



Attribute Reduction of Set-Valued Decision Information System

Jun Hu^(✉), Siyu Huang, and Rui Shao

Chongqing Key Laboratory of Computational Intelligence,
Chongqing University of Posts and Telecommunications,
Chongqing 400065, China
hujun@cqupt.edu.cn

Abstract. In practice, we may obtain data which is set-valued due to the limitation of acquisition means or the requirement of practical problems. In this paper, we focus on how to reduce set-valued decision information systems under the disjunctive semantics. First, a new relation to measure the degree of similarity between two set-valued objects is defined, which overcomes the limitations of the existing measure methods. Second, an attribute reduction algorithm for set-valued decision information systems is proposed. At last, the experimental results demonstrate that the proposed method can simplify set-valued decision information systems and achieve higher classification accuracy than existing methods.

Keywords: Set-valued decision information system · Rough set
Attribute reduction · Uncertainty

1 Introduction

With the development of information technology, the means of data acquisition becomes more and more diverse. Meanwhile, the cost of data storage is getting lower and lower. These make it possible to acquire and store large amount of data, which stimulate the urgent need for automatic data processing. In many real world problems, data uncertainty is pervasive. In the past few years, it can be found that more attention has been paid to uncertain data. Rough set theory [1] is a powerful tool for dealing with uncertainty. Generally, the data processed by rough set theory is complete, accurate and atomic. However, the data in many real problems may be incomplete, inaccurate or non-atomic due to the limitation of acquisition means or the requirement of practical problems. It has become an important issue how to process incomplete data, interval-valued data and set-valued data.

A set-valued information system is that the value for an object on an attribute is not an exact value, but a set containing all possible values. For the processing of set-valued information system, a number of related researches have been studied. Orłowska and Pawlak [3, 4] investigated set-valued information system considering non-deterministic information and introduced the concept of a non-deterministic information system. Sakai et al. [5, 6] set up the theoretical foundations and algorithmic background for adapting rough set methods for the purposes of the analysis of non-deterministic information systems. Yao [7, 8] proposed a number of set-based computation methods

based on set-based information system. In addition, some new relations for set-valued information system and their corresponding attribute reduction methods were proposed. And then, Zhang et al. proposed the concept of set-valued information system and a similarity relation [11, 12]. A tolerance relation was defined and the largest tolerance class was used to divide the universe in that paper. Qian et al. [13] proposed a dominant relation for set-valued information system and a corresponding attribute reduction method. Dai and Tian [14] gave a fuzzy relation, that was used to measure the degree of similarity between two set-valued objects. Wang [15] pointed out that the family of reducts defined by Dai need not be a subset of the family of reducts defined within the standard rough set model for set-valued information system. Bao and Yang [16] proposed a δ -dominance relation and a corresponding attribute reduction approach. Two types of fuzzy rough approximations, and two corresponding relative positive domain reductions were proposed by Wei et al. [17]. Moreover, some researchers transformed incomplete information systems into set-valued information systems to achieve reduction. Lipski [9, 10] discussed set-valued information systems from the view of incomplete information systems under the case of missing values. And as a special case of set values, incomplete information is probabilistically dealt with by some authors [18, 19] in the case of missing values.

It is an important issue how to define a binary relation dealing with set-valued information systems by rough set theory. Generally, there are two different semantic interpretations of set-valued data, namely, the conjunctive semantics and the disjunctive semantics [2]. Many different definitions have been proposed for these two semantic interpretations. However, we find that they are not appropriate when dealing with the set-valued data under the disjunctive semantics. For this reason, we developed a new approach based on probability, which can characterize the relation between two set-valued objects more reasonable under the disjunctive semantics. Then, an attribute reduction algorithm based on keeping positive domain for set-valued decision information systems was proposed and some experiments were conducted to prove the effect of the algorithm.

The study is structured as follows: In Sect. 2, some basic concepts of set-valued information system are reviewed. A new approach based on probability is proposed to measure the degree of the similarity between two set-valued objects under the disjunctive semantics in Sect. 3. In Sect. 4, we put forward an attribute reduction algorithm for set-valued decision information system. In Sect. 5, the experimental results and relative analysis are presented. Section 6 concludes the paper.

2 Preliminaries

In this section, some basic concepts about set-valued information system will be reviewed.

Definition 1 [12]. An information system is defined as (U, A, V, F) , where U is a non-empty finite set of objects called the universe. A is a non-empty finite set of attributes. V is a union of attribute domains ($V = \cup_{a \in A} V_a$), V_a is a set including all possible values for $a \in A$. $F : U \times A \rightarrow V$ is a function that assigns particular values

from attribute domain to objects. For $\forall a \in A, x \in U, F(a, x) \in V_a, F(a, x)$ is the value of a for x . If for any a and $x, F(a, x)$ is a single value, then the information system is called a single-valued information system. Otherwise, it is called a set-valued information system.

There are two different semantic interpretations for set-valued information system [2]. The one is conjunctive semantics, the other is disjunctive semantics. Under the conjunctive semantics, a set value represents all the values of an object to an attribute, and all the values in the set are true. Under the disjunctive semantics, a set value represents all possible values of an object to an attribute, but there is only one true value in the set. For example, for $x \in U$ and $a \in A$, where a means the languages that x can speak. Let $a(x) = \{English, Chinese, French\}$. If $a(x)$ is interpreted conjunctively, it means x can speak *English, Chinese, and French*. If $a(x)$ is interpreted disjunctively, it means x can speak only one of *English, Chinese, and French*. We mainly study set-valued information system under the disjunctive semantics in this paper. In the following, a set-valued information system is under the disjunctive semantics if not otherwise specified.

To characterize the relation between two objects in a set-valued information system, many relations have been developed. Here we cite two important definitions of them.

Definition 2 [10]. Let $S = (U, A, V, F)$ be a set-valued information system. For $\forall b \in A$, a tolerance relation can be defined as follows:

$$R_b^\cap = \{(x, y) \in U \mid b(x) \cap b(y) \neq \emptyset\}. \tag{1}$$

Then, for $B \subseteq A$, a tolerance relation can be defined as follows:

$$R_B^\cap = \{(x, y) \in U \mid \forall b \in B, b(x) \cap b(y) \neq \emptyset\} = \bigcap_{b \in B} R_b^\cap. \tag{2}$$

It is obvious that R_B^\cap is reflexive and symmetric, but not necessarily transitive.

Dai et al. [14] defined a fuzzy relation, then the fuzzy relation is used to measure the similarity between set-valued objects.

Definition 3. Let $S = (U, A, V, F)$ be a set-valued information system. For $\forall b \in A$, a fuzzy relation \widetilde{R}_b can be defined as follows:

$$\mu_{\widetilde{R}_b}(x, y) = \frac{|b(x) \cap b(y)|}{|b(x) \cup b(y)|}. \tag{3}$$

Thus, for $B \subseteq A$, a fuzzy relation can be defined as follows:

$$\mu_{\widetilde{R}_B}(x, y) = \inf_{b \in B} \mu_{\widetilde{R}_b}(x, y). \tag{4}$$

The above two definitions are not reasonable in some real problems. For example, let $a(x) = \{English, Chinese, French\}$, $a(y) = \{German, Japanese, English,$

Chinese}, and $a(z) = \{English, Chinese, French\}$. According to Definition 2, $(y, z) \in R_a^\cap$ because $c(y) \cap c(z) = \{English, Chinese\}$ is not empty. That is to say, y and z are indiscernible with respect to the tolerance relation. In other words, y and z speak the same language. However, they may speak different languages. For example, y can only speak *English* and z can only speak *French*. By Definition 3, we know that $\mu_{R_a}^\sim(x, z) = 1$. Thus, x and z are indiscernible with respect to the fuzzy relation. It means that x and z speak the same language. However, they may speak different language either. For example, x can only speak *English* and z can only speak *French*.

From this example, we find that two objects under the disjunctive semantics satisfied the existing relations only denote that they have possibility to be similar. Therefore, a new relation, which can describe the relation between two set-valued objects under the disjunctive semantics, should be defined.

3 A New Similarity Relation Based on Probability

As the analysis in Sect. 2, the existing methods may get unreasonable results under the disjunctive semantics. In order to solve this problem, a similarity relation based on probability is proposed in this section.

Definition 4. Let $S = (U, A, V, F)$ be a set-valued information system. For $\forall b \in A$, the similarity between x and y on b is defined as follows.

$$\mu_{R_b}(x, y) = \begin{cases} \frac{|b(x) \cap b(y)|}{|b(x)| * |b(y)|}, & x \neq y. \\ 1, & x = y. \end{cases} \tag{5}$$

From the view of probability, $\mu_{R_b}(x, y)$ is the possibility that x and y take the same value. For $B \subseteq A$, the similarity relation R_B between x and y on B can be defined as follows:

$$\mu_{R_B}(x, y) = \prod_{b \in B} \mu_{R_b}(x, y). \tag{6}$$

There are some important properties of the similarity relation defined above:

- (1) R_B is reflective.

Proof. Since $\mu_{R_b}(x, x) = 1$, then $\mu_{R_B}(x, x) = 1$, we know that R_B is reflective.

- (2) R_B is symmetric.

Proof. Since $\mu_{R_b}(y, x) = \frac{|b(y) \cap b(x)|}{|b(y)| * |b(x)|} = \frac{|b(x) \cap b(y)|}{|b(x)| * |b(y)|} = \mu_{R_b}(x, y)$, then $\mu_{R_B}(y, x) = \mu_{R_B}(x, y)$, we know that R_B is symmetric.

For the same example used before, according to Definition 4, the similarity between x and y is $\mu_{R_a}(x, y) = \frac{1}{6}$, and the similarity between x and z is $\mu_{R_a}(x, z) = \frac{1}{3}$. That is to say, the probability of x and y speaking the same language is $\frac{1}{6}$, and the probability of x

and z speaking the same language is $\frac{1}{3}$. Comparing to the existing definitions, the definition we proposed gives a more reasonable and meaningful description between two objects with set values.

Definition 5. Let $S = (U, A, V, F)$ be a set-valued information system, for $B \subseteq A$, $x \in U$, the δ similarity class of x with respect to B can be defined as follows:

$$\delta_B(x) = \{y \in U \mid \mu_{R_B}(x, y) \geq \delta\} (0 \leq \delta \leq 1), \tag{7}$$

where δ is a threshold. We can use δ to control the size of the information granules generated by B . Specifically, the bigger the δ , the smaller size of the information granules generated by B .

Theorem 1. Let $S = (U, A, V, F)$ be a set-valued information system, B and B' be two subsets of A , δ be a threshold. For $x \in U$, if $B \supseteq B'$, then $\delta_B(x) \subseteq \delta_{B'}(x)$.

Proof. Let $B' = \{b_1, b_2, \dots, b_n\}$, $B = B' \cup \{b_{n+1}, b_{n+2}, \dots, b_{n+m}\}$. $\forall x \in U, \forall y \in U$. Then:

$$\begin{aligned} \mu_{R_{B'}}(x, y) &= \prod_{b \in B'} \mu_{R_b}(x, y) = \mu_{R_{b_1}}(x, y) * \mu_{R_{b_2}}(x, y) * \dots * \mu_{R_{b_n}}(x, y) \\ \mu_{R_B}(x, y) &= \prod_{b \in B} \mu_{R_b}(x, y) \\ &= \mu_{R_{b_1}}(x, y) * \mu_{R_{b_2}}(x, y) * \dots * \mu_{R_{b_{n+1}}}(x, y) * \dots * \mu_{R_{b_{n+m}}}(x, y) \\ &= \mu_{R_{B'}}(x, y) * \mu_{R_{b_{n+1}}}(x, y) * \dots * \mu_{R_{b_{n+m}}}(x, y) \end{aligned}$$

Because $\forall b \in A, \mu_{R_b}(x, y) \in [0, 1]$, there must be $\mu_{R_B}(x, y) \leq \mu_{R_{B'}}(x, y)$.

Suppose $\forall x \in \delta_B(x)$, there exist $\mu_{R_B}(x, y) \geq \delta$, then $\delta \leq \mu_{R_B}(x, y) \leq \mu_{R_{B'}}(x, y)$, that is to say, $x \in \delta_{B'}(x)$. Therefore, $\delta_B(x) \subseteq \delta_{B'}(x)$.

Theorem 1 shows that the number of the elements in the similarity class will be changed with the variation of condition attribute set. The smaller the condition attribute set is, the more elements will be in the similarity class.

Theorem 2. Let $S = (U, A, V, F)$ be a set-valued information system, B be a subset of A , δ_1 and δ_2 be two thresholds. For $x \in U$, if $\delta_1 \leq \delta_2$, then $\delta_{1B}(x) \supseteq \delta_{2B}(x)$.

Proof. Suppose $\forall x \in \delta_{2B}(x)$, that is to say, $\mu_{R_B}(x, y) \geq \delta_2$. Because $\delta_1 \leq \delta_2$, then there must be $\mu_{R_B}(x, y) \geq \delta_1$, it can be inferred that $x \in \delta_{1B}(x)$. Therefore, $\delta_{1B}(x) \supseteq \delta_{2B}(x)$.

Theorem 2 indicates that we can control the elements in the similarity class by the threshold. The smaller the threshold is, the more elements will be in the similarity class.

Definition 6. Let $S = (U, A, V, F)$ be a set-valued information system, B be a subset of A . Given an arbitrary set $X \subseteq U$, the δ -upper approximation $\overline{B}_\delta(X)$ and the δ -lower approximation $\underline{B}_\delta(X)$ of X with respect to B are:

$$\overline{B}_\delta(X) = \{x \in U \mid \delta_B(x) \cap X \neq \emptyset\}, \tag{8}$$

$$\underline{B}_\delta(X) = \{x \in U \mid \delta_B(x) \subseteq X\} \tag{9}$$

where $\underline{B}_\delta(X)$ contains all the objects that can be classified into X definitely, and $\overline{B}_\delta(X)$ contains all the objects that can be classified into X approximately.

Theorem 3. Let $S = (U, A, V, F)$ be a set-valued information system, B and B' be two subsets of A , δ be a threshold. For $X \subseteq U$, if $B \supseteq B'$, then $\overline{B}_\delta(X) \subseteq \overline{B}'_\delta(X)$, and $\underline{B}_\delta(X) \supseteq \underline{B}'_\delta(X)$.

Proof. The proof comes directly from Theorem 1, and hence it is omitted here.

Theorem 3 shows that the number of the elements in the upper and lower approximation sets will be changed with the variation of the condition attribute set. The smaller the condition attribute set is, the more elements will be in the δ -upper approximation set and less in the δ -lower approximation set.

Theorem 4. Let $S = (U, A, V, F)$ be a set-valued information system, B be a subset of A , δ_1 and δ_2 be two thresholds. For $X \subseteq U$, if $\delta_1 \leq \delta_2$, then $\overline{B}_{\delta_1}(X) \supseteq \overline{B}_{\delta_2}(X)$, and $\underline{B}_{\delta_1}(X) \subseteq \underline{B}_{\delta_2}(X)$.

Proof. The proof comes directly from Theorem 2, and hence it is omitted here.

Theorem 4 indicates that we can control the elements in the upper and lower approximation sets by changing the threshold. The smaller the threshold is, the more elements will be in the δ -upper approximation set and less in the δ -lower approximation set.

4 Attribute Reduction of Set-Valued Decision Information System

Generally, most of decision information systems have some redundant attributes. These redundant attributes, on the one hand, waste storage space and reduce the efficiency of data processing. On the other hand, they may be our interference to make correct and concise decisions. Next, we will discuss how to reduce a set-valued decision information system.

Definition 7. Let $S = (U, C \cup D, V, F)$ be a set-valued decision information system, where U is a non-empty finite set of objects called the universe. C is a set of condition attribute, D is a decision attribute, F is a function that assigns particular values from attribute domain to objects. $U/D = \{d_1, d_2, \dots, d_m\}$ is a division of U . For $B \subseteq C$, the positive domain of D with respect to B is defined as:

$$POS_B(D) = \bigcup_{i=1}^m \underline{B}_\delta(d_i), \tag{10}$$

and the negative domain of D with respect to B is:

$$NEG_B(D) = U - POS_B(D), \quad (11)$$

where $POS_B(D)$ is the set of objects in U that can be classified into D definitely. $NEG_B(D)$ is the set of objects in U that can not be classified into D definitely.

Theorem 5. Let $S = (U, C \cup D, V, F)$ be a set-valued decision information system, B and B' be two subsets of C . If $B' \subseteq B$, then $POS_{B'}(D) \subseteq POS_B(D)$.

Proof. If $x \in POS_{B'}(D)$, there exist $d_i \in U/D$ such that $\delta_{B'}(x) \subseteq d_i$. According to Theorem 1, we have $\delta_B(x) \subseteq \delta_{B'}(x)$. So, $\delta_B(x) \subseteq d_i$, that is $x \in POS_B(D)$. Thus, $POS_{B'}(D) \subseteq POS_B(D)$.

Theorem 5 shows that the positive domain of D with respect to an attribute subset, is also a subset of the positive domain of D with respect to all the attributes. That means, if some attributes are deleted from a set-valued decision information system, the positive domain may decrease or remain unchanged. If the positive domain remains unchanged after an attribute is deleted, it means this attribute is redundant for keeping the positive domain. In other words, removing redundant attributes does not affect the correct classification ability of the system.

Definition 8. Let $S = (U, C \cup D, V, F)$ be a set-valued decision information system. For $\forall a \in C$, a is reducible in C with respect to D if $POS_{C-\{a\}}(D) = POS_C(D)$. Otherwise, a is irreducible in C with respect to D .

Definition 9. Let $S = (U, C \cup D, V, F)$ be a set-valued decision information system, and B be a subset of C . B is a reduction of C with respect to D if:

- (1) $POS_B(D) = POS_C(D)$, and
- (2) $\forall a \in B, POS_{B-\{a\}}(D) \neq POS_B(D)$.

According to Definitions 8 and 9, it is obvious that B , the reduction of C , has the same classification ability as C , and deleting any attributes from B will decrease the correct classification ability of the system. To obtain a reduction of a set-valued decision information system, we can develop the following algorithm.

Algorithm 1: Attribute Reduction of set-valued decision information system

Require: A set-valued decision information system S

Ensure: A reduction B

1: Let $B=C$

2: for each condition attribute $a \in C$ do:

3: if $POS_{B-\{a\}}(D) = POS_C(D)$ then:

4: Eliminate a from B

5: end if

6: end for

For a given set-valued decision information system, Algorithm 1 check each attribute by the conditions stated in Definition 9. If the conditions are satisfied, then the

attribute will be deleted. Otherwise it will be retained. Finally, we can get a reduction of the set-valued decision information system.

The following example illustrates how to form a reduction using Algorithm 1.

Example 1. For the set-valued decision information system S shown in Table 1, we can compute the reduction of S as follow.

Table 1. A set-valued decision information system

Objects	a_1	a_2	a_3	a_4	a_5	d
x_1	{2}	{1,2}	{1}	{2,3}	{1}	1
x_2	{1,2}	{3}	{1,2}	{2}	{1,2}	3
x_3	{1}	{2,3}	{2}	{1,2}	{1,2}	2
x_4	{1}	{2}	{2}	{2}	{1}	2
x_5	{3}	{2}	{1,2}	{1}	{3}	3
x_6	{1,3}	{2}	{1,2}	{1,2}	{2,3}	3
x_7	{2}	{1,2,3}	{1,2}	{2}	{1,3}	1
x_8	{1}	{3}	{2}	{2}	{2}	3

Let, $\delta = 1/16$, we have:

$$\begin{aligned}
 \delta_C(x_1) &= \{x_1\}, \delta_C(x_2) = \{x_2, x_8\} \\
 \delta_C(x_3) &= \{x_3, x_4, x_8\}, \delta_C(x_4) = \{x_3, x_4\} \\
 \delta_C(x_5) &= \{x_5, x_6\}, \delta_C(x_6) = \{x_5, x_6\} \\
 \delta_C(x_7) &= \{x_7\}, \delta_C(x_8) = \{x_2, x_3, x_8\} \\
 U/D &= \{\{x_1, x_7\}, \{x_3, x_4\}, \{x_2, x_5, x_6, x_8\}\} \\
 POS_C(D) &= \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}
 \end{aligned}$$

After a_1 is deleted, we have:

$$\begin{aligned}
 \delta_{C-a_1}(x_1) &= \{x_1\}, \delta_{C-a_1}(x_2) = \{x_2, x_3, x_8\} \\
 \delta_{C-a_1}(x_3) &= \{x_2, x_3, x_4, x_8\}, \delta_{C-a_1}(x_4) = \{x_3, x_4, x_7\} \\
 \delta_{C-a_1}(x_5) &= \{x_5, x_6\}, \delta_{C-a_1}(x_6) = \{x_5, x_6\} \\
 \delta_{C-a_1}(x_7) &= \{x_4, x_7\}, \delta_{C-a_1}(x_8) = \{x_2, x_3, x_8\}a_3 \\
 POS_{C-a_1}(D) &= \{x_1, x_5, x_6\}
 \end{aligned}$$

It can be found that the positive domain has changed after a_1 is deleted. So, a_1 is irreducible. Likewise, a_2, a_3, a_4 and a_5 are all irreducible. Then, the reduction is $\{a_1, a_2, a_3, a_4, a_5\}$. Let, it can be found that a_1 , and a_5 are irreducible, but and a_4 are reducible. Then, the reduction is $\{a_1, a_2, a_5\}$. According to the above analysis, we can get different reducts by changing the value of δ .

5 Experimental Results

The experiments in this section are to demonstrate the effectiveness of the method proposed in this paper. There are five groups of data sets used in the experiments and the information of all datasets is shown in Table 2.

Table 2. Data description

Source of datasets	Datasets	Number of attributes	Number of samples
weka	vote	17	435
weka	breast-cancer	10	286
UCI	annealing	33	798
UCI	audiology	70	200
UCI	zoo	17	17

In the experiments, we first reduce the five data sets using the tolerance relation, the fuzzy relation and the similarity relation proposed in this paper respectively, then J48 and SMO were used to make comparisons on classification accuracy with the reduction results. The reduction results gotten by different relations are shown in Table 3, and the classification accuracy comparisons are shown as Figs. 1 and 2:

Table 3. Attribute numbers after reduction

Dataset	Tolerance relation	Fuzzy relation	Similarity relation
vote	15	15	16
breast-cancer	9	9	9
annealing	9	10	14
audiology	21	22	63
zoo	13	13	14

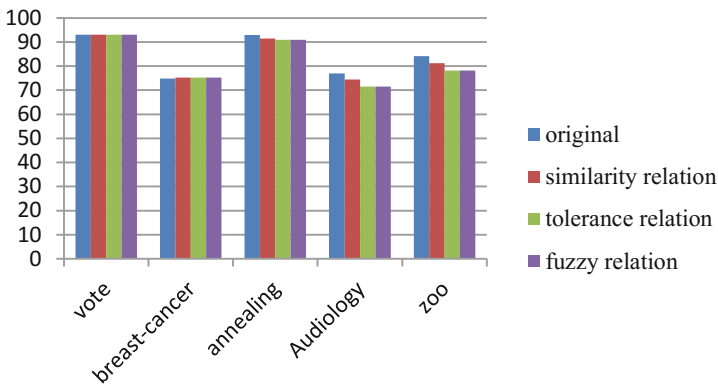


Fig. 1. Comparison of classification accuracy by J48

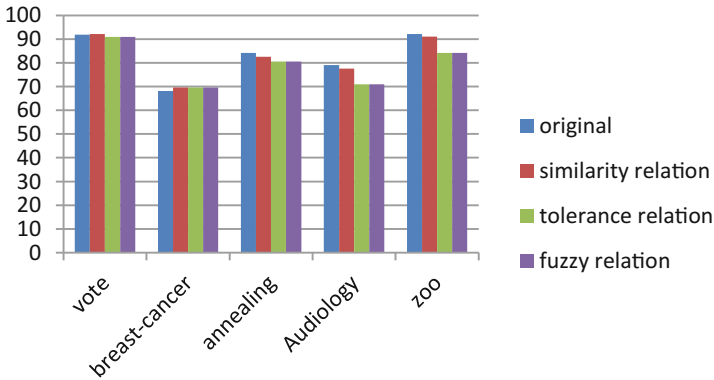


Fig. 2. Comparison of classification accuracy by SMO

It is known that only one or two attributes were removed by attribute reduction in vote and breast-cancer from Table 2, and the classification accuracy gotten by the proposed method and the existing methods has no obvious difference on these two datasets. In other three datasets, the results on attribute reduction are significant. By comparison, the method proposed in this paper retains more attributes than the existing methods. On the other hand, the classification accuracy gotten by the proposed method is almost the same as the classification accuracy of the original data, and is obviously higher than the classification accuracy of the other methods. According to the experimental results, we can draw the conclusion that the method proposed in this paper maintains more attributes than the existing methods, but it ensures that the classification accuracy of the system does not change significantly. Although the existing methods remove more attributes, the classification accuracy of the system is also greatly decreased. That means the existing methods remove some useful attributes, and it is unacceptable in some practical problems.

6 Conclusions

A lot of different methods were proposed to deal with set-valued information system. However, they are not appropriate when dealing with the set-valued data under the disjunctive semantics. To address this problem, a similarity relation based on probability was defined to describe the relation between two set-valued objects. Then, a corresponding attribute reduction algorithm based on keeping positive domain was proposed. In the end, a group of experiments were conducted to demonstrate the effectiveness of the proposed methods. The experimental results indicate that the existing methods can get a smaller reduction, but they may remove some useful attributes, thus reduce the classification accuracy of the system significantly. The proposed method retains more attributes, but all these attributes are useful, and it can always get the classification accuracy close to the original data. Because the threshold has an important influence on the classification accuracy when we used the proposed methods, it will be our future work how to choose an appropriate threshold.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (61472056, 61533020, 61751312, 61379114), the Social Science Foundation of the Chinese Education Commission (15XJA630003), the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1500416), Chongqing Research Program of Basic Research and Frontier Technology (cstc2017jcyjAX0406).

References

1. Pawlak, Z.: Rough set theory and its applications to data analysis. *Cybern. Syst.* **29**(29), 661–688 (2010)
2. Guan, Y.Y., Wang, H.K.: Set-valued information systems. *Inf. Sci.: Int. J.* **176**(17), 2507–2525 (2006)
3. Orłowska, E.: Logic of nondeterministic information. *Stud. Log.* **44**(1), 91–100 (1985)
4. Orłowska, E., Pawlak, Z.: Representation of nondeterministic information. *Theoret. Comput. Sci.* **29**(1), 27–39 (1984)
5. Sakai, H., Ishibashi, R., Nakata, M.: Lower and upper approximations of rules in non-deterministic information systems. In: Chan, C.-C., Grzymala-Busse, J.W., Ziarko, W. P. (eds.) *RSTC 2008. LNCS (LNAI)*, vol. 5306, pp. 299–309. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88425-5_31
6. Sakai, H., Ishibashi, R., Koba, K., Nakata, M.: Rules and apriori algorithm in non-deterministic information systems. In: Peters, J.F., Skowron, A., Rybiński, H. (eds.) *Transactions on Rough Sets IX. LNCS*, vol. 5390, pp. 328–350. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89876-4_18
7. Yao, Y.Y.: Information granulation and rough set approximation. *Int. J. Intell. Syst.* **16**(1), 87–104 (2001)
8. Yao, Y.Y., Noroozi, N.: A unified framework for set-based computations. In: *Proceedings of the 3rd International Workshop on Rough Sets and Soft Computing. The Society for Computer Simulation*, pp. 252–255(1995)
9. Lipski Jr., W.: On semantic issues connected with incomplete information databases. *Trans. Database Syst.* **4**(3), 262–296 (1979)
10. Lipski Jr., W.: On databases with incomplete information. *J. ACM* **28**(1), 41–70 (1981)
11. Zhang, W.X.: *Information Systems and Knowledge Discovery*. Science Press, Beijing (2003)
12. Zhang, W.X., Mi, J.S.: Incomplete information system and its optimal selections. *Comput. Math Appl.* **48**(5), 691–698 (2004)
13. Qian, Y., Dang, C., Liang, J., Tang, D.: Set-valued ordered information systems. *Inf. Sci.* **179**(16), 2809–2832 (2009)
14. Dai, J., Tian, H.: Fuzzy rough set model for set-valued data. *Fuzzy Sets Syst.* **229**, 54–68 (2013)
15. Wang, C.Y.: A note on a fuzzy rough set model for set-valued data. *Fuzzy Sets Syst.* **294**, 44–47 (2016)
16. Bao, Z., Yang, S.: Attribute reduction for set-valued ordered fuzzy decision system. In: *Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 96–99 (2014)

17. Wei, W., Cui, J., Liang, J., Wang, J.: Fuzzy rough approximations for set-valued data. *Inf. Sci.* **360**, 181–201 (2016)
18. Stefanowski, J., Tsoukiàs, A.: Incomplete information tables and rough classification. *Comput. Intell.* **17**(3), 545–566 (2001)
19. Nakata, M., Sakai, H.: Rough sets handling missing values probabilistically interpreted. In: Ślęzak, D., Wang, G., Szczuka, M., Düntsch, I., Yao, Y. (eds.) *RSFDGrC 2005*. LNCS (LNAI), vol. 3641, pp. 325–334. Springer, Heidelberg (2005). https://doi.org/10.1007/11548669_34