \mid B \in A, [u]_B \in U / B\right\}, J\right)$$

In other words, the knowledge base (U, A) can be represented by such background that is derived by the multi-valued background in a formal concept. Based on this proposition, the elements in a conceptual system are formally termed as concept elements.

Definition 3.2 (Concept element and its granular structure): In a knowledge base (U, A), the concept element is $\chi := (G, M, W, I)$, of which $M \subseteq A, G \subseteq U / M, W := \left\{[u]_B \mid u \in G, B \in M\right\}$, and $I \subseteq G \times M \times W$. The granular structure of the concept element is the derived background for its rated scale,

$$\left(G, \left\{\left(B, [u]_B\right) \mid B \in M, [u]_B \in U / B\right\}, J\right)$$

which satisfies $uJ\left(B, [v]_B\right) \Leftrightarrow [u]_B = [v]_B$. Thus, $\forall u, v \in U$ and $u \neq v$.

Definition 3.3 (Relevance of concept element): In a knowledge base (U, A), the two concept elements $\chi := (G, M, W, I)$ and $\chi' := (G', M', W', I')$ are given. These two granular structures and the derived backgrounds of their corresponding rated scales are

$$\left(G, \left\{\left(B, [u]_B\right) \mid B \in M, [u]_B \in W\right\}, J\right)$$

and

$$\left(G', \left\{\left(B', [v]_{B'}\right) \mid B' \in M', [v]_{B'} \in W'\right\}, J'\right),$$

respectively, and $\exists C \subseteq M$, $\exists C' \subseteq M'$ $\exists C \subseteq M$, $\exists C' \subseteq M'$. These concept elements are associated if the following conditions are satisfied:

1. $N \cap N' \neq \varnothing$ and

2. $\bigcup_{Y \in X/C'} \left\{u \in G \mid [u]_C \subseteq Y\right\} \cap \bigcup_{Y' \in X/C} \left\{v \in G' \mid [v]_{C'} \subseteq Y'\right\} \neq \varnothing$.

The nonempty intersection in condition 2 is recorded as $\Lambda(\chi_1, \chi_2)$ and C (C') is defined as the related knowledge set in the concept element χ (χ'), in which the elements are called related knowledge.

To satisfy a certain incidence relation, the concept elements in the conceptual system are combined into a higher-order aggregation or a conceptual granule.

Definition3.4. $C := (G_C, M_C, W_C, I_C)$ and $A_i := (G_{A_i}, M_{A_i}, W_{A_i}, I_{A_i})$ are concept elements with incidence relation of which the granular structures are (G_C, N_C, J_C) and $(G_{A_i}, N_{A_i}, J_{A_i})$, respectively. If s is the mapping from (G_C, N_C, J_C) to $(G_{A_i}, N_{A_i}, J_{A_i})$, then $(s_i)_{i \in I}$ is a conceptual granule with the core of C that is recorded as $(s_i : A_i \to C)_{i \in I}$, where C denotes the conceptual core in the conceptual granule.

4. 4. Constitutive Mechanism of the Conceptual System

The constructed conceptual system is an organic whole with openness and dynamic closure. Its inner composition is divided into different levels, among which the contents in the same level can constitute the elements in the upper level and form the system in the lower level. This paper describes the process of integrating concept elements into conceptual granules. Based on the given definition, this section discusses the formation mechanism of conceptual granules and the related propositions. For convenience, the granular structure of the concept element is expressed as (G, M, I).

Definition 4.1 For a granule structure (G, M, I), if any two objects $g_1, g_2 \in G$ that meet $f(g_1) = f(g_2)$ are $g_1 = g_2$, then (G, M, I) is extension reducible. Similarly, if any two elements that meet $g(m_1) = g(m_2)$ are $m_1 = m_2$, then (G, M, I) is connotation reducible. Granule structure (G, M, I) is purified if the structure is extension and connotation reducible.

According to this definition, the nature of granule structure purification involves searching for a minimum attribute subset that can completely determine the hierarchical structure on the granule structure. Purification facilitates the discovery of implicit knowledge in the granule structure to simplify the knowledge.

Definition 4.2(Cohesion):
Supposing $A_i = (G_i, M_i, I_i)$, $C = (G_C, M_C, I_C)$ is a granular structure in the conceptual granule and C is the core of the conceptual granule. Let s_i $(i = 1, 2, \ldots)$ be the mapping between A_i $(i = 1, 2\ldots)$ and C, while t_i $(i = 1, 2\ldots)$ is the mapping in A_i $(i = 1, 2\ldots)$. If $s_i = s_{i+1} \circ t_i$, then A_i and C are cohesive under the mappings s_i and t_i.

Cohesion ensures that among all concept elements in a discourse domain, only those elements with incidence relation are included in the conceptual system and packaged into the independent conceptual granules. Different conceptual granules in the same layer may have incidence relation, based on which they interact.

The aforementioned cohesion only occurs in a single conceptual granule although a conceptual system has several levels. From a macroscopic perspective, cohesion also occurs at a certain level to represent the relevance between conceptual granules, the formalization of which is defined as follows.

Definition 4.3. If granular structure C' is distinguished from C, then A_i and C show generalized cohesion under the mappings s_i, t_i. If $s'_{i+1} \circ t_i = s'_i$ in the mapping $\left(A_i \xrightarrow{s'_i} C'\right)_{i\geq 1}$, then only one mapping $C \xrightarrow{s} C'$ and $s \circ s_i = s'_i$ is available.

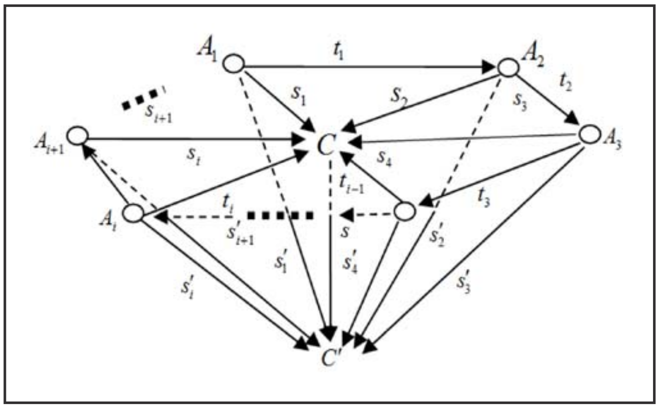

Figure 1. Generalized cohesiveness of the conceptual granules

Theorem 4.1. If the granular structure in a conceptual system is purified, then the system can achieve generalized cohesion.

Proof: First, a granular structure C is constructed as the core of conceptual granules. In a conceptual granule, let $t_i(i=1,2\ldots)$ be the sequence of mapping in $A_i(i=1,2\ldots)$. Supposing $A_i=(G_i,M_i,R_i)$ is purified, then $t_i:=(\eta_i,\xi_i)$ and $\eta_i:G_i\rightarrow G_{i+1}$ is an inclusion mapping. A core granular structure $C=(G,M,I)$ is constructed, where

$$G=\bigcup_{i\geq 1} G_i\ ,\ M=\left\{\left(m_j\right)_{j\geq 1} \mid \forall j\geq 1, m_j\in M_j \text{ and } \xi_j\left(m_{j+1}\right)=m_j\right\}$$
$$I\left(\hat{g},\hat{m}\right)=I_i\left(\hat{g},m_i\right), g\in G_i\ .$$

To differentiate from the g in FCA, $g\in G$ is recorded as $\hat{g}$, while $\left(m_j\right)_{j\geq 1}\in M$is recorded as $\hat{m}$. Therefore, for all $i\geq 1$, the conceptual granule $A_i \xrightarrow{s_i} C$ can be defined, while $\eta_i\left(\hat{g}\right)=\hat{g}$ and $\xi_i\left(\hat{m}\right)=m_i$ for all $\hat{g}\in G_i$ and $\hat{m}\in M$. In addition, according to the construction of C, $\hat{\xi}_i\left(m_{i+1}\right)=m_i$ and $\hat{g}I_{i+1}m_{i+1}=\hat{g}I_i m_i$ can be easily obtained as $\hat{g}I_j m_j=\hat{g}I_i m_i$ ($j\geq i$).

For $\forall\hat{m}\in M$, $\left(\hat{\xi}_i\circ\xi_{i+1}\right)\left(\hat{m}\right)=\hat{\xi}_i\left(m_{i+1}\right)=m_i=\xi_i\left(\hat{m}\right)$, while for $\forall g\in G$, $\left(\eta_{i+1}\circ\eta_i\right)\left(g\right)=g=\eta_i\left(g\right)$. Therefore, $s_{i+1}\circ t_i=s_i$. Based on Definition 4.2, the conceptual granules are determined to be in cohesion under the core C .

Let $\hat{m},\hat{n}\in M$, $m_i,n_i\in M_i$, and suppose that $g\left(\hat{m}\right)=g\left(\hat{n}\right)$ and $\forall\hat{g}\in G_i$ exist in the granular structure. If $\hat{g}I_i m_i=\hat{g}I\hat{m}=\hat{g}I\,\hat{n}=\hat{g}I_i n_i$, then $g\left(m_i\right)=g\left(n_i\right)$. Given that A_i is purified, its connotation is reducible. If $m_i=n_i$, $\hat{m}=\hat{n}$, C is reducible in connotation. Similarly, C can be proven to be reducible in extension. Therefore, the constructed granular structure C is purified.

Let $\left(A_i \xrightarrow{s'_i} C'\right)_{i\geq 1}$ be cohesive, and the core background $C'=\left(G',M',R'\right)$ is different from C. Define the mapping $s:C\rightarrow C'$, where $\hat{\eta}:G\rightarrow G'$. If $\hat{g}\in G_i$, then $\eta\left(\hat{g}\right)=\eta'\left(\hat{g}\right)$. If $\hat{\xi}:M'\rightarrow M$, then $\xi\left(m'\right)=\left(\xi'_k\left(m'\right)\right)_{k\geq 1}$.

The preceding construction background shows $\eta_j\left(\hat{g}\right)=\eta'_i\left(\hat{g}\right)$, where $1\leq i\leq j$, $\forall j\geq 1$ and $\hat{\xi}_j\left(\xi'_{j+1}\left(m'\right)\right)=\xi'_j\left(m'\right)$, $\left(\xi_{km}\left(m'\right)\right)_{k\geq 1}\in M$. Given that $\eta\left(\eta_i\left(\hat{g}\right)\right)=\eta\left(\hat{g}\right)=\eta'_i\left(\hat{g}\right)$and $\xi_i\left(\xi\left(m'\right)\right)=\xi_i\left(\left(\xi'_j\left(m'\right)\right)_{j\geq 1}\right)=\xi'_i\left(m'\right)$, then $s\circ s_i=s'_i$.

Theorem 4.1 extends the generalized cohesion between different conceptual granules in a certain layer within the conceptual system. Similarly, the conceptual granules in different layers also meet the generalized cohesion and constitute an open and expansive system.

5. Rule Discovery Algorithm

Based on the concept granule structure and formation mechanism, the data in the training dataset can be viewed as a granularized conceptual system. The concept granule that comprise done group of samples was used to represent one group of classification rules in the conceptual system. This idea is described as the classification rule discovery algorithm based on concept granule structure (GRD); the flowchart of GRD is shown in Figure.2.

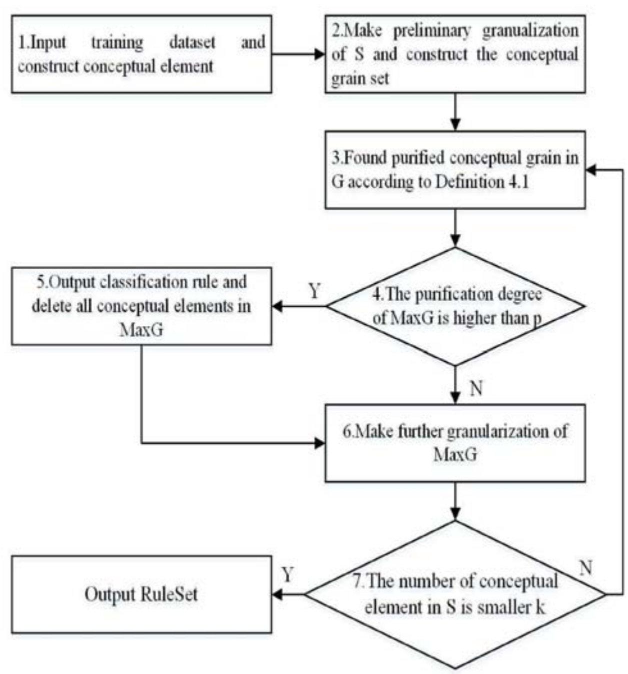

Figure 2. GRD flowchart

Specifically, GRD involves the following steps:

1) Input training dataset and construct concept element: Input all training datasets and construct the conceptual system (S). According to Definition 3.2, every sample in the system was constructed as one concept element in S.

2) Preliminarily granulize S and construct the concept granule set: Concept elements with a proportion of a% are selected randomly. The granule size to which all concept elements belong is divided. Construct the concept granule set (G) according to Definition 3.4.

3) Find the purified concept granule in G according to Definition 4.1: Calculate the category purification degree of all concept elements in the concept granule and arrange them in order. Output the purified concept granule MaxG.

4) If the purification degree of MaxG is higher than p, proceed to Step 5. Otherwise, proceed to Step 6.

5) Output classification rule and delete all concept elements in MaxG: The output classification rule is formatted as follows: If MaxG, thenall concept elements in MaxG belong to the maximum category. Add this rule into RuleSet and delete all concept elements in MaxG. Proceed to Step 7.

6) Further granulize MaxG: Extractm concept elements in MaxG. Obtain a group of subgranules through the further granularization of MaxG and then add these subgranules into G. Delete MaxG from G.

7) If the number of concept elements in S is smaller thank, proceed to Step 8. Otherwise, proceed to Step3.

8) Output RuleSet.

Through this algorithm, all concept granules in the training dataset are extracted and transformed into a classification RuleSet, in which every rule corresponds to one concept granule to help judge the sample categories.

A dataset with 100,000 records was used to verify the GRD algorithm. Each data corresponded to a webpage content that was collected by a Web crawler. In other words, all pure texts with scripts were deleted through HTML reversal coding. The dataset was divided into five categories, namely, Finance, Sports, Society, Life, and Education. The samples of each category were collected from the corresponding topics of large websites. Each category accounted for 20% of the total sample. Each record in this dataset belonged to one of these categories. A comparative analysis of the GRD, ID3 tree, Random Forest (RF), and PART algorithms was performed usingIntel-i5 4590/16G. Java 1.7 was used to apply all algorithms. All experimental data were stored in a My SQL database.

0.1K, 0.5K,1K, 2K,3K,…10K records were selected randomly from the experimental dataset and were divided into 12 groups of training data. The GRD, ID3, RF, and PART algorithms were used to discover the rules. The comparison results of rule number are shown in Table 1 and Figure 3.

	GRD	ID3	RF	PART
0.1	38	26	41	30
0.5	42	47	52	39
1	45	68	71	47
2	67	82	107	72
3	84	140	142	96
4	110	153	179	120
5	115	205	201	132
6	120	204	270	159
7	136	300	326	184
8	150	352	371	206
9	162	370	402	221
10	170	390	430	261

Table 1. Rule number of four algorithms

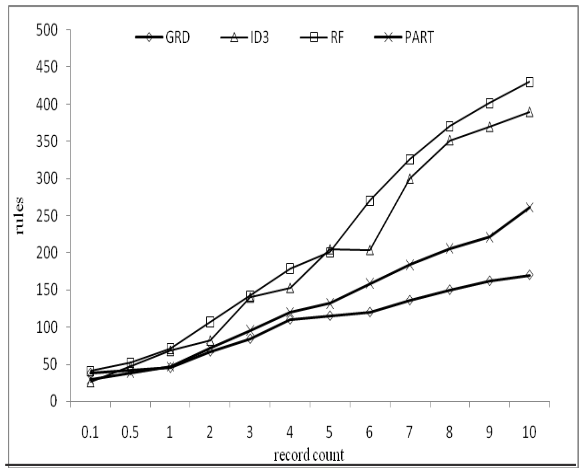

Figure 3. Comparison of rule number

Given that RF uses various trees with each tree having to generate rules, this algorithm generates more rules than the other algorithms. Under 10K samples, the number of extracted rules reached 430. ID3 generates few rules under a small sample size, but the rule number increases quickly as the number of samples increases. Under 10K samples, the number of extracted cases can reach 390. By applying the pruning strategy, the PART algorithm generates fewer rules than the ID3 and RF algorithms and its rule growth is also much slower. The proposed GRD algorithm achieves the slowest rule growth along with increasing sample size. This algorithm has only extracted 170 rules under 10K samples.

The mean numbers of judgment conditions of these algorithms for every rule in the RuleSet are shown in Table 2 and Figure.4.

RF comprises many relatively short subtrees and has fewer judgment conditions than the ID3 and PART algorithms. The ID3 algorithm uses a single tree and has more judgment conditions than the other algorithms except for under a 0.1K sample size. The judgment conditions of the ID3 algorithm under a 10K sample size reach the

	GRD	ID3	RF	PART
0.1	13	12	12	13
0.5	17	27	25	25
1	21	60	33	51
2	27	120	38	63
3	32	167	41	98
4	39	203	55	109
5	43	231	59	127
6	58	245	63	145
7	60	270	79	158
8	67	295	82	185
9	72	304	90	203
10	77	321	92	215

Table 2. Mean numbers of four algorithms

highest value of 321. Despite extracting many rules, the PART algorithm has fewer rules than ID3. The proposed GRD algorithm maintains as low growth of rules and its number of rulesis lower than that of other algorithms when the sample size exceeds 0.5K.Such a number only reaches 77 under a 10K sample size, which indicates that the GRD algorithm hasahigher summary capacity. Moreover, the proposed GRD algorithm has fewer rules and fewer judgment conditions for each rule, which indicates that this algorithm can obtain a simpler RuleSet and can be easily read and understood by users.

A total of 100,000 records were all used as test data. The coverage comparison of the algorithms is presented in Table 3 and Figure.5.

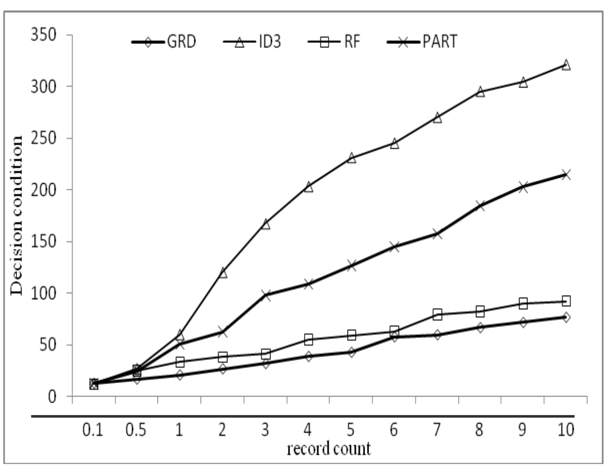

Figure 4. Mean numbers of judgment conditions

The GRD algorithm achieves the highest coverage, which does not decrease significantly along with increasing sample size. On the contrary, the coverage of the other three algorithms is unstable and decreases along with an increasing sample size. In other words, excessive training samples lead to excessive fitting. The higher coverage of the GRD algorithm indicates its stronger generalization of classification rule and better adaptation to more unknown samples. The classification accuracy of thesealgorithms is compared inTable 4 and Figure.6.

	GRD	ID3	RF	PART
0.1	77	65	75	78
0.5	80	74	79	80
1	84	85	80	83
2	88	80	83	81
3	90	82	84	82
4	91	86	85	83
5	93	86	87	86
6	95	85	89	87
7	95	84	89	87
8	97	84	92	89
9	97	82	93	90
10	97	82	90	90

Table 3. Coverage of four algorithms

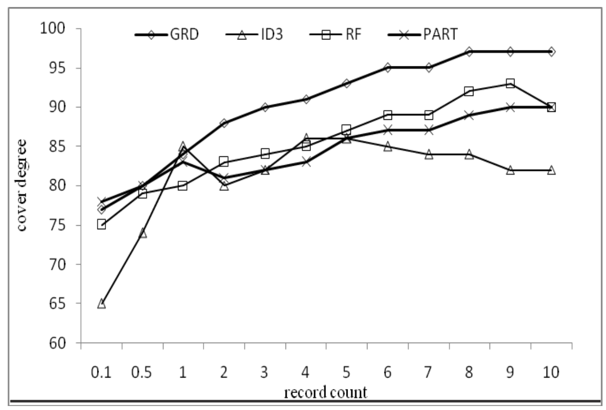

Figure 5. Coverage comparison

	GRD	ID3	RF	PART
0.1	73	62	70	66
0.5	74	65	71	68
1	80	65	74	70
2	82	67	74	70
3	85	68	76	72
4	89	68	78	73
5	90	70	81	75
6	91	70	81	79
7	93	73	83	79
8	94	73	83	80
9	94	73	87	81
10	94	73	89	81

Table 4. Classification accuracy of four algorithms

The ID3 algorithm has the poorest classification accuracy, with only 62% accuracy under a 0.1K sample size and 73% accuracy under a 10K sample size. The PART algorithm shows higher classification accuracy than ID3, obtaining 81% accuracy under a 10K sample size. The classification accuracy of RF under a 10K sample size is even higher and can reach89%. Given its high coverage and simpler rule structure, the proposed GRD algorithm achieves the highest classification accuracy, reaching 73%

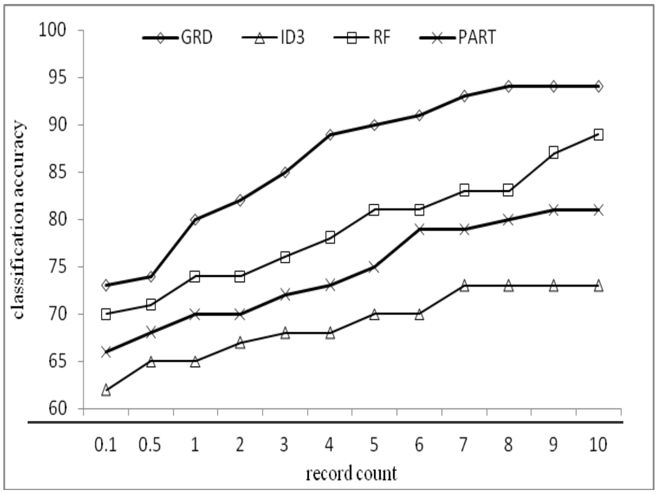

Figure 6. Comparison of classification accuracy

under a 0.1K sample size and 94% under a 10K sample size.

6. Conclusion

This study established a formal representation for a conceptual system based on granular structure. FCA and rough set(RS) constituted concept elements, that is, adding the indiscernibility relation to the multi-value background of FCA to endow the conceptual category with the capability of handling incomplete and inadequate information and the capability to classify concept elements in terms of their different attributes or characteristics. The mechanism of conceptual granule aggregation was also studied based on the incidence relation among concept elements, thereby providing a general constitution form of the conceptual system.

An algorithm for classification rule discovery based on concept granule structure—the GRD algorithm—was proposed to cope with the classification rule mining of big data size. The algorithm considered grain structure as the basic structure to give a formal expression of the conceptual system. Concept elements were constructed based on FCA and RS. In other words, the indiscernibility relation was added under the multi-value background of FCA. Thus, the conceptual category can process incomplete information, that is, classify concept elements according to their attributes or features. According to the test of one group of experimental data, the proposed GRD algorithm is superior to ID3, RF, and PART in the simplification and generalization of classification rules and in classification accuracy. The proposed algorithm can effectively handle a rule discovery task from mass data and has promising application prospects.

In the next step, we will focus on the study of relevance logic among different phenomenal spaces based on correlation. Unlike the analysis logic based on deduction, relevant logic simulates the approximately intuitive thinking pattern adopted by humans in processing some "*non-pure cause and effect*" problems.

Acknowledgement

The study was jointly supported by the National Natural Science Foundation Youth Fund of China (61503044); Foundation of Jilin Province Education Department (2015292); China Postdoctoral Science Foundation (2014M551183); Foundation of Jilin Provincial Science & Technology Department (20140101178JC).

References

[1] López, V., Río, S. D., Benítez J. M. (2015). Cost-sensitive linguistic fuzzy rule based classification systems under the MapReduce framework for imbalanced big data. *Fuzzy Sets & Systems,* 258 (1) 5-38.

[2] Shinde, S., Kulkarni, U. (2015). Extracting classification rules from modified fuzzy min–max neural network for data with mixed attributes. *Applied Soft Computing*, 40 (1) 364-378.

[3] Sharma, P., Saro, J. (2015). Discovery of Classification Rules Using Distributed Genetic Algorithm. *Procedia Computer Science,* (46) 276–284.

[4] Yeh, W. C. (2012). Novel swarm optimization for mining classification rules on thyroid gland data. *Information Sciences,* (197) 65–76.

[5] Hernández-León, R., Carrasco-Ochoa, J. A., Martínez-Trinidad, J. F (2012). Classification based on specific rules and inexact coverage. *Expert Systems with Applications,* 39 (12) 11203–11211.

[6] Elkano, M., Galar, M., Sanz, J. (2016). Fuzzy Rule-Based Classification Systems for multi-class problems using binary decomposition strategies: On the influence of n-dimensional overlap functions in the Fuzzy Reasoning Method. *Information Sciences,* (332) 94-114.

[7] Formica, A. (2012). Semantic Web search based on rough sets and Fuzzy Formal Concept Analysis. *Knowledge-Based Systems,* 26 (1) 40-47.

[8] Kumar, C.A., Ishwarya, M. S., Loo C. K. (2015). Formal concept analysis approach to cognitive functionalities of bidirectional associative memory, *Biologically Inspired Cognitive Architectures,* 12 (1) 20-33

[9] Škopljanac-Maèina, F., Blaškoviæ, B. (2014). Formal Concept Analysis – Overview and Applications. *Procedia Engineering,* 69 (1) 1258-1267

[10] Belohlavek, R., Vychodil V(2009). Formal concept analysis with background knowledge: attribute priorities. *IEEE Trans. Systems, Man, and Cybernetics,* Part C 39(4), (1) 399–409.

[11] Liang, D., Liu, D. (2015). Deriving three-way decisions from intuitionistic fuzzy decision-theoretic rough sets, *Information Sciences,* 300 (1) 28-48.

[12] Poelmans, J., Ignatov, D. I., Kuznetsov, S. O., Dedene,

G. (2013). Formal concept analysis in knowledge processing: A survey on applications. *Expert Systems with Applications,* 40 (16) 6538-6560.

[13] Dubois D., Prade, H.(2012). Possibility theory and formal concept analysis: Characterizing independent sub-contexts. *Fuzzy Sets and Systems*, (196) 4–16.

[14] Kwiatkowska, M., Kielan, K. (2013). Fuzzy logic and semiotic methods in modeling of medical concepts. *Fuzzy Sets and Systems*, 214(1) 25-50.

[15] Pawlak, Z., Skowron, A .(2007). Rough sets and Boolean reasoning. *Information Sciences*, 177 (1) 41-73.

[16] Zhao, J., Liu, L., Hu, L. (2012). Extended Representation of the Concept element in Temporal Context and the Diachronism of the Knowledge System. *Knowledge-Based Systems,* 33 (1) 136-144.

[17] Zhao, J., Liu, L. (2011). Construction of concept granule based on rough set and representation of knowledge-based complex system. *Knowledge-Based Systems*, 24 (6) 809-815.