Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Optimal scale combination selection for inconsistent multi-scale decision tables

**Yingjie Zhu**
Northwest A&F University: Northwest Agriculture and Forestry University
**Bin Yang** ( binyang0906@nwsuaf.edu.cn )
Northwest A & F University https://orcid.org/0000-0002-1707-5119

**Research Article**

**Keywords:** Inconsistent multi-scale decision table, Optimal scale combination, Positive region consistent

**Posted Date:** October 5th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-635004/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

# Optimal scale combination selection for inconsistent multi-scale decision tables

Yingjie Zhu[a], Bin Yang[a,*]

[a]*College of Science, Northwest A & F University, Yangling, PR China*

**Abstract**

Hierarchical structured data are very common for data mining and other tasks in real-life world. How to select the optimal scale combination from a multi-scale decision table is critical for subsequent tasks. At present, the models for calculating the optimal scale combination mainly include lattice model, complement model and stepwise optimal scale selection model, which are mainly based on consistent multi-scale decision tables. The optimal scale selection model for inconsistent multi-scale decision tables has not been given. Based on this, firstly, this paper introduces the concept of complement and lattice model proposed by Li and Hu. Secondly, based on the concept of positive region consistency of inconsistent multi-scale decision tables, the paper proposes complement model and lattice model based on positive region consistent and gives the algorithm. Finally, some numerical experiments are employed to verify that the model has the same properties in processing inconsistent multi-scale decision tables as the complement model and lattice model in processing consistent multi-scale decision tables. And for the consistent multi-scale decision table, the same results can be obtained by using the model based on positive region consistent. However, the lattice model based on positive region consistent is more time-consuming and costly. The model proposed in this paper provides a new theoretical method for the optimal scale combination selection of the inconsistent multi-scale decision table.

*Keywords:* Inconsistent multi-scale decision table; Optimal scale combination; Positive region consistent

## 1. Introduction

Rough set theory was originally proposed by Professor Pawlak in 1982 [8]. Because of the mature mathematical foundation and unnecessity of prior knowledge, it is easy to use and become an effective tool for dealing with various incomplete information such as imprecise, inconsistent information. It is a powerful data analysis method. Rough set theory can, in the absence of prior knowledge, find out the classification of knowledge to determine the upper and lower approximation of the problem by describing the set of the given problem, and then analyze and process the uncertain data.

Based on the theory of Pawlak rough set, Wu and Leung introduced the concept of multi-scale decision table (MSDT) from the perspective of granular computing, and analyzed the knowledge acquisition in it [10]. The general method of processing multi-scale decision tables is to limited multi-scale attributes on a certain scale and then we can obtain a series of single-scale decision tables (SSDT) whose each attribute only has one scale. At last, we can do data mining on a single-scale decision table we choose. In a multi-scale information table, if all the attributes are on the finest scales, then the most information of objects is included, but this process is of high cost. However, if all the attribute are on the coarsest scales, then some useful information may be lost. Therefore, one or several optimal scale combinations which can reduce the cost without losing useful information are existed.

However, Wu and Leung pointed out that their research is based on two assumptions [11]. One of them is that the number of scales of each attribute must be the same. Another one is that only the corresponding single attributes are able to combine into a subsystem in the process of decomposition for subsystems. Under the same assumptions, Gu [2] and She [9] studied the knowledge acquisition and rule induction in multi-scale decision tables.

*Corresponding author.

*Email address:* binyang0906@nwsuaf.edu.cn (Bin Yang)

Later, Li and Hu extended their theory and broke these two assumptions [5, 6]. They proposed lattice model and complement model to calculate the optimal scale combination. Based on the concept of multi-scale attribute significance they introduced, they proposed stepwise optimal scale selection model. For the attribute with different significance, the scale selection should be carried out step by step, which can effectively reduce the time of calculating and get the best results based on the attribute significance.

In 2020, the relationship between rule extraction and feature matrix is further studied [3]. Finally, the matrix is used to describe the scale combination, and the matrix method for optimal scale combination selection and the optimal scale combination keeping the positive region unchanged in the consistent and inconsistent generalized multi-scale decision information system is give respectively.

Nevertheless, the above models are mainly aimed to calculate the optimal scale combination in consistent multi-scale decision tables. The model of selecting all the optimal scale combinations for inconsistent multi-scale decision tables has not been given, which is critical for subsequent tasks. Motivated by these, in this paper, complement model and lattice model based on positive region consistence are proposed. And the algorithms of them are given as well.

The remainder parts of the paper are organized as follows. In Section 2, several basic notions of Pawlak rough set, information tables and decision tables, scale combination and attribute significance are reviewed. In Section 4, the concept of positive region consistent is introduced. And the optimal scale combination selection models for inconsistent multi-scale decision tables and their algorithm are proposed. Some numerical experiments are employed in Section 5. Finally, we conclude the paper with a summary and outlook the further in Section 6.

## 2. Preliminaries

In this section, we review several basic concepts and results of Pawlak rough set, information tables and decision tables, scale combination and attribute significance.

### *2.1. Pawlak rough set*

Let $U$ be a finite and nonempty set called universe of discourse. If $R \subseteq U \times U$ is an equivalence relation on $U$, that is, $R$ is a reflexive, symmetric and transitive binary relation on $U$, then the pair $(U, R)$ is called a Pawlak approximation space [8]. The equivalence relation $R$ partitions the universe of discourse $U$ into disjoint subsets. Such partition is a quotient set of $U$ and denoted by $U/R = \{[x]_R | x \in U\}$, where $[x]_R = \{y \in U | (x, y) \in R\}$ is the $R$ equivalence class containing $x$. The elements in $U/R$ are called elementary sets. For any set $X \in \mathscr{P}(U)$, lower and upper approximations are defined as follows:

**Definition 2.1.** Let $U$ be a finite and nonempty set called universe of discourse. If $X \in \mathscr{P}(U)$, lower and upper approximations of $X$ are defined as:

$$\underline{R}(X) = \cup\{[x]_R | [x]_R \subseteq X\}, \overline{R}(X) = \cup\{[x]_R | [x]_R \cap X \neq \emptyset\}, \tag{1}$$

where $\mathscr{P}(U)$ is the power set of $U$. Obviously, they can be defined by:

$$\underline{R}(X) = \{x \in U | [x]_R \subseteq X\}, \overline{R}(X) = \{x \in U | [x]_R \cap X \neq \emptyset\}. \tag{2}$$

If and only if $\underline{R}(X) \neq \overline{R}(X)$, $X$ cannot be precisely defined by $R$. $(\underline{R}(X), \overline{R}(X))$ is called the Pawlak rough set of $X$ with respect to (w.r.t.) $(U, R)$. The sets $BN_R(X) = \overline{R}(X) - \underline{R}(X)$, $POS_R = \underline{R}(X)$, $NEG_R = U - \overline{R}(X)$ are respectively called the boundary, the positive region and the negation region of $X$ w.r.t. (U,R).

The accuracy of rough set can be defined as [8]:

$$\alpha_R = \frac{|\underline{R}(X)|}{|\overline{R}(X)|}, \tag{3}$$

where $|\cdot|$ is the cardinal number of set. For the empty set $\emptyset$, we define $\alpha_R(\emptyset) = 1$. Obviously, $0 \leqslant \alpha_R(X) \leqslant 1$.

**Definition 2.2.** [11] Let Let $U$ be a finite and nonempty universe of discourse. $\mathbf{P}_1$ and $\mathbf{P}_2$ are two partitions of $U$, For each $A \in \mathbf{P}_1$, if there exists $B \in \mathbf{P}_2$ such that $A \subseteq B$, we say that $\mathbf{P}_1$ is finer than $\mathbf{P}_2$ or $\mathbf{P}_2$ is coarser than $\mathbf{P}_1$, denoted as $\mathbf{P}_1 \sqsubseteq \mathbf{P}_2$. If $A \subset B$, we say $\mathbf{P}_1$ is strictly finer than $\mathbf{P}_2$, denoted as $\mathbf{P}_1 \sqsubset \mathbf{P}_2$.

### 2.2. Information table and decision table

**Definition 2.3.** [10] An information table is a 2-tuple $(U, A)$, where $U = \{x_1, x_2, \cdots, x_n\}$ is a finite and nonempty set called universe of discourse, $A = \{a_1, a_2, \cdots, a_m\}$ is a finite and nonempty set of attributes. For any $a \in A$, there is $a : U \to V_a$, that is, for any $x \in U$, there is $a(x) \in V_a$, where $V_a = (a(x)|x \in U)$ called the domain of $a$.

For each attribute $a \in A$, $a$ is a surjective function from $U$ to $V_a$, it determines an equivalence relation on $U$.

$$R_a = \{(x, y) \in U \times U | a(x) = a(y)\}. \tag{4}$$

**Definition 2.4.** [10] A decision table is a 2-tuple $(U, C \cup \{d\})$, where $(U, C)$ is an information table, $d \notin C$ is a special attribute called decision. In this case, $C$ is called conditional attribute set. $d$ is a map $d : U \to V_d$ from $U$ to $V_d$.

Similarly, we can define the equivalence relation as:

$$R_d = \{(x, y) \in U \times U | d(x) = d(y)\}, \tag{5}$$

Then we obtain a partition $U/R_d$ of $U$. For any $B \subseteq C$, define equivalence $R_B$ as:

$$R_B = \bigcap_{a \in B} R_a = \{(x, y) \in U \times U | a(x) = a(y) \forall a \in B\}, \tag{6}$$

If $R_C \subseteq R_d$, then the decision table $(U, C \cup \{d\})$ is consistent, otherwise it is inconsistent.

For the inconsistent decision table $(U, C \cup \{d\})$, the concept of generalized decision attribute is introduced in [4]. For any $B \subseteq C$, the generalized decision attribute of $x$ w.r.t. $B$, denoted as $\partial_B$, can be defined as:

$$\partial_B(x) = \{d(y)|y \in [x]_B\}, x \in U. \tag{7}$$

According to Eq.7, we know that for any decision table $S = (U, C \cup \{d\})$, although $S$ may be inconsistent, $S' = (U, C \cup \partial_C)$ must be consistent.

Based on the single-scale decision table, Wu and Leung proposed the concept of multi-scale decision table [10]:

**Definition 2.5.** [10] A multi-scale decision table can be denoted as $S = (U, C \cup \{d\})$, where $U$ is finite and nonempty object set called universe of discourse, $A$ is finite and nonenpty set of attribute, $d$ is decision. Each attribute $a_j \in C$ is a multi-scale attribute, that is, for the same object in $U$, attribute $a_j$ can take on different values at different scales.

For each attribute $a_j \in C$, we assume that the higher the level of scale is, the coarser the partition w.r.t. the scale becomes. If the attribute $a_j$ has three levels of scale, its first level of scale $a_j^1$ is finer than its second level of scale $a_j^2$, its second level of scale $a_j^1$ is finer than its third level of scale $a_j^2$.

### 2.3. Scales combinaiton

The general method of processing multi-scale information tables is to limited multi-scale attributes on a certain scale and then we can obtain a series of single-scale information tables whose each attribute only has one scale. At last, we can do data mining on a single-scale information table we choose.

The concept of scales combination and scales collection and some properties was introduced by Li and Hu in [5].

**Definition 2.6.** [5] Let $S = (U, A)$ be a multi-scale information table, where attribute $a_i$ has $I_i$ levels of scale, $i = 1, 2, \cdots, m$. If we restrict attribute $a_1, a_2, \cdots, a_m$ on their $I_i th$ scale respectively, we can obtain a single-scale information table $S^K$, where $K = (l_1; l_2; \cdots; l_m)$. The combination $(l_1; l_2; \cdots; l_m)$ is called the scales combination of $S$ in $S^K$. All the scales combination of $S$ is called scales collection, denoted as $\mathcal{L} = \{(l_1; l_2; \cdots; l_m)|1 \leqslant l_i \leqslant I_i, i = 1, 2, \cdots, m\}$.

**Definition 2.7.** [5] Let $S = (U, A)$ be a multi-scale information table, and $\mathcal{L}$ is the scales collection of $S$. For $K_1, K_2 \in \mathcal{L}$, if the elements of $K_2$ are not less than the corresponding elements of $K_1$, then we say that $K_1$ is weaker than $K_2$ or $K_2$ is stronger than $K_1$, denoted as $K_1 \preceq K_2$.

According to Definition 2.7, we know that $\mathcal{L}$ is an partial order relation. Thus $(\mathcal{L}, \preceq)$ is a partial order set, which is reflexive, antisymmetric and transitive. Furthermore, $(\mathcal{L}, \preceq)$ is a lattice in which every two elements have a unique supremum and a unique infimum. Obviously, we can get the proposition as follow.

**Proposition 2.1.** [5] Let $K_1, K_2 \in \mathcal{L}$ and $K_1 \preceq K_2$, if $S^{K_2} = (U, C^{K_2} \cup \{d\})$ is consistent, then $S^{K_1}$ is also consistent.

According to the Proposition 2.1, we can define the concept of optimal scale combination as follow[5].

**Definition 2.8.** [5] Let $\mathcal{L}$ be a scales collection of a consistent multi-scale decision table $S$, for $K \in \mathcal{L}$, all the $K$ meet the condition that if for all the $K' \in \mathcal{L}$ and $K \preceq K'$, $S^K$ is consistent but $S^{K'}$(if there exists $K'$) is inconsistent are the optimal scale combination of $S$.

Therefore, the consistency of multi-scale decision table can be defined by:

**Definition 2.9.** [10] Let $S = (U, C \cup \{d\})$ be a multi-scale decision table, and $\mathbf{1}_m = (1; 1; \cdots; 1)$. If $S^{\mathbf{1}_m} = (U, \{a_j^1 | j = 1, 2, \cdots, m\} \cup \{d\})$ whose all the attributes are on their finest level of scale is consistent, then the multi-scale decision table $S$ is consistent.

Let $\mathcal{L}$ be the scales collection of $(U, C \cup \{d\})$. For an arbitrary $K \in \mathcal{L}$, the corresponding equivalence relation $R_{A^K}$ can be defined as

$$R_{A^K} = (x, y) \in U \times U | a^k(x) = a^k(y), \forall a \in A, \forall k \in K. \tag{8}$$

U can be partitioned by $R_{A^K}$ into a family of equivalence classes as follow

$$U/R_{A^K} = [x]_{A^K} | x \in U, \tag{9}$$

where $[x]_{A^K} = y \in U | (x, y) \in R_{A^K}$.

According to Eq.8 and Eq.9, we can know the relation between equivalence relation and subsets of attributes.

**Proposition 2.2.** [5] Let $S = (U, C \cup \{d\}) = (U, \{a_j^k | j = 1, 2, \cdots, m, k = 1, 2, \cdots, I_j\} \cup \{d\})$ be a multi-scale decision table. $\mathcal{L}$ is the scales collection of $S$. For $K_0 = (k_1; k_2; \cdots; k_m) \in \mathcal{L}$ and an arbitrary subset $C_1 \subseteq C$, there exists $K_1 \subseteq K_0$ such that the indexes of $K_1$ in $K_0$ are the same as those of $C_1$ in $C$. Similarly, there exist a sequence $C_m \subseteq \cdots \subseteq C_2 \subseteq C_1 \subseteq C_0$ and the corresponding indexes sets $K_m \subseteq \cdots \subseteq K_2 \subseteq K_1 \subseteq K_0$. The following equations hold

$$R_{C_1^{K_1}} \subseteq R_{C_2^{K_2}} \subseteq \cdots \subseteq R_{C_m^{K_m}}, \tag{10}$$

$$[x]_{C_1^{K_1}} \subseteq [x]_{C_2^{K_2}} \subseteq \cdots \subseteq [x]_{C_m^{K_m}}, \quad \forall x \in U, \tag{11}$$

$$U/R_{C_1^{K_1}} \sqsubseteq U/R_{C_2^{K_2}} \sqsubseteq \cdots \sqsubseteq U/R_{C_m^{K_m}}. \tag{12}$$

## 3. Complement model and lattice model

In order to extend the application of multi-scale decision table, Li and Hu proposed complement model and lattice model in [5].

### *3.1. Complement model*

Let $S = (U, C \cup \{d\})$ be a multi-scale decision table, and $S^+$ be its complement system. Let $I_i, i = 1, 2, \cdots, m$ be the number of levels of scales of attribute $a_i$ respectively and they are not necessary to be the same. For some attributes with less number $I_i$, we complement them with several known levels of scales to obtain a new multi-scale decision table, whose attributes have the same number of levels of scales.

Let $I = \max\{I_1, I_2, \cdots, I_m\}$, that is, the maximum of $I_i$ and $p$ be the index of attribute with the largest number of levels of scales. In case of multiple occurrences of the maximum values, the index corresponding to the first occurrence is returned. Firstly, the concept of scale vector is introduced.

**Definition 3.1.** [5] Let $S = (U, C \cup \{d\})$ be a multi-scale decision table. Attribute $a_i \in C$ has $I_i$ levels of scales. $C_i = (1, 2, \cdots, I_i)$ is called the original scale vector of $a_i$ and $C_i^+$ is the corresponding complement scale vector.

In order to ensure that the number of levels of scales of all the attribute are all the same in $S^+$, other complement scale vector $C_i^+(i \neq p)$ should be formed as $(l_{i1}, l_{i2}, \cdots, I_{ij}, l_{iI})$, where $1 \leqslant l_{ij} \leqslant I_i$ and $l_{ij} \leqslant l_{ik}$ when $j \leqslant k$. Moreover, to include more information about $a_i$, the original scale vector $C_i$ should be covered by $C_i^+$. Thus, $C_i^+$ should satisfy the following conditions:

(C1) $l_{i1} = 1, l_{iI} = I_i$

(C2) $dim(C_i^+) = I$

(C3) $0 \leqslant l_{i,j+1} - l_{ij} \leqslant 1, \forall 1 \leqslant j \leqslant I-1$

Therefore, the number of possible choice for scale vector $C_i^+$ of $a_i$ is equivalent to "choose $I - I_i$ from $I_i$ with replacement"[5]. According to [1], there are $\binom{I-1}{I_i-1}$ choices of $C_i^+$. Hence, we can get $\prod_{i=1}^{m} \binom{I-1}{I_i-1}$ different new multi-scale decision tables.

These new multi-scale decision tables can be decomposed into $I$ decision tables with the same decision attribute where $I$ is the number of levels of scales. We can choose the table with optimal scales combination [11].

### 3.2. Lattice model

Let $S = (U, C \cup \{d\})$ be a multi-scale decision table, and $I_i (i = 1, 2, \cdots, m)$ be the number of levels of scales of $a_i$ which are not necessary to be the same. $\mathcal{L}$ is scales collection of $S$, and $|\mathcal{L}| = \prod_{i=1}^{m} I_i$. If $S$ is consistent, then there exists the scales combination $K$ which is the optimal scales combination of $S$. Lattice model is aimed to select all the optimal scales combination.

For a given multi-scale decision table, lattice model can be described via the following procedure:

1. According to Definition 2.6, scales collection $\mathcal{L}$ of $S$ can be calculated.
2. Based on the consistence and the partial order relation between elements in $\mathcal{L}$ of $S$, the set of optimal scales combinations OSC can be obtained according to Definition 2.8.
3. In the subsystem $S^K$ confirmed by the optimal scale combination $K \in OSC$, knowledge acquisition and other tasks can be done.

## 4. Optimal scale combination selection models for inconsistent multi-scale decision tables

### 4.1. Consistence of positive region

Let $S = (U, C \cup \{d\})$ be a multi-scale decision table, where $U = \{x_1, x_2, \cdots, x_n\}$, $C = \{a_1, a_2, \cdots, a_m\}$ and $I_i$ is the number of levels of scales of $a_i, (i = 1, 2, \cdots, m)$. If $\mathbf{1}_m \preceq K_1 \preceq K_2 \preceq (I_1; I_2; \cdots; I_m)$ and $S^{K_1}$ is an inconsistent decision table, according to Proposition 2.1, $S^{K_2}$ is also inconsistent. Hence, if $S$ is inconsistent, that is $S^{\mathbf{1}_m}$ is inconsistent, then $S^K$ is also inconsistent for any $K \in \mathcal{L}$.

For $K \in \mathcal{L}$, there is an equivalence relation as follow:

$$R_{C^K} = \{(x, y) \in U \times U | a^k(x) = a^k(y), \forall a \in C, \forall k \in K\}. \tag{13}$$

For $X \in U$, the upper and lower approximations of $X$ are shown as:

$$\underline{R_{C^K}}(X) = \{x \in U | [x]_{C^K} \subseteq X\}, \tag{14}$$

$$\overline{R_{C^K}}(X) = \{x \in U | [x]_{C^K} \cap X \neq \emptyset\}, \tag{15}$$

where $[x]_{C^K} = \{y \in U | (x, y) \in R_{C^K}\}$.

The positive region under scale combination $K$ in $S$ is defined as[6, 11]:

$$POS_{C^K}(d) = \bigcup_{X \subseteq U/d} \underline{R_{C^K}}(X), \tag{16}$$

where $U/d = \{D_1, D_2, ..., D_r\}$.

**Definition 4.1.** [6] Let $S = (U, C \cup \{d\})$ be a multi-scale decision table, where $U = \{x_1, x_2, \cdots, x_n\}$, $C = \{a_1, a_2, \cdots, a_m\}$ and $I_i$ is the number of levels of scales of $a_i, (i = 1, 2, \cdots, m)$. For $K \in \mathcal{L}$, if $POS_{C^K}(d) = POS_{C^{\mathbf{1}_m}}(d)$, $S$ is said to be positive region consistent. If $S^K$ is positive region consistent and $S^{K'}$(if there exists $K'$,$K \preceq K'$ and $K' \in \mathcal{L}$) is not positive region consistent, then $K$ is the positive region scale combination of $S$.

The algorithm of judging whether or not a given subsystem of a multi-scale decision table is positive region consistent has been given in [6].

### 4.2. Complement model and lattice model based on positive region consistence

We extend the application of complement model and lattice model in this subsection.

The complement model and lattice model proposed by Li and Hu is aimed to process the consistent multi-scale decision table. Therefore, we can combine positive region consistence with these two models to obtain the complement model and lattice model based on positive region consistence which can deal with inconsistent multi-scale decision tables.

In order to deal with multi-scale decision table by using complement model and lattice model based on positive region consistence, we only need to replace the judgement of consistence of subsystem with the judgement of positive region consistence in complement model and lattice model.

And for $K \in \mathcal{L}$, "$S^K$ is consistent" is included by "$S^K$ is positive region consistent"[6]. Hence, complement model and lattice model are included by the complement model and lattice model based on positive region consistence. The new models can deal with all kinds of multi-scale decision tables.

Then we propose the algorithm of complement model and lattice model based on positive region consistence.

**Algorithm 1:** Complement model based on positive region consistence

```
Input: any multi-scale decision table S = (U, C ∪ {d}) and the numbers of levels of scales(I_1, I_2, ..., I_m)
Output: the set of the positive region optimal scales combinations
I = max(I_1, I_2, ..., I_m);
m = len(I_1, I_2, ..., I_m);
for i in range(1, m + 1) do
    obtain a set of complemented scale vectors A_i of a_i;
end
SC = set(); // SC is the set of all the scales combination;
for (e_1, e_2, ..., e_m) in A_1 × A_2 × ... × A_m do
    //A_1 × A_2 × ... × A_m is the Cartesian product;
    for i in range(I) do
        sc = (e_1[i], e_2[i], ..., e_m[i]);
        //(e_1[i], e_2[i], ..., e_m[i]) is a scales combination;
        SC = SC ∪ {sc};
    end
end
CSC = set(); // CSC is the set of all the positive region scales combination ;
for sc in SC do
    if S^sc is Positive Region Consistent then
        CSC = CSC ∪ {sc};
    end
end
OSC = CSC; // OSC is the set of all the positive region optimal scales combination;
for osc in OSC do
    for csc in CSC do
        if osc ≤ csc then
            OSC = OSC − {osc};
            break;
        end
    end
end
return OSC;
```

**Algorithm 2:** Lattice model based on positive region consistence

**Input**: any multi-scale decision table $S = (U, C \cup \{d\})$ and the numbers of levels of scales$(I_1, I_2, ..., I_m)$
**Output**: the set of the positive region optimal scales combinations
$m$ =len$(I_1, I_2, ..., I_m)$;
**for** $i$ *in range*$(1, m + 1)$ **do**
    $C_i = (1, 2, ..., I_i)$;
**end**
$SC = set()$; // $SC$ is the set of all the scales combination;
**for** $(e_1, e_2, ..., e_m)$ *in* $C_1 \times C_2 \times ... \times C_m$ **do**
    //$C_1 \times C_2 \times ... \times C_m$ is the Cartesian product;
    $SC = SC \cup \{(e_1, e_2, ..., e_m)\}$;
    //$(e_1, e_2, ..., e_m)$ is a scales combination;
**end**
$CSC = set()$; // $CSC$ is the set of all the positive region scales combination ;
**for** $sc$ *in* $SC$ **do**
    **if** $S^{sc}$***Positive Region Consistent*** **then**
        $CSC = CSC \cup \{sc\}$;
    **end**
**end**
$OSC = CSC$; // $OSC$ is the set of all the positive region optimal scales;
**for** $osc$ *in* $OSC$ **do**
    **for** $csc$ *in* $CSC$ **do**
        **if** $osc \preceq csc$ **then**
            $OSC = OSC - \{osc\}$;
            break;
        **end**
    **end**
**end**
**return** $OSC$;

## 5. Numerical experiments

In order to verify the feasibility of complement model(CM-PR) and lattice model(LM-PR) based on positive region consistence, some numerial experiments are employed in this section. And we compare the results of them with the result of stepwise optimal scale selection based on positive region consistence(SOSS-PR) proposed in [6].

**Example 5.1.** Table 1 is an inconsistent multi-decision table $S = (U, C \cup \{d\})$, where $U = \{x_1, x_2, \cdots, x_{20}\}$, $C = \{a_1, a_2, a_3, a_4\}$. We can notice that $x_4$ and $x_6$ are indistinguishable w.r.t. $R_C$, but $d(x_4) \neq d(x_6)$. The results obtained by using CM-PR, LM-PR and SOSS-PR respectively are shown in Table 2.

In order to evaluate the above algorithms more objectively, two data sets are collected from the University of California, Irvine (UCI) Machine Learning Repository [7]. These two decision tables are single-scale decision tables. Thus we use the method in [5] to obtain their corresponding multi-scale decision tables. There are four steps in that method, but we only do the first three steps. The decision value of object $x$ in multi-scale decision table is not change to $\partial_{C^{(1;1;...;1)}}(x)$, that is, it keep the original decision value. Then the multi-scale decision tables we obtain are not inconsistent. And the details of them and the results are shown in Table 3 and Table 4 respectively.

Through the numerical experiments, it can be found that the optimal scale combination of CM-PR are weaker than that of LM-PR and the result of SOSS-PR is one of the results of LM-PR. Moreover, the running time of SOSS-PR is shorter than that of LM-PR. These conclusions are similar to the conclusions of the model deal with consistent multi-scale decision tables summerized by Li and Hu in [5, 6].

Table 1: An inconsistent multi-scale decision table

| $U$ | $a_1^1$ | $a_1^2$ | $a_2^1$ | $a_2^2$ | $a_2^3$ | $a_3^1$ | $a_3^2$ | $a_3^3$ | $a_4^1$ | $a_4^2$ | $d$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0 | 0 | 2 | 2 | 2+ | 1 | 1 | 1 | 3 | 2+ | 1 |
| $x_2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| $x_3$ | 1 | 1 | 3 | 3+ | 2+ | 2 | 2 | 2+ | 1 | 1 | 1 |
| $x_4$ | 0 | 0 | 2 | 2 | 2+ | 1 | 1 | 1 | 2 | 2+ | 1 |
| $x_5$ | 1 | 1 | 4 | 3+ | 2+ | 2 | 2 | 2+ | 2 | 2+ | 1 |
| $x_6$ | 0 | 0 | 2 | 2 | 2+ | 1 | 1 | 1 | 2 | 2+ | 2 |
| $x_7$ | 0 | 0 | 2 | 2 | 2+ | 1 | 1 | 1 | 0 | 0 | 1 |
| $x_8$ | 0 | 0 | 3 | 3+ | 2+ | 1 | 1 | 1 | 3 | 2+ | 1 |
| $x_9$ | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2+ | 0 | 0 | 2 |
| $x_{10}$ | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2+ | 0 | 0 | 2 |
| $x_{11}$ | 0 | 0 | 2 | 2 | 2+ | 2 | 2 | 2+ | 1 | 1 | 2 |
| $x_{12}$ | 0 | 0 | 2 | 2 | 2+ | 2 | 2 | 2+ | 1 | 1 | 2 |
| $x_{13}$ | 1 | 1 | 2 | 2 | 2+ | 3 | 3+ | 2+ | 1 | 1 | 3 |
| $x_{14}$ | 2 | 2+ | 1 | 1 | 1 | 3 | 3+ | 2+ | 1 | 1 | 3 |
| $x_{15}$ | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 3 |
| $x_{16}$ | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2+ | 0 | 0 | 3 |
| $x_{17}$ | 3 | 2+ | 2 | 2 | 2+ | 2 | 2 | 2+ | 0 | 0 | 3 |
| $x_{18}$ | 3 | 2+ | 3 | 3+ | 2+ | 4 | 3+ | 2+ | 0 | 0 | 3 |
| $x_{19}$ | 2 | 2+ | 1 | 1 | 1 | 2 | 2 | 2+ | 0 | 0 | 3 |
| $x_{20}$ | 1 | 1 | 3 | 3+ | 2+ | 3 | 3+ | 2+ | 1 | 1 | 3 |

Table 2: The results of Table 1(The models in table are based on positive region consistence)

| Table | CM-PR | | LM-PR | | SOSS-PR | |
|---|---|---|---|---|---|---|
| | $OSC_1$ | running time(s) | $OSC_2$ | running time(s) | $OSC_3$ | running time(s) |
| 4-1 | (2;2;2;1) | 0.0156 | (2;3;2;1) | 0.1249 | (2;3;2;1) | 0.0625 |

In order to test the performance of CM-PR and LM-PR in consistent multi-scale decision tables, we use complement model, lattice model, stepwise optimal scale selection, CM-PR, LM-PR, SOSS-PR to deal with some consistent multi-scale decision tables respectively. Table 5, Table 6 and Table 7 are three consistent multi-scale decision tables collected from [5]. The results are shown in Table 8 and Table 9

For Table 5, the set of optimal scales combination via complement model and CM-PR is {(3;2;3;2),(3;3;3;1)}, the set of optimal scales combination via lattice model and LM-PR is {(4;3;1;2),(4;2;2;2),(1;3;3;2),(4;3;2;1),(3;2;3;2),(3;3;3;1),(4;1;3;2)}, and the optimal scales combination via stepwise optimal scale selection and SOSS-PR is (4;3;1;2).

For Table 6, the set of optimal scales combination via complement model and CM-PR is {(2;1;3;3;3),(2;2;2;2;2),(1;2;3;3;3)}, the set of optimal scales combination via lattice model and LM-PR is {(2;2;4;1;3),(1;2;4;3;3),(2;1;4;3;3),(1;2;3;3;4),(2;2;4;3;2),(2;2;3;1;4),(2;2;2;3;3),(2;1;3;3;4)}, and the optimal scales combination via stepwise optimal scale selection and SOSS-PR is (2;2;4;3;2).

For Table 7, the set of optimal scales combination via complement model and CM-PR is {(2;2;2;2;4)}, the set of optimal scales combination via lattice model and LM-PR is {(2;2;2;2;4)}, and the optimal scales combination via

Table 3: The details of inconsistent multi-scale decision tables

| Data sets | Instances | Features | $I_1 \times I_2 \times ... \times I_m$ | Classes |
|---|---|---|---|---|
| Auto-MPG | 392 | 7 | $2 \times 1 \times 2 \times 3 \times 3 \times 3 \times 2$ | 3 |
| Seeds | 210 | 7 | $2 \times 2 \times 3 \times 3 \times 2 \times 3 \times 3$ | 3 |

Table 4: The results of data sets(inconsistent)(The models in table are based on positive region consistence)

| Data sets | CM-PR | | LM-PR | | SOSS-PR | |
|---|---|---|---|---|---|---|
| | $OSC_1$ | running time(s) | $OSC_2$ | running time(s) | $OSC_3$ | running time(s) |
| Auto-MPG | (1;1;2;2;2;2;1) | 0.8904 | (1;1;2;3;3;2;1) | 17.0384 | (1;1;2;3;3;2;1) | 1.7604 |
| Seeds | (1;1;1;1;1;1;1) | 0.5001 | (2;1;1;3;1;3;3)<br>(2;2;1;3;1;3;2) | 27.3886 | (2;2;1;3;1;3;2) | 1.0316 |

Table 5: An multi-scale decision table based on a general information system.

| $U$ | $a_1^1$ | $a_1^2$ | $a_1^3$ | $a_1^4$ | $a_2^1$ | $a_2^2$ | $a_2^3$ | $a_3^1$ | $a_3^2$ | $a_3^3$ | $a_4^1$ | $a_4^2$ | $d$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 1 | E | S | Y | 1 | E | Y | 1 | S | Y | 1 | S | + |
| $x_2$ | 2 | G | S | Y | 2 | E | Y | 1 | S | Y | 1 | S | + |
| $x_3$ | 3 | G | S | Y | 3 | G | Y | 2 | S | Y | 2 | S | + |
| $x_4$ | 4 | F | M | N | 4 | F | N | 3 | M | N | 3 | M | - |
| $x_5$ | 5 | B | L | N | 5 | F | N | 4 | L | N | 4 | L | + |
| $x_6$ | 6 | B | L | N | 6 | B | N | 5 | L | N | 4 | L | + |
| $x_7$ | 4 | F | M | N | 4 | F | N | 1 | S | Y | 1 | S | - |
| $x_8$ | 5 | B | L | N | 5 | F | N | 1 | S | Y | 1 | S | - |
| $x_9$ | 6 | B | L | N | 6 | B | N | 2 | S | Y | 2 | S | + |
| $x_{10}$ | 4 | F | M | N | 4 | F | N | 3 | M | N | 1 | S | - |
| $x_{11}$ | 5 | B | L | N | 5 | F | N | 4 | L | N | 1 | S | + |
| $x_{12}$ | 6 | B | L | N | 6 | B | N | 5 | L | N | 2 | S | + |

stepwise optimal scale selection and SOSS-PR is (2;2;2;2;4).

The results are shown in Table 8 and Table 9

Moreover, for the data sets described in Table 3, we use the method in [5] to obtain their corresponding consistent multi-scale decision tables. And optimal scales combination on these two consistent multi-scale decision tables using three models based on consistence and three models based on positive region consistence are shown in Table 10 and Table 11 respectively.

Compared Table 8 with Table 9 and compared Table 10 with Table 11, some facts are verified. The running times of complement model and CM-PR have no static relationship and the running time of LM-PR is about two times longer than that of lattice model. The running time of SOSS-PR is slightly slower than that of stepwise optimal scale selection. When dealing with the consistent multi-scale decision table, the model based on the positive region

Table 6: An multi-scale decision table based on an interval information system.

| $U$ | $a_1^1$ | $a_1^2$ | $a_2^1$ | $a_2^2$ | $a_3^1$ | $a_3^2$ | $a_3^3$ | $a_3^4$ | $a_4^1$ | $a_4^2$ | $a_4^3$ | $a_5^1$ | $a_5^2$ | $a_5^3$ | $a_5^4$ | $d$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | [3.0,5.0] | B | [1.9,2.0] | A | [0.9,4.5] | C | M | Y | [0.9,4.5] | C | M | [1.8,2.3] | A | S | Y | - |
| $x_2$ | [1.4,2.5] | A | [3.8,4.4] | B | [3.8,4.4] | B | S | Y | [1.7,2.5] | A | S | [4.1,4.4] | B | S | Y | + |
| $x_3$1 | [1.9,2.2] | A | [3.0,5.0] | B | [3.8,4.4] | B | S | Y | [1.6,2.0] | A | S | [1.2,4.2] | C | M | Y | - |
| $x_4$ | 4.0 | B | [1.3,3.7] | C | [0.9,4.5] | C | M | Y | 4.0 | B | M | [1.4,2.3] | A | S | Y | + |
| $x_5$ | 4.0 | B | [1.3,3.7] | C | [1.3,2.5] | A | S | Y | 4.0 | B | M | [7.4,7.8] | E | M | Y | - |
| $x_6$ | [1.1,2.3] | A | [3.2,4.5] | B | [1.6,2.2] | A | S | Y | [1.3,2.2] | A | S | [1.8,2.3] | A | S | Y | - |
| $x_7$ | [3.2,4.6] | B | [1.8,2.3] | A | [1.3,2.5] | A | S | Y | [1.2,4.2] | C | M | [4.2,6.7] | D | M | Y | + |
| $x_8$ | [3.8,4.3] | B | [1.4,2.5] | A | [5.2,6.3] | D | M | Y | [1.3,3.9] | C | M | [4.6,4.9] | B | S | Y | - |
| $x_9$ | [1.3,3.9] | C | [1.4,2.5] | A | [3.4,7.3] | G | L | N | [1.6,2.6] | A | S | [8.2,9.1] | F | L | N | + |
| $x_{10}$ | [1.3,2.1] | A | [3.0,4.5] | B | [7.4,7.8] | E | L | N | [3.4,4.5] | B | M | [5.2,6.3] | S | L | N | - |
| $x_{11}$ | [3.2,4.6] | B | [0.9,4.2] | C | [3.8,4.4] | B | S | Y | [4.6,4.9] | B | M | [3.4,4.1] | B | S | Y | - |
| $x_{12}$ | [1.8,2.0] | A | [3.0,5.0] | B | [7.4,7.8] | E | L | N | [3.0,5.0] | B | M | [3.4,7.3] | G | L | N | - |
| $x_{13}$ | [1.2,4.2] | C | [1.8,2.3] | A | [8.2,9.1] | F | L | N | [1.7,2.3] | A | S | [7.1,7.3] | E | L | N | + |
| $x_{14}$ | [1.3,2.1] | A | [3.8,4.4] | B | [8.2,9.1] | F | L | N | [3.8,4.4] | B | M | [4.2,6.7] | E | L | N | - |
| $x_{15}$ | [0.9,4.5] | C | [1.9,2.0] | A | [7.4,7.8] | E | L | N | [1.2,2.8] | A | S | [4.6,8.9] | G | L | N | + |

Table 7: An multi-scale decision table based on an intuitionistic information system

| $U$ | $a_1^1$ | $a_1^2$ | $a_2^1$ | $a_2^2$ | $a_3^1$ | $a_3^2$ | $a_4^1$ | $a_4^2$ | $a_5^1$ | $a_5^2$ | $a_5^3$ | $a_5^4$ | $d$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | (0.4,0.2) | A | (0.5,0.2) | B | (0.4,0.3) | A | (0.4,0.6) | A | (0.7,0.2) | C | M | Y | - |
| $x_2$ | (0.5,0.2) | B | (0.7,0.2) | C | (0.7,0.2) | C | (0.5,0.4) | B | (0.3,0.4) | A | S | Y | + |
| $x_3$ | (0.5,0.2) | B | (0.7,0.3) | C | (0.8,0.2) | C | (0.5,0.4) | B | (0.1,0.8) | E | M | Y | + |
| $x_4$ | (0.4,0.6) | A | (0.5,0.2) | B | (0.4,0.5) | A | (0.3,0.4) | A | (0.6,0.4) | B | S | Y | - |
| $x_5$ | (0.6,0.1) | B | (0.8,0.2) | C | (0.8,0.2) | C | (0.6,0.2) | B | (0.5,0.4) | B | S | Y | + |
| $x_6$ | (0.6,0.2) | B | (0.4,0.3) | A | (0.8,0.2) | C | (0.6,0.1) | B | (0.4,0.3) | A | S | Y | - |
| $x_7$ | (0.3,0.5) | A | (0.6,0.4) | B | (0.3,0.3) | A | (0.5,0.2) | B | (0.1,0.7) | D | L | N | - |
| $x_8$ | (0.3,0.6) | A | (0.5,0.4) | B | (0.3,0.5) | A | (0.6,0.1) | B | (0.2,0.8) | E | L | N | - |
| $x_9$ | (0.7,0.2) | C | (0.3,0.3) | A | (0.4,0.5) | B | (0.4,0.5) | A | (0.1,0.8) | E | L | N | + |
| $x_{10}$ | (0.7,0.1) | C | (0.4,0.5) | A | (0.4,0.1) | B | (0.3,0.5) | A | (0.1,0.9) | F | L | N | + |
| $x_{11}$ | (0.5,0.1) | B | (0.3,0.3) | A | (0.7,0.1) | C | (0.6,0.2) | B | (0.2,0.7) | D | M | Y | - |
| $x_{12}$ | (0.6,0.2) | B | (0.4,0.5) | A | (0.8,0.1) | C | (0.5,0.1) | B | (0.6,0.3) | B | S | Y | - |

Table 8: The results of models based on consistence

| Table | $I_1 \times I_2 \times ... \times I_m$ | complement model | | lattice model | | stepwise optimal scale | |
|---|---|---|---|---|---|---|---|
| | | $\|OSC_1\|$ | running time | $\|OSC_2\|$ | running time | $\|OSC_3\|$ | running time |
| 3-6 | $4 \times 3 \times 3 \times 2$ | 2 | 0.0343 | 7 | 0.1062 | 1 | 0.0343 |
| 3-7 | $2 \times 2 \times 4 \times 3 \times 4$ | 3 | 0.0375 | 8 | 0.2649 | 1 | 0.0421 |
| 3-8 | $2 \times 2 \times 2 \times 2 \times 4$ | 1 | 12.3857 | 1 | 0.0891 | 1 | 0.0328 |

Table 9: The results of models based on positive region consistence

| Table | $I_1 \times I_2 \times ... \times I_m$ | CM-PR | | LM-PR | | SOSS-PR | |
|---|---|---|---|---|---|---|---|
| | | $\|OSC_1\|$ | running time(s) | $\|OSC_2\|$ | running time(s) | $\|OSC_3\|$ | running time(s) |
| 3-6 | $4 \times 3 \times 3 \times 2$ | 2 | 0.0562 | 7 | 0.2031 | 1 | 0.0484 |
| 3-7 | $2 \times 2 \times 4 \times 3 \times 4$ | 3 | 0.0609 | 8 | 0.5089 | 1 | 0.0515 |
| 3-8 | $2 \times 2 \times 2 \times 2 \times 4$ | 1 | 12.2326 | 1 | 0.1594 | 1 | 0.0375 |

Table 10: The results of data sets(consistent)(The models in table are based on consistence)

| Data sets | complement model | | lattice model | | stepwise optimal scale | |
|---|---|---|---|---|---|---|
| | $OSC_1$ | running time(s) | $OSC_2$ | running time(s) | $OSC_3$ | running time(s) |
| Auto-MPG | (1;1;2;2;2;2;1) | 0.4514 | (1;1;2;3;3;2;1) | 8.2009 | (1;1;2;3;3;2;1) | 1.3350 |
| Seeds | (1;1;1;1;1;1;1) | 0.2168 | (2;1;1;3;1;3;3)<br>(2;2;1;3;1;3;2) | 10.5442 | (2;2;1;3;1;3;2) | 0.6372 |

Table 11: The results of data sets(consistent)(The models in table are based on poitive region consistence)

| Data sets | CM-PR | | LM-PR | | SOSS-PR | |
|---|---|---|---|---|---|---|
| | $OSC_1$ | running time(s) | $OSC_2$ | running time(s) | $OSC_3$ | running time(s) |
| Auto-MPG | (1;1;2;2;2;2;1) | 0.8814 | (1;1;2;3;3;2;1) | 17.4130 | (1;1;2;3;3;2;1) | 1.7846 |
| Seeds | (1;1;1;1;1;1;1) | 0.4081 | (2;1;1;3;1;3;3)<br>(2;2;1;3;1;3;2) | 22.2005 | (2;2;1;3;1;3;2) | 0.8357 |

consistence has the same results as the model based on consistence. Thus CM-PR, LM-PR and SOSS-PR are also able to deal with consistent multi-scale decision tables efficiently.

In a word, for the general multi-scale decision tables, we can directly use the model based on positive region consistence to deal with them. For single-scale decision tables, we can only do the first three steps in the method in [5] to obtain their corresponding multi-scale decision tables. Generalized decision values are not need to calculate. Finally, the same results can be obtained by using the models based on positive region consistent. Moreover, the optimal scales combination obtained after converting single scale decision table to multi-scale decision table often has excellent performance in classification experiments[6].

## 6. Conclusions

At first, Wu and Leung introduced the concept of multi-scale information table from the perspective of granular computing and analyzed the knowledge acquisition in multi-scale decision tables under different scales. However, their research is based on two assumptions which may limited the theory and application of multi-scale information systems. Later, Li and Hu extended their theories and proposed complement and lattice model and stepwise optimal scale selection model to compute the optimal scale combination.

However, the optimal scale selection model for inconsistent multi-scale decision tables has not been given. Motivated by these, in this paper, complement model and lattice model based on positive region consistence are proposed. And the algorithms of them are given as well.

Finally, some numerical experiments are employed to verify that the new models have the same properties as the complement model and lattice model in dealing with the inconsistent multi-scale decision table. And for consistent multi-scale decision tables, the model based on positive region consistence can also get the same results, but the lattice model based on positive region consistence is more time-consuming and costly.

In fact, "the consistence of the single-scale subsystem" is included by the "positive region consistence of it". Therefore the model based on positive region consistent can efficiently solve the problem of optimal scale selection in the consistent and inconsistent multi-scale decision tables.

## Acknowledgements

This research was supported by Chinese Universities Scientific Fund (Grant no. 2452018054) and the College Students' Innovation and Entrepreneurship Training Program (Grant no. X202010712184).

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## Author contributions

Bin Yang contributed significantly to analysis and manuscript preparation; Yingjie Zhu performed the data analyses and wrote the manuscript.

## References

[1] R. A. Brualdi. *Introductory Combinatorics, 5th Edition*. Pearson Education, 2010.

[2] S.-M. Gu and W.-Z. Wu. On knowledge acquisition in multi-scale decision systems. *International Journal of Machine Learning and Cybernetics*, 4:477–486, 2013.

[3] J.-X. Huang, W.-K. Li, X.-P. Zhang, and J.-J. Li. Knowledege acquisition and matrix method of generalized multi-scale information system. *Journal of Shanxi University(Nat. Sci. Ed.)*, 43(4):878–887, 2020.

[4] J. Komorowski, Z. Pawlak, L. Polkowski, and A. Skowron. *Rough sets: tutorial*, pages 3–98. Springer-Verlag, Berlin, 1999.

[5] F. Li and B. Q. Hu. A new approach of optimal scale selection to multi-scale decision tables. *Information Sciences*, 381:193–208, 2017.

[6] F. Li, B. Q. Hu, and J. Wang. Stepwise optimal scale selection for multi-scale decision tables via attribute significance. *Knowledge-Based Systems*, 129:4–16, 2017.

[7] M. Lichman. Uci machine learning repository. http://archive.ics.uci.edu/ml.

[8] Z. Pawlak. *Rough Sets:Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publisher, 1992.

[9] Y. She, J. Li, and H. Yang. Yang, a local approach to rule induction in multi-scale decision tables. *Knowledge-Based Systems*, 89:398–410, 2015.

[10] W.-Z. Wu and Y. Leung. Theory and applications of granular labelled partitions in multi-scale decision tables. *Information Sciences*, 181:3878–3897, 2011.

[11] W.-Z. Wu and Y. Leung. Optimal scale selection for multi-scale decision tables. *International Journal of Approximate Reasoning*, 54:1107–1129, 2013.