Risk Decision Making Based on Decision-theoretic Rough Set: A Three-way View Decision Model

Huaxiong Li [1,2] , **Xianzhong Zhou** [1,2] *

[1] *School of Management and Engineering, Nanjing University,*
No. 22 Hankou Road, Nanjing, Jiangsu 210093, P.R.China
E-mail: huaxiongli@nju.edu.cn, zhouxz@nju.edu.cn

[2] *State Key Laboratory for Novel Software Technology, Nanjing University,*
No. 22 Hankou Road, Nanjing, Jiangsu 210093, P.R.China

Abstract

Rough set theory has witnessed great success in data mining and knowledge discovery, which provides a good support for decision making on a certain data. However, a practical decision problem always shows diversity under the same circumstance according to different personality of the decision makers. A simplex decision model can not provide a full description on such diverse decisions. In this article, a review of Pawlak rough set models and probabilistic rough set models is presented, and a three-way view decision model based on decision-theoretic rough set model is proposed, in which optimistic decision, pessimistic decision, and equable decision are provided according to the cost of misclassification. The thresholds of probabilistic inclusion are calculated based on minimization of risk cost under respective decision bias. The study not only presents a new theoretic decision model considering the different personality of the decision makers, but also provides a practical explanation and an illustrative example on diverse risk bias decision.

Keywords: decision-theoretic rough set; three-way view decision; risk decision making; Bayesian decision

1. Introduction

In many decision making problems, data mining can be regarded as a useful strategy to discover knowledge structure from the data sets, which helps to promote decision support. It mainly focuses on a computerized technology that uses diverse algorithms to find relationships and trends in large databases. Proposed by Pawlak in the early 1980s [1,2], rough set theory has been widely conceived as a mathematical tool to deal with vague or imprecise information [3]. In recent decades, rough set theory has become a focus topic in the research area of data mining and knowledge discovery, and witnessed great success in practical applications on intelligence information process, knowledge reduction and decision making support system. Many examples of applications of the rough set are frequently mentioned in literatures [5,6,7,8,9,10].

A fundamental application of rough set is to induce rules based on lower and upper approximation. Rules induced from lower approximation are certain

---

*Corresponding author.

rules, and rules induced from upper approximation are possible rules. All these rules can be used for decision making. In practice, however, a decision problem may show diversity under the same circumstance according to different personality of the decision makers. Decision makers may induce different decision rules based on the same data set according to their personal risk bias, and they may take different action on the same description of a risk state. A simplex decision model can not provide a full description on such diverse decisions, and a three-way view decision model is necessary.

Researches on classical rough set theory mainly focus on Pawlak rough set model [1,2], in which the lower and upper approximation operators are defined by core and support of a rough membership function represent only two extreme cases. They may be regarded as qualitative approximations of a set, which are restricted to induce certain decisions without flexible parameter. The Pawlak rough set model can be used to induce decision rules, which provide support for decision making. In Pawlak rough set model, the set inclusion must be fully correct or certain, and the actual degree of membership is not taken into consideration, which makes the rough set approach to be very sensitive to the accuracy of input data and not suitable to process noisy data.

In order to overcome the disadvantage of Pawlak rough set model, an extended probabilistic rough set model has been suggested. By introducing membership functions of set inclusion with statistical information, Yao proposes decision-theoretic rough set model(DTRS) [12,13,14,15] and Ziarko proposes variable precision rough set model(VPRS) [16]. DTRS model is a typical probabilistic rough set models, in which thresholds of the conditional probability for classification are introduced, and an available method to determine the thresholds is proposed based on Bayesian cost theory. In DTRS model, whether or not an object belongs to positive region depends on conditional probability and the thresholds, and the thresholds are precalculated by minimizing the decision cost according to Bayesian theory. Such extension of set inclusion allows for a controlled degree of misclassification in its formalism

which, in turn, leads to more general notions of concept approximations. DTRS model is a generalized rough set model, in which α-cut fuzzy set model, Pawlak rough set model and other type of probabilistic rough set models such as 0.5 probabilistic model and VPRS model can be derived [12]. When compared to DTRS model, VPRS model can be regarded as a special case of DTRS model, and a main disadvantage of VPRS is that the thresholds in VPRS are not presented with a certain method. One may find it is hard to select an appropriate value of threshold in VPRS model.

This paper is an extended version of the paper published in the proceedings of RSKT'2009 [17]. A main objective of this paper is to propose a three-way view decision model based on DTRS, which may present a full description on diverse decisions according to different risk bias of decision makers. As mentioned above, practical decision problem often shows diverse characteristic. For a given description, different decision makers often take different attitude under a certain decision background. DTRS can be regarded as a suitable mathematical tool to describable the diverse characteristic of decisions, which will be illustrated detailedly in the following discussion. Based on DTRS, optimistic decision, pessimistic decision, and equable decision are proposed according to the cost of misclassification, which are well explained from both practical example and mathematical theory.

2. A Review of Rough Set Models

In this section, we will review two categories of rough set models, which play an fundamental important role in the development of rough set theory. One category is Pawlak rough set models, and the other is probabilistic rough set models.

2.1. *Pawlak rough set models*

In Pawlak rough set models, upper approximation and lower approximation are defined by extreme cases of set inclusion. Let U denote a finite and non-empty set called the universe, and let $\Re \subseteq U \times U$ denote an equivalence relation on U. The partition of

the universe by $\mathfrak{R}$ is denoted by $U/\mathfrak{R}$. In information table, $\mathfrak{R}$ can be induced by a subset of attribute $A \subseteq C$, where C is the entire set of attributes. Suppose $X \subseteq U$ is a subset of U, the lower and upper approximation are respectively defined as follows:

$$
\begin{aligned}
\underline{apr}(X) &= \{x \mid x \in U, [x]_\mathfrak{R} \subseteq X\}; \\
\overline{apr}(X) &= \{x \mid x \in U, [x]_\mathfrak{R} \cap X \neq \emptyset\}.
\end{aligned} \quad (1)
$$

Based on the lower and upper approximations of a set $X \subseteq U$, the universe U can be divided into three disjoint regions, the positive region $POS(X)$, the negative region $NEG(X)$, and the boundary region $BND(X)$, which are respectively defined as follows:

$$
\begin{aligned}
POS(X) &= \underline{apr}(X); \\
NEG(X) &= U - \overline{apr}(X); \\
BND(X) &= \overline{apr}(X) - \underline{apr}(X).
\end{aligned} \quad (2)
$$

One may induce certain rules based on positive region, i.e., all elements $x \in POS(X)$ are certainly the members of X, and all elements $x \in POS(X)$ are impossible the members of X. In boundary region, one can not decide with certainty whether or not an element $x \in BND(X)$ is a member of X. From (1) and (2), we have $\overline{apr}(X) = POS(X) \cup BND(X)$. For element $x \in \overline{apr}(X)$, one can only conclude that x possibly belongs to X.

2.2. *Probabilistic rough set models*

In Pawlak rough set models, all elements with non-zero and non-full membership values will be classified into boundary region. However, in practice, a looser classification may be more useful. An object may be classified into the positive region if the conditional probability is sufficiently large. Similarly, an object may be classified into the negative region if the conditional probability is sufficiently small. Based on the statistical information of membership function, two fundamental probabilistic rough set models are proposed. One is DTRS model, and the other is VPRS model. As mentioned in section 1, DTRS is superior to VPRS in that the thresholds in DTRS are calculated according to Baysion

theory while VPRS does not provide a method to determine the values of thresholds. Furthermore, VPRS is just a special case of DTRS, and VPRS can be directly derived from DTRS when the decision costs are equal to some certain values [12]. Therefore, DTRS can be regarded as a representative model for probabilistic rough set.

Let us review the DTRS in detail. Suppose $\Omega = \{w_1, w_2, \ldots, w_s\}$ is a finite set of s states, $\mathscr{A} = \{a_1, a_2, \ldots, a_m\}$ is a finite set of m possible actions, and $P(w_j|\mathbf{x})$ is the conditional probability of an object x being in state w_j given that the object is described by $\mathbf{x}$. Let $\lambda(a_i|w_j)$ denote the loss, or cost, for taking action a_i when the state is w_j. For an object with description $\mathbf{x}$, suppose action a_i is taken. The expected loss associated with taking action a_i can be calculated by:

$$
R(a_i|\mathbf{x}) = \sum_{j=1}^{s} \lambda(a_i|w_j) P(w_j|\mathbf{x}) \quad (3)
$$

In general, a decision rule can be conceived as a function $\tau(\mathbf{x})$ that specifies which action to take, and the overall risk $\mathbf{R}$ of a decision rule is calculated by:

$$
\mathbf{R} = \sum_{\mathbf{x}} R(\tau(\mathbf{x})|\mathbf{x}) P(\mathbf{x}) \quad (4)
$$

Let us consider the object classification problem with the approximation operators. The set of states is given by $\Omega = \{X, \neg X\}$ indicating that an element is in X and not in X. The set of actions is given by $\mathscr{A} = \{a_P, a_N, a_B\}$, represent the three actions in classifying an object, deciding $POS(X)$, deciding $NEG(X)$, and deciding $BND(X)$, respectively. When an object belongs to X, let λ_{PP}, λ_{BP} and λ_{NP} denote the costs of taking actions a_P, a_B and a_N, respectively [18]. When an object does not belong to X, let λ_{PN}, λ_{BN}, and λ_{NN} denote the costs of taking the same three actions. Then the expected loss $R(a_i|[x]_\mathfrak{R})$ associated with taking the individual actions can be expressed as:

$$
\begin{aligned}
R(a_P|[x]_\mathfrak{R}) &= \lambda_{PP}P(X|[x]_\mathfrak{R}) + \lambda_{PN}P(\neg X|[x]_\mathfrak{R}); \\
R(a_N|[x]_\mathfrak{R}) &= \lambda_{NP}P(X|[x]_\mathfrak{R}) + \lambda_{NN}P(\neg X|[x]_\mathfrak{R}); \\
R(a_B|[x]_\mathfrak{R}) &= \lambda_{BP}P(X|[x]_\mathfrak{R}) + \lambda_{BN}P(\neg X|[x]_\mathfrak{R}).
\end{aligned}
$$
$$(5)$$

The Bayesian decision procedure leads to the following minimum-risk decision rules:

$$
\begin{aligned}
\text{If } R(a_P|[x]_{\mathfrak{R}}) &\leqslant & R(a_N|[x]_{\mathfrak{R}}) \\
R(a_P|[x]_{\mathfrak{R}}) &\leqslant & R(a_B|[x]_{\mathfrak{R}}), \\
\text{decide} & POS(X) & ; \\
\text{If } R(a_N|[x]_{\mathfrak{R}}) &\leqslant & R(a_P|[x]_{\mathfrak{R}}) \\
R(a_N|[x]_{\mathfrak{R}}) &\leqslant & R(a_B|[x]_{\mathfrak{R}}), \\
\text{decide} & NEG(X) & ; \\
\text{If } R(a_B|[x]_{\mathfrak{R}}) &\leqslant & R(a_P|[x]_{\mathfrak{R}}) \\
R(a_B|[x]_{\mathfrak{R}}) &\leqslant & R(a_N|[x]_{\mathfrak{R}}), \\
\text{decide} & BND(X) & .
\end{aligned} \tag{6}
$$

Consider a reasonable kind of loss functions with $\lambda_{PP} \leqslant \lambda_{BP} \leqslant \lambda_{NP}$ and $\lambda_{NN} \leqslant \lambda_{BN} \leqslant \lambda_{PN}$, one can induce following decision rules based on the fact that $P(X|[x]_{\mathfrak{R}}) + P(\neg X|[x]_{\mathfrak{R}}) = 1$:

$$
\begin{aligned}
\text{If } P(X|[x]_{\mathfrak{R}}) &\geqslant \gamma & \text{and } P(X|[x]_{\mathfrak{R}}) \geqslant \alpha, \\
\text{decide} & POS(X) & ; \\
\text{If } P(X|[x]_{\mathfrak{R}}) &\leqslant \beta & \text{and } P(X|[x]_{\mathfrak{R}}) \leqslant \gamma, \\
\text{decide} & NEG(X) & ; \\
\text{If } P(X|[x]_{\mathfrak{R}}) &\geqslant \beta & \text{and } P(X|[x]_{\mathfrak{R}}) \leqslant \alpha, \\
\text{decide} & BND(X) & ,
\end{aligned} \tag{7}
$$

where

$$
\begin{aligned}
\alpha &= \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}; \\
\gamma &= \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}; \\
\beta &= \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}.
\end{aligned} \tag{8}
$$

When $(\lambda_{PN} - \lambda_{BN})(\lambda_{NP} - \lambda_{BP}) > (\lambda_{BP} - \lambda_{PP})(\lambda_{BN} - \lambda_{NN})$, we have $\alpha > \beta$, thus $\alpha > \gamma > \beta$, and induce following decision rules:

$$
\begin{aligned}
\text{If} & \quad P(X|[x]_{\mathfrak{R}}) \geqslant \alpha, & \text{decide } POS(X); \\
\text{If} & \quad P(X|[x]_{\mathfrak{R}}) \leqslant \beta, & \text{decide } NEG(X); \\
\text{If} & \quad \beta < P(X|[x]_{\mathfrak{R}}) < \alpha, & \text{decide } BND(X).
\end{aligned} \tag{9}
$$

As a special case of DTRS and another type of probabilistic rough set model, VPRS also introduce statistical information for set approximation.

A parameter $\beta \in (0.5, 1]$ is introduced to set thresholds for lower approximation upper approximation, which is provided by the user. A generalized notion of β-lower approximation and β-upper approximation can be obtained as follows:

$$
\begin{aligned}
\underline{apr}_{\beta}(X) &= \{x \in U | P(X|[x]_R) \geqslant \beta\}; \\
\overline{apr}_{\beta}(X) &= \{x \in U | P(X|[x]_R) > 1 - \beta\}.
\end{aligned} \tag{10}
$$

With the definition of β-lower approximation and β-upper approximation, one can also partition the entire universe into three regions based on the value of β:

$$
\begin{aligned}
POS_{\beta}(X) &= \underline{apr}_{\beta}(X); \\
NEG_{\beta}(X) &= U - \overline{apr}_{\beta}(X); \\
BND_{\beta}(X) &= \overline{apr}_{\beta}(X) - \underline{apr}_{\beta}(X).
\end{aligned} \tag{11}
$$

However, a fundamental important problem still remains in VPRS. How to determine the threshold values for deciding the three regions? Users have to determine the thresholds according to their intuition, which may not be reliable. That is a main disadvantage of VPRS when compared to DTRS.

3. A three-way view decision model based on DTRS

DTRS is a typical probabilistic rough set model, in which a reasonable method to determine the threshold for distinguishing the three regions is provided, and the method is proposed based on the reliable mathematics foundation: Bayesian decision theory. The successful applications of the DTRS in a variety of problems have amply demonstrated its usefulness and versatility [19,20,21,22,23]. In this section, we will propose a new three-way view decision model based on DTRS.

3.1. A three-way view decision model

In a practical decision problem, we may find diverse characteristics between the types of decisions. Different attitudes towards a single decision can be seen among separate groups of people. Some people always take optimistic decision, while other people may take pessimistic decision or equable decision.

For example, in medical examination, some symptoms indicate that a person may get illness with a certain possibility. For an optimistic person, he or she will think the possibility of get illness is not so high, and the illness may be self-recover even without treatment, therefore, he or she makes the decision not to see the doctor, while a pessimistic person will be afraid of getting illness and think that the illness may cause an aggravation without treatment, therefore he or she will make the decision to see the doctor immediately.

In general, people will take different type decision according their personal character. When considering the difference among different people, it is necessary to develop a fully descriptive decision model in which diverse types of decision are embodied. To this end, we propose a three-way view decision model based on DTRS. It can be argued that optimistic decision, pessimistic decision, and equable decision are different in that they adopt different values on the costs.

Let us take the example mentioned above for illustration. Suppose there are two states after a series of diagnoses for a disease: X is a confirmed disease state and $\neg X$ is a confirmed no-disease state, thus they are two complementary states. There are three actions regarding the three regions of the disease decision: a_P is to take some treatments to a patient, a_B is an amphibolous action, which means to wait-and-see, and a_N is to cancel any further treatment. For each pair of action-state, there is a corresponding cost λ_{ij}, where $i = P$, N, or B, and $j = P$, or N. For optimistic people, the cost of taking action of canceling any treatment for a confirmed disease state, λ_{NP}, is not so high, because they have optimistic character and do not care so much about getting disease. In comparison, pessimistic people will always be anxiety and afraid of getting any disease. Therefore, they adopt a very high value on the cost λ_{NP}.

In addition, there are some other people whom take middle-of-the road, i.e., the values of cost λ_{NP} is medium. Likewise, for pessimistic people, the cost of taking action of treatment for a confirmed no-disease state, λ_{PN}, is not so high as that of optimistic people. That is, for pessimistic people, λ_{PN} is

low and λ_{NP} is high. Based on the analysis above, for a given description $[x]_{\mathfrak{R}}$, we provide three type of decision: optimistic decision, pessimistic decision, and equable decision, which are defined according to the values of cost λ_{PP}, λ_{BP}, λ_{NP}, λ_{PN}, λ_{BN}, and λ_{NN}. All these three types of decision are based on following assumption:

$$
\begin{aligned}
&(i) \quad \alpha > \beta; \quad (ii)\lambda_{PP} = \lambda_{NN} = 0; \\
&(iii) \quad \lambda_{BP} = \sigma\lambda_{NP}, \lambda_{BN} = \sigma\lambda_{PN}, 0 < \sigma < 1.
\end{aligned}
\tag{12}
$$

Based on formula (8) and assumption (12), we have:

$$
\begin{aligned}
\alpha &= \frac{(1-\sigma)\lambda_{PN}}{(1-\sigma)\lambda_{PN} + \sigma\lambda_{NP}}; \\
\beta &= \frac{\sigma\lambda_{PN}}{\sigma\lambda_{PN} + (1-\sigma)\lambda_{NP}}.
\end{aligned}
\tag{13}
$$

Let $\sigma \in (0,1)$, and the three decisions are presented as follows:

Optimistic decision: Select a **Low** λ_{NP} and a **High** λ_{PN}, then calculate the α_O and β_O based on formula (13),

$$
\begin{aligned}
&\text{If} \quad P(X|[x]_{\mathfrak{R}}) \geqslant \alpha_O, \text{ decide } POS(X); \\
&\text{If} \quad P(X|[x]_{\mathfrak{R}}) \leqslant \beta_O, \text{ decide } NEG(X); \\
&\text{If} \quad \beta_O < P(X|[x]_{\mathfrak{R}}) < \alpha_O, \text{ decide } BND(X).
\end{aligned}
\tag{14}
$$

Pessimistic decision: Select a **High** λ_{NP} and a **Low** λ_{PN}, then calculate the α_P and β_P based on formula (13),

$$
\begin{aligned}
&\text{If} \quad P(X|[x]_{\mathfrak{R}}) \geqslant \alpha_P, \text{ decide } POS(X); \\
&\text{If} \quad P(X|[x]_{\mathfrak{R}}) \leqslant \beta_P, \text{ decide } NEG(X); \\
&\text{If} \quad \beta_P < P(X|[x]_{\mathfrak{R}}) < \alpha_P, \text{ decide } BND(X).
\end{aligned}
\tag{15}
$$

Equable decision: Select a **Medium** λ_{NP} and a **Medium** λ_{PN}, then calculate the α_E and β_E based on formula (13),

$$
\begin{aligned}
&\text{If} \quad P(X|[x]_{\mathfrak{R}}) \geqslant \alpha_E, \text{ decide } POS(X); \\
&\text{If} \quad P(X|[x]_{\mathfrak{R}}) \leqslant \beta_E, \text{ decide } NEG(X); \\
&\text{If} \quad \beta_E < P(X|[x]_{\mathfrak{R}}) < \alpha_E, \text{ decide } BND(X).
\end{aligned}
\tag{16}
$$

For simplicity, all these three decisions can be presented as a simply uniform expression:

$$\text{If} \quad P(X|[x]_{\mathfrak{R}}) \geqslant \alpha_i, \text{ decide } POS(X);$$
$$\text{If} \quad P(X|[x]_{\mathfrak{R}}) \leqslant \beta_i, \text{ decide } NEG(X); \quad (17)$$
$$\text{If} \quad \beta_i < P(X|[x]_{\mathfrak{R}}) < \alpha_i, \text{ decide } BND(X),$$

where $i \in \{$optimistic, pessimistic, equable$\}$, and $\alpha_i \in \{\alpha_{optimistic}, \alpha_{pessimistic}, \alpha_{equable}\}$. The values of α_i is determined according to the selected λ_{NP} and λ_{PN}:
Optimistic decision: *Low λ_{NP} and High λ_{PN}*;
Pessimistic decision: *High λ_{NP} and Low λ_{PN}*;
Equable decision: *Medium λ_{NP} and Medium λ_{PN}*.

The three decision models are different in the selected values of λ_{NP} and λ_{PN}, which indicate the different risk bias of decision-makers. An optimistic decision-maker always shows low cost of λ_{NP} and high cost of λ_{PN}, and a pessimistic decision shows quite the contrary: high cost of λ_{NP} and low cost of λ_{PN}. It should be noted that this description is based on the fact that POS(X) is an unpleasant event for all decision-makers, while NEG(X) is a pleasant event for all decision-makers. The decision models will take opposite result if POS(X) is an pleasant event for all decision-makers and NEG(X) is an unpleasant event. We may illustrate it in detail by an example of medical diagnoses.

For example, diagnoses for a disease and decision for treatment are unpleasant event for all decision-makers, and an optimistic decision-maker is overly optimistic when judging whether unpleasant events are more or less likely to bring them unhappiness. This optimistic bias often stems from their perception that they have control over unpleasant events. They believe the disease is just a matter of no importance, therefore the cost of λ_{NP} is low. On the contrary, they think the wrong diagnoses for a disease in the case that they actually have not got the disease will bring more trouble or cost, therefore the cost of λ_{PN} is high. For a pessimistic decision-maker, they are pessimistic when judging whether unpleasant events are more or less likely to bring them unhappiness. The pessimistic bias stems from their perception that they can not control over unpleasant events, and they believe the disease is an unpleasant and uncontrollable event, which should

be treated as soon as possible, otherwise it will be much dangerous, therefore the cost of λ_{NP} is high and λ_{PN} is relatively low.

However, the cost λ_{NP} and λ_{PN} in the decision model will go upside down when POS(X) is a pleasant event. In the example mentioned above, suppose POS(X) denote that the diagnoses result is healthy and decide no treatment, and NEG(X) denote that the diagnoses result is diseased and decide for further treatment. In this case, POS(X) is a pleasant event for all decision-makers, and NEG(X) is an unpleasant event, and the cost of taking positive decision(no treatment) under the negative state is denoted by λ'_{PN}, and the cost of taking negative decision(treatment) under the positive state is denoted by λ'_{NP}. Compared with previous case, we have $\lambda'_{NP} = \lambda_{PN}$, $\lambda'_{PN} = \lambda_{NP}$, and the selected cost λ'_{PN} and λ'_{NP} with regard to decision type are modified as:
Optimistic decision: *High λ'_{NP} and Low λ'_{PN}*;
Pessimistic decision: *Low λ'_{NP} and High λ'_{PN}*;
Equable decision: *Medium λ'_{NP} and Medium λ'_{PN}*.

3.2. *An explanation on three-way view decision model*

The three-way view decision model given in last subsection is based on assumption (12), and it selects λ_{NP} and λ_{PN} instead of other costs as parameters to define the type of decision. In this subsection, we will explain why we adopt such an assumption and why λ_{NP} and λ_{PN} instead of other costs are key factors to classify the decisions.

In general, the cost of a correct classification of a description $[x]_{\mathfrak{R}}$ is always lower than that of a wrong classification or amphibolous classification. Therefore, we have $\lambda_{PP} \leqslant \lambda_{BP} \leqslant \lambda_{NP}$ and $\lambda_{NN} \leqslant \lambda_{BN} \leqslant \lambda_{PN}$. For simplification, suppose the ratio of λ_{BP} to λ_{NP} and the ratio of λ_{BP} to λ_{NP} are constant and take the same value σ, then we have $0 < \sigma < 1$. Furthermore, the boundary region of a given state is not always empty, which indicates that $\alpha > \beta$ holds. In the six costs, λ_{PP} and λ_{NN} are the costs of correct classification, therefore, it is reasonable to set them as 0.

Based on the assumption, the key factors to classify the types of decision lies in the costs of wrong

classification: λ_{NP} and λ_{PN}. The relationship of λ_{NP}, λ_{PN} and α, β is depicted in Fig.1, where $\sigma = 1/3$ and x-axis, y-axis are λ_{NP} and λ_{PN} respectively, and the values of α and β are presented in z-axis. In Fig.1, it shows that with the increase of λ_{NP}, the threshold of α and β are decreased, and a large λ_{NP} and a low α means a pessimistic decision and a low λ_{NP} and a high α means optimistic decision. Likewise, with the increase of λ_{PN}, the threshold of α and β are increased, and a large λ_{PN} and a high α means a optimistic decision and a low λ_{PN} and a low α means pessimistic decision. For the medical example mentioned above, a pessimistic person always adopts a high value of cost λ_{NP}, and the corresponding α is low, which means a low possibility of getting disease will lead a pessimistic people to make the decision to take some treatment, but for optimistic persons, they always adopts a low value of cost λ_{NP}, and the corresponding α is high, which means that the optimistic people will not take some treatment until the possibility of getting disease is very high. We may get similar conclusion when considering the relation between λ_{PN} and α.
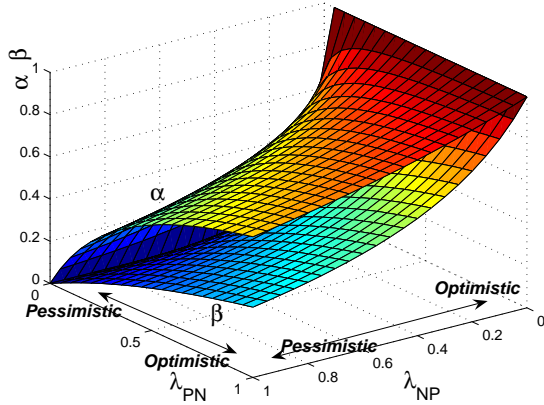

Fig. 1. Curved surface of α and β with regard to λ_{NP}, λ_{PN}
.

3.3. An illustrative example

We use an example to illustrate the main idea of three-way view decision based on decision-theoretic rough set. The example is a medical diagnose case given in Table 1, which is also called a decision table.

The columns of the Table 1 represent cases, which include the symptoms of patients and the diagnosis results for whether or not the patients have got flu. There are 9 columns in the table respectively represent 9 patients denoted by $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, \}$. The rows of the decision table are labeled by variables, which are called attribute, abbreviated by Att, and the attribute set of this decision is $Att = \{temperature, cough, nausea, flu\}$, and the first three attributes are condition attributes, which are respectively abbreviated and denoted by $C = \{Tem, Cou, Nau\}$. The last attribute flu is decision attribute: $d = \{flu\}$, which denotes the diagnosis result. The symbol "+" indicates that the patient has got flu, and "-" indicates that the patient has not got flu.

Table 1. An decision table for flu diagnose

Att	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
Tem	high	low	high	low	high	high	low	low	high
Cou	yes	yes	no	yes	no	yes	yes	yes	yes
Nau	no	yes	no	yes	no	no	yes	yes	no
Flu	+	-	+	-	-	+	+	-	+

In Table 1, the set of condition attributes C determine an equivalence relation $\mathfrak{R}$:

$$\mathfrak{R} = \{(x,y)|I_a(x) = I_a(y), \forall a \in C\},$$

where $I_a(x)$ is the value of x on attribute a. A partition will be constructed with regard to equivalence relation $\mathfrak{R}$:

$$U/\mathfrak{R} = U/C = \{\{x_1, x_6, x_9\}, \{x_2, x_4, x_7, x_8\}, \{x_3, x_5\}\}$$
$$= \{[x_1]_{\mathfrak{R}}, [x_2]_{\mathfrak{R}}, [x_3]_{\mathfrak{R}}\},$$

where $[x_i]_{\mathfrak{R}}$ $(i = 1, 2, 3)$ is the equivalence class of x_i:

$$[x_i]_{\mathfrak{R}} = \{y \in U|(x_i, y) \in \mathfrak{R}\}.$$

Similarly, the decision set $d = \{flu\}$ constructs a partition as:

$$U/d = \{\{x_1, x_3, x_6, x_7, x_9\}, \{x_2, x_4, x_5, x_8\}\}$$
$$= \{X, \neg X\},$$

where $X = \{x_1, x_3, x_6, x_7, x_9\}$ represents the state "flu" and $\neg X = \{x_2, x_4, x_5, x_8\}$ represents the state "no flu".

For each equivalence class in $U/\mathfrak{R}$, we may calculate the conditional probability as follows:

$$
\begin{aligned}
P(X|[x_1]_{\mathfrak{R}}) &= \frac{|[x_1]_{\mathfrak{R}} \cap X|}{|[x_1]_{\mathfrak{R}}|} \\
&= \frac{|\{x_1, x_6, x_9\} \cap \{x_1, x_3, x_6, x_7, x_9\}|}{|\{x_1, x_6, x_9\}|} \\
&= 1, \quad\quad\quad (18)
\end{aligned}
$$

$$
\begin{aligned}
P(X|[x_2]_{\mathfrak{R}}) &= \frac{|[x_2]_{\mathfrak{R}} \cap X|}{|[x_2]_{\mathfrak{R}}|} \\
&= \frac{|\{x_2, x_4, x_7, x_8\} \cap \{x_1, x_3, x_6, x_7, x_9\}|}{|\{x_2, x_4, x_7, x_8\}|} \\
&= 0.25, \quad\quad\quad (19)
\end{aligned}
$$

$$
\begin{aligned}
P(X|[x_3]_{\mathfrak{R}}) &= \frac{|[x_3]_{\mathfrak{R}} \cap X|}{|[x_32]_{\mathfrak{R}}|} \\
&= \frac{|\{x_3, x_5\} \cap \{x_1, x_3, x_6, x_7, x_9\}|}{|\{x_3, x_5\}|} \\
&= 0.5, \quad\quad\quad (20)
\end{aligned}
$$

where $|.|$ denotes the cardinality of a set. The conditional probability $P(X|[x_i]_{\mathfrak{R}})$ describes the probability of getting flu under the description $[x_i]_{\mathfrak{R}}$. The higher the conditional probability, the more possible getting flu. But optimistic decision maker, pessimistic decision maker and equable decision maker may take different attitude on the same conditional probability, and a three-way view decision model is necessary.

According to three-way view decision model presented in subsection 3.1, the values of α_i is determined by λ_{NP} and λ_{PN}. For optimistic decision, λ_{NP} is greater than that in pessimistic decision, and λ_{PN} is less than that in pessimistic decision, while equable decision takes middle way. For simplicity, let λ_{NP} and λ_{PN} be a integer range from 0 to 10. The values of λ_{NP} and λ_{PN} for each decision are listed as following payoff matrix (Table 2).

Table 2. A payoff matrix of three-way view decision model

Cost	Optimistic	Pessimistic	Equable
λ_{NP}	2	7	4
λ_{PN}	7	2	4

Let $\sigma = 0.4$, the threshold α_i and β_i are determined based on formula (13):

$$
\begin{aligned}
\alpha_{optimistic} &= \frac{(1-0.4) \times 7}{(1-0.4) \times 7 + 0.4 \times 2} = 0.84 \\
\beta_{optimistic} &= \frac{0.4 \times 7}{0.4 \times 7 + (1-0.4) \times 2} = 0.7 \\
\alpha_{pessimistic} &= \frac{(1-0.4) \times 2}{(1-0.4) \times 2 + 0.4 \times 7} = 0.3 \\
\beta_{pessimistic} &= \frac{0.4 \times 2}{0.4 \times 2 + (1-0.4) \times 7} = 0.16 \\
\alpha_{equable} &= \frac{(1-0.4) \times 4}{(1-0.4) \times 4 + 0.4 \times 4} = 0.6 \\
\beta_{equable} &= \frac{0.4 \times 4}{0.4 \times 4 + (1-0.4) \times 4} = 0.4
\end{aligned}
$$

The three-way view decision can be described as:
Optimistic decision:

If $\quad P(X|[x]_{\mathfrak{R}}) \geqslant 0.84$, decide $POS(X)$;
If $\quad P(X|[x]_{\mathfrak{R}}) \leqslant 0.7$, decide $NEG(X)$; $\quad$ (21)
If $\quad 0.7 < P(X|[x]_{\mathfrak{R}}) < 0.84$, decide $BND(X)$.

Pessimistic decision:

If $\quad P(X|[x]_{\mathfrak{R}}) \geqslant 0.3$, decide $POS(X)$;
If $\quad P(X|[x]_{\mathfrak{R}}) \leqslant 0.16$, decide $NEG(X)$; $\quad$ (22)
If $\quad 0.16 < P(X|[x]_{\mathfrak{R}}) < 0.3$, decide $BND(X)$.

Equable decision:

If $\quad P(X|[x]_{\mathfrak{R}}) \geqslant 0.6$, decide $POS(X)$;
If $\quad P(X|[x]_{\mathfrak{R}}) \leqslant 0.4$, decide $NEG(X)$; $\quad$ (23)
If $\quad 0.4 < P(X|[x]_{\mathfrak{R}}) < 0.6$, decide $BND(X)$.

In Table 1, all cases in the universe U are partitioned to three equivalence classes by equivalence relation $\mathfrak{R}$, i.e., $[x_1]_{\mathfrak{R}} = \{\{x_1, x_6, x_9\}, [x_2]_{\mathfrak{R}} =$

$\{x_2, x_4, x_7, x_8\}$, $[x_3]_{\mathfrak{R}} = \{x_3, x_5\}$, and all equivalence classes are respectively associated with logic descriptions as:

$[x_1]_{\mathfrak{R}}:$ $\quad Tem = high \wedge Cou = yes \wedge Nau = no,$

$[x_2]_{\mathfrak{R}}:$ $\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$[x_3]_{\mathfrak{R}}:$ $\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

and the decisions of POS(X), NEG(X), and BND(X) are respectively interpreted as:

Decide $\quad POS(X):$ $\quad Flu = yes$ (*treatment*),

Decide $\quad NEG(X):$ $\quad Flu = no:$ (*no treatment*),

Decide $\quad BND(X):$ $\quad Flu = uncertain$ (*wait and see*).

According to (18)$\sim$(23), we may get optimistic decision rules, pessimistic decision rules and equable decision rules. For optimistic decision rules, we may induce rules based on (21): $P(X|[x_1]_{\mathfrak{R}}) = 1 \geqslant 0.84$, then decide POS(X), and $P(X|[x_2]_{\mathfrak{R}}) = 0.25 \leqslant 0.7$, then decide NEG(X), and $P(X|[x_3]_{\mathfrak{R}}) = 0.5 \leqslant 0.7$, then decide NEG(X). Therefore, optimistic decision rules are:

$\quad Tem = high \wedge Cou = yes \wedge Nau = no,$

$\implies Flu = yes$ (*treatment*),

$\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$\implies Flu = no$ (*no treatment*),

$\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$\implies Flu = no$ (*no treatment*).

For pessimistic decision rules, we may induce rules based on (22): $P(X|[x_1]_{\mathfrak{R}}) = 1 \geqslant 0.3$, then decide POS(X), and $0.16 < P(X|[x_2]_{\mathfrak{R}}) = 0.25 < 0.3$, then decide NEG(X), and $P(X|[x_3]_{\mathfrak{R}}) = 0.5 \geqslant 0.3$, then decide POS(X). Therefore, pessimistic decision rules are:

$\quad Tem = high \wedge Cou = yes \wedge Nau = no,$

$\implies Flu = yes$ (*treatment*),

$\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$\implies Flu = uncertain$ (*wait and see*),

$\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$\implies Flu = yes$ (*treatment*),

For equable decision rules, we may induce rules based on (23): $P(X|[x_1]_{\mathfrak{R}}) = 1 \geqslant 0.6$, then decide POS(X), and $P(X|[x_2]_{\mathfrak{R}}) = 0.25 \leqslant 0.4$, then decide NEG(X), and $0.4 < P(X|[x_3]_{\mathfrak{R}}) = 0.5 < 0.6$, then decide BND(X).

$\quad Tem = high \wedge Cou = yes \wedge Nau = no,$

$\implies Flu = yes$ (*treatment*),

$\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$\implies Flu = no$ (*no treatment*),

$\quad Tem = no \wedge Cou = yes \wedge Nau = yes,$

$\implies Flu = uncertain$ (*wait and see*).

For a specific description $[x_i]_{\mathfrak{R}}$, the three-way view decision model presents a general decision strategy considering diversiform risk bias, which are listed in Table 3 for comparison. In Table 3, "treatment","no treatment" and "wait and see" are respectively abbreviated as "T", "NT" and "WAS".

Table 3. A comparative analysis on different risk bias

Description	Optimistic	Pessimistic	Equable
$[x_1]_{\mathfrak{R}}$	T	T	T
$[x_2]_{\mathfrak{R}}$	NT	WAS	NT
$[x_3]_{\mathfrak{R}}$	NT	T	WAS

We may find the decision rules induced by three-way view decision model shows a good accordance with the practical decision-making. In Table 3, both optimistic decision makers and pessimistic decision makers take the same decision on the description $[x_1]_{\mathfrak{R}}$ since the conditional probability of getting flu ($P(X|[x_1]_{\mathfrak{R}}) = 1$) is very high, but they take different decision on the description $[x_2]_{\mathfrak{R}}$ and $[x_3]_{\mathfrak{R}}$. For optimistic decision makers, they take "no treatment" on both $[x_2]_{\mathfrak{R}}$ and description $[x_3]_{\mathfrak{R}}$ because they think the probability of getting flu is not so high, whereas pessimistic decision makers decide "treatment" on the description $[x_3]_{\mathfrak{R}}$ since the probability of getting flu under the description $[x_3]_{\mathfrak{R}}$ beyond their security threshold, and they can not make a resolute decision on the description $[x_2]_{\mathfrak{R}}$ even though the probability of getting flu under the description $[x_2]_{\mathfrak{R}}$ is very low. In addition, equable decision makers always show

a moderation attitude towards the three description, which is more optimistic than pessimistic decision makers and more pessimistic than optimistic decision makers.

4. Conclusion

Rough sets theory has proved to be a useful tool for decision making based on data sets. A practical decision making problem always shows diversity according to the personal risk bias of the decision makers. A simplex decision model can not provide a full description on diverse decisions, and a three-way view decision model is suggested, in which diverse risk bias of the decision makers are embodied. For this purpose, a review of Pawlak rough set models and probabilistic rough set models is presented in this paper, and a three-way view decision model based on decision-theoretic rough set is proposed, in which optimistic decision, pessimistic decision, and equable decision are provided according to the cost of misclassification. The thresholds of probabilistic inclusion are calculated based on minimization of risk cost under respective decision bias. All three decisions in the model are different in their risk cost of taking positive action under the negative state, and risk cost of taking negative action under the positive state, which result in different thresholds of probabilistic inclusion for decision making. The proposed decision model truthfully reflects the diversity of decision making according to different risk bias. The study not only presents a new theoretic decision model considering the different personality of the decision makers, but also provides a practical explanation and an illustrative example on diverse risk bias decision.

Acknowledgments

References

1. Z. Pawlak, "Rough sets," *International Journal of Computer and Information Science*, **11**, 341–356 (1982).
2. Z. Pawlak, "Rough classification," *International Journal of Man-Machine Studies*, **20**, 469–483 (1984).
3. Z. Pawlak, J. G. Grzymala-Busse, R. Slowinski, and W. Ziarko, "Rough sets," *Communications of the ACM*, **38**, 89–95 (1995).
4. L. Polkowski, and P. Artiemjew, "On knowledge granulation and applications to classifier induction in the framework of rough mereology," *International Journal of Computational Intelligence Systems*, **2**, 315–331 (2009).
5. A. Mrozek and L. Plonka, "Rough sets in industrial applications," *Rough Sets in Knowledge Discovery*, **2**, 214–237 (1998).
6. R. Jensen and Q. Shen, "Semantics-preserving dimensionality reduction: rough and fuzzy-rough-based approaches," *IEEE Transactions on Knowledge and Data Engineering*, **16**, 1457–1471 (2004).
7. P. Lingras, M. Chen, and D. Q. Miao, "Rough cluster quality index based on decision theory," *IEEE Transactions on Knowledge and Data Engineering*, **21**, 1014–1026 (2009).
8. Y. Leung, W. Z. Wu, and W. X. Zhang, "Knowledge acquisition in incomplete information systems: a rough set approach," *European Journal of Operational Research*, **168**, 164–180 (2006).
9. D. Q. Miao, Y. Zhao, Y. Y. Yao, H. X. Li, and F. F. Xu, "Relative reducts in consistent and inconsistent decision tables of the Pawlak rough set model," *Information Sciences*, **179**, 4140–4150 (2009).
10. H. X. Li, Y. Y. Yao, X. Z. Zhou, and B. Huang, "A two-phase model for learning rules from incomplete data," *Fundamenta Informaticae*, **94**, 219–232 (2009).
11. S. Upadhyaya, A. Arora, and R. Jain, "Reduct driven pattern extraction from clusters," *International Journal of Computational Intelligence Systems*, **2**, 10–16 (2009).
12. Y. Y. Yao, "Decision-theoretic rough set models," *Proc. RSKT'2007, LNAI* **4481**, Springer, Berlin, 1–12(2007).
13. Y. Y. Yao and S. K. M. Wong, "A decision theoretic framework for approximating concepts," *International Journal of Man-machine Studies*, **37**, 793–809 (1992).
14. Y. Y. Yao, S. K. M. Wong, and P. Lingras, "A decision-theoretic rough set model," *Methodologies for Intelligent Systems*, **5**, North-Holland, New York, 17–24 (1990).

15. Y. Y. Yao, "Probabilistic approaches to rough sets," *Expert Systems*, **20**, 287–297 (2003).
16. W. Ziarko, "Variable precision rough set model," *Journal of Computer and System Sciences*, **46**, 39–59 (1993).
17. X. Z. Zhou and H. X. Li, "A multi-view decision model based on decision-theoretic rough set," *LNCS*, **5589**, Springer, Berlin, 650–657 (2009).
18. Y. Y. Yao and Y. Zhao, "Attribute reduction in decision-teoretic rough set models," *Information Sciences*, **178**, 3356–3373 (2008).
19. J. P. Herbert and J. T. Yao, "Game-theoretic risk analysis in decision-theoretic rough sets," *Proc. RSKT'2008, LNAI*, **5009**, Springer, Berlin, 132–139 (2008).
20. J. P. Herbert and J. T. Yao, "Rough set model selection for practical decision making," *Proc. FSKD'2007*, IEEE Press, 203–207 (2007).
21. Y. Li, C. Zhang, and J. R. Swanb, "Rough set based model in information retrieval and filtering," *Proc. ISAS'1999*, 398–403 (1999).
22. S. Tsumoto, "Accuracy and coverage in rough set rule induction," *Proc. RSCTC'2002, LNAI*, **2475**, Springer, Berlin, 373–380 (2002).
23. J. T. Yao and J. P. Herbert, "Web-based support systems with rough set analysis," *Proc. RSEISP'2007, LNAI* **4585**, Springer, Berlin, 360–370 (2007).