



Constructing the Optimal Approximation Sets of Rough Sets in Multi-granularity Spaces

Qinghua Zhang^{1,2}(✉), Fan Zhao^{1,2}, Xu Yubin², and Jie Yang¹

¹ Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
zhangqh@cqupt.edu.cn

² School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract. Rough set theory is an important tool to solve the uncertain problems. How to use the existing knowledge granules to approximately describe an uncertain target concept X has been a key issue. However, current research on theories and methods is still not comprehensive enough. $R_{0.5}(X)$, a kind of approximation sets of an uncertain concept, was proposed and analyzed in detail in our previous research work. However, whether $R_{0.5}(X)$ is the optimal approximation set of an uncertain concept X is still unable to determine. As a result, in this paper, based on the approximation of an uncertain concept, the existence of the optimal approximation set is explored. Then an optimal approximation set $R_{Best}(X)$ is proposed and discussed. At first, the definition of $R_{Best}(X)$ is defined. Then several comparative analysis between $R_{Best}(X)$ and other approximation sets is carried out. Next, operation properties of $R_{Best}(X)$ are presented and proved respectively. Finally, with changing knowledge granularity spaces, the change rules of the similarity between an uncertain set X and its $R_{Best}(X)$ are revealed.

Keywords: Rough sets · Uncertain concept · Similarity · Knowledge granularity · Granular computing

1 Introduction

Recently, computer technology and automatic control technology have rapidly developed, and research on the uncertain information system has attracted more and more researchers' attention [18, 22]. Fuzzy set theory, rough set theory and quotient space theory are three basic granular computing models which have been successfully applied to process uncertain information. As a simple computing model, rough set theory [4, 13] is an important method for handling uncertain problems as well as probability theory, fuzzy set theory and evidence theory. In the view of Pawlak's rough sets, people usually research how to acquire decision rules from the upper approximation set $\overline{R}(X)$ and the lower approximation

set $\underline{R}(X)$. Furthermore, many extended rough set models are proposed to deal with the real-life uncertain information, such as variable precision rough set model [26], probability rough set model [11], game-theoretic rough sets [2] and so on [10, 12, 25]. Pawlak and Skworn analyzed and summarized these extended models referred to [11]. Mi analyzed the variable precision rough set model and discussed how to use this model to obtain attribute reduction [8]. Yao and Ziarko et al., combining probability and inclusion degree, proposed probability rough set model and obtained many related results [17, 19, 20]. However, these methods mainly focus on constructing the extended approximation operators of traditional rough set model. There is little research on how to use the existing knowledge granules in knowledge base to construct an approximation set of X . Could we construct an approximation set which is more approximate to X than $\overline{R}(X)$ or $\underline{R}(X)$? And does an optimal approximation set exist? The first problem is solved in our previous work [21], and the second problem is our main motivations in this paper.

Based on above assumptions, the related models and results on the approximation set of an uncertain set were proposed in our other paper referred to [21, 23]. In these papers, the basic idea is translating rough sets into fuzzy sets according to the different membership degree of elements in boundary region and constructing an approximation set of an uncertain concept by using cut-set of the fuzzy set with some thresholds. With this construction method, the approximation sets with the existing knowledge granules can be obtained directly. In the literature [21], a general approximation set was constructed and it had many good properties. Experimental results show that $R_{0.5}(X)$ is a better model dealing with uncertain information systems. Better classification results could be obtained with $R_{0.5}(X)$. The amount of correct classification objects increases and amount of uncertain classification objects reduces. However, that $R_{0.5}(X)$ is the optimal approximation set of an uncertain set X is still unable to determine, and the related concepts and results on the optimal approximation were not presented in [21]. It is difficult to search for the optimal approximation set of an uncertain set directly. Based on the research referred to [21], through minimizing similarity between the target concept and its approximation sets, the optimal approximation set $R_{Best}(X)$ is defined, and an algorithm for constructing the optimal approximation set $R_{Best}(X)$ is proposed in this paper. And several comparative analysis between $R_{Best}(X)$ and other approximation sets is carried out. In addition, the operations properties of $R_{Best}(X)$ is analyzed. Finally we discuss the change rules of the similarity degree between X and its $R_{Best}(X)$ in different knowledge granularity levels.

The rest of this paper is organized as follows. In Sect. 2, the related basic concepts and preliminary knowledge are reviewed. The $R_{Best}(X)$ of an uncertain set X in rough approximation spaces is proposed in Sect. 3. Besides, several comparative analysis between $R_{Best}(X)$ and other approximation sets and many operation rules related the $R_{Best}(X)$ are given in Sect. 3. The change rules of the similarity degree between X and its $R_{Best}(X)$ in the different knowledge granularity levels are discussed in Sect. 4. Finally, the conclusions are drawn in Sect. 5.

2 Preliminaries

In order to better present the context of this paper, many preliminary concepts, definitions and results related to rough set and uncertainty measurement are reviewed as follows.

Definition 1 (Information table of knowledge system [9,14]). A knowledge system can be described as $S = \langle U, A, V, f \rangle$. U is the domain. $A = C \cup D$ is the set of all attributes. Subset C is the set of conditional attributes, and D is the set of decision attributes. $V = \cup_{r \in A} V_r$ is the set of attribute values. V_r describes the range of attribute values r where $r \in A$. $f : U \times A \rightarrow V$ is a function which describes attribute values of object x in U .

Definition 2 (Indiscernibility Relation [9,14]). For any attribute set $R \subseteq A$, an indiscernibility relation is defined as

$$IND(R) = \{(x, y) | (x, y) \in U^2 \wedge \forall b \in R (b(x) = b(y))\}.$$

Definition 3 (Upper Approximation Set and Lower Approximation Set [9,14]). Let $S = \langle U, A, V, f \rangle$ be a knowledge System, for any $X \subseteq U$ and $R \subseteq A$, the upper approximation set $\overline{R}(X)$ and the lower approximation set $\underline{R}(X)$ of X are defined as follows,

$$\begin{aligned} \overline{R}(X) &= \cup \{Y_i | Y_i \in U/IND(R) \wedge Y_i \cap X \neq \emptyset\}, \\ \underline{R}(X) &= \cup \{Y_i | Y_i \in U/IND(R) \wedge Y_i \subseteq X\}, \end{aligned}$$

where $U/IND(R) = \{X | X \subseteq U \wedge \forall x \in X, y \in Y, b \in R (b(x) = b(y))\}$ is a partition of equivalence relation R on U . The upper approximation set and the lower approximation set of X on R can be defined in another form as follows,

$$\begin{aligned} \overline{R}(X) &= \{x | x \in U \wedge [x]_R \cap X \neq \emptyset\}, \\ \underline{R}(X) &= \{x | x \in U \wedge [x]_R \subseteq X\}, \end{aligned}$$

where $[x]_R \in U/IND(R)$, and $[x]_R$ is an equivalence class of x on relation R . $\underline{R}(X)$ is a set of objects which certainly belong to U according to knowledge R ; $\overline{R}(X)$ is a set of objects which possibly belong to U according to knowledge R . $BND_R(X) = \overline{R}(X) - \underline{R}(X)$ is called as **Boundary region** of the target concept X on attribute set R . $POS_R(X) = \underline{R}(X)$ is called as **Positive region** of target concept X on attribute set R . $NEG_R(X) = U - \overline{R}(X)$ is called as **Negative region** of target concept X on attribute set R . $BND_R(X)$ is a set of objects which just possibly belong to target concept X .

Let U be a finite domain. Let $X \subseteq U$, $x \in U$, and the membership degree of x belong to set X is defined as

$$\mu_X^R(x) = \frac{|X \cap [x]_R|}{|[x]_R|},$$

obviously, $0 \leq \mu_X^R(x) \leq 1$.

Definition 4 (λ -Approximation Sets of X [21]). Let X be a subset (the target concept) of U , let

$$R_\lambda(X) = \{x|x \in U \wedge \mu_X^R(x) \geq \lambda\} \quad (1 \geq \lambda > 0),$$

then $R_\lambda(X)$ is called as λ -approximation sets of X . Let

$$R_{\cdot\lambda}(X) = \{x|x \in U \wedge \mu_X^R(x) > \lambda\} \quad (1 \geq \lambda > 0),$$

then $R_{\cdot\lambda}(X)$ is called as λ -strong approximation sets of X .

So when $\lambda = 0.5$, $R_{0.5}(X)$ is called as 0.5-approximation sets of X and $R_{\cdot 0.5}(X)$ is called as 0.5-strong approximation sets of X .

Definition 5 [14]. Let $U = \{x_1, x_2, \dots, x_n\}$ be a non-empty finite set, $P' = \{P'_1, P'_2, \dots, P'_l\}$ and $P'' = \{P''_1, P''_2, \dots, P''_m\}$ be two partition spaces on U . If $\forall P'_i \in P' (\exists P''_j \in P'' (P'_i \subseteq P''_j))$, then P' is finer than P'' , denoted by $P' \preceq P''$.

Definition 6 [14]. Let $U = \{x_1, x_2, \dots, x_n\}$ be a non-empty finite set, $P' = \{P'_1, P'_2, \dots, P'_l\}$ and $P'' = \{P''_1, P''_2, \dots, P''_m\}$ be two partition spaces on U . If $P' \preceq P''$, and $\exists P'_i \in P' (\exists P''_j \in P'' (P'_i \subset P''_j))(P'_i \subset P''_j)$, then P' is strictly finer than P'' , denoted by $P' \prec P''$.

Definition 7 (Similarity Degree [21]). Let A and B be two subsets of U , the mapping: $S : U \times U \rightarrow [0, 1]$. $S(A, B)$ is called as **similarity degree** between A and B , if and only if $S(A, B)$ satisfy the following properties:

- (1) For any $A, B \subseteq U$, $0 \leq S(A, B) \leq 1$ (Boundedness),
- (2) For any $A, B \subseteq U$, $S(A, B) = S(B, A)$ (Symmetry),
- (3) For any $A \subseteq U$, $S(A, A) = 1$; $S(A, B) = 0$ if and only if $A \cap B = \emptyset$.

Any formula that satisfy above (1), (2) and (3) is a similarity degree formula between two sets. In similarity measurement of rough sets, because of its universality and effectiveness, most experts and scholars have adopted a similarity degree formula in reference [7] as

$$S(A, B) = \frac{|A \cap B|}{|A \cup B|},$$

where $|\cdot|$ represents cardinality of elements in finite subset. Obviously, this formula satisfies above (1), (2) and (3).

3 Optimal Approximation Set of Rough Set

In our previous works, we find that $R_{0.5}(X)$, as an approximation set of X , has many excellent properties. However, whether $R_{0.5}(X)$ is the optimal approximation set of X when $0 \leq \lambda < 0.5$? Let us analyze according to the following example.

Table 1. Decision information table

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
a	1	1	1	1	1	1	1	1	1
b	1	1	0	1	1	0	1	1	0
c	1	0	0	1	0	0	1	0	0
d	1	1	1	0	0	0	0	0	0

Example 1. In a decision information table (Table 1), let $U = \{x_1, x_2, \dots, x_9\}$, the condition attribute set $C = \{a, b, c\}$ and the decision attribute set $D = \{d\}$.

According to rough set theory, the following partitions can be obtained easily,

$$IND(C) = \{\{x_1, x_4, x_7\}, \{x_2, x_5, x_8\}, \{x_3, x_6, x_9\}\},$$

$$IND(D) = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7, x_8, x_9\}\}.$$

Let $X_1 = \{x_1, x_2, x_3\}$, $X_2 = \{x_4, x_5, x_6, x_7, x_8, x_9\}$. According to the definition $\mu_X^R(x) = \frac{|[x]_R \cap X|}{|[x]_R|}$, a fuzzy set can be obtained as

$$F_{X_1}(U) = \left\{ \frac{1/3}{x_1}, \frac{1/3}{x_2}, \frac{1/3}{x_3}, \frac{1/3}{x_4}, \frac{1/3}{x_5}, \frac{1/3}{x_6}, \frac{1/3}{x_7}, \frac{1/3}{x_8}, \frac{1/3}{x_9} \right\}.$$

Then

$$R_{0.5}(X_1) = \emptyset, \quad R_{0.3}(X_1) = U.$$

While

$$S(X_1, R_{0.5}(X_1)) = \frac{|X_1 \cap R_{0.5}(X_1)|}{|X_1 \cup R_{0.5}(X_1)|} = \frac{0}{3} = 0, \quad S(X_1, R_{0.3}(X_1)) = \frac{1}{3}.$$

So, $S(X_1, R_{0.3}(X_1)) > S(X_1, R_{0.5}(X_1))$.

In the same way, the following fuzzy set can be obtained,

$$F_{X_2}(U) = \left\{ \frac{2/3}{x_1}, \frac{2/3}{x_2}, \frac{2/3}{x_3}, \frac{2/3}{x_4}, \frac{2/3}{x_5}, \frac{2/3}{x_6}, \frac{2/3}{x_7}, \frac{2/3}{x_8}, \frac{2/3}{x_9} \right\}.$$

Then we have $R_{0.5}(X_2) = U$, $R_{0.3}(X_2) = U$. Here we have $S(X_2, R_{0.3}(X_2)) = S(X_2, R_{0.5}(X_2))$.

So, we find that the approximation set $R_{0.5}(X_1)$ is not the optimal approximation set of X_1 , on the contrary, $R_{0.5}(X_2)$ is the optimal approximation set of X_2 . Therefore, the approximation set $R_{0.5}(X)$ may not be the optimal approximation set of X and an optimal approximation set of X in rough approximation spaces must exist. Thus, based on the membership degree function $\mu_X^R(x)$, the optimal approximation set of X in rough approximation space is proposed through minimizing similarity between the target concept and its approximation sets.

Definition 8. (Optimal approximation set) Let X be a subset (target concept) of U , $R_\lambda(X)$ be a λ - approximation set of X , and $S_{Best} = \max_{0 < \lambda \leq 1} \{S(X, R_\lambda(X))\}$. For any $\lambda(0 < \lambda \leq 1)$, if $S(X, R_\lambda(X)) = S_{Best}$, then the $R_\lambda(X)$ is called the optimal approximation set of X , denoted by $R_{Best}(X)$. Namely, $S(X, R_{Best}(X)) = S_{Best}$.

According to Definition 8, we know if $R_\lambda(X)$ is the $R_{Best}(X)$, λ must be in the interval $(0, 0.5]$. In order to more clearly show the Definition 8, an example with a decision information table (Table 2) is presented.

Example 2. Let domain $U = \{x_1, x_2, \dots, x_{15}\}$, the condition attribute set $C = \{a, b, c\}$ and the decision attribute set $D = \{d\}$.

Table 2. Decision information table

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}
a	1	1	1	1	2	2	1	1	1	2	1	2	1	2	2
b	2	2	0	0	1	2	2	0	0	1	2	2	0	1	2
c	0	0	2	2	1	1	0	2	2	1	0	1	2	1	1
d	1	1	1	1	1	2	2	2	2	2	3	3	3	3	3

According to rough set theory, the following partition is obtained easily,

$$U/IND(D) = \{\{x_1, x_2, x_3, x_4, x_5\}, \{x_6, x_7, x_8, x_9, x_{10}\}, \{x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\}\}.$$

Here, three decision concepts induced by decision attribute set are generated, and they are $X_1 = \{x_1, x_2, x_3, x_4, x_5\}$, $X_2 = \{x_6, x_7, x_8, x_9, x_{10}\}$ and $X_3 = \{x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\}$ respectively. For the decision concept X_1 , computing $U/IND(C)$,

$$U/IND(C) = \{\{x_1, x_2, x_7, x_{11}\}, \{x_3, x_4, x_8, x_9, x_{13}\}, \{x_5, x_{10}, x_{14}\}, \{x_6, x_{12}, x_{15}\}\}.$$

Then let $Y_1 = \{x_1, x_2, x_7, x_{11}\}$, $Y_2 = \{x_3, x_4, x_8, x_9, x_{13}\}$, $Y_3 = \{x_5, x_{10}, x_{14}\}$, $Y_4 = \{x_6, x_{12}, x_{15}\}$. Computing the membership degree $\mu(x)$ of $x (x \in U)$, where $\mu(x) = \frac{|Y_i \cap X|}{|Y_i|}$, that is to say, every object in equivalence class Y_i has same membership degree. For X_1 , the membership degrees are shown as follows,

- (1) For equivalence class Y_1 , the membership degree is $1/2$;
- (2) For equivalence class Y_2 , the membership degree is $2/5$;
- (3) For equivalence class Y_3 , the membership degree is $1/3$;
- (4) For equivalence class Y_4 , the membership degree is 0 .

Computing $R_{0.5}(X_1)$ and $R_{\mu_i}(X_1)$ and sorting the μ_i , then we can get

$$R_{1/3}(X_1) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{13}, x_{14}\};$$

$$R_{2/5}(X_1) = \{x_1, x_2, x_3, x_4, x_7, x_8, x_9, x_{11}, x_{13}\}; R_{0.5}(X_1) = \{x_1, x_2, x_7, x_{11}\}.$$

Further, we can obtain:

$$S(X_1, R_{1/3}(X_1)) = \frac{5}{12}, S(X_1, R_{2/5}(X_1)) = \frac{2}{5}, S(X_1, R_{0.5}(X_1)) = \frac{2}{7}.$$

We find that $S_{1/3}(X_1)$ is maximum value, that is to say, the approximation set $R_{1/3}(X_1)$, is the optimal approximation set of X_1 . Then we can obtain $R_{Best}(X_1)$ as follows,

$$R_{Best}(X_1) = R_{1/3}(X_1) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{13}, x_{14}\}.$$

Similarly,

$$R_{Best}(X_2) = \{x_3, x_4, x_5, x_6, x_8, x_9, x_{10}, x_{12}, x_{13}, x_{14}, x_{15}\};$$

$$R_{Best}(X_3) = \{x_5, x_6, x_{10}, x_{12}, x_{14}, x_{15}\}.$$

The purpose of selecting $R_{Best}(X)$ is to characterize a target concept and further acquire rules. So compared with $\bar{R}(X)$, $\underline{R}(X)$ and $R_{0.5}(X)$, what advantages does the $R_{Best}(X)$ have? Then, relative analysis is shown as follows:

(Continue) Example 3. In Table 2, we can get three decision concepts $X_1 = \{x_1, x_2, x_3, x_4, x_5\}$, $X_2 = \{x_6, x_7, x_8, x_9, x_{10}\}$ and $X_3 = \{x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\}$. With these decision concepts, we can obtain,

$$\underline{R}(X_1) = \emptyset, \underline{R}(X_2) = \emptyset, \underline{R}(X_3) = \emptyset;$$

$$\bar{R}(X_1) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{13}, x_{14}\},$$

$$\bar{R}(X_2) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\},$$

$$\bar{R}(X_3) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\};$$

$$R_{0.5}(X_1) = \{x_1, x_2, x_7, x_{11}\}, R_{0.5}(X_2) = \emptyset, R_{0.5}(X_3) = \emptyset;$$

$$R_{Best}(X_1) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}, x_{11}, x_{13}, x_{14}\},$$

$$R_{Best}(X_2) = \{x_3, x_4, x_5, x_6, x_8, x_9, x_{10}, x_{12}, x_{13}, x_{14}, x_{15}\},$$

$$R_{Best}(X_3) = \{x_5, x_6, x_{10}, x_{12}, x_{14}, x_{15}\}.$$

From the decision information Table 2 we can acquire many decision rules based on $\underline{R}(X)$ are shown as follows,

- (1) For X_1 , the corresponding approximation set is \emptyset ;
- (2) For X_2 , the corresponding approximation set is \emptyset ;
- (3) For X_3 , the corresponding approximation set is \emptyset .

We can acquire many decision rules based on $\bar{R}(X)$ are shown as follows,

- (1) For X_1 , the decision rule is $(a = 1 \wedge b = 2 \wedge c = 0) \vee (a = 1 \wedge b = 0 \wedge c = 2) \vee (a = 2 \wedge b = 1 \wedge c = 1) \rightarrow d = 1$;
- (2) For X_2 , the decision rule is $(a = 2 \wedge b = 2 \wedge c = 1) \vee (a = 1 \wedge b = 2 \wedge c = 0) \vee (a = 1 \wedge b = 0 \wedge c = 2) \vee (a = 2 \wedge b = 1 \wedge c = 1) \rightarrow d = 2$;
- (3) For X_3 , the decision rule is $(a = 1 \wedge b = 2 \wedge c = 0) \vee (a = 2 \wedge b = 2 \wedge c = 1) \vee (a = 1 \wedge b = 0 \wedge c = 2) \vee (a = 2 \wedge b = 1 \wedge c = 1) \rightarrow d = 3$.

We can acquire many decision rules based on $R_{0.5}(X)$ are shown as follows,

- (1) For X_1 , the decision rule is $(a = 1 \wedge b = 2 \wedge c = 0) \rightarrow d = 1$;
- (2) For X_2 , the corresponding approximate set is \emptyset ;
- (3) For X_3 , the corresponding approximate set is \emptyset .

We can acquire many decision rules based on $R_{Best}(X)$ are shown as follows,

- (1) For X_1 , the decision rule is $(a = 1 \wedge b = 2 \wedge c = 0) \vee (a = 1 \wedge b = 0 \wedge c = 2) \vee (a = 2 \wedge b = 2 \wedge c = 1) \vee (a = 2 \wedge b = 1 \wedge c = 1) \rightarrow d = 1$;
- (2) For X_2 , the decision rule is $(a = 2 \wedge b = 1 \wedge c = 1) \vee (a = 1 \wedge b = 0 \wedge c = 2) \vee (a = 2 \wedge b = 2 \wedge c = 1) \rightarrow d = 2$;
- (3) For X_3 , the decision rule is $(a = 2 \wedge b = 1 \wedge c = 1) \vee (a = 2 \wedge b = 2 \wedge c = 1) \rightarrow d = 3$.

Table 3. Comparative analysis

		Supporting amount	Wrong amount	Unrecognized amount
$\underline{R}(X)$	X_1	0	0	5
	X_2	0	0	5
	X_3	0	0	5
$\overline{R}(X)$	X_1	5	7	0
	X_2	5	10	0
	X_3	5	10	0
$R_{0.5}(X)$	X_1	2	2	3
	X_2	0	0	5
	X_3	0	0	5
$R_{Best}(X)$	X_1	5	7	0
	X_2	4	7	1
	X_3	3	3	2

A comparative analysis Table 3 is constructed according to the above decision rules. From these above rules acquired from $\underline{R}(X)$, $\overline{R}(X)$, $R_{0.5}(X)$ and $R_{Best}(X)$ and Table 3, the qualitative and quantitative comparisons could be made. It obvious that many objects can not determine decision classification if the decision rules are acquired based on $R_{0.5}(X)$ and $\underline{R}(X)$, and this is not an expected result in actual decision problems. Though each object belongs to a certain decision classification if the decision rules are acquired based on $\overline{R}(X)$ the objects in the boundary cannot be assigned to a certain decision classification, and the error classifications will be produced. Compared with $R_{0.5}(X)$ and $\underline{R}(X)$, the amount of correct classification objects increases, and amount of uncertain classification objects reduces if the decision rules are acquired based on $R_{Best}(X)$. Compared with $\overline{R}(X)$, although the amount of correct classification objects weakly declines, the amount of error classification objects also reduces if the decision rules are acquired based on $R_{Best}(X)$.

According to above comparison results, we find that the decision rules based on $R_{Best}(X)$ have more powerful classification ability than the decision rules

based on $\underline{R}(X)$, $\overline{R}(X)$ and $R_{0.5}(X)$. The $R_{Best}(X)$ provide a novel perspective for approximate characterization of a target concept in multi-granularity spaces. Furthermore, it would be an effective method that could be suitable to real-life knowledge discovery from the uncertain information systems.

4 Operation Properties of $R_{Best}(X)$

It is well known that $\underline{R}(X)$ and $\overline{R}(X)$ have many important operation properties as literature [8, 24]. Now, we will prove that the optimal approximation set $R_{Best}(X)$ has many similar operation properties with the upper approximation set and lower approximation set also. For convenience, let $R_{Best}(X) = R_k(X)$, ($0 < k \leq 0.5$), U be a finite domain, and X, Y be two subsets on U , we have

- (1) $R_k(\sim X) = \sim R_{1-k}(X)$, $\overline{R}_k(\sim X) = \sim \overline{R}_{1-k}(X)$;
- (2) if $X \subseteq Y$, then $R_k(X) \subseteq R_k(Y)$, $\overline{R}_k(X) \subseteq \overline{R}_k(Y)$;
- (3) $R_k(X \cap Y) \subseteq R_k(X) \cap R_k(Y)$, $\overline{R}_k(X \cap Y) \subseteq \overline{R}_k(X) \cap \overline{R}_k(Y)$;
- (4) $R_k(X \cup Y) \supseteq R_k(X) \cup R_k(Y)$, $\overline{R}_k(X \cup Y) \supseteq \overline{R}_k(X) \cup \overline{R}_k(Y)$.

Proof. (1) Because

$$\begin{aligned} R_k(\sim X) &= \{x | \mu_{\sim X}^R(x) \geq k\} = \left\{x \mid \frac{|[x]_R \cap (\sim X)|}{|[x]_R|} \geq k\right\} = \left\{x \mid \frac{|[x]_R \cap (U - X)|}{|[x]_R|} \geq k\right\} \\ &= \left\{x \mid 1 - \frac{|[x]_R \cap X|}{|[x]_R|} \geq k\right\} = \left\{x \mid \frac{|[x]_R \cap X|}{|[x]_R|} \leq 1 - k\right\} = \sim R_{1-k}(X), \end{aligned}$$

Similarly, $\overline{R}_k(\sim X) = \sim \overline{R}_{1-k}(X)$ holds. Hence, the proposition (1) is proved.

(2) $\forall x \in R_k(X)$, we have $[x]_R$ satisfying $\frac{|[x]_R \cap X|}{|[x]_R|} \geq k$; because $X \subseteq Y$, we have $|[x]_R \cap X| \leq |[x]_R \cap Y|$, then $\frac{|[x]_R \cap X|}{|[x]_R|} \leq \frac{|[x]_R \cap Y|}{|[x]_R|}$, then $\frac{|[x]_R \cap Y|}{|[x]_R|} \geq k$. So we can get $x \in R_k(Y)$, and then $R_k(X) \subseteq R_k(Y)$. Similarly, $\forall x \in \overline{R}_k(X)$, we have $[x]_R$ satisfying $\frac{|[x]_R \cap X|}{|[x]_R|} > k$; because $X \subseteq Y$, we have $|[x]_R \cap X| \leq |[x]_R \cap Y|$, then $\frac{|[x]_R \cap X|}{|[x]_R|} \leq \frac{|[x]_R \cap Y|}{|[x]_R|}$, and we can get $\frac{|[x]_R \cap Y|}{|[x]_R|} > k$. Therefore, we have $x \in \overline{R}_k(Y)$, and then $\overline{R}_k(X) \subseteq \overline{R}_k(Y)$. So, the proposition (2) is proved successfully.

(3) Because $X \cap Y \subseteq X$ and $X \cap Y \subseteq Y$, according to proposition (2) we have $R_k(X \cap Y) \subseteq R_k(X)$ and $R_k(X \cap Y) \subseteq R_k(Y)$, and we have $R_k(X \cap Y) \subseteq R_k(X) \cap R_k(Y)$. Similarly, we can get $\overline{R}_k(X \cap Y) \subseteq \overline{R}_k(X) \cap \overline{R}_k(Y)$. So the proposition (3) holds.

(4) Because $X \subseteq X \cup Y$ and $Y \subseteq X \cup Y$, according to proposition (2), $R_k(X) \subseteq R_k(X \cup Y)$ and $R_k(Y) \subseteq R_k(X \cup Y)$, so $R_k(X \cup Y) \supseteq R_k(X) \cup R_k(Y)$ is held. Similarly because $X \subseteq X \cup Y$ and $Y \subseteq X \cup Y$, both $\overline{R}_k(X) \subseteq \overline{R}_k(X \cup Y)$ and $\overline{R}_k(Y) \subseteq \overline{R}_k(X \cup Y)$ are held, so we have $\overline{R}_k(X \cup Y) \supseteq \overline{R}_k(X) \cup \overline{R}_k(Y)$. Then the proposition (4) is proved successfully.

5 Change Rules of $R_{Best}(X)$ with Changing Knowledge Granularity

Currently, in different knowledge granularity levels, change rules of rough set uncertainty is one of important issues to measure the uncertainty of knowledge

[1, 3, 5, 6, 15, 16]. Therefore, in different knowledge granularity levels, change rules of $S(X, R_{Best}(X))$ are also focus on our attention. Next, we will discuss the change rules in detail. Firstly, suppose a, b, c, d, e and f be all real number, and some basic results and lemmas are reviewed in order to discuss the relevant theorems easily.

Lemma 1. [21] *If $0 < a < b$ and $0 < c < d$, then $a/b < (a + d)/(b + c)$.*

Lemma 2. [21] *If $f/e = (b + d)/(a + c)$ and $b/a < f/e$, then $d/c > f/e$.*

Lemma 3. [19] *Let $0 < c < a, 0 < d < b$. If $a/b \geq c/d$, then $a/b \leq (a - c)/(b - d)$. On the contrary, if $a/b \leq c/d$, then $a/b \geq (a - c)/(b - d)$.*

Next, we would discuss the relationship between $S(X, R_{Best}(X))$ and $S(X, R'_{Best}(X))$ in different knowledge granularity levels. Let $R_{Best}(X) = R_\lambda(X)$ then $(0 < \lambda \leq 0.5)$. And let $[x_{i_1}]_R, [x_{i_2}]_R, \dots, [x_{i_k}]_R$ be the equivalence classes induced by an equivalence relation R on U , and $[x_{i_1}]_{R'}, [x_{i_2}]_{R'}, \dots, [x_{i_k}]_{R'}$ be the equivalence classes induced by another equivalence relation R' on U . If the partition U/R' is finer than U/R namely $U/R' \preceq U/R$, for any $x \in U$, there is $[x]_{R'} \subseteq [x]_R$. For convenience, let $R_\lambda(X) = R(X) \cup [x_{i_1}]_R \cup [x_{i_2}]_R \cup \dots \cup [x_{i_k}]_R$, and $BND_R(X) = [x_{i_1}]_R \cup [x_{i_2}]_R \cup \dots \cup [x_{i_m}]_R$. The equivalence classes divided into finer equivalence classes (sub-granules) may be in $NEG_R(X), POS_R(X)$ or $BND_R(X)$. When the equivalence classes divided into finer equivalence classes are in $NEG_R(X)$ or $POS_R(X)$, $S(X, R_\lambda(X)) = S(X, R'_\lambda(X))$ is held. Next we will focus on $S(X, R'_\lambda(X))$ when the equivalence classes divided into finer equivalence classes are in $BND_R(X)$. For simplicity, suppose there is only one granule subdivided to two disjoint granules and others remain unchanged. That is to say, suppose $[x_{i_t}]_R = [x_{i_t}^1]_{R'} \cup [x_{i_t}^2]_{R'}$. This situation will be discussed as follows.

Theorem 1. *If $\lambda = 0.5, 1 \leq t \leq k$, that is $[x_{i_t}]_R \subset R_\lambda(X)$:*
 (1) *If $[x_{i_t}^1]_{R'} \subset R'_\lambda(X), [x_{i_t}^2]_{R'} \subset R'_\lambda(X)$, then $S(X, R_\lambda(X)) = S(X, R'_\lambda(X))$;*
 (2) *If $[x_{i_t}^1]_{R'} \subset R'_\lambda(X), [x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$, then $S(X, R_\lambda(X)) \leq S(X, R'_\lambda(X))$.*

Proof. $\forall x \in R_{0.5}(X)$, we have $\mu_X^R(x) = \frac{|[x]_R \cap X|}{|[x]_R|} \geq 0.5$. Then we can obtain,

$$R_{0.5}(X) = \{x | \mu_X^R(x) \geq 0.5\} = \{x | \mu_X^R(x) = 1\} \cup \{x | 0.5 \leq \mu_X^R(x) < 1\}.$$

Obviously, $\{x | \mu_X^R(x) = 1\} = \underline{R}(X)$, and then Let $\{x | 0.5 \leq \mu_X^R(x) < 1\} = [x_{i_1}]_R \cup [x_{i_2}]_R \cup \dots \cup [x_{i_k}]_R$, we have

$$X \cap R_{0.5}(X) = X \cap (\underline{R}(X) \cup [x_{i_1}]_R \cup [x_{i_2}]_R \cup \dots \cup [x_{i_k}]_R).$$

Since the intersection sets between any two elements in $\underline{R}(X), [x_{i_1}]_R, [x_{i_2}]_R, \dots, [x_{i_k}]_R$ is empty set, we have

$$\begin{aligned} |X \cap R_{0.5}(X)| &= |X \cap \underline{R}(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_n}]_R| \\ &= |\underline{R}(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_n}]_R|. \end{aligned}$$

Since $X \cup R_\lambda(X) = X \cup ([x_{i_1}]_R - X) \cup ([x_{i_2}]_R - X) \cup \dots \cup ([x_{i_k}]_R - X)$ and the intersection between any two elements in X , $([x_{i_1}]_R - X)$, $([x_{i_2}]_R - X)$, \dots , $([x_{i_k}]_R - X)$ is empty set, we have

$$|X \cup R_{0.5}(X)| = |X| + |([x_{i_1}]_R - X)| + |([x_{i_2}]_R - X)| + \dots + |([x_{i_k}]_R - X)|.$$

Therefore,

$$S(X, R_{0.5}(X)) = \frac{|R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|}.$$

While $[x_{i_t}^1]_{R'} \subset R'_\lambda(X)$ and $[x_{i_t}^2]_{R'} \subset R'_\lambda(X)$, we can get

$$\begin{aligned} & S(X, R'_{0.5}(X)) \\ &= \frac{|R'(X)| + |X \cap [x_{i_1}]_{R'}| + \dots + |X \cap [x_{i_t}^1]_{R'}| + |X \cap [x_{i_t}^2]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + \dots + |[x_{i_t}^1]_{R'} - X| + |[x_{i_t}^2]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X|} \\ &= \frac{|R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|} \\ &= S(X, R_\lambda(X)). \end{aligned}$$

So the part (1) is proved successfully.

For the part (2) when $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$, we can have the equality as follows,

$$S(X, R'_{0.5}(X)) = \frac{|R'(X)| + |X \cap [x_{i_1}]_{R'}| + \dots + |X \cap [x_{i_t}^1]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + \dots + |[x_{i_t}^1]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X|}.$$

With Lemma 2, we have the inequality $\frac{|X \cap [x_{i_t}^1]_{R'}|}{|[x_{i_t}^1]_{R'} - X|} > \frac{|X \cap [x_{i_t}]_{R'}|}{|[x_{i_t}]_{R'} - X|}$, let $\frac{|X \cap [x_{i_t}^1]_{R'}|}{|[x_{i_t}^1]_{R'} - X|} = \frac{|X \cap [x_{i_t}]_{R'}| + p}{|[x_{i_t}]_{R'} - X| + q}$, then $\frac{p}{q} > \frac{|X \cap [x_{i_t}]_{R'}|}{|[x_{i_t}]_{R'} - X|}$, then we can get

$$\begin{aligned} & S(X, R'_\lambda(X)) \\ &= \frac{|R'(X)| + |X \cap [x_{i_1}]_{R'}| + |X \cap [x_{i_2}]_{R'}| + \dots + |X \cap [x_{i_t}^1]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + |[x_{i_2}]_{R'} - X| + \dots + |[x_{i_t}^1]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X|} \\ &= \frac{|R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R| + p}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X| + q}. \end{aligned}$$

We know

$$S(X, R_{0.5}(X)) = \frac{|R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|},$$

according to Definition 6 we have $0 \leq S(X, R_{0.5}(X)) \leq 1$, then we can easily get the inequality as follow,

$$\begin{aligned} & |R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R| \\ & \leq |X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|. \end{aligned}$$

Since $\frac{p}{q} > \frac{|X \cap [x_{i_t}]_{R'}|}{|[x_{i_t}]_{R'} - X|}$ we have $p > q$. With Lemma 1,

$$\frac{|\underline{R}(X)| + |X \cap [x_{i_1}]_R| + \dots + |X \cap [x_{i_k}]_R| + p}{|X| + |[x_{i_1}]_R - X| + \dots + |[x_{i_k}]_R - X| + q} \geq \frac{|R(X)| + |X \cap [x_{i_1}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + \dots + |[x_{i_k}]_R - X|},$$

namely $S(X, R_{0.5}(X)) \leq S(X, R'_{0.5}(X))$. Hence the part (2) is proved successfully.

Theorem 1 show that when $\lambda = 0.5$, no matter the equivalence classes divided into many sub-granules are in $NEG_R(X)$, $POS_R(X)$ or $BND_R(X)$, similarity degree $S(X, R'_\lambda(X))$ is not smaller than $S(X, R_\lambda(X))$.

Theorem 2. *If $0 < \lambda < 0.5$, $k < t \leq m$, that is $[x_{i_t}]_R \not\subset R_\lambda(X)$:*

(1) *If $[x_{i_t}^1]_{R'} \subset R'_\lambda(X)$, $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$ and $S(X, R_\lambda(X)) \leq |X \cap [x_{i_t}^1]_{R'}| / |[x_{i_t}^1]_{R'} - X|$, then $S(X, R_\lambda(X)) \leq S(X, R'_\lambda(X))$;*

(2) *If $[x_{i_t}^1]_{R'} \not\subset R'_\lambda(X)$, $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$, then $S(X, R_\lambda(X)) = S(X, R'_\lambda(X))$.*

Proof. According to the Theorem 1 we can get the equality as follow,

$$S(X, R_\lambda(X)) = \frac{|R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|}.$$

Since $[x_{i_t}^1]_{R'} \subset R'_\lambda(X)$ and $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$ we can obtain the following equality,

$$\begin{aligned} S(X, R'_\lambda(X)) &= \frac{|X \cap R'_\lambda(X)|}{|X \cup R'_\lambda(X)|} \\ &= \frac{|R(X)| + |X \cap [x_{i_1}]_{R'}| + |X \cap [x_{i_2}]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}| + |X \cap [x_{i_t}^1]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + |[x_{i_2}]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X| + |[x_{i_t}^1]_{R'} - X|} \\ &= \frac{|R(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R| + |X \cap [x_{i_t}^1]_{R'}|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X| + |[x_{i_t}^1]_{R'} - X|}. \end{aligned}$$

Let $S(X, R_\lambda(X)) = k$, $|R(X)| + |X \cap [x_{i_1}]_R| + \dots + |X \cap [x_{i_k}]_R| = k_1$ and $|X| + |[x_{i_1}]_R - X| + \dots + |[x_{i_k}]_R - X| = k_2$, we can get the equality as follow,

$$S(X, R'_\lambda(X)) = \frac{k_1 + |X \cap [x_{i_t}^1]_{R'}|}{k_2 + |[x_{i_t}^1]_{R'} - X|}.$$

For $S(X, R_\lambda(X)) = k = \frac{k_1}{k_2} \leq \frac{|X \cap [x_{i_t}^1]_{R'}|}{|[x_{i_t}^1]_{R'} - X|}$, and according to Lemma 2 we can easily obtain the following inequality,

$$S(X, R'_\lambda(X)) = \frac{k_1 + |X \cap [x_{i_t}^1]_{R'}|}{k_2 + |[x_{i_t}^1]_{R'} - X|} \geq \frac{k_1 + k_1}{k_2 + k_2} = S(X, R_\lambda(X)).$$

Therefore the part (1) is proved.

For the part (2), since $[x_{i_t}^1]_{R'} \not\subset R'_\lambda(X)$ and $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$, we can easily get the equality,

$$\begin{aligned} S(X, R'_\lambda(X)) &= \frac{|X \cap R'_\lambda(X)|}{|X \cup R'_\lambda(X)|} \\ &= \frac{|\underline{R}'(X)| + |X \cap [x_{i_1}]_{R'}| + |X \cap [x_{i_2}]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + |[x_{i_2}]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X|} \\ &= \frac{|\underline{R}(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|} \\ &= S(X, R_\lambda(X)). \end{aligned}$$

Therefore the part (2) is proved completely.

Note: since $[x_{i_t}]_R \not\subset R_\lambda(X)$, according to the Lemma 2, the inclusion relations $[x_{i_t}^1]_{R'} \subset R'_\lambda(X)$ and $[x_{i_t}^2]_{R'} \subset R'_\lambda(X)$ are not existed.

Theorem 3. *If $0 < \lambda < 0.5$, $k < t \leq m$, that is $[x_{i_t}]_R \subset R_\lambda(X)$:*

(1) *If $[x_{i_t}^1]_{R'} \subset \underline{R}'_\lambda(X)$, $[x_{i_t}^2]_{R'} \subset \underline{R}'_\lambda(X)$, then $S(X, R_\lambda(X)) = S(X, R'_\lambda(X))$.*

(2) *If $[x_{i_t}^1]_{R'} \subset R'_\lambda(X)$, $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$, as well as $S(X, R_\lambda(X)) \geq |X \cap [x_{i_t}^2]_{R'}| / |[x_{i_t}^2]_{R'} - X|$, then $S(X, R_\lambda(X)) \leq S(X, R'_\lambda(X))$.*

Proof. For the part (1), since $[x_{i_t}^1]_{R'} \subset R'_\lambda(X)$, and $[x_{i_t}^2]_{R'} \subset R'_\lambda(X)$, we can easily get the equality as follow,

$$\begin{aligned} S(X, R'_\lambda(X)) &= \frac{|X \cap R'_\lambda(X)|}{|X \cup R'_\lambda(X)|} \\ &= \frac{|\underline{R}'(X)| + |X \cap [x_{i_1}]_{R'}| + |X \cap [x_{i_2}]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + |[x_{i_2}]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X|} \\ &= \frac{|\underline{R}(X)| + |X \cap [x_{i_1}]_R| + |X \cap [x_{i_2}]_R| + \dots + |X \cap [x_{i_k}]_R|}{|X| + |[x_{i_1}]_R - X| + |[x_{i_2}]_R - X| + \dots + |[x_{i_k}]_R - X|} = S(X, R_\lambda(X)). \end{aligned}$$

Hence the part (1) is proved.

For the part (2), since $[x_{i_t}^2]_{R'} \not\subset R'_\lambda(X)$, we can easily get the equality as follow,

$$\begin{aligned} S(X, R'_\lambda(X)) &= \frac{|X \cap R'_\lambda(X)|}{|X \cup R'_\lambda(X)|} \\ &= \frac{|\underline{R}'(X)| + |X \cap [x_{i_1}]_{R'}| + |X \cap [x_{i_2}]_{R'}| + \dots + |X \cap [x_{i_t}^1]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + |[x_{i_2}]_{R'} - X| + \dots + |[x_{i_t}^1]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X|} \\ &= \frac{|\underline{R}'(X)| + |X \cap [x_{i_1}]_{R'}| + \dots + |X \cap [x_{i_t}]_{R'}| + \dots + |X \cap [x_{i_k}]_{R'}| - |X \cap [x_{i_t}^2]_{R'}|}{|X| + |[x_{i_1}]_{R'} - X| + \dots + |[x_{i_t}]_{R'} - X| + \dots + |[x_{i_k}]_{R'} - X| - |[x_{i_t}^2]_{R'} - X|} \\ &= \frac{|X \cap R_\lambda(X)| - |X \cap [x_{i_t}^2]_{R'}|}{|X \cup R_\lambda(X)| - |[x_{i_t}^2]_{R'} - X|}. \end{aligned}$$

Since $S(X, R_\lambda(X)) \geq |X \cap [x_{i_t}^2]_{R'}| / |[x_{i_t}^2]_{R'} - X|$, then according to Lemma 3, we can get $S(X, R_\lambda(X)) \leq S(X, R'_\lambda(X))$. So the part (2) is proved perfectly.

Note: Since $[x_{it}]_R \subset R_\lambda(X)$, according to the Lemma 2, the inclusion relations $[x_{it}^1]_{R'} \not\subset R'_\lambda(X)$ and $[x_{it}^2]_{R'} \not\subset R'_\lambda(X)$ are not held either.

Theorems 2 and 3 show that when $0 < \lambda < 0.5$ and the equivalence classes are subdivided into many finer equivalence classes (sub-granules) by R' , the similarity degree between $R'_\lambda(X)$ and X is not generally lower than the similarity degree between $R_\lambda(X)$ and X .

6 Conclusions

Since rough set theory was proposed in 1982, it has developed more than 30 years. Many scholars have made some improvements for the traditional models and obtained many extended rough set models which overcome some shortcomings of the traditional models. Combining with the fuzzy set theory, we have constructed an approximation set of an uncertain set X with the cut-set and proposed a general approximation model $R_{0.5}(X)$, but the optimal approximation set of X still is not established. In order to solve this problem, in this paper the optimal approximation set through minimizing similarity between the uncertain concept and its approximation sets is defined. Then comparative analysis between $R_{Best}(X)$ and other approximation sets is given. Next, the operation properties of $R_{Best}(X)$ are presented and proved successfully. Finally, change rules of the similarity degree between X and $R_{Best}(X)$ in different knowledge granularity levels are discussed in detail. These research presents a computational method for establishing or searching an optimal approximation set of X from the perspective of similarity. We hope these works can expand the range of rough set theory model to deal with uncertain problems in the real world and promote the development of uncertainty artificial intelligence.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant 61876201, in part by the National Key Research and Development Program of China under Grant 2017YFC0 804002, in part by the Chongqing Postgraduate Scientific Research and Innovation Project under Grant CYS18244.

References

1. Guan, L.H., Wang, G.Y., Yu, H.: Incremental algorithm of Pawlak reduction based on attribute order. *J. Southwest Jiaotong Univ.* **46**(3), 461–468 (2011)
2. Herbert, J.P., Yao, J.T.: Game-theoretic rough sets. *Fundamenta Informaticae* **108**(3–4), 267–286 (2011)
3. Hu, F., Wang, G.Y.: Quick algorithm for certain rule acquisition based on divide and conquer method. *Pattern Recognit. Artif. Intell.* **23**(3), 349–356 (2010)
4. Li, T.Y., et al.: Rough sets and knowledge technology. *Fundamenta Informaticae* **74**14(2–3), I–II (2008)
5. Liang, J.Y., Qian, Y.H.: Granulation monotonicity of entropy measure in information systems. *J. Shanxi Univ.* **30**(2), 156–162 (2007)

6. Liang, J.Y., Shi, Z.Z.: The information entropy, rough entropy and knowledge granulation in rough set theory. *Int. J. Uncertainty Fuzziness Knowl. Based Syst.* **12**(1), 37–46 (2011)
7. Liu, X.C.: Entropy, distance measure and similarity measure of fuzzy sets and their relations. Elsevier North-Holland Inc. (1992)
8. Mi, J.S., Wu, W.Z., Zhang, W.X.: Approaches to knowledge reduction based on variable precision rough set model. *Inf. Sci.* **159**(3–4), 255–272 (2004)
9. Pawlak, Z.: Rough sets. *Int. J. Comput. Inf. Sci.* **11**(5), 341–365 (1982)
10. Pawlak, Z., Skowron, A.: Rough sets: some extensions. *Inf. Sci.* **177**(1), 28–40 (2007)
11. Pawlak, Z., Wong, S.K.M., Ziarko, W.: Rough sets: probabilistic versus deterministic approach*. *Int. J. Man-Mach. Stud.* **29**(1), 81–95 (1988)
12. Qian, Y.H., Zhang, H., Sang, Y.L., Liang, J.Y.: Multigranulation decision-theoretic rough sets. *Int. J. Approximate Reasoning* **55**(1), 225–237 (2014)
13. Tiwari, S.P., Srivastava, A.K.: Fuzzy rough sets, fuzzy preorders and fuzzy topologies. Elsevier North-Holland Inc. (2013)
14. Wang, G.Y.: Rough set theory and knowledge discovery. China Xi'an Jiaotong University Press (2007)
15. Wang, G.Y., Zhang, Q.H.: Uncertainty of rough sets in different knowledge granularities: uncertainty of rough sets in different knowledge granularities. *Chin. J. Comput.* **31**(9), 1588–1598 (2008)
16. Yao, J.T., Vasilakos, A.V., Pedrycz, W.: Granular computing: perspectives and challenges. *IEEE Trans. Cybern.* **43**(6), 1977–1989 (2013)
17. Yao, Y.Y.: Two semantic issues in a probabilistic rough set model. *Fundamenta Informaticae* **108**(3), 249–265 (2011)
18. Yao, Y.Y.: Rough-set concept analysis: interpreting RS-definable concepts based on ideas from formal concept analysis. *Inf. Sci.* **346**, 442–462 (2016)
19. Yao, Y.Y., Li, X.N.: Comparison of rough-set and interval-set models for uncertain reasoning. *Fundamenta Informaticae* **27**(2), 289–298 (1996)
20. Yao, Y.Y., She, Y.H.: Rough set models in multigranulation spaces. *Inf. Sci.* **327**, 40 (2016)
21. Zhang, Q.H., Wang, G.Y., Xiao, Y.: Approximation sets of rough sets. *J. Softw.* **23**(7), 1745–1759 (2012)
22. Zhang, Q.H., Xie, Q., Wang, G.Y.: A survey on rough set theory and its applications. *CAAI Trans. Intell. Technol.* **1**(4), 323–333 (2016)
23. Zhang, Q.H., Zhang, P., Wang, G.Y.: Research on approximation set of rough set based on fuzzy similarity. *J. Intell. Fuzzy Syst.* **32**(3), 2549–2562 (2017)
24. Zhang, W.X., Wu, W.Z.: An introduction and a survey for the studies of rough set theory. *Fuzzy Syst. Math.* **14**(4), 1–12 (2000)
25. Zhang, X., Mo, Z.: Variable precision rough sets. *Pattern Recognit. Artif. Intell.* **17**(2), 151–155 (2004)
26. Ziarko, W.: Variable precision rough set model. *J. Comput. Syst. Sci.* **46**(1), 39–59 (1993)