# An Incremental Approach Based on Hierarchical Classification in Multikernel Fuzzy Rough Sets Under the Variation of Object Set

Wei Fan[1,2], Chunlin He[1], Anping Zeng[2(✉)], and Ke Lin[1]

[1] China West Normal University, Nanchong 637000, Sichuan, China
[2] Yibin University, Yibin 644007, Sichuan, China
zengap@126.com

**Abstract.** In the context of the era of big data, hybrid data are multimodality, including numerical, images, audio, etc., and even attribute values may have unknown values. Multikernel fuzzy rough sets can effectively solve large-scale multimodality attributes. At the same time, the decision attribute values may have hierarchical structure relationships, and the multikernel fuzzy rough sets based on hierarchical classification can solve the hierarchical relationships among decision attribute values. In real life, data often change dynamically. The article discusses the updating method when one object changes in the multimodality incomplete decision system based on hierarchical classification, and according to the variation of the tree-based hierarchical class structure, the upper and lower approximations are updated. Finally the deduction is carried out through relevant examples.

**Keywords:** Multikernel fuzzy rough sets · Hierarchical classification · Multimodality incomplete decision system · Variation of object set

## 1 Introduction

Fuzzy rough sets skillfully combined the rough sets that deal with classification uncertainty and the fuzzy sets that deal with boundary uncertainty to handle various classes of data types [1]. Subsequently, many scholars expanded and developed fuzzy rough sets [2–7]. To deal with multimodality attributes, Hu proposed multikernel fuzzy rough sets [8], different kernel functions used to process multimodality attributes.

Actually, there are semantic hierarchies among most data types [9]. Chen and others constructed a decision tree among classes with a tree-like structure [10]. Inspired by fuzzy rough sets theory, Wang proposed deep fuzzy trees [11]. Zhao embed the hierarchical structure into fuzzy rough sets [12]. Qiu proposed a fuzzy rough sets method using the Hausdorff distance of the sample set for hierarchical feature selection [13].

The expansion and updating of data information will also cause data loss. Therefore, in the multimodality incomplete information system, the incremental updating algorithm of the approximations is particularly important, Zeng proposed an incremental

updating algorithm for the variation of object set [14] and the attribute value based on the hybrid distance [6]. Dong L proposed an incremental algorithm for attribute reduction when samples and attributes change simultaneously [15]. Huang proposed a multi-source hybrid rough set (MCRS) [16], under variation of the object, attributes set and attribute values, a matrix-based incremental mechanism studied. However, multikernel fuzzy rough sets incremental algorithm based on hierarchical classification is not considered now.

This paper is organized as follows: Sect. 2 introduces the basic knowledge. In Sect. 3, the hierarchical structure among decision attribute values is considered into the multi-kernel fuzzy rough sets, the tree-based hierarchical class structure is imported, and the upper and lower approximations of all nodes is proposed. In Sect. 4, an incremental updating algorithm of upper and lower approximations for the immigration and emmigration of single object. Finally a corresponding example is given for deduction. The paper ends with conclusions and further research topics in Sect. 5.

## 2   An Introduction of Fuzzy Rough Sets

In this section, some related content of fuzzy rough sets will be briefly introduced.

**Definition 1** [17]. Given an fuzzy approximate space $\{U, R\}, \forall x, y, z \in U$. If $R$ satisfies:
Reflexivity: $R(x, x) = 1$;
Symmetry: $R(x, y) = R(y, x)$;
Min-max transitivity: $min(R(x, y), R(y, z)) \leq R(x, z)$.
Then $R$ is said to be a fuzzy equivalence relation on $U$.

**Definition 2** [4]. Given a fuzzy approximate space $\{U, R\}$, $R$ is a fuzzy equivalence relation on the $U$, and the fuzzy lower and upper approximations are respectively defined as follows:

$$\begin{cases} \underline{R_s}X(x) = \inf_{y \in u} S(N(R(x, y)), X(y)) \\ \overline{R_T}X(x) = \sup_{y \in u} T(R(x, y), X(y)) \end{cases} \tag{1}$$

$T$ and $S$ are the triangular mode. $N$ is a monotonically decreasing mapping function.

In a multimodality information system, the attributes of samples are multimodality, and multikernel learning is an effective method, often using different kernel functions to extract information from different attributes [8]. At the same time, there may also be unknown value in the attribute values. In this paper, the case where the attribute values exist unknown value is considered into the multimodality information system.

**Definition 3** [6]. Given a multimodality incomplete information system $\{U, MC\}$, $MC$ is a multimodality conditional attribute, $\forall x, y \in U, \forall M \in MC$, and it exists unknown value. $M(x), M(y)$ are the $M$ attribute values of $x$ and $y$, respectively. Unknown values are marked as "?". The similarity relationships extracted from this data using a matching kernel.

$$K(M(x), M(y)) = 1 \tag{2}$$

It is easy to prove that the kernel function satisfies Definition 1. Therefore, the similarity relation calculated is fuzzy equivalence relation.

**Definition 4** [8]. Given a $\{U, MC\}$, $MC$ divides the multimodality information system into $p$ subsets with different attributes, referred to as $MC/U = \{M_1, M_2, ..., M_P\}$, $M_i \in MC$, $K_i$ is a fuzzy similarity relation computed with single attribute. $\forall x, y \in U$, the fuzzy similarity relation based on combination kernels is defined as follows:

$$K_{T_{\cos}}(x, y) = max\left(\prod_{i=1}^{p} K_i(x, y) - \prod_{i=1}^{p} \sqrt{1 - K_i(x, y)^2}, 0\right) \tag{3}$$

*Example 1.* Given a $\{U, MC, D\}$, there are seven objects, $MC = \{K_1, K_2, K_3, K_4\}$, $D = \{d_1, d_2, d_3, d_4, d_5\}$, The details are shown in Table 1.

**Table 1.** Dataset used in the example

| $X$ | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $D$ |
|---|---|---|---|---|---|
| $x_1$ | 26 | 1.56 | 102 | Yes | $d_4$ |
| $x_2$ | 25 | 1.83 | 121 | ? | $d_2$ |
| $x_3$ | 27 | 1.80 | 87 | No | $d_3$ |
| $x_4$ | 22 | 1.78 | 89.8 | No | $d_1$ |
| $x_5$ | 26 | 1.85 | 105 | ? | $d_5$ |
| $x_6$ | 27 | 1.93 | 102 | Yes | $d_1$ |
| $x_7$ | 52 | 1.69 | 87 | Yes | $d_4$ |
| $x_8$ | 25 | 1.6 | 105 | Yes | $d_4$ |

In Table 1, there are three types of conditional attributes, numerical type, categorical type, and unknown values. It can be regarded as a multimodality incomplete decision system, and different kernel functions is used to extract the fuzzy similarity relationships of different attributes.

$x_1$ and $x_2$ as an example, calculating the fuzzy similarity relationships based on combination kernels:

$$K_1(x_1, x_2) = 0.990, K_2(x_1, x_2) = 0.930, K_3(x_1, x_2) = 0.965, K_4(x_1, x_2) = 1.$$

$$K_{T_{\cos}}(x_1, x_2) = max(\prod_{i=1}^{4} K_i(x_1, x_2) - \prod_{i=1}^{4} \sqrt{1 - K_i(x_1, x_2)^2}, 0) = 0.888.$$

In the same way, the fuzzy similarity relationships based on combination kernels among other objects can be obtained, and the fuzzy similarity relationships matrix

$K_{T_{\cos}}(x, y)$ can be obtained as follows:

$$
K_{T_{\cos}}(x, y) =
\begin{pmatrix}
1 & 0.888 & 0 & 0 & 0.918 & 0.863 & 0.001 & 0.973 \\
0.888 & 1 & 0.855 & 0.824 & 0.965 & 0.918 & 0.001 & 0.956 \\
0 & 0.855 & 1 & 0.779 & 0.956 & 0 & 0 & 0 \\
0 & 0.824 & 0.779 & 1 & 0.827 & 0 & 0 & 0 \\
0.918 & 0.965 & 0.956 & 0.827 & 1 & 0.983 & 0.001 & 0.965 \\
0.863 & 0.918 & 0 & 0 & 0.983 & 1 & 0.002 & 0.906 \\
0.001 & 0.001 & 0 & 0 & 0.001 & 0.002 & 1 & 0.001 \\
0.973 & 0.956 & 0 & 0 & 0.965 & 0.906 & 0.001 & 1
\end{pmatrix}
$$

## 3  Multikernel Fuzzy Rough Sets Based on Hierarchical Classifification

Given a multimodality incomplete decision system, in addition to multimodality attributes of objects, there may be hierarchical relationships among decision attribute values. The hierarchical classification is mostly based on a tree structure. A multikernel fuzzy rough sets based on hierarchical classification takes the hierarchical relationships of decision attribute values into account in the fuzzy rough sets.

**Definition 5** [4]. Given a $\{U, MC\}$. $K_{T_{\cos}}$ is a fuzzy equivalence relation based on combination kernels, $X$ is a fuzzy subset of $U$, and the approximations are respectively defined as as follows:

$$
\begin{cases}
\underline{K_T}X(x) = \inf_{y \in u} S\big(N\big(K_{T_{\cos}}(x, y)\big), X(y)\big) \\
\overline{K_T}X(x) = \sup_{y \in u} T\big(K_{T_{\cos}}(x, y), X(y)\big)
\end{cases}
\tag{4}
$$

**Definition 6** [12]. $\{U, MC, D_{tree}\}$ is a multimodality decision system based on hierarchical classification, $D_{Tree}$ is decision attribute based on hierarchical classification, and divides $U$ into $q$ subsets, referred to as $U/D_{Tree} = \{d_1, \ldots\ldots, d_p\}$, $sib(d_i)$ represents the sibling node of $d_i.\forall x \in U$, if $x \in sib(d_i)$, then $d_i(x) = 0$, else $d_i(x) = 1$. The approximations of decision class $d_i$ is defined as follows:

$$
\begin{cases}
\underline{K_T}_{sibling}d_i(x) = \inf_{y \in \{sib(d_i)\}} \{\sqrt{1 - K^2_{T_{\cos}}(x, y)}\} \\
\overline{K_T}_{sibling}d_i(x) = \sup_{y \in d_i} \{K_{T_{\cos}}(x, y)\}
\end{cases}
\tag{5}
$$

**Proposition 1.** This paper extends the lower approximations algorithm of the decision class to any node. When the decision class is a non-leaf node, the upper approximations is obtained by finding the least upper bound of the upper approximations of its child nodes. Leaf(d) represents no child node, the child node of $d_i$ is marked as $d_{ich} = \{d_{i1}, d_{i1}, \ldots, d_{ik}\}$, where $k$ is the number of child nodes, there are:

$$
\underline{K_T}_{sibling}d_i(x) =
\begin{cases}
\inf_{y \in \{sib(d_i)\}} \left\{\sqrt{1 - K^2_{T_{\cos}}(x, y)}\right\} & sib(d_i) \neq \emptyset \\
0 & else
\end{cases}
\tag{6}
$$

$$\overline{K_{T\,sibling}}d_i(x) = \begin{cases} \sup\limits_{y \in d_i}\{K_{T_{\cos}}(x, y)\} & d_i \in leaf(d) \\ \sup\{\overline{K_{T\,sibling}}d_{ich}(x)\} & else \end{cases} \qquad (7)$$

Given a $\{U, MC, D_{tree}\}$, the algorithm for the lower and upper approximations is designed in Algorithm 1.

---

**Algorithm 1:** The algorithm for the lower and upper approximations based on hierarchical classification

---

**Input:** $\{U, MC, D_{tree}\}$ and combination kernels $K_{T_{\cos}}(x, y)$

**Output:** The approximations $\underline{K_{T\,sibling}}$ and $\overline{K_{T\,sibling}}$

**for** *each* $d_i \in U / D_{Tree}$ $\psi$**do**

    **for** *each* $d \in U / D_{Tree}$ $\psi$**do**

        **if** $d \in leaf(d)$ **then**

            **for** *each* $x \in U$ **do**

                $\overline{K_{T\,sibling}}d(x) = max\left(K_{T_{\cos}}(x, y)\right)$

            **end**

            //Calculate the upper approximations of leaf nodes.

        **end**

    **end**

    $\overline{K_{T\,sibling}}d(x) = max\left(\overline{K_{T\,sibling}}d_{ch}(x)\right)$

    //Calculate the upper approximations of non-leaf nodes.

    **for** *each* $x \in U$ $\psi$**do**

        **for** *each* $y \in sib(d)$ **do**

            $\underline{K_{T\,sibling}} = inf\left\{\sqrt{1 - K_{T_{\cos}}^2(x, y)}\right\}$

            //Calculate the lower approximations.

        **end**

    **end**

**end**

---

*Example 2.* On the basis of Example 1, the decision attributes are divided into five subsets, namely $d_1, d_2, d_3, d_4, d_5$. From Fig. 1, $d_1 = \{x_4, x_6\}$, and from Table 1, $sib(d_1) = d_4 = \{x_1, x_8\}$, according to Proposition 1, the lower and upper approximations of the decision class are calculated as follows:

$$\underline{K_{T\,sibling}}d_1(x_2) = \inf_{y \in \{x_1, x_8\}}\left\{\sqrt{1 - K_{T_{\cos}}^2(x_2, y)}\right\} = min\{0.475, 0.293\} = 0.293$$

$$\underline{K_{T\,sibling}}d_1(x_1) = 0, \underline{K_{T\,sibling}}d_1(x_3) = 0, \underline{K_{T\,sibling}}d_1(x_4) = 0, \underline{K_{T\,sibling}}d_1(x_5) = 0.262.$$

$$\underline{K_{T\,sibling}}d_1(x_6) = 0.423, \underline{K_{T\,sibling}}d_1(x_7) = 1, \underline{K_{T\,sibling}}d_1(x_8) = 0.$$

Because $d_1$ is a non-leaf node, according to Proposition 1 there are:

$$\overline{K_{T\,sibling}}d_1(x_1) = \sup_{y \in d_3}\left\{\overline{K_{T\,sibling}}d_3(x_1)\right\} = \sup\left\{K_{T_{\cos}}(x_1, x_3)\right\} = 0$$

$$\overline{K_{T\,sibling}}d_1(x_2) = 0.855, \overline{K_{T\,sibling}}d_1(x_3) = 1, \overline{K_{T\,sibling}}d_1(x_4) = 0.778.$$

$$\overline{K_{T\,sibling}}d_1(x_5) = 0.956, \overline{K_{T\,sibling}}d_1(x_6) = 0, \overline{K_{T\,sibling}}d_1(x_7) = 0, \overline{K_{T\,sibling}}d_1(x_8) = 0.$$

So the lower and upper approximations of of $d_1$ are:

$$\underline{K_{T\,sibling}}d_1 = \{0.293/x_2, 0.262/x_5, 0.423/x_6, 1/x_7\}$$

$$\overline{K_{T\,sibling}}d_1 = \{0.885/x_2, 1/x_3, 0.778/x_4, 0.956/x_5\}$$

On the basis of Table 1, the tree-based hierarchical class structure is established in Fig. 1.
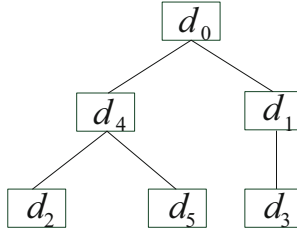


**Fig. 1.** The tree of decision class

## 4   Incremental Updating for Lower and Upper Approximations Under the Variation of Single Object

A $\{U, MC, D_{tree}\}$ at time $t$ is given. $\underline{K}_{T_{sibling}}^{(t)}X$ represents lower approximations, and $\overline{K}_{T_{sibling}}^{(t)}X$ represents upper approximations. Given $\{\overline{U}, \overline{MC}, \overline{D}_{tree}\}$ represents a multi-modality decision system based on hierarchical classification at time $t + 1$, $x^+, x^-$ represents immigration and emmigration of one object, respectively. The fuzzy upper and lower approximations at time $t + 1$ are denoted by $\overline{K}_{T_{sibling}}^{(t+1)}X$ and $\underline{K}_{T_{sibling}}^{(t+1)}X$, respectively.

### 4.1   Immigration of Single Object

The $x^+$ immigrates into the $\{\overline{U}, \overline{MC}, \overline{D}_{tree}\}$ at time $t + 1$, in which $\overline{U} = U \cup \{x^+\}$. If no new decision class is generated at time $t + 1$, tree-based hierarchical class structure does not need to be updated. Otherwise, the tree will be updated. Next, the incremental updating is discussed by whether to generate a new decision class.

**Proposition 2.** $\forall d_i \in \overline{U}/\overline{D}_{tree}, x \in \overline{U}, x^+$ will generate a new decision class, and the new class is marked as $\overline{d}_{n+1}^+$. The labeled class $\overline{d}_{n+1}^+$ needs to be inserted into the tree-based hierarchical class structure. Then the approximations updates of the decision class is shown:

$$
\underline{K}_{T_{sibling}}^{(t+1)} d_i(x) =
\begin{cases}
\inf\limits_{y=x^+} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)}, \underline{K}_{T_{sibling}}^{(t)} d_i(x) \right\} & x^+ \in \{sib(d_i)\} \text{ and} \\[2ex]
& \{sib(d_i)\} \neq \overline{d}_{n+1}^+ \\
\underline{K}_{T_{sibling}}^{(t)} d_i(x) & x^+ \notin \{sib(d_i)\} \text{ and } x \neq x^+ \\[3ex]
\inf\limits_{y \in \{sib(d_i)\}} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\} & else
\end{cases}
\tag{8}
$$

$$
\overline{K}_{T_{sibling}}^{(t+1)} d_i(x) =
\begin{cases}
\overline{K}_{T_{sibling}}^{(t)} d_i(x) & x \neq x^+ \text{ and } d_i \in leaf(d) \\
\sup\limits_{y \in d_i} K_{T_{\cos}}(x, y) & x = x^+ \text{ and } d_i \in leaf(d) \\
\sup\left\{ \overline{K}_{T_{sibling}}^{(t)} d_{ick}(x) \right\} & else
\end{cases}
\tag{9}
$$

*Proof.* For the lower approximations of $d_i$, the approximations at time $t + 1$ is determined by the objects belonging to the sibling nodes of $d_i$. When $x \neq x^+$ and $x^+ \notin \{sib(d_i)\}$, the lower approximations are the same as time $t$; when $x = x^+$ or the sibling nodes of at time $t + 1$ are newly decision class, calculating its lower approximations directly according to Proposition 1; when $x^+ \in \{sib(d_i)\}$ and all sibling nodes of $d_i$ are not newly classes, $\forall x \neq x^+$, there is:

$$
\begin{aligned}
\underline{K}_{T_{sibling}}^{(t+1)} d_i(x) &= \inf\limits_{y \in \{sib(d_i)\}} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\} \\
&= \inf\limits_{y \in \{sib(d_i) - x^+\} \cup x^+} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\} \\
&= \underline{K}_{T_{sibling}}^{(t)} d_i(x) \wedge \inf\limits_{y=x^+} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\} \\
&= \inf\limits_{y=x^+} \left\{ \underline{K}_{T_{sibling}}^{(t)} d_i(x), \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\}
\end{aligned}
$$

For the upper approximations of $d_i$, the upper approximations is determined by the object that belongs to $d_i$. When $d_i$ is a leaf node and $x \neq x^+$ at time $t + 1$, that is the newly decision subset not equal to $d_i$, and its upper approximations is the same as time $t$; When $d_i$ is a leaf and $x = x^+$ at time $t + 1$, direct computational approximations according to Proposition 1; In the tree-based hierarchical class structure, the decision class with the same parents node belongs to a major class, so $d_i$ and dich belong to the same major class, When $d_i$ is not a leaf node, we directly find the least upper bound for all child nodes of the decision class $d_i$, and obtain updating of the upper approximations.

**Proposition 3.** $\forall d_i \in \overline{U}/\overline{D}_{tree}$, $x \in \overline{U}$, $x^+$ will not generate a new decision class, and then the tree-based hierarchical class structure does not need to be updated. The approximations updating of the decision class is shown:

$$
\underline{K}_{T_{sibling}}^{(t+1)} d_i(x) = \begin{cases} \inf_{y=x^+} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)}, \underline{K}_{T_{sibling}}^{(t)} d_i(x) \right\} & x \neq x^+ \text{ and} \\ & x^+ \in \{sib(d_i)\} \\ \underline{K}_{T_{sibling}}^{(t)} d_i(x) & x \neq x^+ \text{ and} \\ & x^+ \notin \{sib(d_i)\} \\ \inf_{y \in \{sib(d_i)\}} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\} & else \end{cases} \tag{10}
$$

$$
\overline{K}_{T_{sibling}}^{(t+1)} d_i(x) = \begin{cases} \overline{K}_{T_{sibling}}^{(t)} d_i(x) & x \neq x^+, x^+ \notin d_i \\ & \text{and } d_i \in leaf(d) \\ \sup_{y=x^+} \left\{ K_{T_{\cos}}(x, y), \overline{K}_{T_{sibling}}^{(t)} d_i(x) \right\} & x \neq x^+, x^+ \in d_i \\ & \text{and } d_i \in leaf(d) \\ \sup_{y \in d_i} \left\{ K_{T_{\cos}}(x, y) \right\} & x = x^+ \\ & \text{and } d_i \in leaf(d) \\ \sup \left\{ \overline{K}_{T_{sibling}}^{(t)} d_{ich}(x) \right\} & else \end{cases} \tag{11}
$$

*Proof.* The proof of the lower approximations updating is similar to Proposition 2; for the upper approximations, at time $t + 1$, when $d_i$ is a non-leaf node, directly find the least upper bound for the child nodes of $d_i$, and the approximations updating is obtained. When $d_i$ is a leaf node, the upper approximations at time $t + 1$ is determined by the object that belongs to $d_i$; When $x = x^+$, the upper approximations is directly calculated according to Proposition 1; When $x \neq x^+$ and $x^+ \notin \{d_i\}$ the approximations is the same as time t; when $x \neq x^+$ and $x^+ \in \{d_i\}$, there is:

$$
\overline{K}_{T_{sibling}}^{(t+1)} d_i(x) = \sup_{y \in d_i} \left\{ K_{T_{\cos}}(x, y) \right\}
$$

$$
= \inf_{y \in \{sib(d_i) - x^+\} \cup x^+} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x, y)} \right\}
$$

$$
= \sup_{y=x^+} \left\{ \overline{K}_{T_{sibling}}^{(t)} d_i(x), K_{T_{\cos}}(x, y) \right\}
$$

Given a $\{\overline{U}, \overline{MC}, \overline{D}_{tree}\}$, when immigration of one object, the algorithm for the lower and upper approximations is designed in Algorithm 2.

**Algorithm 2:** Incremental Algorithm Based on Hierarchical Classification For the Immigration of single Object

**Input:** $\left\{\overline{U}, \overline{MC}, \overline{D}_{tree}\right\}$ , $\underline{K}_{T_{sibling}}^{(t)}$ and $\overline{K}_{T_{sibling}}^{(t)}$ . //Already obtained in Algorithm 1.

**Output:** The approximations $\underline{K}_{T_{sibling}}^{(t+1)}$ and $\overline{K}_{T_{sibling}}^{(t+1)}$

$K_{T_{cos}}(x,y) \Leftarrow K_{T_{cos}}(x^+,y)$

**for** *each* $d \in U / D_{Tree}$ $\psi$**do**

    **for** *each* $x \in \overline{U}$ $\psi$**do**

        **if** $\left\{\overline{d}^+_{n+1} \neq sib(d) \ and \ x^+ \in sib(d)\right\}$ *or* $\left\{x \neq x^+ \ and \ x^+ \in sib(d)\right\}$ **then**

            $\underline{K}_{T_{sibling}}^{(t+1)} d(x) = \inf_{y=x^+} \left\{\sqrt{1-K^2_{T_{cos}}(x,y)}, \underline{K}_{T_{sibling}}^{(t)} d(x)\right\}$

        **end**

        **if** $x \neq x^+ \ and \ x^+ \notin sib(d)$ **then**

            $\underline{K}_{T_{sibling}}^{(t+1)} d(x) = \underline{K}_{T_{sibling}}^{(t)} d(x)$

        **end**

        **else**

            $\underline{K}_{T_{sibling}}^{(t+1)} d(x) = \inf_{y \in \{sib(d)\}} \left\{\sqrt{1-K^2_{T_{cos}}(x,y)}\right\}$

        **end**

        //Low approximations.

        **if** $\overline{d}^+_{n+1}$ **then**

            **if** $d \in leaf \ and \ x = x^+$ **then**

                $\overline{K}_{T_{sibling}}^{(t+1)} d(x) = \sup_{y \in d} K_{T_{cos}}(x,y)$

            **end**

            **if** $d \in leaf \ and \ x \neq x^+$ **then**

                $\overline{K}_{T_{sibling}}^{(t+1)} d(x) = \overline{K}_{T_{sibling}}^{(t)} d(x)$

            **end**

            **else**

                $\overline{K}_{T_{sibling}}^{(t+1)} d_i(x) = \sup \left\{\overline{K}_{T_{sibling}}^{(t)} d_{ick}(x)\right\}$

            **end**

            //Upper approximations when new decision classes are generated.

        **end**

        **else**

            **if** $d \in leaf \ and \ x = x^+$ **then**

                $\overline{K}_{T_{sibling}}^{(t+1)} d(x) = \sup_{y \in d} K_{T_{cos}}(x,y)$

**end**

**if** $d \in leaf$, $x^+ \notin d$ *and* $x = x^+$ **then**

$$\overline{K}_{T_{sibling}}^{(t+1)} d(x) = \overline{K}_{T_{sibling}}^{(t)} d(x)$$

**end**

**if** $d \in leaf$, $x^+ \in d$ *and* $x = x^+$ **then**

$$\overline{K}_{T_{sibling}}^{(t+1)} d(x) = \sup_{y=x^+}\left\{ K_{T_{cos}}(x,y), \overline{K}_{T_{sibling}}^{(t)} d(x) \right\}$$

**end**

**else**

$$\overline{K}_{T_{sibling}}^{(t+1)} d_i(x) = \sup\left\{ \overline{K}_{T_{sibling}}^{(t)} d_{ick}(x) \right\}$$

**end**

**end**

//Upper approximations when no new decision classes are generated

**end**

**end**

---

**Table 2.** Information about immigration of single bject

| $X$ | $K_1$ | $K_2$ | $K_3$ | $K_4$ | $D$ |
|-----|-------|-------|-------|-------|-----|
| $x_9^+$ | 17 | 1.58 | 50 | no | Normal_Weight |

*Example 3.* On the basis of Table 1, one object $x_9^+$ immigrates into system, and its information is shown in Table 2. A new decision class $d_6$ is generated. first the fuzzy similarity relationships with other objects based on combination kernels are calculated according to Definition 6, as follows:

$$K_{T_{cos}}(x_9^+, x) = (\,0\ 0.299\ 0.306\ 0.643\ 0.306\ 0\ 0\ 0\ 1\,)$$

On the basis of Fig. 1, inserting the newly decision class $d_6$ into the tree, the tree is update as follows:
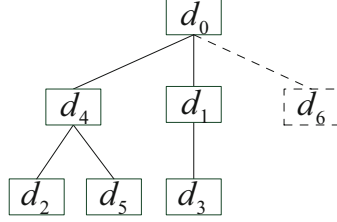
**Fig. 2.** The tree of decision class

As shown in Fig. 2, $d_6 \in leaf$, $sib(d_1) = \{d_4, d_6\} = \{x_1, x_8, x_9^+\}$.
According to Proposition 2:

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1(x_1) = \inf_{y=x_9^+} \left\{ \sqrt{1 - K_{T_{\cos}}^2(x_1, y)}, \underline{K}_{T_{sibling}}^{(t)} d_1(x_1) \right\} = min(0, 0) = 0$$

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1(x_2) = 0.293, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_3) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_4) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_5) = 0.263.$$

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1(x_6) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_7) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_8) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_9) = 0.$$

From Fig. 2, $d_1$ can be obtained as a non-leaf node, according Proposition 2:

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1(x_1) = \sup_{y \in d_3} \left\{ \overline{K}_{T_{sibling}}^{(t+1)} d_{1ch}(x_1) \right\} = \sup_{y \in d_3} \left\{ \overline{K}_{T_{sibling}}^{(t)} d_3(x_1) \right\} = 0$$

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1(x_2) = 0.855, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_3) = 1, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_4) = 0.778.$$

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1(x_5) = 0.956, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_6) = 0, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_7) = 0, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_8) = 0.$$

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1(x_9) = 0.306.$$

So the lower and upper approximations of $d_1$ are:

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1 = \{0.293/x_2, 0.262/x_5\}$$

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1 = \{0.885/x_2, \ 1/x_3, \ 0.778/x_4, \ 0.956/x_5, \ 0.306/x_9\}$$

From Example 3, obviously, when immigration of one object in the multimodality decision system based on hierarchical classification, it will affect the upper and lower approximations of other decision classes. Only a small amount of update operations need to be performed according to properties 2 and 3, which greatly reduces the amount of calculation and time cost.

## 4.2   Emmigration of Single Object

$x^-$ emmigrates from the $\{\overline{U}, \overline{MC}, \overline{D}_{tree}\}$ at time $t + 1$, in which $\overline{U} = U - \{x^-\}$. If no decision class is removed at time $t + 1$, the tree-based hierarchical class structure does not need to be updated, otherwise, the tree is updated. Next, the incremental updating will be discussed by whether to remove a decision class.

**Proposition 4.** $\forall d_i \in \overline{U}/\overline{D}_{tree}, x \in \overline{U}, x^-$ emmigrates form system, which leads to the decision class is removed, marked as $\overline{d}_l$. The tree-based hierarchical class structure will be updated, and the $\overline{d}_l$ will be removed from the tree. Then the approximations updating of the decision class is shown:

$$\underline{K}^{(t+1)}_{T_{sibling}} d_i(x) = \begin{cases} \inf_{y \in sib(d_i)} \left\{ \sqrt{1 - K^2_{T_{\cos}}(x, y)}, \underline{K}^{(t)}_{T_{sibling}} d_i(x) \right\} & x^- \in sib(d_i) \text{ and} \\ & sib(d_i) \neq \emptyset \\ \underline{K}^{(t)}_{T_{sibling}} d_i(x) & x^- \notin sib\{d_i\} \\ 0 & else \end{cases} \tag{12}$$

$$\overline{K}^{(t+1)}_{T_{sibling}} d_i(x) = \overline{K}^{(t)}_{T_{sibling}} d_i(x) \tag{13}$$

*Proof.* At time $t+1$, the removed decision class has no upper and lower approximations, but $x^-$ will have the impact on the upper and lower approximations of other decision classes. If $x^-$ is the closest object to the sibling node to which the object $x$ belongs, the lower approximations need to be recalculated, otherwise, the lower approximations remains unchanged; if the sibling node of $d_i$ does not exist after $x^-$ emmigrates, the lower approximations is 0. For a decision class, the upper approximations is only related to the objects belonging to the decision class. At time $t+1$, $x^- \notin d_i$, so the upper approximations remains the same.

**Proposition 5.** $\forall d_i \in \overline{U}/\overline{D}_{tree}, x \in \overline{U}, x^-$ will not cause the decision class to be removed, and then the tree-based hierarchical class structure does not need to be updated. Then the approximations updating of the decision class is shown:

$$\underline{K}^{(t+1)}_{T_{sibling}} d_i(x) = \begin{cases} \inf_{y \in sib(d_i)} \left\{ \sqrt{1 - K^2_{T_{\cos}}(x, y)}, \underline{K}^{(t)}_{T_{sibling}} d_i(x) \right\} & x^- \in sib(d_i) \\ \underline{K}^{(t)}_{T_{sibling}} d_i(x) & else \end{cases} \tag{14}$$

$$\overline{K}^{(t+1)}_{T_{sibling}} d_i(x) = \begin{cases} \sup_{y=x^-} \left\{ K_{T_{\cos}}(x, y), \overline{K}^{(t)}_{T_{sibling}} d_i(x) \right\} & x^- \in d_i \text{ and } d_i \in leaf(d) \\ \overline{K}^{(t)}_{T_{sibling}} d_i(x) & x^- \notin d_i \text{ and } d_i \in leaf(d) \\ \sup \left\{ \overline{K}^{(t)}_{T_{sibling}} d_{ich}(x) \right\} & else \end{cases} \tag{15}$$

The proof process is similar.

Given a $\{\overline{U}, \overline{MC}, \overline{D}_{tree}\}$, when one object emmigrates from system, the algorithm for the lower and upper approximations is designed in Algorithm 3.

---

**Algorithm 3:** Incremental Algorithm Based on Hierarchical Classification For the Emmigration of single Object

---

**Input:** $\left\{\overline{U}, \overline{MC}, \overline{D}_{tree}\right\}$, $\underline{K}^{(t)}_{T_{sibling}}$ and $\overline{K}^{(t)}_{T_{sibling}}$ . **//**Already obtained in Algorithm 1..

**Output:** The approximations $\underline{K}^{(t+1)}_{T_{sibling}}$ and $\overline{K}^{(t+1)}_{T_{sibling}} X$

**for** *each* $d \in U / D_{Tree}$ $\psi$**do**

    **for** *each* $x \in \overline{U}$ $\psi$**do**

        **if** $\overline{d}_l$ **then**

            **if** $x^- \in sib(d)$ *and* $sib(d) \neq \varnothing$ **then**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \inf_{y \in sib(d)} \left\{\sqrt{1 - K^2_{T_{cos}}(x,y)}, \underline{K}^{(t)}_{T_{sibling}} d(x)\right\}$$

            **end**

            **if** $x^- \notin sib(d)$ **then**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \underline{K}^{(t)}_{T_{sibling}} d(x)$$

            **end**

            **else**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = 0$$

            **end**

            **//**Lower approximations without decision class removal.

$$\overline{K}^{(t+1)}_{T_{sibling}} d(x) = \overline{K}^{(t)}_{T_{sibling}} d(x)$$

        **end**

        **else**

            **if** $x^- \in sib(d)$ **then**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \inf_{y \in sib(d)} \left\{\sqrt{1 - K^2_{T_{cos}}(x,y)}, \underline{K}^{(t)}_{T_{sibling}} d(x)\right\}$$

            **end**

            **else**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \underline{K}^{(t)}_{T_{sibling}} d(x)$$

            **end//**Lower approximations with decision class removal.

            **If** $x^- \in d$ *and* $d \in$ l eaf $(d)$ **then**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \sup_{y = x^-} \left\{K_{T_{cos}}(x,y), \overline{K}^{(t)}_{T_{sibling}} d(x)\right\}$$

            **end**

            **if** $x^- \notin d$ *and* $d \in$ l eaf $(d)$ **then**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \underline{K}^{(t)}_{T_{sibling}} d(x)$$

            **end**

            **else**

$$\underline{K}^{(t+1)}_{T_{sibling}} d(x) = \underline{K}^{(t)}_{T_{sibling}} d(x)$$
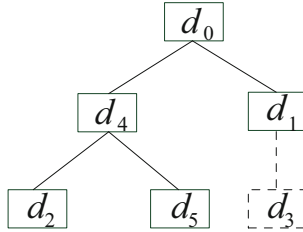
            **end**

        **end**

        **end**

    **end**

**end**

---

*Example 4.* The object $x_3$ emmigrates from the system on the basis of Table 1, which will cause the decision class $d_3$ to be removed, and the tree will be updated as follows:



**Fig. 3.** The new tree for emmigration of one object $x_3$

It can be seen from Fig. 3 that the child node of $d_3$ has been removed, and $d_1$ is a leaf node. According to Proposition 4, it can be obtained:

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1(x_1) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_2) = 0.293, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_3) = 0, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_4) = 0.$$

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1(x_5) = 0.262, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_6) = 0.423, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_7) = 1, \underline{K}_{T_{sibling}}^{(t+1)} d_1(x_8) = 0.$$

Because $d_1$ become a leaf node, where are:

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1(x_1) = 0, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_2) = 0.855, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_3) = 0, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_5) = 0.956.$$

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1(x_6) = 0, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_7) = 0, \overline{K}_{T_{sibling}}^{(t+1)} d_1(x_8) = 0.$$

So the lower and upper approximations of $d_1$ are:

$$\underline{K}_{T_{sibling}}^{(t+1)} d_1 = \{0.293/x_2, 0.262/x_5, 0.423/x_6, 1/x_7\}$$

$$\overline{K}_{T_{sibling}}^{(t+1)} d_1 = \{0.885/x_2, 0.778/x_4, 0.956/x_5\}$$

From Example 4, it can be seen that when one object emmigrate form the multi-modality decision system based on hierarchical class, the emmigration of one objects have an impact on the upper and lower approximations of other decision classes, and only a small amount of update operations need to be performed according to Proposition 4 and 5, which greatly reduces the calculation and time cost.

## 5 Conclusions and Further Research

In the multimodality decision system based on hierarchical classification, attributes of simples are multimodality, decision attribute values often have a hierarchical structure, and data changes frequently. The paper takes into account the fact that the attribute

values are not known in the multimodality information system based on hierarchical classification, and proposes an incremental updating algorithm in multikernel fuzzy rough sets based on hierarchical classification. The specific process of this algorithm is demonstrated through relevant examples. This algorithm can effectively reduce the time cost caused by object set changes. In future work, the approximations updating algorithm of more objects changing in multikernel fuzzy rough sets based on hierarchical classification and the performance of the algorithm to test the algorithm using the UCI dataset will be the focus of the research.

# References

1. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. Int. J. Gen. Syst. **17**, 191–209 (1990)
2. Morsi, N., Yakout, M.: Axiomatics for fuzzy rough sets. Fuzzy Sets Syst. **100**, 327–342 (1998)
3. Mi, J., Zhang, W.: An axiomatic characterization of a fuzzy generalization of rough sets. Inf. Sci. **160**, 235–249 (2004)
4. Yeung, D., Chen, D., Tsang, E.C.: On the generalization of fuzzy rough sets. IEEE Trans. Fuzzy Syst. **13**, 343–361 (2005)
5. Hu, Q., Yu, D., Pedrycz, W., Chen, D.: Kernelized fuzzy rough sets and their applications. IEEE Trans. Knowl. Data Eng. **23**, 1649–1667 (2011)
6. Zeng, A., Li, T., Luo, C., Hu, J.: Incremental updating fuzzy rough approximations for dynamic hybrid data under the variation of attribute values. In: 2015 International Conference on Machine Learning and Cybernetics (ICMLC), pp. 157–162 (2015)
7. Zeng, K., She, K.: Fuzzy rough set model based on multi-Kernelized granulation. J. Univ. Electron. Sci. Technol. China **5**, 717–723 (2014)
8. Hu, Q., Zhang, L., Zhou, Y.: Large-scale multimodality attribute reduction with multi-kernel fuzzy rough sets. IEEE Trans. Fuzzy Syst. **26**, 226–238 (2018)
9. Zhao, H., Cao, J., Chen, Q.: Methods for hierarchical multi-label text classification. J. Chin. Comput. Syst. **43**, 673–683 (2022)
10. Ceci, M., Malerba, D.: Classifying web documents in a hierarchy of categories: a comprehensive study. J. Intell. Inf. Syst. **28**, 37–78 (2007). https://doi.org/10.1007/s10844-006-0003-2
11. Wang, Y., Hu, Q., Zhu, P., Li, L., Lu, B.: Deep fuzzy tree for large-scale hierarchical visual classification. IEEE Trans. Fuzzy Syst. **28**, 1395–1406 (2019)
12. Zhao, H., Wang, P., Hu, Q.: Fuzzy rough set based feature selection for large-scale hierarchical classification. IEEE Trans. Fuzzy Syst. **27**, 1891–1903 (2019)
13. Qiu, Z., Zhao, H.: A fuzzy rough set approach to hierarchical feature selection based on Hausdorff distance. Appl. Intell., 1–14 (2022). https://doi.org/10.1007/s10489-021-03028-4
14. Zeng, A., Li, T., Luo, C.: Incremental approach for updating approximations of gaussian kernelized fuzzy rough sets under variation of object set. Comput. Sci. **6**, 173–178 (2013)
15. Dong, L., Chen, D.: Incremental attribute reduction with rough set for dynamic datasets with simultaneously increasing samples and attributes. Int. J. Mach. Learn. Cybern. **11**(6), 1339–1355 (2020). https://doi.org/10.1007/s13042-020-01065-y
16. Huang, Y., Li, T., Luo, C.: Dynamic maintenance of rough approximations in multi-source hybrid information systems. Inf. Sci. **530**, 108–127 (2020)
17. Chen, S., Li, J., W, X.: Fuzzy Set Theory and Its Application, pp. 100–123. Science Press (2006)