Mathematics and Statistics 10(4): 759-772, 2022 http://www.hrpub.org
DOI: 10.13189/ms.2022.100407

Uncertainty Optimization-Based Rough Set for Incomplete Information Systems

Arvind Kumar Sinha, **Pradeep Shende***,
Department of Mathematics, National Institute of Technology Raipur, Chhattisgarh, India

Received February 22, 2022; Revised May 18, 2022; Accepted June 21, 2022

Cite This Paper in the following Citation Styles

***(a)**: [1] Arvind Kumar Sinha, Pradeep Shende, "Uncertainty Optimization-Based Rough Set for Incomplete Information Systems," Mathematics and Statistics, Vol.10, No.4, pp. 759-772, 2022. DOI: 10.13189/ms.2022.100407*

***(b)**: Arvind Kumar Sinha, Pradeep Shende, (2022). Uncertainty Optimization-Based Rough Set for Incomplete Information Systems. Mathematics and Statistics, 10(4), 759-772 DOI: 10.13189/ms.2022.100407*

Abstract Often the information in the surrounding world is incomplete, and such incomplete information gives rise to uncertainties. Pawlak's rough set model is an approach to approximation under uncertainty. It uses a tolerance relation to obtain single granulation of the incomplete information system for approximation. In this work, we extend the single granulation rough set for the incomplete information system to an uncertainty optimization-based rough set (UOBRS). The proposed approach is used to minimize the uncertainty using multiple tolerance relations. We list properties of the UOBRS for incomplete information systems. We compare UOBRS with the classical single granulation rough set (SGRS) and multi-granular rough set (MGRS). We list the basic properties of the UOBMGRS. We introduce the application of the UOBRS for attribute subset selection in case of incomplete information system. We use the measure of approximation quality to assess the uncertainties of the attributes. We compare the approximation quality of the attributes using UOBRS with the approximation quality using SGRS and MGRS. We get higher approximation quality with the less number of attributes using UOBRS as compared to SGRS and MGRS. The proposed method is a novel approach to dealing with incomplete information systems for more effective dataset analysis.

Keywords Incomplete Information System, Rough Set, Uncertainty Optimization, Approximation Quality, Feature Subset Selection

1 Introduction

Decision making under uncertainty is a problem of concern in many applications related to artificial intelligence. Missing data lead to uncertainties in the dataset and may lead to wrong decisions; therefore, optimization under uncertainty is imperative for expert systems. The attribute set is considered as the granular space [1] and is the main reason of the granular structure of the knowledge. Partitioning the universe of discourse, knowledge granulation and set approximating are the extensively used methods of human's reasoning under uncertainty [2, 3]. Pawlak's rough set theory [4, 5, 6] is a well established tool to handle uncertainty involved in many problems of science and technology. The rough set framework is a very useful tool to handle uncertainty and is used in features extraction [7], noise reduction [8] and pattern recognition [9].

It is the general fact that every attribute preserves a different amount of objects information and hence every attribute set is considered a granular space that brings about different granular structure of the knowledge. The rough set methodology uses objects-attribute relation in the form of table known as information system (IS). Mathematically, the IS is an ordered touple $(U, At \cup \{D\}, f)$, where U is the non-empty finite set of objects, At is the attribute set (conditional features), D is the decision attribute (decision feature) and $f_A : U \to V_A$ for any $A \in At$, where V_A is the attribute value set of A.

There are many extentions of the rough set model available in the literature including variable precision rough set model [10], multi-granular rough set model [1], neighborhood-based multigranulation rough set model [11], covering based rough set

U	A_1	A_2	A_3	D
ω_1	1	2	1	1
ω_2	1	1	*	1
ω_3	2	2	2	1
ω_4	*	1	1	1
ω_5	3	*	3	1
ω_6	4	1	2	2
ω_7	2	2	3	2
ω_8	3	*	1	2

Table 1. Incomplete information system

model [12], fuzzy-rough set and rough-fuzzy set model [13], probabilistic rough set model [14], uncertainty optimization-based multigranular rough set for complete information system [15]. Pawlak [4, 5, 6] uses the granular structure of the knowledge induced by single relation (such as equivalence relations, tolerance relation) on the universe in the characterization of the target concepts via the so-called lower and upper approximations. The main assumption considered in the approximation of target concept is if ς and ϑ are two conditional attribute sets and $\chi \subseteq U$ is the target set, then the approximation to χ is derived from the quotient set $U/(\varsigma \cup \vartheta)$. The following formula gives the quotient set:

$$U/(\varsigma \cup \vartheta) = \{\varsigma_i \cap \vartheta_i : \varsigma_i \in U/\varsigma,\ \vartheta_j \in U/\vartheta\}. \tag{1}$$

This assumption makes sure that the attribute sets ς and ϑ are independent to each other, $\forall \varsigma_i \in U/\varsigma$ and $B_j \in U/\vartheta$, the intersection $\varsigma_i \cap \vartheta_j$ can be carried out, and the target concept is approximated by attribute set $A \cup B$ as a single set.

It is the general fact that each attribute preserves a different amount of information. Lack of information leads to the uncertainties in the information system. An information system containing incomplete attribute values is termed as incomplete information system. We often have to deal with information systems containing incomplete attribute values. Many studies are proposed to deal with incomplete information system [16, 17, 18, 19, 20]. The attribute values incompetent of describing a particular happening with a certainty is called as boundary region or the uncertainty region of the rough set [4]. The single granulation rough set finds the uncertainty region but does not provide a way for its minimization. Also, in the single granulation rough set, we find the uncertainty region using the entire attribute set as a result it does not consider the individual attributes preserved information. In this work, we used the multi-granulation induced by multiple attribute subsets to find and minimize the uncertainty region of the rough set in case of incomplete information systems. We propose an uncertainty optimization-based multi-granular rough set for the incomplete information systems to minimize uncertainty region of the rough set in the case of incomplete information system. The proposed work is a novel approach to dealing with incomplete information systems' uncertainties.

2 Preliminaries

Let attribute set At be the set of conditional attributes and D be the decision attribute. For the IS $I = (U, At \cup \{D\}, V)$, the attribute domain V_A may contain special symbol $*$ to show missing attribute value for the attribute $A \in At$. If any attribute domain V_A contains the $*$ value, it means that some of the attribute values are missing. The attribute values other than $*$ are called regular (known) values [21, 22]. If V_A contains no $*$ values $\forall\ A \in At$, then the IS $I = (U, At \cup \{d\}, V)$ is called complete IS. The IS $S = (U, At \cup \{D\}, V)$ is termed as incomplete IS if for any $A \in At$, the domain V_A contains $*$ value and all the values of the domain V_D are regular.

Throughout this paper, we assume that U is the nonempty finite set of objects. Let $I = (U, At \cup \{D\}, f)$ be the incomplete information system, $A \in At$, and for any $\alpha \in U$, $A(\alpha)$ denotes the attribute value of A for the object α.

Example 1: Table 1 is the example of incomplete information system as it contains some missing attribute values. For the incomplete information system given in Table 1, we have $U = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8\}$, $At = \{A_1, A_2, A_3\}$, and the decision attribute id $\{D\}$. The attribute value set for the attribute A_1 is $V_{A_1} = \{1, 2, 3, 4\}$, the attribute value set for the attribute A_2 is $V_{A_2} = \{1, 2\}$, and the attribute value set for the attribute A_3 is $V_{A_3} = \{1, 2, 3\}$. The attribute value set for the decision attribute D is $V_D = \{1, 2\}$. Each value of the decision attribute gives rise to a concept in the information system and is also termed as decision class. For the incomplete information system given in Table 1, we have two decision classes $\chi_1 = (D, 1) = \{1, 2, 3, 4, 5\}$ and $\chi_2 = (D, 2) = \{6, 7, 8\}$. The attribute value set for the incomplete information system represented in Table 1 is given by $V = V_{A_1} \cup V_{A_2} \cup V_{A_3} \cup V_D$.

Let $\varsigma \subseteq At$, then define a relation $SIML(\varsigma)$ on U as follows:

$$SIML(\varsigma) = \left\{ \begin{array}{r} (\alpha, \beta) \in U \times U :\ \forall\ B \in \varsigma,\ B(\alpha) = B(\beta) \\ \text{or } B(\alpha) = * \text{ or } B(\beta) = * \end{array} \right\} \tag{2}$$

The relation $SIML(U)$ is the tolerance relation [5]. The maximal set of objects indiscernible with any $\alpha \in U$ using the attribute set $\varsigma \subseteq At$ is denoted as $I_\varsigma(\alpha)$ and is given as follows:

$$I_\varsigma(\alpha) = \{\beta \in U : (\alpha, \beta) \in SIML(\varsigma)\} \tag{3}$$

The classification of the U induced by $\varsigma \subseteq At$ is denoted as $U/SIML(\varsigma)$ and is given as follows:

$$U/SIML(\varsigma) = \{I_\varsigma(\alpha) : \alpha \in U\} \tag{4}$$

It is clear that, $U/SIML(\varsigma)$ is the partition of the object set U by the relation $SIML(\varsigma)$. The set $U/SIML(\varsigma)$ is called the cover of U.

Let $\varsigma, \vartheta \subseteq At$, then the lower approximation and the upper approximation of any concept $\chi \subseteq U$ using the single tolerance relation $SIML(\varsigma \cup \vartheta)$ over U are given as follows:

$$\underline{(\varsigma \cup \vartheta)}(\chi) = \left\{\alpha \ : \ I_{(\varsigma\cup\vartheta)}(\alpha) \underset{U}{\subseteq} \chi\right\} \tag{5}$$

$$\overline{(\varsigma \cup \vartheta)}(\chi) = \{\alpha \ : \ I_{(\varsigma\cup\vartheta)}(\alpha) \cap \chi \neq \phi\} \tag{6}$$

The single granulation rough set is the pair $\langle\underline{(\varsigma \cup \vartheta)}(\chi),\ \overline{(\varsigma \cup \vartheta)}(\chi)\rangle$. The boudary region of the single granulation rough set $\langle\underline{(\varsigma \cup \vartheta)}(\chi),\ \overline{(\varsigma \cup \vartheta)}(\chi)\rangle$ is denoted as $BN_{(\varsigma\cup\vartheta)}(\chi)$ and is defined as follows:

$$BN_{(\varsigma\cup\vartheta)}(\chi) = \overline{(\varsigma \cup \vartheta)}(\chi) - \underline{(\varsigma \cup \vartheta)}(\chi) \tag{7}$$

$$==\sim \underline{(\varsigma \cup \vartheta)}(\sim \chi) \tag{8}$$

The boundary region $BN_{(\varsigma\cup\vartheta)}(\chi)$ is the set of objects that are not classified into χ with a certainty.

For any $\varsigma \subseteq At$, an element $I_\varsigma(\alpha)$ of the set $U/SIML(\varsigma)$ is called an information granule. The lower approximation and the upper approximation of any concept $\chi \subseteq U$ using the multi-granulation of the information system induced by the covers $U/SIML(\varsigma)$ and $U/SIML(\vartheta)$ over U are given as follows:

$$\underline{(\varsigma + \vartheta)}(\chi) = \left\{u \ : \ I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \text{ or } I_\vartheta(\alpha) \underset{U}{\subseteq} \chi\right\} \tag{9}$$

$$\overline{(\varsigma + \vartheta)}(\chi) =\sim \underline{(\varsigma + \vartheta)}(\sim \chi) \tag{10}$$

(11)

The multi-granular rough set is the pair $\langle\underline{(\varsigma + \vartheta)}(\chi),\ \overline{(\varsigma + \vartheta)}(\chi)\rangle$. The boundary region of the multi-granular rough set $\langle\underline{(\varsigma + \vartheta)}(\chi),\ \overline{(\varsigma + \vartheta)}$ is denoted as $BN_{(\varsigma+\vartheta)}(\chi)$ and is defined as follows:

$$BN_{(\varsigma+\vartheta)}(\chi) = \overline{(\varsigma + \vartheta)}(\chi) - \underline{(\varsigma + \vartheta)}(\chi) \tag{12}$$

3 UOBRS for the incomplete information system

Let $\chi \subseteq U$ denote a decision class in the information system $I = (U, At \cup \{D\}, f)$, and $\varsigma \subseteq At$, then the lower approximation $\underline{\varsigma}(\chi)$ finds the set of objects that are classified into χ with 100% certainty. The complement of $\underline{\varsigma}(\chi)$ is the set of objects that are not classified into χ with a certainty.

Definition 1: For the attribute set ς, define a tolerance relation I_{ς^*} as follows:

$$I_{\varsigma^*}(\alpha) = \{\beta \in U : \forall\ A \in \varsigma,\ A(\alpha) = A(\beta)\}. \tag{13}$$

The cover of U by the relation I_{ς^*} is denoted as $U/SIML(\varsigma^*)$ and is defined as follows:

$$U/SIML(\varsigma^*) = \{I_{\varsigma^*}(\alpha) : \alpha \in U\}. \tag{14}$$

Definition 2: Let D be the decision class in the information system $I = (U, At \cup \{D\}, f)$ and $U/SIML(\mathcal{D})$ be the cover of U by the relation I_D, then the uncertainty region of the attribute set $\varsigma \subseteq At$ is denoted as $Unct(\varsigma)$ and is defined as follows:

$$Unct(\varsigma) = U - \bigcup_{\chi \in U/SIML(\mathcal{D})} \underline{\varsigma^*}(\chi) \tag{15}$$

where, $\underline{\varsigma^*}(\chi) = \{u \in U : I_{\varsigma^*}(u) \subseteq \chi\}$.

It is clear that the uncertainty region $Unct(\varsigma)$ is the set of objects that are not classified into any of the decision classes with certainty using the attribute set ς. The certainty region of ς is denoted as $Cert(\varsigma)$ and is defined as follows:

$$Cert(\varsigma) = \bigcup_{\chi \in U/SIML(\mathcal{D})} \underline{\mathcal{A}^*}(\chi) \tag{16}$$

The certainty region is the set of objects that are classified into any of the decision classes with a 100% certainty.

Here we introduce the uncertainty optimization-based multi-granular rough set that minimizes the uncertainty region of the rough set.

Definition 3: Let χ be any concept in the information system $I = (U, At \cup \{D\}, f)$ and $\varsigma, \vartheta \subseteq At$ such that $Cert(\vartheta) \subseteq Unct(\varsigma)$. Let $U/SIML(\varsigma)$ be the cover of U induced by the relation $SIML(\varsigma)$ over U, and $Unct(\varsigma)/SIML(\vartheta)$ be the cover of $Unct(\varsigma)$ induced by the relation $SIML(\vartheta)$ over $Unct(\varsigma)$, then the lower approximation of χ, ϑ given ς is denoted as $\underline{(\vartheta|\varsigma)}(\chi)$ and is defined as follows:

$$\underline{(\vartheta|\varsigma)}(\chi) = \left\{ \alpha \in U : I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \text{ or } I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi \right\} \tag{17}$$

The upper approximation of the concept χ, ϑ given ς is denoted as $\overline{(\vartheta|\varsigma)}(\chi)$ and is defined as follows:

$$\overline{(\vartheta|\varsigma)}(\chi) = \sim \underline{(\vartheta|\varsigma)}(\sim \chi) \tag{18}$$

The lower approximation of χ, ϑ given ς has two parts. The first part is the set $\left\{ \alpha \in U : I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \right\}$ which is nothing but the set $\underline{\varsigma}(\chi)$ in the region U. Hence the first part finds the objects that are classified into χ with a 100% certainty using attribute set ς. The second part of the lower approximation of χ, ϑ given ς is the set set $\left\{ u \in U : I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi \right\}$ which is nothing but the set $\underline{\vartheta}(\chi)$ in the region $Unct(\varsigma)$. It means that, the second part finds the objects of the region $Unct(\varsigma)$ that are classified into χ using the attribute set ϑ. Thus, $\underline{(\vartheta|\varsigma)}(\chi)$ minimizes the uncertainty region of the attribute set ς using the certainty region of the attribute set ϑ.

The uncertainty optimization-based rough set (UOBRS) is the pair $\langle \underline{(\vartheta|\varsigma)}(\chi),\ \overline{(\vartheta|\varsigma)}(\chi)$. The boundary region of the UOB-MGRS $\langle \underline{(\vartheta|\varsigma)}(\chi),\ \overline{(\vartheta|\varsigma)}(\chi)$ is denoted as $BN_{(\vartheta|\varsigma)}(\chi)$ and is defined as follows:

$$BN_{(\vartheta|\varsigma)}(\chi) = \overline{(\vartheta|\varsigma)}(\chi) - \underline{(\vartheta|\varsigma)}(\chi) \tag{19}$$

Example 2: *(Continued from Example 1)* Let $\chi = (D, 1) = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$. Let $\varsigma = \{A_1\}$ and $\vartheta = \{A_3\}$, then the covers induced by the attribute sets ς, ϑ and $(\varsigma \cup \vartheta)$ are as follows:

$$U/SIML(\varsigma) = \left\{ \begin{array}{l} \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_3, \omega_4, \omega_7\}, \\ \{\omega_1 - \omega_8\}, \{\omega_4, \omega_5, \omega_8\}, \{\omega_4, \omega_6\}, \\ \{\omega_3, \omega_4, \omega_7\}, \{\omega_4, \omega_5, \omega_8\} \end{array} \right\}$$

$$U/SIML(\vartheta) = \left\{ \begin{array}{l} \{\omega_1, \omega_2, \omega_4, \omega_8\}, \{\omega_1 - \omega_8\}, \{\omega_2, \omega_3, \omega_6\}, \\ \{\omega_1, \omega_2, \omega_4, \omega_8\}, \{\omega_2, \omega_5, \omega_7\}, \{\omega_2, \omega_3, \omega_6\}, \\ \{\omega_2, \omega_5, \omega_7\}, \{\omega_1, \omega_2, \omega_4, \omega_8\} \end{array} \right\}$$

$$U/SIML(\varsigma \cup \vartheta) = \left\{ \begin{array}{l} \{\omega_1, \omega_2, \omega_4\}, \{\omega_1, \omega_2, \omega_4\}, \{\omega_3\}, \\ \{\omega_1, \omega_2, \omega_4, \omega_8\}, \{\omega_5\}, \{\omega_6\}, \{\omega_7\}, \{\omega_8\} \end{array} \right\}$$

Let $\chi = (D, 1)$ then $\chi = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$. By calculations, we have following results:

$$\begin{aligned} \underline{(\varsigma \cup \vartheta)(\chi)} &= \{\omega_1, \omega_2, \omega_4\} \cup \{\omega_3\} \cup \{\omega_5\} \\ &= \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\} \quad (\because (5)) \\ \underline{(\varsigma + \vartheta)(\chi)} &= \{\omega_1, \omega_2, \omega_4\} \cup \phi = \{\omega_1, \omega_2, \omega_4\} \quad (\because (9)) \end{aligned}$$

The cover of U by the relation $I_{\varsigma*}$ is obtained as follows:

$$U/SIML(\varsigma^*) = \left\{ \begin{array}{l} \{\omega_1, \omega_2\}, \{\omega_1, \omega_2\}, \{\omega_3, \omega_7\}, \{\omega_4\}, \{\omega_5, \omega_8\}, \\ \{\omega_6\}, \{\omega_3, \omega_7\}, \{\omega_5, \omega_8\} \end{array} \right\}$$

The decision classes for the incomplete IS given in Table 1 are $\chi = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$, and $\sim \chi = \{\omega_6, \omega_7, \omega_8\}$. Thus by (15), we have $Unct(\varsigma) = U - \{\{\omega_1, \omega_2\} \cup \{\omega_4\} \cup \{\omega_6\}\} = U - \{\omega_1, \omega_2, \omega_4, \omega_6\} = \{\omega_3, \omega_5, \omega_7, \omega_8\}$, and by (18), $Cert(\varsigma) = \{\omega_1, \omega_2, \omega_4, \omega_6\}$. The cover of $Unct(\varsigma)$ by the relation I_{ϑ^*} is obtained as follows:

$$Unct(\varsigma)/SIML(\vartheta) = \{\{\omega_3\}, \{\omega_5, \omega_7\}, \{\omega_5, \omega_7\}, \{\omega_8\}\}$$

Therefore, by (17) we get

$$\begin{aligned}
\underline{(\vartheta|\varsigma)}(\chi) &= \left\{\alpha \in U : I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \text{ or } I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi\right\} \\
&= \left\{\alpha \in U : I_\varsigma(\alpha) \underset{U}{\subseteq} \chi\right\} \bigcup \left\{\alpha \in U : I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi\right\} \\
&= \{\omega_1, \omega_2, \omega_4\} \cup \{3\} \\
&= \{\omega_1, \omega_2, \omega_3, \omega_4\}
\end{aligned}$$

From above example, it is clear that $\underline{(\varsigma + \vartheta)}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi) \subseteq \underline{(\varsigma \cup \vartheta)}(\chi)$.

The above example motivates us to following results.

Proposition 3.1: Let $I = (U, At \cup \{D\}, f)$ be the incomplete IS, χ be any decision class in I and $\varsigma, \vartheta \subseteq At$ such that $Cert(\vartheta) \subseteq Unct(\varsigma)$, then

1. $\underline{(\varsigma + \vartheta)}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi) \subseteq \underline{(\varsigma \cup \vartheta)}(\chi)$ (20)
2. $\overline{(\varsigma \cup \vartheta)}(\chi) \subseteq \overline{(\vartheta|\varsigma)}(\chi) \subseteq \overline{(\varsigma + \vartheta)}(\chi)$ (21)

Proof:

(1a) Let any $\alpha \in \underline{(\varsigma + \vartheta)}(\chi)$ then by (9)

$$\begin{aligned}
&\alpha \in \underline{(\varsigma + \vartheta)}(\chi) \\
&\Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \ \ or \ \ I_\vartheta(\alpha) \underset{U}{\subseteq} \chi \\
&\Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \ or \ I_\vartheta(\alpha) \underset{P}{\subseteq} \chi, \quad \forall P \subseteq U \\
&\Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \ or \ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi, \ (\because Unct(\varsigma) \subseteq U) \\
&\Rightarrow \alpha \in \underline{(\vartheta|\varsigma)}(\chi), \ (\because (17)
\end{aligned}$$

Therefore, $\underline{(\varsigma + \vartheta)}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi)$.

(1b) Let any $\alpha \in \underline{(\vartheta|\varsigma)}(\chi)$. By equation (17),

$$\begin{aligned}
&\alpha \in \underline{(\vartheta|\varsigma)}(\chi) \\
&\Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \ \ or \ \ I_\vartheta(\alpha) \underset{Unc(\varsigma)}{\subseteq} \chi \\
&\Rightarrow I_\varsigma(\alpha) \cap I_\vartheta(\alpha) \underset{U}{\subseteq} \chi \\
&\Rightarrow \alpha \in \underline{(\varsigma \cup \vartheta)}(\chi) \\
&\Big(\because \underline{(\varsigma \cup \vartheta)}(\chi) = \{A \cap B : A \in U/SIML(\varsigma), \\
&\qquad\qquad B \in U/SIML(\vartheta), A \cap B \underset{U}{\subseteq} \chi\}\Big)
\end{aligned}$$

Therefore $\underline{(\vartheta|\varsigma)}(\chi) \subseteq \underline{(\varsigma \cup \vartheta)}(\chi)$.

(2) By part (1), $\underline{(\varsigma + \vartheta)}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi) \subseteq \underline{(\varsigma \cup \vartheta)}(\chi)$. Replacing χ by $\sim \chi$, we get $\underline{(\varsigma + \vartheta)}(\sim \chi) \subseteq \underline{(\vartheta|\varsigma)}(\sim \chi) \subseteq \underline{(\varsigma \cup \vartheta)}(\sim \chi)$. Taking complement of each side, we get $\sim \underline{(\varsigma \cup \vartheta)}(\sim \chi) \subseteq \sim \underline{(\vartheta|\varsigma)}(\sim \chi) \subseteq \sim \underline{(\varsigma + \vartheta)}(\sim \chi)$. By definition of upper approximation, we get $\overline{(\varsigma \cup \vartheta)}(\chi) \subseteq \overline{(\vartheta|\varsigma)}(\chi) \subseteq \overline{(\varsigma + \vartheta)}(\chi)$.

Proposition 3.2: Let $BN_{At}(\chi)$ denote the boundary region of the rough set $\langle \underline{At}(\chi),\ \overline{At}(\chi)\rangle$, then

$$BN_{(\varsigma\cup\vartheta)}(\chi) \subseteq BN_{(\vartheta|\varsigma)}(\chi) \subseteq BN_{(\varsigma+\vartheta)}(\chi). \tag{22}$$

Proof: By Proposition 3.1, we have $\underline{(\varsigma+\vartheta)}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi) \subseteq \underline{(\varsigma\cup\vartheta)}(\chi)$ and $\overline{(\varsigma\cup\vartheta)}(\chi) \subseteq \overline{(\vartheta|\varsigma)}(\chi) \subseteq \overline{(\varsigma+\vartheta)}(\chi)$. But we know that, $\underline{(\varsigma\cup\vartheta)}(\chi) \subseteq \overline{(\varsigma\cup\vartheta)}(\chi)$. Thus $\underline{(\varsigma+\vartheta)}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi) \subseteq \underline{(\varsigma\cup\vartheta)}(\chi) \subseteq \overline{(\varsigma\cup\vartheta)}(\chi) \subseteq \overline{(\vartheta|\varsigma)}(\chi) \subseteq \overline{(\varsigma+\vartheta)}(\chi)$. Obviously, $\overline{(\varsigma\cup\vartheta)}(\chi) - \underline{(\varsigma\cup\vartheta)}(\chi) \subseteq \overline{(\vartheta|\varsigma)}(\chi) - \underline{(\vartheta|\varsigma)}(\chi) \subseteq \overline{(\varsigma+\vartheta)}(\chi) - \underline{(\varsigma+\vartheta)}(\chi)$. Thus, by definition of the boundary region of the rough set, $BN_{(\varsigma\cup\vartheta)}(\chi) \subseteq BN_{(\vartheta|\varsigma)}(\chi) \subseteq BN_{(\varsigma+\vartheta)}(\chi)$.

Proposition 3.3: Let $I = (U, At \cup \{D\}, f)$ be the incomplete IS, χ be the decision class in I and $\varsigma, \vartheta \subseteq At$, then

1. $\underline{(\vartheta|\varsigma)}(\chi) \subseteq \chi \subseteq \overline{(\vartheta|\varsigma)}(\chi)$ (23)
2. $\underline{(\vartheta|\varsigma)}(\phi) = \overline{(\vartheta|\varsigma)}(\phi) = \phi$
   and $\underline{(\vartheta|\varsigma)}(U) = \overline{(\vartheta|\varsigma)}(U) = U$ (24)
3. $\underline{(\vartheta|\varsigma)}((\sim\chi)) =\sim \overline{(\vartheta|\varsigma)}(\chi)$
   and $\overline{(\vartheta|\varsigma)}((\sim\chi)) =\sim \underline{(\vartheta|\varsigma)}(\chi)$ (25)
4. $\underline{\varsigma}(\chi) \cup \underline{\vartheta}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi)$ (26)

Proof: If $\varsigma = \vartheta (\varsigma, \vartheta \subseteq At)$, then $\left\{\alpha \in U : I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi\right\} = \phi$. Thus $\underline{(\vartheta|\varsigma)}(\chi)$ degenerates to $\underline{\varsigma)}(\chi)$. Hence by Pawlak's rough set theory [5], all the results are true.

Let us suppose that $\varsigma \neq \vartheta$.

(1a) Let any $\alpha \in \underline{(\vartheta|\varsigma)}(\chi) \Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi$ or $I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi$. This implies that $\alpha \in \chi$. Thus $\underline{(\vartheta|\varsigma)}(\chi) \subseteq \chi$.

(1b) We know that, $\overline{(\vartheta|\varsigma)}(\chi) =\sim \underline{(\vartheta|\varsigma)}(\sim\chi)$. Taking complement of both sides, we get $\sim \overline{(\vartheta|\varsigma)}(\chi) = \underline{(\vartheta|\varsigma)}(\sim\chi))$. But $\underline{(\vartheta|\varsigma)}(\sim\chi) \subseteq\sim \chi$. Therefore $\sim \overline{(\vartheta|\varsigma)}(\chi) \subseteq\sim \chi$. Again taking complement of both sides, we get $\chi \subseteq \overline{(\vartheta|\varsigma)}(\chi)$.

(2a) By part (1), we have $\underline{(\vartheta|\varsigma)}(\phi) \subseteq \phi$. But $\phi \subseteq \underline{(\vartheta|\varsigma)}(\phi)$. Therefore, $\underline{(\vartheta|\varsigma)}(\phi) = \phi$.

(2b) Assume that, $\overline{(\vartheta|\varsigma)}(\phi) \neq \phi$. It means that, there exists some $\alpha \in \overline{(\vartheta|\varsigma)}(\phi)$. This implies that, $I_\varsigma(\alpha) \underset{U}{\subseteq} \phi$ or $I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \phi$. This is a contradiction as ϕ is an empty set. Therefore $\overline{(\vartheta|\varsigma)}(\phi) = \phi$.

(2c) By part (1), we have $\underline{(\vartheta|\varsigma)}(U) \subseteq U$. Now, let any $\alpha \in U$, then obviously $I_\varsigma(\alpha) \underset{U}{\subseteq} U$ or $I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} U$. Thus, by definition 17, $\alpha \in \underline{(\vartheta|\varsigma)}(U)$. Therefore $\underline{(\vartheta|\varsigma)}(U) = U$.

(2d) From part (1), $U \subseteq \overline{(\vartheta|\varsigma)}(U)$. Also we know that, $\overline{(\vartheta|\varsigma)}(U) \subseteq U$. Therefore, we have $\overline{(\vartheta|\varsigma)}(U) = U$.

(3a) By definition of upper approximation, we know that, $\overline{(\vartheta|\varsigma)}(\chi_) =\sim \underline{(\vartheta|\varsigma)}((\sim\chi))$. Taking complement of both sides, we get $\underline{(\vartheta|\varsigma)}((\sim\chi)) =\sim \overline{(\vartheta|\varsigma)}(\chi)$.

(3b) We have $\overline{(\vartheta|\varsigma)}(\chi) =\sim \underline{(\vartheta|\varsigma)}((\sim\chi))$. Replacing χ by $\sim\chi$, we get $\overline{(\vartheta|\varsigma)}((\sim\chi)) =\sim \underline{(\vartheta|\varsigma)}((\sim(\sim\chi)))$. Therefore, $\overline{(\vartheta|\varsigma)}((\sim\chi)) =\sim \underline{(\vartheta|\varsigma)}(\chi)$.

(4) We know that, for any $\alpha \in \underline{\varsigma}(\chi) \Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi$. This implies that, $\alpha \in \underline{(\varsigma|\vartheta)}(\chi)$. Thus $\underline{\varsigma}(\chi) \subseteq \underline{(\varsigma|\vartheta)}(\chi)$. Next, for any $\alpha \in \underline{\vartheta}(\chi) \Rightarrow I_\vartheta(\alpha) \underset{U}{\subseteq} \chi$. But we know that, if $I_\vartheta(\alpha) \underset{U}{\subseteq} \chi$ then $I_\vartheta(\alpha) \underset{P}{\subseteq} \chi, \forall P \subseteq U$. As $Unct(\varsigma) \subseteq U$, we have $I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi$. Therefore, $\alpha \in \underline{(\vartheta|\varsigma)}(\chi)$. This gives $\underline{\vartheta}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi)$. Hence the result $\underline{\varsigma}(\chi) \cup \underline{\vartheta}(\chi) \subseteq \underline{(\vartheta|\varsigma)}(\chi)$.

Proposition 3.4: Let χ_1 and χ_2 be any two decision classes in the information system $(U, At \cup \{D\}, f)$ and $\varsigma, \vartheta \subseteq At$ then

$$\underline{(\vartheta|\varsigma)}(\chi_1) \cup \underline{(\vartheta|\varsigma)}(\chi_2) \subseteq \underline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_2) \tag{27}$$

Proof: Note that, for any $\alpha \in U$, attribute set $\varsigma \subseteq At$, and decision class χ, if $I_\varsigma(\alpha) \underset{U}{\subseteq} \chi \Rightarrow I_\varsigma(\alpha) \underset{P}{\subseteq} \chi, \ \forall P \subseteq U$

$$\begin{aligned}
&\text{Let any } \alpha \in \underline{(\vartheta|\varsigma)}(\chi_1)\\
&\Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_1 \ or \ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_1\\
&\Rightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} (\chi_1 \cup \chi_2) \ or \ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} (\chi_1 \cup \chi_2)\\
&\qquad\qquad\qquad\qquad (\because \chi_1 \subseteq (\chi_1 \cup \chi_2))\\
&\Rightarrow \alpha \in \underline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_2)
\end{aligned}$$

Thus $\underline{(\vartheta|\varsigma)}(\chi_1) \subseteq \underline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_2)$. Similarly, we can prove that $\underline{(\vartheta|\varsigma)}(\chi_2) \subseteq \underline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_2)$. Hence $\underline{(\vartheta|\varsigma)}(\chi_1) \cup \underline{(\vartheta|\varsigma)}(\chi_2) \subseteq \underline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_2)$.

Proposition 3.5: Let χ_1 and χ_2 be any two decision classes in the incomplete IS $(U, At \cup \{D\}, f)$ and $\varsigma, \vartheta \subseteq At$ then,

$$\underline{(\vartheta|\varsigma)}(\chi_1 \cap \chi_2) \subseteq \underline{(\vartheta|\varsigma)}(\chi_1) \cap \underline{(\vartheta|\varsigma)}(\chi_2) \tag{28}$$

Proof:

$$\begin{aligned}
&\text{Let any } \alpha \in \underline{(\vartheta|\varsigma)}(\chi_1 \cap \chi_2)\\
&\Rightarrow \left\{I_\varsigma(\alpha) \underset{U}{\subseteq} (\chi_1 \cap \chi_2)\right\} \vee \left\{I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} (\chi_1 \cap \chi_2)\right\}\\
&\Rightarrow \{I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_1 \ \wedge \ I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_2\}\\
&\qquad \vee \left\{I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_1 \wedge I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_2\right\}\\
&\Rightarrow \left\{\{I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_1\} \vee \left\{I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_1\right\}\right\}\\
&\qquad \wedge \left\{\{I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_2\} \vee \left\{I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_2\right\}\right\}\\
&\qquad \wedge \left\{\{I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_1\} \vee \left\{I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_2\right\}\right\}\\
&\qquad \wedge \left\{\{I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_2\} \vee \left\{I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_1\right\}\right\}\\
&\Rightarrow \alpha \in \left\{\left\{\underline{(\vartheta|\varsigma)}(\chi_1) \cap \underline{(\vartheta|\varsigma)}(\chi_2)\right\}\right.\\
&\qquad \cap \left\{\alpha \in U : I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_1 \bigvee I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_2\right\}\\
&\qquad \left. \cap \left\{\alpha \in U : I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_2 \bigvee I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_1\right\}\right\}\\
&\Rightarrow \alpha \in \underline{(\vartheta|\varsigma)}(\chi_1) \cap \underline{(\vartheta|\varsigma)}(\chi_2)
\end{aligned}$$

$\therefore \underline{(\vartheta|\varsigma)}(\chi_1 \cap \chi_2) \subseteq \underline{(\vartheta|\varsigma)}(\chi_1) \cap \underline{(\vartheta|\varsigma)}(\chi_2)$

Proposition 3.6: Let χ_1 and χ_2 be any two decision classes in the incomplete IS $(U, At \cup \{D\}, f)$ and $\varsigma, \vartheta \subseteq At$ then,

$$\underline{(\vartheta|\varsigma)}\left(\sim (\chi_1 \cup \chi_2)\right) = \underline{(\vartheta|\varsigma)}\left((\sim \chi_1) \cap (\sim \chi_2)\right) \tag{29}$$

Proof:

$$\begin{aligned}
&\text{Let any } \alpha \in \underline{(\vartheta|\varsigma)}\,(\sim(\chi_1 \cup \chi_2)) \\
&\iff I_\varsigma(\alpha) \underset{U}{\subseteq} \sim(\chi_1 \cup \chi_2) \ \bigvee\ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \sim(\chi_1 \cup \chi_2) \\
&\iff I_\varsigma(\alpha) \underset{U}{\subseteq} (\sim\chi_1) \ \cap\ (\sim\chi_2) \\
&\qquad \bigvee\ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} (\sim\chi_1) \ \cap\ (\sim\chi_2) \\
&\iff \alpha \in \underline{(\vartheta|\varsigma)}\,((\sim\chi_1) \cap (\sim\chi_2))\ (\because \text{By17})
\end{aligned}$$

$\therefore \underline{(\vartheta|\varsigma)}\,(\sim(\chi_1 \cup \chi_2)) = \underline{(\vartheta|\varsigma)}\,((\sim\chi_1) \cap (\sim\chi_2))$

Proposition 3.7: Let χ_1 and χ_2 be any two decision classes in the incomplete IS $(U, At \cup \{D\}, f)$ and $\varsigma, \vartheta \subseteq At$ then,

$$\overline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_2) \supseteq \overline{(\vartheta|\varsigma)}(\chi_1) \cup \overline{(\vartheta|\varsigma)}(\chi_2) \tag{30}$$

Proof: By Proposition 3.5, we have

$$\underline{(\vartheta|\varsigma)}\,((\sim\chi_1) \cap (\sim\chi_2)) \subseteq \underline{(\vartheta|\varsigma)}((\sim\chi_1)) \cap \underline{(\vartheta|\varsigma)}(\sim\chi_2)$$

Taking complement of both sides, we get

$$\sim\left\{\underline{(\vartheta|\varsigma)}(\sim\chi_1) \cap \underline{(\vartheta|\varsigma)}(\sim\chi_2)\right\} \subseteq \sim\left\{\underline{(\vartheta|\varsigma)}((\sim\chi_1) \cap (\sim\chi_2))\right\}$$

This implies that,

$$\sim\underline{(\vartheta|\varsigma)}(\sim\chi_1) \cup \sim\underline{(\vartheta|\varsigma)}(\sim\chi_2) \subseteq \sim\underline{(\vartheta|\varsigma)}(\sim(\chi_1 \cup \chi_2))$$

Using the definition of the upper approximation, we get

$$\overline{(\vartheta|\varsigma)}(\chi_1) \cup \overline{(\vartheta|\varsigma)}(\chi_2) \subseteq \overline{(\vartheta|\varsigma)}(\chi_1 \cup \chi_{2)}.$$

Proposition 3.8: Let χ_1 and χ_2 be any two decision classes in the IS $(U, At \cup \{d\}, V)$ and $\varsigma, \vartheta \subseteq At$, then

1. $\chi_1 \subseteq \chi_2 \Longrightarrow \underline{(\vartheta|\varsigma)}(\chi_1) \subseteq \underline{(\vartheta|\varsigma)}(\chi_2)$ (31)
2. $\chi_1 \subseteq \chi_2 \Longrightarrow \overline{(\vartheta|\varsigma)}(\chi_1) \subseteq \overline{(\vartheta|\varsigma)}(\chi_2)$ (32)

Proof:

1. Let $\chi_1 \subseteq \chi_2$ then,

$$\begin{aligned}
&\alpha \in \underline{(\vartheta|\varsigma)}(\chi_1) \\
&\Longrightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_1 \ \vee\ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_1 \\
&\Longrightarrow I_\varsigma(\alpha) \underset{U}{\subseteq} \chi_2 \ \vee\ I_\vartheta(\alpha) \underset{Unct(\varsigma)}{\subseteq} \chi_2 \ \ (\because \chi_1 \subseteq \chi_2) \\
&\Longrightarrow \alpha \in \underline{(\vartheta|\varsigma)}(\chi_2)\ (\because \text{by equation}(17))
\end{aligned}$$

   Therefore, $\underline{(\vartheta|\varsigma)}(\chi_1) \subseteq \underline{(\vartheta|\varsigma)}(\chi_2)$

2. Let $\chi_1 \subseteq \chi_2$ then,

$$\begin{aligned}
\chi_1 \subseteq \chi_2 &\Rightarrow \sim\chi_2 \subseteq \sim\chi_1 \\
&\Rightarrow \underline{(\vartheta|\varsigma)}(\sim\chi_2) \subseteq \underline{(\vartheta|\varsigma)}(\sim\chi_1)(\because \text{by part}(1)) \\
&\Rightarrow \sim\underline{(\vartheta|\varsigma)}(\sim\chi_1) \subseteq \sim\underline{(\vartheta|\varsigma)}(\sim\chi_2) \\
&\Rightarrow \overline{(\vartheta|\varsigma)}(\chi_1) \subseteq \overline{(\vartheta|\varsigma)}(\chi_2)(\because \text{by equation}(18))
\end{aligned}$$

To apply the proposed method in practice, here we give an algorithm that finds the lower approximation of the decision class in the incomplete IS corresponding to one attribute set given the other attribute set. Let $\varsigma, \vartheta \subseteq At$ and At be any concept in the information system.

Algorithm 1 For the incomplete IS $(U, At \cup \{D\}, f)$, attribute sets $\varsigma, \vartheta \subseteq At$ and decision class $\chi \subset U$, this algorithm finds $\underline{(\vartheta|\varsigma)}(\chi)$

Assign $Unct(\varsigma) \longleftarrow \phi, Cert(\varsigma) \longleftarrow \phi, \underline{\varsigma}(\chi) \longleftarrow \phi, \underline{(\vartheta|\varsigma)}(\chi) \longleftarrow \phi$
Compute cover $U/SIML(\varsigma) = \{a_1, a_2, ...a_m\}$
Compute cover $U/SIML(D) = \{\chi_1, \chi_2, ..., \chi_k\}$
for i=1:m **do**
  if $\frac{|a_i \cap \chi|}{|a_i|} = 1$ **then**
    $\underline{\varsigma}(\chi) = \underline{\varsigma}(\chi) \cup a_i$
  end if
end for
$\underline{(\vartheta|\varsigma)}(\chi) = \underline{\varsigma}(\chi)$
Compute cover $U/SIML(\varsigma^*) = \{a_1^*, a_2^*, ...a_n^*\}$
for i=1:n **do**
  for j=1:k **do**
    if $\frac{|a_i^* \cap \chi_j|}{|a_i^*|} = 1$ **then**
      $Cert(\varsigma) = Cert(\varsigma) \cup a_i^*$
    end if
  end for
end for
$Unct(\varsigma) = U - Cert(\varsigma)$
$U = Unc(\varsigma)$
Find the cover $U/SIML(\vartheta) = \{b_1, b_2, ..., b_l\}$
for $i = 1 : l$ **do**
  if $\frac{|b_i \cap \chi|}{|b_i|} = 1$ **then**
    $\underline{(\vartheta|\varsigma)}(\chi) = \underline{(\vartheta|\varsigma)}(\chi) \cup b_i$
  end if
end for

4 Uncertainty optimization-based feature subset selection

In this section, we give the application of the uncertainty optimization-based rough set for the incomplete IS. Let D be the decision attribute in the incomplete information system $I = (U, At \cup \{d\}, f)$, then the positive region [5] of the incomplete IS for the decision attribute D using the attribute set $\varsigma \subseteq At$ is defined as follows:

$$Pos_{\varsigma}(D) = \bigcup_{\chi \in U/SIML(\mathcal{D})} \underline{\mathcal{A}^*}(\chi) \tag{33}$$

It should be noted that, $Pos_{\varsigma}(D) = Cert(\varsigma)$ Clearly, the attribute set with biggest certainty region is the most informative attribute set regarding the decision attribute. We shall make use of the positive region of the decision table to find the most informative feature subsets of the given attribute set.

Definition 4: The approximation quality of the decision attribute D by the given attribute set ς is defined as follows:

$$\gamma(\varsigma, D)) = \frac{|Cert(\varsigma)|}{|\mathbb{U}|} \tag{34}$$

The attribute whose approximation quality is the maximum preserves the maximum information and is considered as the most informative feature in the attribute set.

Here, we specify the uncertainty minimization problem for finding the most informative feature subset of the given attribute set. Let $\mathcal{S} \subseteq At$, and we have to minimize the uncertainty region of S using the attributes of At. If $\mathcal{S}$ is empty, then the uncertainty region of S is the complete set U. The uncertainty minimization problem is specified below.

$$\text{Minimize } \; Unct(\mathcal{S}) \tag{35}$$

Subject to,

$$Cert(\varsigma|\mathcal{S}) = \bigcup_{\chi\in U/SIML(D)} \left\{u \in U : I_{\mathcal{S}}(u) \underset{U}{\subseteq} \chi \text{ or } I_{\varsigma}(u) \underset{Unct(\varsigma)}{\subseteq} \chi\right\}, \quad \varsigma \in At \tag{36}$$

and

$$Cert(\mathcal{S},\varsigma) = \bigcup_{\chi\in U/SIML(D)} \left\{u \in U : I_{\mathcal{S}\cup\varsigma}(u) \underset{U}{\subseteq} \chi\right\}, \quad \varsigma \in At \tag{37}$$

Algorithm 2 finds the most informative feature subset using uncertainty optimization. We find the approximation quality of each individual feature of the given attribute set. The feature with the maximum approximation quality is the most informative in the given attribute set. It is desirable to minimize the uncertainty region of the most informative feature so that we can find the feature subset with increased approximation quality. Hence we minimize the uncertainty region of the most informative attributes for feature subset selection.

Algorithm 2 For the incomplete IS $(U, At \cup \{D\}, f)$, this algorithm finds the most informative feature subset $\mathcal{S}$ of At using uncertainty optimization

```
Assign S ⟵ ϕ, Unct(S) ⟵ U,
Compute cover U/SIML(D) = {χ_1, χ_2, ..., χ_l}
while γ(S, D)) ≠ 1 do
    for i=1:|At| do
        Compute cover U/SIML(ς_i) = {a^i_1, a^i_2, ...a^i_m}
        for j=1:m do
            for k=1:l do
                if |a^i_j ∩ χ_k| / |a^i_j| = 1 then
                    Cert(ς_i) = Cert(ς_i) ∪ a^j_i
                end if
            end for
        end for
    end for
    A = max(|Cert(A_1)|, |Cert(A_2)|, ..., |Cert(A_|At|)|)
    R = ϕ, Cert(R) = ϕ
    for p = 1 : |At| do
        for m = 1 : |At| do
            if |Cert(A_m)| = A then
                R = A_m
                Cert(R) = Cert(A_m)
            end if
        end for
    end for
    Compute Cert(S, R)
    Cert(S) = Cert(S, R)
    S = S ∪ R
    At = At − S
    U = U − Cert(S)
    γ(S, D) = |Cert(S)| / |U|
end while
```

4.1 Experimental analysis

To check the effectiveness of the proposed algorithm for feature subset selection, we have used three publicly available datasets. As our methods use discrete attribute values, we transformed the continuous attribute values into discrete ones using equal frequency discretization. The discription of the used datasets is given in Table 2.

Dataset	objects	Attributes	Classes
Breast-cancer wisconsin	699	9	2
Hepatitis	155	19	2
crx	690	15	2

Table 2. Description of the used aatasets

In Table 3, we give the iteration wise feature subset formed using Algorithm 2. The dataset used for this purpose is the breast-cancer wisconsin. It has nine discrete attributes and the binary decision classes. We considered the sequence of attributes as $f_1, f_2, ..., f_n$.

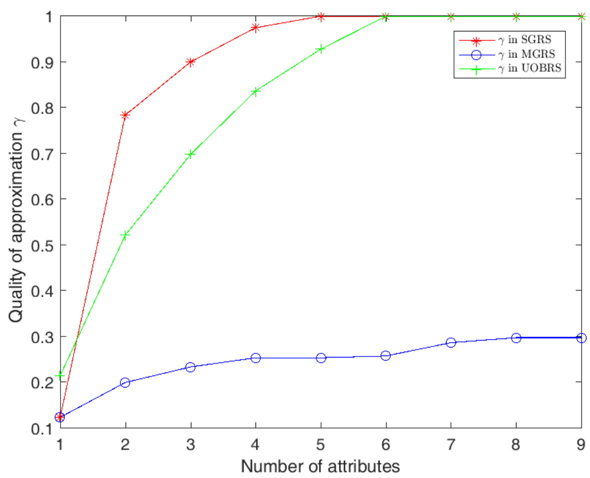

Figure 1. Quality of approximations for BCW dataset

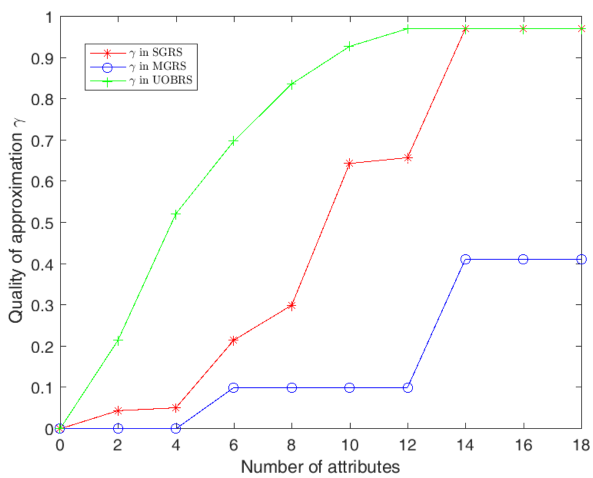

Figure 2. Quality of approximations for hepatitis dataset

From Table 3, it is seen that the most informative attribute is the f_5, and its approximation quality is 0.215. Hence we minimized the $Unct(f_5)$ using the remaining attributes of At. In iteration 2, the attribute with the maximum value of approximation quality is obtained as f_1. So feature subset formed in iteration 2 is $\{f_5, f_1\}$. In iteration 3, we minimized $Unct(f_1, f_5)$ and so on until we get approximation quality equals to 1. In this way, the feature subset selection for the breast-cancer Wisconsin dataset using uncertainty optimization is obtained as $\{f_5, f_1, f_9, f_3, f_7, f_8, f_2\}$. We have compared the approximation quality of the feature subsets formed using uncertainty optimization with the approximation quality obtained using a single granulation rough set and multi-granulation rough set. Table 4-6 and Figure 1-3 compare the approximation quality obtained for the dataset using different approaches.

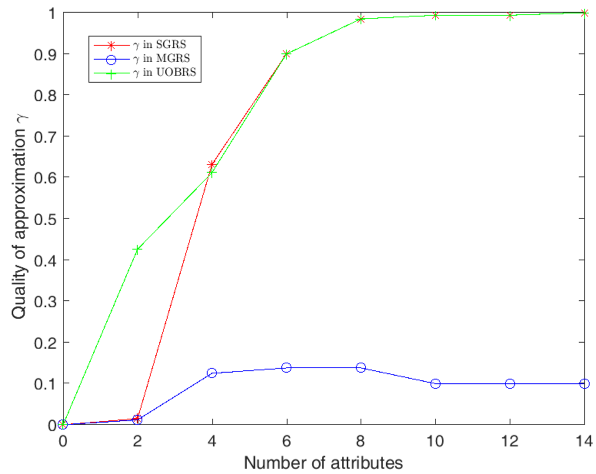

Figure 3. Quality of approximations for crx dataset

The breast-cancer Wisconsin dataset has nine attributes and the binary decision classes. It contains the information of 699 breast cancer patients using nine attributes. The decision classes are the types of breast cancer known as malignant and benign. The variation of approximation quality against number of attributes for the breast-cancer Wisconsin dataset using different approaches is shown in Table 4. For the breast-cancer Wisconsin dataset, we get 100% approximation quality with five attributes using SGRS, whereas the UOBRS requires seven attributes to get 100% approximation quality. The MGRS fails to get 100% approximation quality with all the given attributes. With MGRS, we get approximately 29% approximation quality using nine attributes. Thus, UOBRS gives better approximation quality with less number of attributes as compared to MGRS. The SGRS approximation quality is the highest among the used approaches.

The hepatitis dataset contains the information of 155 patients using 19 attributes. It has binary decision classes and 13 attributes are discrete, whereas six are continuous. As rough set uses only discrete attribute values, we need to discretize the continuous attribute values into discrete ones before applying the rough set methodology. We discretized the continuous attribute values using equal frequency binning to completely discrete attribute values. Table 5 shows the approximation quality variation against the number of attributes for the hepatitis dataset. From Table 5, we can see that the SGRS requires 14 attributes to get about 97% approximation quality, whereas UOBRS requires only 12 attributes to get 97% approximation quality. The highest approximation quality with the MGRS is approximately 41%. Thus, UOBRS gives a higher approximation quality with fewer attributes than SGRS and MGRS.

	Region	$\gamma(f_1, D)$	$\gamma(f_2, D)$	$\gamma(f_3, D)$	$\gamma(f_4, D)$	$\gamma(f_5, D)$	$\gamma(f_6, D)$	$\gamma(f_7, D)$	$\gamma(f_8, D)$	$\gamma(f_9, D)$	At	$\gamma(AT, D)$
Itr 1	U	0	0	0	0	0.215	0	0	0	0	$\{f_5\}$	0.215
Itr 2	$Unct(At)$	0.388	0.135	0.049	0.091	-	0.072	0	0	0.114	$\{f_5, f_1\}$	0.522
Itr 3	$Unct(At)$	-	0.079	0.036	0	-	0.038	0.103	0.036	0.178	$\{f_5, f_1, f_9\}$	0.698
Itr 4	$Unct(At)$	-	0	0.138	0	-	0.082	0.059	0.047	-	$\{f_5, f_1, f_9, f_3\}$	0.837
Itr 5	$Unct(At)$	-	0	-	0.046	-	0.024	0.092	0	-	$\{f_5, f_1, f_9, f_3, f_7\}$	0.928
Itr 6	$Unct(At)$	-	0	-	0	-	0	-	0.044	-	$\{f_5, f_1, f_9, f_3, f_7, f_8\}$	0.971
Itr 7	$Unct(At)$	-	0.031	-	0	-		0	-	-	$\{f_5, f_1, f_9, f_3, f_7, f_8, f_2\}$	1

Table 3. Iteration wise feature subset form for the breast-cancer wisconsin dataset

$\gamma(At, D)$	Number of attributes								
	1	2	3	4	5	6	7	8	9
γ in SGRS [23]	0.124	0.784	0.90	0.975	1	1	1	1	1
γ in MGRS [23]	0.124	0.2	0.234	0.254	0.254	0.254	0.258	0.287	0.298
γ in UOMGRS	0.215	0.522	0.698	0.837	0.928	0.971	1	1	1

Table 4. Quality of approximation with different approaches for dataset breast-cancer wisconsin

$\gamma(At, D)$	Number of attributes								
	2	4	6	8	10	12	14	16	18
γ in SGRS [23]	0.044	0.051	0.215	0.3	0.644	0.658	0.971	0.971	0.971
γ in MGRS [23]	0	0	0.1	0.1	0.1	0.1	0.412	0.412	0.412
γ in UOMGRS	0.214	0.315	0.542	0.625	0.812	0.971	0.971	0.971	0.971

Table 5. Quality of approximation with different approaches for dataset hepatitis

The third dataset used in the study is the crx dataset. The crx dataset contains 690 objects information using 15 attributes and has binary decision classes. It has six continuous attributes and nine discrete attributes. Hence we discretized the continuous attribute values of the crx dataset using equal frequency binning. After discretization, we applied the UOBRS methodology to

$\gamma(At, D)$	Number of attributes						
	2	4	6	8	10	12	14
γ in SGRS [23]	0.0125	0.632	0.9	0.985	0.994	0.994	1
γ in MGRS [23]	0	0.125	0.012	0.138	0.138	0.1	0.1
γ in UOMGRS	0.426	0.612	0.9	0.985	0.994	0.994	1

Table 6. Quality of approximation with different approaches for dataset crx

get the approximation quality of the attributes. The variation of approximation quality against number of attributes for the crx dataset using different approaches is shown in Table 6. In our study, we found that the approximation quality of the UOBRS is about 98.5% with eight attributes and 100% with the use of 14 attributes. The SGRS also gives the same approximation quality as UOBRS, but the difference is that we get higher approximation quality with fewer attributes. The Approximation quality of the MGRS with the complete attribute set is 10%. Thus MGRS fails to find higher approximation quality of the attributes.

5 Discussion and Conclusions

In the incomplete information system, the attribute values lead to uncertainty in the classification of the objects. In this work, we have defined the certainty and uncertainty region of the attributes in the incomplete IS. The certainty region is the set of attribute values that signify a particular happening with certainty. We have given a mathematical framework to find the certainty and uncertainty region of the attributes using the rough set. As a result, we have proposed a novel concept of uncertainty optimization-based rough set (UOBRS) for the incomplete IS. The UOBRS is the extension of incomplete multi-granulation rough set in the context of uncertainty optimization. We have listed important results regarding the approximation using UOBRS. The uncertainty optimization problem for minimization of attribute uncertainty region is formulated. The UOBRS gives higher approximation quality with lesser number of attributes as compared to SGRS and MGRS. The present work gives a novel approach to deal with incomplete information systems for maximum information retrieval.

REFERENCES

[1] Qian, Y., Liang, J., Yao, Y., Dang, C., ”MGRS: A multi-granulation rough set,” Information sciences, vol 180, no. 6, pp. 949-970, 2010. doi:https://doi.org/10.1016/j.ins.2009.11.023

[2] Yao, Y., ”Information granulation and rough set approximation,” International Journal of Intelligent Systems, Vol 16, no. 1, pp. 87-104, 2001. doi:https://doi.org/10.1002/1098-111X(200101)16

[3] Liang, J., Shi, Z., ”The information entropy, rough entropy and knowledge granulation in rough set theory,” International journal of uncertainty, fuzziness and knowledge-based systems, vol. 12, no. 1, pp.37-46, 2004. doi:https://doi.org/10.1142/S0218488504002631

[4] Pawlak, Z., ”Rough set theory and its applications to data analysis,” Cybernetics & Systems, vol. 29, no. 7, pp. 661-688, 1998. doi:https://doi.org/10.1080/019697298125470

[5] Pawlak, Z., ”Rough set theory and its applications,” Journal of Telecommunications and information technology, pp.7-10, 2002

[6] Pawlak, Z., ”Vagueness and uncertainty: a rough set perspective,” Computational intelligence, vol. 11, no. 2, pp.227-232, 1995. doi: https://doi.org/10.1111/j.1467-8640.1995.tb00029.x

[7] Zhao, D., Song, H., Li, H., ”Fuzzy integrated rough set theory situation feature extraction of network security,” Journal of Intelligent & Fuzzy Systems, (Preprint), pp.1-12, 2021. doi:10.3233/JIFS-189664

[8] Lei, L., ”Wavelet neural network prediction method of stock price trend based on rough set attribute reduction,” Applied Soft Computing, vol. 62, pp.923-932, 2018. doi:https://doi.org/10.1016/j.asoc.2017.09.029

[9] Sinha, A.K., Namdev, N., Kumar, A., ”Rough set method accurately predicts unknown protein class/family of leishmania donovani membrane proteome,” Mathematical Biosciences, vol. 301, pp.37-49, 2018. doi:https://doi.org/10.1016/j.mbs.2018.03.027

[10] Ziarko, W., ”Variable precision rough set model,” Journal of computer and system sciences, vol. 46, no. 1, pp. 39-59, 1993. doi:https://doi.org/10.1016/0022-0000(93)90048-2

[11] Lin, G., Qian, Y. Li, J., ”NMGRS: Neighborhood-based multigranulation rough sets,” International Journal of Approximate Reasoning, vol. 53, no. 7, pp.1080-1093, 2012. https://doi.org/10.1016/j.ijar.2012.05.004

[12] Yao, Y., Yao, B., ”Covering based rough set approximations,” Information Sciences, vol. 200, pp.91-107, 2012. doi:https://doi.org/10.1016/j.ins.2012.02.065

[13] Dubois, D., Prade, H., ”Rough fuzzy sets and fuzzy rough sets,” International Journal of General System, vol. 17, no. 2-3, pp.191-209, 1990. doi:https://doi.org/10.1080/03081079008935107

[14] Yao, Y., ”Probabilistic rough set approximations,” International Journal of approximate reasoning, vol. 49, no. 2, pp.255-271, 2008. doi:https://doi.org/10.1016/j.ijar.2007.05.019

[15] Shende, P., Sinha, A.K., ”A Novel Concept of Uncertainty Optimization Based Multi-Granular Rough Set and Its Application,” Mathematics and Statistics, vol. 9, no. 4, pp. 608-616, 2021. doi:10.13189/ms.2021.090420

[16] Abdel-Basset, M. Mohamed, M., ”The role of single-valued neutrosophic sets and rough sets in the smart city: Imperfect and incomplete information systems,” Measurement, vol. 124, no. 10, pp. 47-55, 2018. doi:https://doi.org/10.1016/j.measurement.2018.04.001

[17] Yang, X., Li, T., Tan, A., ”Three-way decisions in fuzzy incomplete information systems,” International Journal of Machine Learning and Cybernetics, vol. 11, no. 3, pp.667-674, 2020. doi:https://doi.org/10.1007/s13042-019-01025-1

[18] Hamed, A., Sobhy, A., Nassar, H., ”Distributed approach for computing rough set approximations of big incomplete information systems,” Information Sciences, vol. 547, pp.427-449, 2021. doi:https://doi.org/10.1016/j.ins.2020.08.049

[19] Luo, C., Li, T., Huang, Y., Fujita, H., ”Updating three-way decisions in incomplete multi-scale information systems,” Information sciences, vol. 476, pp.274-289, 2019. doi:https://doi.org/10.1016/j.ins.2018.10.012

[20] Zhang, C., Li, J., Lin, Y., ”Knowledge reduction of pessimistic multigranulation rough sets in incomplete information systems,” Soft Computing, vol. 25, no. 20, pp.12825-12838, 2021. doi:https://doi.org/10.1007/s00500-021-06081-w

[21] Kryszkiewicz, M., ”Rough set approach to incomplete information systems,” Information sciences, vol. 112, no. 1-4, pp.39-49, 1998. doi:https://doi.org/10.1016/S0020-0255(98)10019-1

[22] Kryszkiewicz, M., ”Rules in incomplete information systems,” Information sciences, vol. 113, no. 3-4, pp.271-292, 1999. doi:https://doi.org/10.1016/S0020-0255(98)10065-8

[23] Qian, Y., Liang, J., Dang, C., ”Incomplete multigranulation rough set,” IEEE transactions on systems, man, and cybernetics-part a: systems and humans, vol. 40, no. 2, pp.420-431, 2009. doi: 10.1109/TSMCA.2009.2035436.