Multi-Attribute Decision Making Based on Attribute Importance Degree and Case-Based Reasoning

Zhong Wu, Ruixia Yan

School of Management, Shanghai University of Engineering Science, Shanghai201620, China
Emails: wuzhong8008@163.com   yanruixia@gmail.com

Abstract: To tackle a multi-attribute decision making problem, rough set and case-based reasoning are often combined. However, the reduction in a rough set is always complex. In this paper we provide a new relative importance measure about the unitary attributes values by ranking the relative importance of the attributes in the rough set theory. A new rough set model based on ranking the relative importance of the attributes is built and its properties are studied. Then unitary attributes values are utilized to compute the similarity of rules in case-based reasoning, for there might be incompletely match or miss values. A new multi-attribute decision making based on case-based reasoning and a rough set based on the ranking relative importance of the attributes is constructed, which obtains rules, avoiding reduction and rule extraction.

Keywords: Multi-attribute decision-making, case-based reasoning, rough set, relative position region, attribute importance.

1. Introduction

Decision making is choosing a strategy among many different projects in order to achieve some purposes. With different attitude of the decision makers for different types of decision models, the decision criteria are formulated as five different models: optimistic-decision criterion, pessimistic-decision criterion, evenness-decision criterion, minimum-risk-decision criterion, compromise-decision criterion. According to various decision criteria, the decision making problem is formulated

as three different models: high risk decision, usual risk decision and low risk decision. These different decision problems can be handled with a different method.

Case-Based Reasoning (CBR) is the process of solving new problems (cases) by retrieving the most relevant ones from an existing knowledge-base (called case-base) and adapting them to fit new situations [1]. CBR systems have been used in a wide variety of fields and applications. One example is its use in prediction and classification. CBR systems are well known for their ability to successfully tackle rich and complex domains. CBR is often used when generalized knowledge is lacking. Nevertheless, a CBR system is sensitive to noisy and unreliable data which may contribute negatively to its classification accuracy [2]. In fact, this problem may appear in CBR even if the domain contains few features and/or cases. Additionally, CBR classifiers suffer from the curse of dimensionality problem [3]. This is also the case for other learning approaches; nevertheless they showed that CBR is more dependent on the actual sample distribution than on the dimensions of the problem.

In CBR literature, these problems have been faced by two areas of research: feature selection and instance selection. The first one alleviates these problems by identifying as much of the irrelevant descriptive information (features) of a case as possible. On the other hand, instance selection – known as case-base maintenance in CBR – aims at reducing the number of unnecessary or redundant cases.

Some works focus on feature weighting and instance selection methods based on Rough Set Theory (RST) [4-6]. They have been proven to offer a good trade-off between reduction and problem solving efficiency.

Nevertheless, some data sets may also contain irrelevant features (attributes in CBR) in addition to irrelevant cases. Thus, it becomes necessary to reduce the number of features (rather than just weighting them) that is considered when solving new cases [7, 8].

Many algorithms within Artificial Intelligence literature deal with feature selection. These algorithms can be placed in two main categories: wrappers and filters. The wrapper methods use the performance algorithm itself as an evaluation function to estimate the accuracy of attribute subsets. Thus, wrappers tend to be computationally expensive because the learning algorithm is called on repeatedly. On the other hand, filter methods filter out undesirable attributes before learning takes place. Therefore, filters have been proven to be much faster than wrappers and hence they can be efficiently applied to large data sets containing many attributes.

In these 4 steps of case-based reasoning, "retrieve" is the most important step in complete systems. Here a complete system means that all solutions of the problems are in the system.

Traditionally, researchers always obtain retrieve cases utilizing the nearest neighbour subtraction algorithm. The nearest neighbour subtraction algorithm firstly computes the similarity between every attribute. Then it obtains unitary similarity by adding all similarities of the attributes. However, every attribute has a different role in the relations. So that we should consider the weight of an attribute when obtaining unitary similarity. The weight of an attribute is often presented by people, which causes subjective judgments.

A rough set is a methodology built on data structures, which avoids subjectiveness. Rough set theory [9, 10], a new mathematical approach to deal with inexact, uncertain or vague knowledge, has recently received wide attention in the research areas both of real-life applications and theory itself. Real-life applications speed up the theory research about a rough set. Rough set theory is an extension of set theory, in which a subset of a universe is described by a pair of ordinary sets called the lower and upper approximations. Rough set theory is emerging as a powerful theory dealing with imperfect data. It is an expanding research area which stimulates explorations on both real-world applications and on the theory itself. It has found practical applications in many areas, such as knowledge discovery, machine learning, data analysis, approximate classification, conflict analysis, and so on. The theory of rough sets has been successfully applied to diverse areas, such as pattern recognition, artificial intelligence, machine learning, knowledge acquisition, economy forecast, data mining and so on [11, 12]. Rough set theory adopts the concept of equivalence classes to partition the training instances according to some criteria. Two kinds of partitions are formed in the mining process: lower approximations and upper approximations, from which certain and possible rules are easily derived. It operates only on the data and does not require any added information; it is completely data-driven.

But there are still some defects in Pawlak rough set. Classification must be absolutely correct in a Pawlak rough set model, so the classical model cannot deal efficiently with datasets which have noisy data. Some latent useful knowledge may be abandoned. Researchers have put forward many extended rough set models combined with other soft computing theories, such as dominance-based rough set [13], rough fuzzy set [14] and fuzzy rough set [15], etc. Probabilistic rough set [16], variable precision rough set [17], and Bayesian rough set [18] are one of the most important branches. A variable precision rough set model was aimed at handling uncertain and noisy information and was directly derived from the original rough set model without any additional assumptions [17]. It integrated the concept of rough inclusion relation into Pawlak rough set model, thus being able to allow some degree of misclassification in the mining process.

Probabilistic rough set approximations can be formulated based on the notions of rough membership functions [19] and rough inclusion [20]. Also, we have a research rough set over dual-universe studied in [21, 22].

Since introduction of the extended models, they have been successfully used in many research fields. In these models, decisions about new problems are made according to the rules extracted. However, some attributes of the new problems are inconsistent with the rules completely. In this situation, how to make a decision might well repay investigation.

In this paper we combine a rough set and case-based reasoning to build an Multi-Attribute Decision Making (MADM) model. Considering the complexity of attribute reduction, we compute the sum of some attributes from specific rules. Firstly, we discuss the progress and research status of MADM. Secondly, we present the basic concept of a rough set and case-based reasoning. Then we provide a new concept about the importance of the attribute and a new MADM model is

built. An example is proposed to illustrate the practicability and efficiency of this MADM model.

2. Preliminaries

2.1. Pawlak rough set

The rough set theory, firstly introduced by Pawlak in 1982, is a valuable mathematical tool for dealing with vagueness and uncertainty [10, 11]. A rough set is a formal approximation of a crisp set (i.e., a conventional set) in terms of a pair of sets which give the lower and the upper approximation of the original set.

Let $I = (U, A)$ be an information system (attribute-value system), where U is a non-empty set of finite objects (the universe) and A is a non-empty, finite set of attributes, such that $a : U \rightarrow V_a$ for every $a \in A$; V_a is the set of values that attribute a may take. The information table assigns a value $a(x)$ from V_a to each attribute a and object x in the universe U. With any $P \subseteq A$ there is an associated equivalence relation

$$(1) \qquad \text{IND}(P) = \{(x, y) \in U^2 \mid \forall a \in P, a(x) = a(y)\}.$$

The relation $\text{IND}(P)$ is called a P-indiscernibility relation. The partition of U is a family of all equivalence classes of $\text{IND}(P)$ and is denoted by $U/\text{IND}(P)$ (or U/P). If $(x, y) \in \text{IND}(P)$, then x and y are indiscernible (or indistinguishable) by attributes from P.

Let $X \subseteq U$ be a target set that we wish to represent using the attribute subset P; that is, we are told that an arbitrary set of objects X comprises a single class, and we wish to express this class (i.e., this subset), using the equivalence classes induced by attribute P.

However, the target set X can be approximated using only the information contained within P by constructing the P-lower and P-upper approximations of X:

$$(2) \qquad \underline{R}X = \{x \in U : [x]_R \subseteq X\} = \bigcup\{[x]_R : [x]_R \subseteq X\},$$

$$(3) \qquad \overline{R}X = \{x \in U : [x]_R \cap X \neq \varnothing\} = \bigcup\{[x]_R : [x]_R X \neq \varnothing\}.$$

The *P*-lower approximation, or positive region, is the union of all equivalence classes in $[x]_P$ which are contained by (i.e., are subsets of) the target set. The lower approximation is the complete set of objects in *U/P* that can be positively (i.e., unambiguously) classified as belonging to target set *X*.

The *P*-upper approximation is the union of all equivalence classes in $[x]_P$ which have a non-empty intersection with the target set. The upper approximation is the complete set of objects that in *U/P* cannot be positively (i.e., unambiguously) classified as belonging to the complement (*X*) of the target set *X*. In other words, the upper approximation is the complete set of objects that are *P*.

In summary, the lower approximation of a target set is a conservative approximation consisting of only those objects which can positively be identified as

members of the set. (These objects have no indiscernible "clones" which are excluded by the target set.) The upper approximation is a liberal approximation which includes all objects that might be members of the target set. (Some objects in the upper approximation may not be members of the target set.) From the perspective of U/P, the lower approximation contains objects that are members of the target set with certainty (probability = 1), while the upper approximation contains objects that are members of the target set with non-zero probability (probability > 0).

An interesting question is whether there are attributes in the information system (attribute-value table) which is more important to the knowledge represented in the equivalence class structure than other attributes. Often we wonder whether there is a subset of attributes which can, by itself, fully characterize the knowledge in the database; such an attribute set is called a reduction.

Formally, a reduction is a subset of attributes $\mathrm{RED} \subseteq P$, such that

(1) $[x]_{\mathrm{RED}} = [x]_P$, that is, the equivalence classes induced by the reduced attribute set RED are the same as the equivalence class structure induced by the full attribute set P.

(2) the attribute set RED is minimal, in the sense that $[x]_{(\mathrm{RED}-\{a\})} \neq [x]_P$ for any attribute $a \in \mathrm{RED}$; in other words, no attribute can be removed from set RED without changing the equivalence classes $[x]_P$.

A reduction can be thought of as a sufficient set of features – sufficient, that is, to represent the category structure.

One of the most important aspects of database analysis or data acquisition is the discovery of attribute dependencies; that is, we wish to discover which variables are strongly related to which other variables. Generally, it is these strong relationships that will warrant further investigation, and that will ultimately be of use in predictive modelling.

In rough set theory, the notion of dependency is defined very simply. Let us take two (disjoint) sets of attributes, set P and set Q, and inquire what degree of dependency is obtained between them. Each attribute set induces an (indiscernibility) equivalence class structure, the equivalence classes induced by P given by $[x]_P$, and the equivalence classes induced by Q given by $[x]_Q$.

Let $[x]_Q = \{Q_1, Q_2, \cdots, Q_N\}$, where Q_i is a given equivalence class from the equivalence-class structure induced by attribute set Q. Then, the relative position region is given by $\mathrm{Pos}_P Q = \bigcup_{X \in U/Q} \underline{P}X$.

The dependency of attribute set Q on attribute set P is defined as

$$\gamma_P(Q) = \frac{|\mathrm{Pos}_P Q|}{|U|} \leq 1.$$

2.2. Case-based reasoning

Case-based reasoning has been formalized for the purposes of computer reasoning as a four-step process.

(1) Retrieve: Given a target problem, retrieve from the memory cases the relevant ones to solve it. A case consists of a problem, its solution, and, typically, annotations about how the solution was derived. For example, suppose Fred wants to prepare blueberry pancakes. Being a novice cook, the most relevant experience he can recall is the one, in which he successfully made plain pancakes. The procedure he followed for making the plain pancakes, together with justifications for decisions made along the way, constitutes Fred's retrieved case.

(2) Reuse: Map the solution from the previous case to the target problem. This may involve adapting the solution as needed to fit the new situation. In the pancake example, Fred must adapt his retrieved solution to include the addition of blueberries.

(3) Revise: Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, revise. Suppose Fred adapted his pancake solution by adding blueberries to the batter. After mixing, he discovers that the batter has turned blue – an undesired effect. This suggests the following revision: delay the addition of blueberries until after the batter has been ladled into the pan.

(4) Retain: After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory. Fred, accordingly, records his new-found procedure for making blueberry pancakes, thereby enriching his set of stored experiences, and better preparing him for future pancake-making demands.

3. A new attribute importance degree

Different attributes have variable affection on the problems. We consider the importance of an attribute in a different way from reduction.

Definition 1. Relative importance degree of attribute. In information system $I = (U, A)$, $A = C \bigcup D$, $R \subseteq C$. $\forall a \in C - R$, the importance degree of a is defined as

(4)
$$\mathrm{sig}(a, R, D) = \begin{cases} \gamma(R \bigcup \{a\}, D) - \gamma(R, D), R \neq \varnothing \\ \gamma(\{a\}, D), R = \varnothing \end{cases}.$$

In information system $I = (U, A)$, $A = C \bigcup D$, $\forall c_i \in C$. We just consider the relative importance of an attribute: $\mathrm{sig}(c_i, C, D)$ (for short).

From the definition of the relative position region, we know $\mathrm{Pos}_C D = \bigcup_{X \in U/D} \underline{C}X$. That is, for each equivalence class Y_i in $[x]_D$, we add up the size of its lower approximation by the attributes in C, i.e., $\underline{C}Y_i$. This approximation (as above, for the arbitrary set Y) is the number of objects which on

attribute set C can be positively identified as belonging to target set Y_i. Added across all equivalence classes in $[x]_D$, the numerator above represents the total number of objects which – based on the attribute set C – can be positively categorized according to the classification induced by attributes D. The dependency ratio therefore expresses the proportion (within the entire universe) of such classifiable objects. The dependency $\gamma_C(D)$ "can be interpreted as a proportion of such objects in the information system for which it suffices to know the values of attributes in C to determine the values of attributes in D". $0 \le \gamma(C, D) \le 1$.

Similarly, we comprehend $\gamma(R \cup \{a\}, D)$ and $\gamma(R, D)$.

Then $\text{sig}(a, R, D)$ is a relative importance degree of a.

Theorem 1. In information system $I = (U, A)$, $A = C \cup D$, $R \subseteq C$. $\forall a \in C - R$, the importance degree of a $\text{sig}(a, R, D) \ge 0$.

P r o o f: (1) If $R = \varnothing$, $\text{sig}(a, R, D) = \gamma(\{a\}, D) = \dfrac{\left| \text{Pos}_{\{a\}} D \right|}{|U|} \ge 0$

obviously.

(2) If $R \ne \varnothing$, let assume proof $\gamma(R \cup \{a\}, D) \ge \gamma(R, D)$, which equals to $\left| \text{Pos}_{R \cup \{a\}} D \right| \ge \left| \text{Pos}_R D \right|$.

For $a \in C - R$, $R \subset R \cup \{a\}$.

$\forall X \in U\big/_{R \cup \{a\}}$ $\exists Y \in U\big/_R$ and $X \subseteq Y$.

Based on definition of $\text{Pos}_C D$, we obtain $\text{Pos}_R D = \bigcup\limits_{Z \in U/_D} \underline{R} Z = \bigcup\limits_{Z \in U/_D} \bigcup\limits_{Y \in U/_R} (Y \subseteq Z)$ and

$\text{Pos}_{R \cup \{a\}} D = \bigcup\limits_{Z \in U/_D} \underline{R \cup \{a\}} Z = \bigcup\limits_{Z \in U/_D} \bigcup\limits_{X \in U/_{R \cup \{a\}}} (X \subseteq Z)$.

For $X \subseteq Y$, we have $\text{Pos}_R D \subseteq \text{Pos}_{R \cup \{a\}} D$.

Then $\left| \text{Pos}_{R \cup \{a\}} D \right| \ge \left| \text{Pos}_R D \right|$ and $\text{sig}(a, R, D) \ge 0$.

Definition 2. In information system $I = (U, A)$, $A = C \cup D$. Let $C = \{c_1, c_2, \cdots, c_n\}$, support $\text{sig}(c_1) \ge \text{sig}(c_2) \ge \cdots \ge \text{sig}(c_n)$. Define

(5)
$$R_i = \begin{cases} \{c_1\}, i = 1 \\ R_{i-1} \cup c_i, i > 1 \end{cases}.$$

That is, $R_1 = \{c_1\}$, $R_2 = \{c_1, c_2\}, \cdots, R_{n-1} = \{c_1, c_2, \cdots, c_{n-1}\}$, $R_n = C$.

Definition 3. In information system $I = (U, A)$, $A = C \cup D$, $R_i \subseteq C$. $\forall X \subseteq U$, we define

(6) $\underline{R_i} X = \{x \in U \mid [x]_{R_i} \subseteq X\} = \bigcup \{[x]_{R_i} \mid [x]_{R_i} \subseteq X\}$,

(7) $\quad \overline{R_i}X = \{x \in U \mid [x]_{R_i} \cap X \neq \varnothing\} = \bigcup\{[x]_{R_i} \mid [x]_{R_i} \cap X \neq \varnothing\}.$

$\underline{R_i}X$ and $\overline{R_i}X$ are lower and upper approximation of X based on R_i.

Theorem 2. In information system $I = (U, A)$, $A = C \bigcup D$, $R_i \subseteq C$. $X, X_1, X_2 \subseteq U$,

(1) $\underline{R_i}\varnothing = \overline{R_i}\varnothing$.

(2) $\underline{R_i}U = \overline{R_i}U$.

(3) $\underline{R_i}X \subseteq X \subseteq \overline{R_i}X$.

(4) If $X_1 \subseteq X_2$, then $\underline{R_i}X_1 \subseteq \underline{R_i}X_2, \overline{R_i}X_1 \subseteq \overline{R_i}X_2$.

(5) $\underline{R_i}(X_1 \bigcap X_2) = \underline{R_i}X_1 \bigcap \underline{R_i}X_2$.

(6) $\overline{R_i}(X_1 \bigcup X_2) = \overline{R_i}X_1 \bigcup \overline{R_i}X_2$.

P r o o f: (1) From Definition 2 and Definition 3, we know $\underline{R_i}\varnothing = \overline{R_i}\varnothing$ absolutely.

(2) From Definition 2 and Definition 3, $\underline{R_i}U = \overline{R_i}U$ absolutely.

(3) For $\forall x \in \underline{R_i}X \Rightarrow [x]_{R_i} \subseteq X \Rightarrow x \subseteq X$, then $\underline{R_i}X \subseteq X$.

For $\forall x \in X \Rightarrow [x]_{R_i} \cap X \neq \varnothing \Rightarrow x \subseteq \overline{R_i}X$, then $X \subseteq \overline{R_i}X$.

That is, $\underline{R_i}X \subseteq X \subseteq \overline{R_i}X$.

(4) For $X_1 \subseteq X_2$, then $\forall x \in \underline{R_i}X_1 \Rightarrow [x]_{R_i} \subseteq X_1 \Rightarrow [x]_{R_i} \subseteq X_2 \Rightarrow x \in \underline{R_i}X_2$.

For $X_1 \subseteq X_2$, then $\forall x \in \overline{R_i}X_1 \Rightarrow [x]_{R_i} \cap X_1 \neq \varnothing \Rightarrow [x]_{R_i} \cap X_2 \neq \varnothing \Rightarrow x \in \overline{R_i}X_2$.

That is, $X_1 \subseteq X_2 \Rightarrow \underline{R_i}X_1 \subseteq \underline{R_i}X_2, \overline{R_i}X_1 \subseteq \overline{R_i}X_2$.

(5) For $X_1 \bigcap X_2 \subseteq X_1 \bigcap X_2$ and $\underline{R_i}(X_1 \bigcap X_2) \subseteq \underline{R_i}X_1 \bigcap \underline{R_i}X_2$ from Theorem 1.

$\forall x \in \underline{R_i}X_1 \bigcap \underline{R_i}X_2 \Rightarrow x \in \underline{R_i}X_1 \wedge x \in \underline{R_i}X_2 \Rightarrow [x]_{R_i} \subseteq X_1 \wedge [x]_{R_i} \subseteq X_2 \Rightarrow$ $[x]_{R_i} \subseteq X_1 \bigcap X_2 \Rightarrow x \in \underline{R_i}(X_1 \bigcap X_2)$.

That is, $\underline{R_i}(X_1 \bigcap X_2) = \underline{R_i}X_1 \bigcap \underline{R_i}X_2$.

(6) For $X_1 \bigcup X_2 \supseteq X_1 \wedge X_1 \bigcup X_2 \supseteq X_2, \overline{R_i}(X_1 \bigcup X_2) \supseteq \overline{R_i}X_1 \bigcup \overline{R_i}X_2$ from Theorem 1.

$\forall x \in \overline{R_i}(X_1 \bigcup X_2) \Rightarrow [x]_{R_i} \cap (X_1 \bigcup X_2) \neq \varnothing \Rightarrow ([x]_{R_i} \cap X_1) \bigcup ([x]_{R_i} \cap X_2) \neq \varnothing \Rightarrow$ $([x]_{R_i} \cap X_1) \vee ([x]_{R_i} \cap X_2) \neq \varnothing \Rightarrow x \in \overline{R_i}X_1 \vee x \in \overline{R_i}X_2 \Rightarrow x \in (\overline{R_i}X_1 \bigcup \overline{R_i}X_2)$.

That is, $\overline{R_i}(X_1 \bigcup X_2) = \overline{R_i}X_1 \bigcup \overline{R_i}X_2$.

Theorem 3. In information system $I = (U, A)$, $A = C \bigcup D$, $R_i \subseteq C$.

(1) Define $f_{R_i}(x) = \{y \mid (x,y) \in c_j, \forall c_j \in R_i\}$ and $f_{R_i} = \bigcap_{c_j \in R_i} c_j$.

(2) Let $\alpha \in (0,1]$, define

$$R_{i\alpha} = \{(x,y) \mid \forall c_j \in R_i \subseteq C, (x,y) \in c_j \wedge \frac{i}{n} \geq \alpha\} \, (7) \text{ as } \alpha \text{-importance based}$$

on R_i.

(3) Let $\alpha \in (0,1]$, α-importance based on R_i can be denoted as

$$R_{i\alpha} = \{(x,y) \mid (x,y) \in R_i \wedge \frac{i}{n} \geq \alpha\}.$$

(4) $\underline{R_{1\alpha}}X \supseteq \underline{R_{2\alpha}}X \supseteq \cdots \supseteq \underline{R_{n\alpha}}X$ and $\overline{R_{1\alpha}}X \supseteq \overline{R_{2\alpha}}X \supseteq \cdots \supseteq \overline{R_{n\alpha}}X$.

4. A decision-making model

Multi-attribute decision making (MADM) provides a structured approach to decision making. MADM approach requires that the selection must be made among decision alternatives described by their attributes. It assumes that the problem has a predetermined number of decision alternatives.

Let us support that the decision making system is a complete information system. A complete information system means that all decision rules can be found in the rule table. We consider mismatched values in decision alternatives and construct a new MADM model. Considering the case-based reasoning theory, the steps of decision making are five.

Step 1. Normalize the information system by discretizing continuous attributes and clean repetitive lines.

Step 2. Compute the relative importance of the attributes and rank them.

Step 3. Set α. Compute $\underline{R_i}X$ ($i = 1, \ldots, j$, $j \leq n$). Obtain $\underline{R_{j\alpha}}X$.

Step 4. The elements in $\underline{R_{j\alpha}}X$ are similarity elements with X. Compute the similarity of rules based on case-based reasoning.

Step 5. Judge $\underline{R_{j\alpha}}X$ from an expert and make a decision.

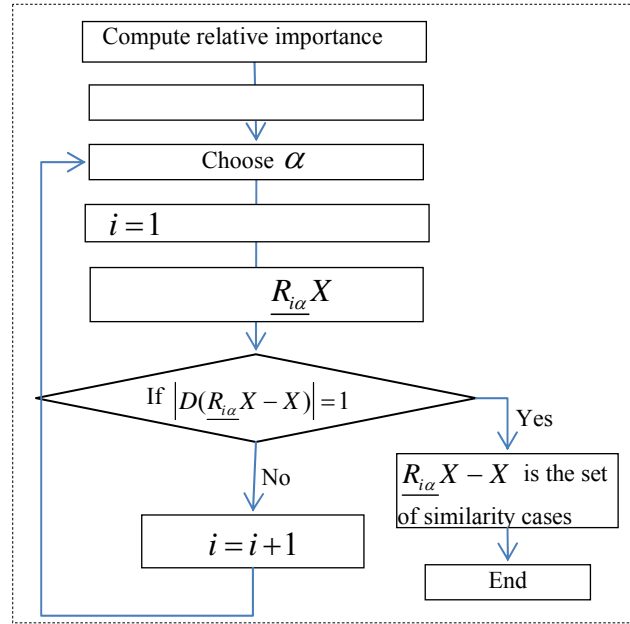Steps 2-4 are presented in the following flow chart.

Fig. 1. Flow chart of reasoning

Example 1. In information system $I = (U, A)$, $A = C \cup D$, $U = \{u_1, \cdots, u_{10}\}$, $C = \{c_1, c_2, c_3, c_4\}$ and $D = \{d\}$. Table 1 is an information system of $I = (U, A)$.

Table 1. Information system

U	c_1	c_2	c_3	c_4	d
u_1	2	1	2	1	1
u_2	1	2	1	1	1
u_3	1	1	2	2	2
u_4	1	2	1	1	1
u_5	1	1	2	1	1
u_6	1	2	1	2	2
u_7	2	1	1	1	2
u_8	1	2	2	1	1
u_9	2	1	1	1	2
u_{10}	1	2	2	2	2

Question. There is a new object, which is denoted by u_{11} and its values are 2, 2, 2, 1 under attributes c_1, c_2, c_3, c_4 respectively. What decision value should u_{11} have?

Let us handle this question by the modelling we provided above.

From the rough set theory we know

71

$$U\big/_{\{d\}} = \{\{u_1,u_2,u_4,u_5,u_8\},\{u_3,u_6,u_7,u_9,u_{10}\}\},$$

$$U\big/_{C} = \{\{u_1\},\{u_2,u_4\},\{u_3\},\{u_5\},\{u_6\},\{u_7\},\{u_8\},\{u_9\},\{u_{10}\}\},$$

$$U\big/_{C-\{c_1\}} = \{\{u_1,u_5\},\{u_2,u_4\},\{u_3\},\{u_6\},\{u_7,u_9\},\{u_8\},\{u_{10}\}\},$$

$$U\big/_{C-\{c_2\}} = \{\{u_1\},\{u_2,u_4\},\{u_3,u_{10}\},\{u_6\},\{u_7,u_9\},\{u_5,u_8\}\},$$

$$U\big/_{C-\{c_3\}} = \{\{u_1,u_7,u_9\},\{u_2,u_4,u_8\},\{u_3\},\{u_5\},\{u_{10}\},\{u_6\}\},$$

$$U\big/_{C-\{c_4\}} = \{\{u_1\},\{u_2,u_4,u_6\},\{u_3,u_5\},\{u_7,u_9\}\ \{u_8,u_{10}\}\}.$$

Similarly, $U\big/_{C-\{c_i,c_j\}}$ and $U\big/_{C-\{c_i,c_j,c_k\}}$ can also be computed as above, $i,j,k=1,2,3,4$, $i \neq j \neq k$.

Let us compute the relative set of condition attributes and the decision attribute,

$$Pos_C D = \{u_1,u_2,u_3,u_4,u_5,u_6,u_7,u_8,u_9,u_{10}\}.$$
$$Pos_{C-\{c_1\}} D = \{u_1,u_2,u_3,u_4,u_5,u_6,u_7,u_8,u_9,u_{10}\}.$$
$$Pos_{C-\{c_2\}} D = \{u_1,u_2,u_4,u_5,u_6,u_7,u_8,u_9\}.$$
$$Pos_{C-\{c_3\}} D = \{u_2,u_3,u_4,u_5,u_6,u_8,u_{10}\}.$$
$$Pos_{C-\{c_4\}} D = \{u_1,u_7,u_9\}.$$

From (4):
$$W_{c_1} D = \text{sig}(c_1,R,D) = \gamma(C,D) - \gamma(C-\{c_1\},D) = 0.$$
$$W_{c_2} D = \text{sig}(c_2,R,D) = \gamma(C,D) - \gamma(C-\{c_2\},D) = 0.2.$$
$$W_{c_3} D = \text{sig}(c_3,R,D) = \gamma(C,D) - \gamma(C-\{c_3\},D) = 0.3.$$
$$W_{c_4} D = \text{sig}(c_4,R,D) = \gamma(C,D) - \gamma(C-\{c_4\},D) = 0.7.$$

We rank the attributes $C' = \{c'_1,c'_2,c'_3,c'_4\} = \{c_4,c_3,c_2,c_1\}$ According to the size of $W_{c_4} D > W_{c_3} D > W_{c_2} D > W_{c_1} D$.

From the decision model above, let $R_1 = \{c_1\}$, $X = u_{11}$.

From (6): $\underline{R_1}X = \{u_1,u_2,u_4,u_5,u_7,u_8,u_9,u_{11}\}$, $\underline{R_2}X = \{u_1,u_5,u_{11}\}$, $\underline{R_3}X = \{u_{11}\}$.

Let $\alpha = 0.5$. From Equation (7), $R_{i\alpha} = R_2$.

In $\underline{R_2}X$, for $\left| D(\underline{R_2}X - \{u_{11}\}) \right| = 1$, $\underline{R_{i\alpha}}X - X$ are similarity cases with X.

Here, $D(u_{11}) = D(u_1) = D(u_5) = 1$.

The answer of the question is "if u_{11}'s values are 2, 2, 2, 1 under attributes c_1,c_2,c_3,c_4 respectively, then its decision value is 1".

Concerning this question, we did not operate the attribute reduction or extract the decision rules, which simplify the decision-making reasoning process.

5. Conclusion

The MADM model in this paper has the following advantages in a few points.

(1) This is a new rough set model, which considers the order based on the relative importance degree.

From the Example in this paper, we know that the new rough set model considering the order based on the relative importance degree has important significance to deal with uncertainty information.

(2) Decision making model based on attribute importance degree and case-based reasoning avoids the complexity of reduction in rough set theory, which facilitates the decision process.

References

1. R i e s b e c k, C. K., R. C. S c h a n k. Inside Case-Based Reasoning. London, Psychology Press, 2013.
2. K u m a r, K. A., Y. S i n g h, S. S a n y a l. Hybrid Approach Using Case-Based Reasoning and Rule-Based Reasoning for Domain Independent Clinical Decision Support in ICU. – Expert Systems with Applications, Vol. **36**, 2009, No 1, 65-71.
3. Z h u a n g, Z. Y., L. C h u r i l o v, F. B u r s t e i n et al. Combining Data Mining and Case-Based Reasoning for Intelligent Decision Support for Pathology Ordering by General Practitioners. – European Journal of Operational Research, Vol. **5**, 2009, No 3, 662-675.
4. D a i, J. Rough Set Approach to Incomplete Numerical Data. – Information Sciences, 2013, No 241, 43-57.
5. Y e, M., X. W u, X. H u, et al. Anonymizing Classification Data Using Rough Set Theory. – Knowledge-Based Systems, 2013, No 43, 82-94.
6. C h e n, D., S. K w o n g, Q. H e et al. Geometrical Interpretation and Applications of Membership Functions With Fuzzy Rough Sets. – Fuzzy Sets and Systems, 2012, No 193, 122-135.
7. S a l a m ó, M, E. G o l o b a r d e s. Unifying Weighting and Case Reduction Methods Based on Rough Sets to Improve Retrieval. Case-Based Reasoning Research and Development. Berlin, Heidelberg, Springer, 2003, 494-508.
8. S a l a m ó, M, M. L ó p e z-S á n c h e z. Rough Set Based Approaches to Feature Selection for Case-Based Reasoning Classifiers. – Pattern Recognition Letters, Vol. **32**, 2011, No 2, 280-292.
9. P a w l a k, Z. Rough Sets. – International Journal of International Sciences, 1982, No 11, 341-356.
10. P a w l a k, Z. Rough Sets Theoretical Aspects of Reasoning About Data. Dordrecht, Kluwer Academic Publishers, 1991.
11. Y a n t o, I. T. R., P. V i t a s a r i, T. H e r a w a n et al. Applying Variable Precision Rough Set Model for Clustering Student Suffering Study's Anxiety. – Expert Systems with Applications, Vol. **39**, 2012, No 1, 452-459.
12. C h a k h a r, S., I. S a a d. Dominance-Based Rough Set Approach for Groups in Multicriteria Classification Problems. – Decision Support Systems, Vol. **54**, 2012, No 1, 372-380.
13. L i o u, J. J. H., G. H. T z e n g. A Dominance-Based Rough Set Approach to Customer Behavior in the Airline Market. – Information Sciences, Vol. **180**, 2010, No 11, 2230-2238.
14. Q i a n, Y., J. L i a n g, W. W e i. Consistency-Preserving Attribute Reduction in Fuzzy Rough Set Framework. – International Journal of Machine Learning and Cybernetics, 2012, 1-13.

15. R a d z i k o w s k a, A. M., E. E. K e r r e. A Comparative Study of Fuzzy Rough Sets. – Fuzzy Sets and Systems, Vol. **126**, 2002, No 2, 137-155.
16. Z i a r k o, W. Probabilistic Rough Sets [M]. Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. – Berlin, Heidelberg, Springer, 2005, 283-293.
17. Z i a r k o, W. Variable Precision Rough Set Model. – Journal of Computer and System Sciences, Vol. **46**, 1993, No 1, 39-59.
18. S u n, B., W. M a, Z. G o n g. Dominance-Based Rough Set Theory over Interval-Valued Information Systems. – Expert Systems, 2013, No 5, 54-67.
19. B ł a s z c z y ń s k i, J., R. S ł o w i ń s k i, M. S z e l ą g. Sequential Covering Rule Induction Algorithm for Variable Consistency Rough Set Approaches. – Information Sciences, Vol. **181**, 2011, No 5, 987-1002.
20. Z h a n g, X., Z. M o, F. X i o n g, et al. Comparative Study of Variable Precision Rough Set Model and Graded Rough Set Model. – International Journal of Approximate Reasoning, Vol. **53**, 2012, No 1, 104-116.
21. Y a n, R., J. Z h e n g, J. L i u, Y. Z h a i. Research on the Model of Rough Set over Dual-Universes. – Knowledge-Based Systems, Vol. **23**, 2010, No 8, 817-822.
22. Y a n, R., J. Z h e n g, J. L i u, C. Q i n. Rough Set Model over Dual-Universes in Fuzzy Approximation. – Iranian Journal of Fuzzy Systems, Vol. **9**, 2012, No 3, 79-91.