Attribute Reduction in Stochastic Information Systems Based on α-Dominance

Emel KIZILKAYA AYDOĞAN, Mihrimah ÖZMEN

Department of industrial engineering, Erciyes University, TURKEY
emelkizilkaya@gmail.com, mihrimah@erciyes.edu.tr
(Geliş/Received:03.11.2016; Kabul/Accepted:22.01.2017)
DOI: 10.17671/gazibtd.309305

Abstract— Rough set has been commonly taken part in literature to examine inadequate and incomplete information systems. The efficiency of rough set with stochastic data observed for developing convenience and scalability. In this study, we use a ranking approach for attribute reduction in stochastic information systems and generalized this via presenting a dominance relation. We obtained the rough set approach of attribute reduction in stochastic information systems by establishing the dominance degrees. Furthermore, attribute reduction methods are studied by considering discernibility matrix and this approach is applied to explanatory examples to demonstrate its validity. Also this research proposes many research fields and new application areas show a tendency to concerning rough set approach to stochastic information systems.

Keywords- *Rough Sets, Stochastic Information Systems, Attribute Reduction, Dominance Relation*

1. INTRODUCTION

As a soft computing technique, rough set theory is proposed by Pawlak and it is an important theory to deal with insufficient information [1,2]. The theory is an expanded of the classical set theory for modelling uncertainty or imprecision information. It has conceived as a powerful mathematical tool for knowledge discovery. Particularly it can be used in evaluation the significance of attributes and derive decision rules and also to describe the dependencies among attributes [2,3]. It is useful in many fields such as machine learning, data mining and pattern recognition [4-8]. The information systems, the basis of rough set based data analysis; contain data about objects of interest, characterized by a finite set of attributes [9,10].

The classical rough set theory does not examine the problem of ordering objects. However, note that indiscernibility-based rough set model is powerless in many real situations that are the ordering of objects attributes with preference-ordered domains. To solve this problem, Greco, Matarazzo, and Slowinski have given to the literature, the dominance-based rough set approach (DRSA) which is an extension of rough set theory [11-14]. DRSA considers the ordering characteristics of attributes. Many researches have been studied about DRSA, lately [15-17].

Stochastic data is to define objects with uncertain judgements. It is also influenced by varied kinds of errors such as measurement, computation errors etc. Also stochastic information systems ensures us a powerful tool for data analyses and they are found almost everywhere in real world applications. Researchers from all over the world interested in stochastic data from among so many complex data. So the stochastic extension of DRSA was generalized as the notion of lower approximations to the stochastic case by estimating the class intervals for each object [18].

Nowadays, researchers are proposed approaches using stochastic dominance (SD) rules [19-28]. Some of these methods are to rank alternatives by using the determined SD relations based on rough set theory [26-27]. Moreover, the approach depends on stochastic dominance degree (SDD) has been also proposed [28].

The aim of this paper is to future researches rough set in stochastic information systems based on dominance relation. Nevertheless, to solve applicable problem, just only one knowledge granulation may not be fine enough. For this reason the issues of multi granulation have been absolutely discussed in many rough set studies [29-34]. From this aspect, we will propose a new dominance approach into stochastic information systems. This mechanism is to apply a parameter α, which checks the size of dominance-based information granulation.

The remaining of this study is consisted of as follows. Dominance relation in stochastic information systems are shortly reminded in Section 2. In Section 3, attribute reduction in stochastic information systems based on discernibility matrix is introduced and some important properties are discussed. In Section4, illustrative examples for the two more cases are investigated to show the possible implementation of the proposed approach. And lastly, Section 5 summarizes and emphasizes the main properties of this study.

2. α-DOMINANCE RELATION IN STOCHASTIC INFORMATION SYSTEMS

In this section we present a dominance relation to a stochastic information system by using SDD. Stochastic dominance is a comparing technique for uncertain alternatives in decision making [36]. And also SDD was presented as an approach to solve the stochastic multiple criteria decision making problem, where results of alternatives according to criteria are demonstrated by random variables with probability distributions by Liu et. al [35].

In decision making, two kinds of problems are examined: to obtain ranking based on information aggregation, and to obtain dominance rules with relations. In this study, we only work on to rank all objects with the dominance relation in a stochastic information system. In their study, Zhang and Qiu describe a dominance degree notion to rank all objects in classical ordered information systems [37]. Hereinafter, we define a dominance degree between two objects and a whole dominance degree of an object.

A stochastic information system (SIS) is a quadruple $S=(U,AT,V,D)$ where U is a finite non-empty set of objects and AT is a finite non-empty set of attributes, $V=\cup_{a\in AT} V_a$ and V_a is a domain of attribute a, $D: f_1(x)\times f_2(x)\times AT\rightarrow V$ is a total function such that $f_1(x)\ and\ f_2(x)$ are probability distributions of two objects x_1 and x_2 respectively, $D(f_1(x), f_2(x), a)\in V_a$ for every $a\in AT, x_1\ and\ x_2\in U$ called an information function where V_a is a set of stochastic data.

In comparing two probability distributions, there are two kinds of random variables, i.e. continuous (probability density functions) and discrete (probability mass functions) random variables. So, two probability distributions are compared with three possible states. i.e., (1) two continuous probability distributions, (2) two discrete probability distributions and (3) a continuous probability distribution and a discrete probability distribution [35]. Dominance degree between two objects with respect to the dominance relation is defined as:

Definition 1 (Stochastic Dominance Relation Between Two Continuous Probability Distribution)

Let $X_1\ and\ X_2$ be two independent continuous random variables with probability distributions $f_1(x)\ and\ f_2(x)$, respectively, where $\int_{-\infty}^{+\infty} f_1(x)dx=1\ and\ \int_{-\infty}^{+\infty} f_2(x)dx=1$. Then the dominance degree of $f_1(x)\ over\ f_2(x)$ (noted as $D_{f_1\succ f_2}$) is given by

$$D_{f_1\succ f_2}=\int_{-\infty}^{+\infty}\int_{-\infty}^{x_1} f_1(x_1)\, f_2(x_2)dx_2dx_1, \quad (1)$$

and accordingly, the dominance degree of $f_2(x)$ over $f_1(x)$ (noted as $D_{f_2\succ f_1}$) is given by

$$where\ D_{f_1\succ f_2}+D_{f_2\succ f_1}=1,$$

$$D_{f_2\succ f_1}=\int_{-\infty}^{+\infty}\int_{x_1}^{+\infty} f_1(x_1)\, f_2(x_2)dx_2dx_1. \quad (2)$$

Definition 2 (Stochastic Dominance Relation Between Two Discrete Probability Distribution)

Let $Y_1\ and\ Y_2$ be two independent discrete random variables with probability distributions $g_1(y)\ and\ g_2(y)$, respectively, where $\sum_{y=-\infty}^{+\infty} g_1(y)=1\ and\ \sum_{y=-\infty}^{+\infty} g_2(y)=1$. Then the dominance degree of $g_1(y)\ over\ g_2(y)$ (noted as $D_{g_1\succ g_2}$) is given by

$$D_{g_1\succ g_2}=\sum_{y_1=-\infty}^{+\infty}\sum_{y_2=-\infty}^{y_1} g_1(y_1)g_2(y_2)-0.5\sum_{y_1=-\infty}^{+\infty} g_1(y_1)g_2(y_1), \quad (3)$$

and accordingly, the dominance degree of $g_2(y)$ over $g_1(y)$ (noted as $D_{g_2\succ g_1}$) is given by

$$where\ D_{g_1\succ g_2}+D_{g_2\succ g_1}=1,$$

$$D_{g_2\succ g_1}=\sum_{y_1=-\infty}^{+\infty}\sum_{y_2=y_1}^{+\infty} g_1(y_1)g_2(y_2)-0.5\sum_{y_1=-\infty}^{+\infty} g_1(y_1)g_2(y_1). \quad (4)$$

Definition 3 (Stochastic Dominance Relation Between Continuous Probability Distribution and Discrete Probability Distribution)

Let X be an independent continuous random variable with probability distribution $f(x)$, and Y be an independent discrete random variable with probability distribution $g(y)$, where $\int_{-\infty}^{+\infty} f(x)dx=1\ and \sum_{y=-\infty}^{+\infty} g(y)=1$. Then the dominance degree of $f(x)$ over $g(y)$ (noted as $D_{f\succ g}$) is given by

$$D_{f\succ g}=\sum_{y=-\infty}^{+\infty}\left[g(y)\int_{y}^{+\infty} f(x)dx\right], \quad (5)$$

and accordingly, the dominance degree of $g(y)$ over $f(x)$ (noted as $D_{g\succ f}$) is given by

$$where\ D_{f\succ g}+D_{g\succ f}=1,$$

$$D_{g\succ f}=\sum_{y=-\infty}^{+\infty}\left[g(y)\int_{-\infty}^{y} f(x)dx\right]. \quad (6)$$

Definition 4 (Stochastic Dominance Relation Between Two Normal Probability Distributions)

The normal probability distribution is used extensively in that it can well model the additive effect of many independent factors [38-42]. In the following, a case that continuous probability distributions are normal ones was analyzed [35].

Let $Z_1\ and\ Z_2$ be two independent normal random variables with probability distributions $h_1(z)=\frac{1}{\sigma_1\sqrt{2\pi}}e^{-(z-\mu_1)^2/2\sigma_1^2}$ and $h_2(z)=$

$\frac{1}{\sigma_2\sqrt{2\pi}}e^{-(z-\mu_2)^2/2\sigma_2^2}$ respectively, where μ_1 *and* μ_2 are means, and σ_1 *and* σ_2 are standard deviations. Then the dominance degree of $h_1(z)$ over $h_2(z)$ (noted as $D_{h_1 \succ h_2}$) is given by

$$D_{h_1 \succ h_2} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha} e^{-t^2/2}dt, \quad (7)$$

where $\alpha = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$

Remark 1. Consider arbitrary two objects A_1 *and* A_2 in decision analysis. Let $f_1(x)$ *and* $f_2(x)$ be probability distributions on consequences of A_1 *and* A_2, respectively. $D_{f_1 \succ f_2}$ can be also regarded as the dominance degree of A_1 *over* A_2 (noted as $D_{A_1 \succ A_2}$), i.e., $D_{f_1 \succ f_2} \Leftrightarrow D_{A_1 \succ A_2}$. The greater $D_{f_1 \succ f_2}$ is, the greater the dominance degree of object A_1 *over* A_2 will be.

By Equations (1), (2), (3), (4), (5), (6) and (7), the dominance degree matrix D_j of object pairwise comparisons with respect to criterion a_j can be constructed, i.e.,

$$D_j = [D_{ikj}]_{m\times m} = \begin{bmatrix} D_{11j} & D_{12j} & \cdots & D_{1mj} \\ D_{21j} & D_{22j} & \cdots & D_{2mj} \\ \vdots & \vdots & \ddots & \vdots \\ D_{m1j} & D_{m2j} & \dots & D_{mmj} \end{bmatrix},$$

$j = 1,2,\dots,n,$

where D_{ikj} denotes the dominance degree of x_i over x_k with respect to attribute a_j, and $D_{ikj} + D_{kij} = 1$.

First, overall dominance degree matrix D of object pairwise comparisons is constructed, i.e.,

$$D = [D_{ik}]_{m\times m} = \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1m} \\ D_{21} & D_{22} & \cdots & D_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ D_{m1} & D_{m2} & \dots & D_{mm} \end{bmatrix},$$

where D_{ik} denotes the overall dominance degree of x_i over x_k and is calculated by

$$D_{ik} = \frac{\sum_{j=1}^{n} D_{ikj}}{|AT|}, \; i,k = 1,2,\dots,m. \quad (8)$$

From Eq. (8), it can be seen that $D_{ik} + D_{ki} = 1$ and $0 \leq D_{ik} \leq 1$.

Let $(x_i, x_k) \in U \times U$, thus a dominance relation matrix with respect to AT can be constructed with the dominance relation. Also the whole dominance degree of each object can be obtained from this matrix with this formula

$$D(i) = \frac{1}{|U|-1}\sum_{i\neq k} D_{ik}, \quad x_i, x_k \in U. \quad (9)$$

According to these two concepts, a ranking approach for all objects can also be generated. Through the number of D_i, whole dominance degree of each object on the universe, all objects can be rank. Therefore larger value indicates that an object better. This concept can be understood with the next example.

In the following, we will give an example for the comparing of two discrete probability distributions.

Example 1.1

This example is adapted from Zaras and Martel (1994) and Liu et al. (2011) [24,35]. A stochastic information system is given in Table 1, where $U = \{x_1, x_2, \dots, x_{10}\}, AT = \{a_1, a_2, a_3, a_4\}$ and there are 7 decision makers.

Firstly the types of stochastic dominance relations are defined for each pair of objects according to each attribute. The larger values of attributes are preferable one than the smaller ones.

The evaluations of objects with respect to attributes are defined in the form of probability distributions (Table 1). For example, three experts in the seven give their considerations on project x_1 with respect to criterion a_1 using score 2, then the 'probability' that the consideration on project x_1 is score 2 is regarded as 3/7 (see Table 1). Firstly in Table 1, the probability distribution of random variable on each object according to each attribute is obtained. Then, the dominance degree of one object over another according to each attribute can be determined using Equation (3) and (4). So, four dominance degree matrices D_1, D_2, D_3 *and* D_4 are respectively established, i.e.,

Table 1. The evaluations provided by experts

Attribute	Values	Objects									
		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
a_1	1	0	0	0	1/7	0	1/7	1/7	1/7	0	0
	2	3/7	1/7	0	0	0	0	0	2/7	0	1/7
	3	1/7	0	0	0	1/7	0	0	2/7	0	2/7
	4	0	2/7	0	0	0	0	0	1/7	0	2/7
	5	2/7	1/7	3/7	1/7	0	0	3/7	1/7	2/7	1/7
	6	0	2/7	1/7	0	2/7	0	1/7	0	1/7	0
	7	1/7	0	1/7	0	2/7	1/7	0	0	3/7	1/7
	8	0	1/7	2/7	1/7	0	3/7	1/7	0	1/7	0
	9	0	0	0	4/7	2/7	0	0	0	0	0
	10	0	0	0	0	0	2/7	1/7	0	0	0
a_2	1	0	1/7	1/7	0	0	0	1/7	3/7	0	0
	2	2/7	0	0	0	0	0	3/7	3/7	0	1/7
	3	1/7	0	0	1/7	0	4/7	1/7	0	1/7	0
	4	0	0	0	1/7	0	0	0	1/7	1/7	0
	5	2/7	0	0	0	1/7	0	1/7	0	0	0
	6	0	1/7	1/7	1/7	2/7	0	1/7	0	1/7	0
	7	0	1/7	0	0	1/7	1/7	0	0	4/7	2/7
	8	1/7	1/7	2/7	3/7	2/7	2/7	0	0	0	3/7
	9	1/7	3/7	1/7	1/7	1/7	0	0	0	0	0
	10	0	0	2/7	0	0	0	0	0	0	1/7
a_3	1	0	0	1/7	0	1/7	0	0	2/7	0	1/7
	2	0	0	0	0	0	0	3/7	1/7	0	2/7
	3	1/7	0	0	1/7	0	0	1/7	4/7	1/7	0
	4	3/7	0	0	0	0	1/7	1/7	0	2/7	0
	5	0	1/7	0	0	0	1/7	2/7	0	2/7	0
	6	1/7	0	0	0	0	0	0	0	0	2/7
	7	0	1/7	0	1/7	0	0	0	0	2/7	2/7
	8	1/7	2/7	0	2/7	3/7	2/7	0	0	0	0
	9	1/7	3/7	2/7	1/7	1/7	1/7	0	0	0	0

	10	0	0	4/7	2/7	2/7	2/7	0	0	0	0
a4	1	0	1/7	0	1/7	0	0	0	2/7	0	0
	2	0	0	0	0	0	0	0	0	1/7	0
	3	3/7	0	0	0	0	0	1/7	0	0	0
	4	0	0	0	0	0	0	0	1/7	1/7	0
	5	2/7	0	0	0	0	1/7	1/7	2/7	0	0
	6	0	0	0	0	1/7	1/7	0	1/7	3/7	3/7
	7	0	0	1/7	0	1/7	1/7	0	0	0	1/7
	8	1/7	2/7	4/7	0	3/7	2/7	3/7	1/7	1/7	1/7
	9	0	2/7	0	1/7	1/7	1/7	1/7	0	0	1/7
	10	1/7	2/7	2/7	5/7	1/7	1/7	1/7	0	1/7	1/7

The dominance degree of one object over another with respect to each attribute is obtained according to Liu et. al [35]. Thus four dominance degree matrices D_1, D_2, D_3 *and* D_4 are constructed.

$$D_1=\begin{bmatrix} 0.5 & 0.3163 & 0.1531 & 0.1837 & 0.1327 & 0.1531 & 0.2857 & 0.6122 & 0.1327 & 0.4286 \\ 0.6837 & 0.5 & 0.2959 & 0.2245 & 0,2449 & 0,1939 & 0,4082 & 0,8265 & 0,2551 & 0,6735 \\ 0.8469 & 0,7041 & 0,5 & 0,2755 & 0,3878 & 0,2551 & 0,5714 & 0,9694 & 0,4694 & 0,8776 \\ 0,8163 & 0,7755 & 0,7245 & 0,5 & 0,6122 & 0,5102 & 0,6633 & 0.8571 & 0.7245 & 0.8265 \\ 0,8673 & 0,7551 & 0,6122 & 0,3878 & 0,5 & 0,3265 & 0,6531 & 0,9388 & 0,5714 & 0,8367 \\ 0,8469 & 0,8061 & 0,7449 & 0,4898 & 0,6735 & 0,5 & 0,7143 & 0,8673 & 0,7755 & 0,8469 \\ 0,7143 & 0,5918 & 0,4286 & 0,3367 & 0,3469 & 0,2857 & 0.5 & 0.8367 & 0.3878 & 0.7449 \\ 0,3878 & 0,1735 & 0,0306 & 0,1429 & 0,0612 & 0,1327 & 0,1633 & 0,5 & 0,0204 & 0,2959 \\ 0,8673 & 0,7449 & 0,5306 & 0,2755 & 0,4286 & 0,2245 & 0,6122 & 0,9796 & 0,5 & 0,8878 \\ 0,5714 & 0,3265 & 0,1224 & 0,1735 & 0,1633 & 0,1531 & 0,2551 & 0,7041 & 0,1122 & 0,5 \end{bmatrix}$$

$$D_2=\begin{bmatrix} 0,5 & 0,2857 & 0,2551 & 0,3163 & 0,2551 & 0,4694 & 0,7041 & 0,8776 & 0,3776 & 0,2959 \\ 0,7143 & 0,5 & 0,4184 & 0,6020 & 0,6939 & 0,7245 & 0,8571 & 0,8878 & 0,7245 & 0,5204 \\ 0,7449 & 0,5816 & 0,5 & 0,6530 & 0,6020 & 0,7551 & 0,8571 & 0,8878 & 0,7653 & 0,5918 \\ 0,6837 & 0,3980 & 0,3470 & 0,5 & 0,4796 & 0,7143 & 0,8980 & 0,9694 & 0,6633 & 0,4592 \\ 0,7449 & 0,3061 & 0,3980 & 0,5204 & 0,5 & 0,7245 & 0,9490 & 1 & 0,6735 & 0,4082 \\ 0,5306 & 0,2755 & 0,2449 & 0,2857 & 0,2755 & 0,5 & 0,7959 & 0,9184 & 0,4286 & 0,3061 \\ 0,2959 & 0,1429 & 0,1429 & 0,1020 & 0,0510 & 0,2041 & 0,5 & 0,7143 & 0,1020 & 0,0918 \\ 0,1224 & 0,1122 & 0,1122 & 0,0306 & 0 & 0,0816 & 0,2857 & 0,5 & 0,0306 & 0,0510 \\ 0,6224 & 0,2755 & 0,2347 & 0,3367 & 0,3265 & 0,5714 & 0,8980 & 0,9694 & 0,5 & 0,2245 \\ 0,7041 & 0,4796 & 0,4082 & 0,5408 & 0,5918 & 0,6939 & 0,9082 & 0,9490 & 0,7755 & 0,5 \end{bmatrix}$$

$$D_3=\begin{bmatrix} 0,5 & 0,1939 & 0,1633 & 0,2449 & 0,2449 & 0,2245 & 0,7755 & 0,9592 & 0,5204 & 0,6122 \\ 0,8061 & 0,5 & 0,2041 & 0,4490 & 0,4184 & 0,4694 & 0,9796 & 1 & 0,9184 & 0,8980 \\ 0,8367 & 0,7959 & 0,5 & 0,6735 & 0,6837 & 0,6735 & 0,8571 & 0,8776 & 0,8571 & 0,8673 \\ 0,7551 & 0,5510 & 0,3265 & 0,5 & 0,4796 & 0,5 & 0,9286 & 0,9592 & 0,8469 & 0,8980 \\ 0,7551 & 0,5816 & 0,3163 & 0,5204 & 0,5 & 0,5204 & 0,8571 & 0,8776 & 0,8571 & 0,8673 \\ 0,7755 & 0,5306 & 0,3265 & 0,5 & 0,4796 & 0,5 & 0,9286 & 1 & 0,8367 & 0,8367 \\ 0,2245 & 0,0204 & 0,1429 & 0,0714 & 0,1429 & 0,0714 & 0,5 & 0,6837 & 0,2143 & 0,3673 \\ 0,0408 & 0 & 0,1224 & 0,0408 & 0,1224 & 0 & 0,3163 & 0,5 & 0,0408 & 0,3061 \\ 0,4796 & 0,0816 & 0,1429 & 0,1531 & 0,1429 & 0,1633 & 0,7857 & 0,9592 & 0,5 & 0,5510 \\ 0,3878 & 0,1020 & 0,1327 & 0,1020 & 0,1327 & 0,1633 & 0,6327 & 0,6939 & 0,4490 & 0,5 \end{bmatrix}$$

$$D_4=\begin{bmatrix} 0,5 & 0,2653 & 0,1837 & 0,2143 & 0,2041 & 0,2347 & 0,2959 & 0,5612 & 0,3878 & 0,2245 \\ 0,7347 & 0,5 & 0,5714 & 0,2959 & 0,6327 & 0,6531 & 0,6327 & 0,8571 & 0,7347 & 0,6735 \\ 0,8163 & 0,4286 & 0,5 & 0,2857 & 0,5816 & 0,6429 & 0,5918 & 0,9388 & 0,8163 & 0,7041 \\ 0,7857 & 0,7041 & 0,7143 & 0,5 & 0,7755 & 0,7755 & 0,7755 & 0,8776 & 0,7857 & 0,7755 \\ 0,7959 & 0,3673 & 0,4184 & 0,2245 & 0,5 & 0,5714 & 0,5408 & 0,9184 & 0,7653 & 0,6224 \\ 0,7653 & 0,3469 & 0,3571 & 0,2245 & 0,4286 & 0,5 & 0,5 & 0,8673 & 0,6939 & 0,5306 \\ 0,7041 & 0,3673 & 0,4082 & 0,2245 & 0,4592 & 0,5 & 0,5 & 0,8061 & 0,6531 & 0,5204 \\ 0,4388 & 0,1429 & 0,0612 & 0,1224 & 0,0816 & 0,1327 & 0,1939 & 0,5 & 0,2959 & 0,1224 \\ 0,6122 & 0,2653 & 0,1837 & 0,2143 & 0,2347 & 0,3061 & 0,3469 & 0,7041 & 0,5 & 0,3163 \\ 0,7755 & 0,3265 & 0,2959 & 0,2245 & 0,3776 & 0,4694 & 0,4796 & 0,8776 & 0,6937 & 0,5 \end{bmatrix}$$

From the definition of dominance degree, the dominance relation matrix can be obtained as

The overall dominance degree D matrix of object is obtained.

$$D=\begin{bmatrix} 0.5 & 0.2653 & 0.1888 & 0.2398 & 0.2092 & 0.2704 & 0.5153 & 0.7526 & 0.3546 & 0.3903 \\ 0.7347 & 0.5 & 0.3725 & 0.3929 & 0.4975 & 0.5102 & 0.7194 & 0.8929 & 0.6582 & 0.6914 \\ 0.8112 & 0.6276 & 0.5 & 0.4719 & 0.5638 & 0.5817 & 0.7194 & 0.9184 & 0.770 & 0.7602 \\ 0.7602 & 0.6072 & 0.5281 & 0.5 & 0.5867 & 0.6250 & 0.8164 & 0.9158 & 0.7551 & 0.7398 \\ 0.7908 & 0.5025 & 0.4362 & 0.4133 & 0.5 & 0.5357 & 0.75 & 0.9337 & 0.7168 & 0.6837 \\ 0.7296 & 0.4898 & 0.4184 & 0.3750 & 0.4643 & 0.5 & 0.7347 & 0.9133 & 0.6837 & 0.6301 \\ 0.4847 & 0.2806 & 0.2807 & 0.1837 & 0.25 & 0.2653 & 0.5 & 0.7602 & 0.3393 & 0.4311 \\ 0.2475 & 0.1072 & 0.0816 & 0.0842 & 0.0663 & 0.0868 & 0.2398 & 0.5 & 0.0969 & 0.1939 \\ 0.6454 & 0.3418 & 0.2730 & 0.2449 & 0.2832 & 0.3163 & 0.6607 & 0.9031 & 0.5 & 0.4949 \\ 0.6097 & 0.3087 & 0.2398 & 0.2602 & 0.3164 & 0.3699 & 0.5689 & 0.8062 & 0.5051 & 0.5 \end{bmatrix}$$

$D(1)$=0.65 $D(2)$=0.4, $D(3)$=0.32, $D(4)$=0.3, $D(5)$=0.36, $D(6)$=0.4, $D(7)$=0.64, $D(8)$=0.87, $D(9)$=0.54, $D(10)$=0.56.

Afterwards, all objects ranks are constructed with respect to the number of $D(i)$. An object with larger number shows a better object.

$$X_8 \succcurlyeq X_1 \succcurlyeq X_7 \succcurlyeq X_{10} \succcurlyeq X_9 \succcurlyeq \begin{pmatrix} X_2 \\ X_6 \end{pmatrix} \succcurlyeq X_5 \succcurlyeq X_3 \succcurlyeq X_4.$$

Qian et. al. proposed a ranking approach based on dominance classes and establish a dominance based rough set approach to interval information systems [43]. Yang et. al. presented lower and upper approximate reducts into α-dominance based rough set approach to interval-valued decision systems [44]. The information granulation is the most significant and initial stage, to rough set data analysis. In this part, the stochastic dominance degree which is a parameter that indicates how a stochastic data dominates another one will be taken in stochastic information systems,

Definition 5 (Stochastic α-Dominance Relation)

Let $S=(U,AT,V,D)$ be a stochastic information system, $\forall a \in AT$, a α-dominance relation in terms of a is defined as:

$$DOM_a^{\alpha}(x_i,x_k)=\begin{cases} 1: & i=k \\ 1: & \{(x_i,x_k)\epsilon U^2: D_{ik} \geq \alpha\}, \\ 0: & otherwise \end{cases} \quad (10)$$

where D_{ik} denotes the dominance degree of x_i over x_k with respect to attribute a,

$$DOM_a^{\alpha}=\{(x_i,x_k)\epsilon U^2: D_{ik} \geq \alpha\}, \quad (11)$$

$\forall A \subseteq AT$, a α-dominance relation in terms of A is defined as:

$$DOM_A^{\alpha}=\{(x_i,x_k) \in U^2: \forall a \in A, (x_i,x_k) \in DOM_a^{\alpha}\}, \quad (12)$$

where $\alpha \in [0,1]$ is a given threshold of dominance degree.

$DOM_a^{\alpha}(x_i,x_k)$ is a crisp binary relation and α can get any value in interval [0, 1]. This is mainly because the value of $DOM_a^{\alpha}(x_i,x_k)$ is between 0 and 1. For a given α, if $(x_i,x_k) \in DOM_a^{\alpha}(x_i,x_k)$, then x_i is considered as dominating x_k on attributes a with the at least degree of α. In this paper, we denoted it by

$$f(x_i,a) \succeq_{\alpha} f(x_k,a) \ or \ f(x_k,a) \preceq_{\alpha} f(x_i,a) \quad (13)$$

Example 1.2 (Continued from Example 1.1) If we consider threshold of dominance degree 0.6 and the stochastic information system dedicated in Table 1, then a Boolean matrix $DOM_{AT}^{0.6}$ can be established:

$$DOM_{AT}^{0.6}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Take for instance x_2, under the given dominance degree $\alpha = 0.6$, objects x_1, x_2, x_7, x_8, x_9 *and* x_{10} are regarded as dominated by x_2 with respect to set of attributes AT.

3. ATTRIBUTE REDUCTION IN STOCHASTIC INFORMATION SYSTEMS BASED ON DISCERNIBILITY MATRIX

Searching significant subsets of attributes which ensure the identical or more meaning for the whole set of attributes, is one of the main research subject. These subsets are named reducts. Nowadays, many attribute reduction approaches have been studied in literature for different dominance relations. For instance, Shao and Zhang proposed an expansion of reduct dominance relation for incomplete information systems [45]. Yang et al. proposed a similarity dominance relation to form all reducts [46].

In this this part, a framework of attribute reduction with respect to α-dominance relation in stochastic information systems is established and an example is employed to demonstrate its effectiveness.

Definition 6. Let $S = (U, AT, V, D)$ be a stochastic information system, $\forall A \subseteq AT$ and $\alpha \in [0,1]$,

1. if $DOM_A^{\alpha} = DOM_{AT}^{\alpha}$, then A is referred to as a α-dominance consistent set in S;

2. if A is a α-dominance consistent set in S and $A - \{a\}$ is not a α-dominance consistent set for each $a \in A$, i.e.,

$DOM_{A-\{a\}}^{\alpha} \neq DOM_{AT}^{\alpha}$, then A is referred to as a α-dominance reduct in S;

3. the set of all α-dominance reducts in S is denoted by $Red^{\alpha}(S)$;

4. the intersection of all α-dominance reducts in S is referred to as the α-dominance core.

A α-dominance consistent set in S is a subset of full attributes which preserves α-dominance relation DOM_{AT}^{α}, while α-dominance reduct in S is a *minimal* α-dominance consistent set which preserves α-dominance relation DOM_{AT}^{α}.

In the following, the discernibility matrix based approach to compute all α-dominance reducts in S is presented.

$\forall x_i, x_k \in U$, let us denote

$$D_{AT}^{\alpha}(x_i, x_k) = \begin{cases} \{a \in AT: (x_i, x_k) \notin DOM_a^{\alpha}\} & (x_i, x_k) \notin DOM_{AT}^{\alpha} \\ \emptyset & other\ wise \end{cases}$$

$D_{AT}^{\alpha}(x_i, x_k)$ is showed to as the α-discernibility set for objects pair (x_i, x_k) for the given dominance degree α. Describe, furthermore, the α-dominance matrix such that

$$M^{\alpha} = \{D_{AT}^{\alpha}(x_i, x_k): \forall x_i, x_k \in U\}, \quad (14)$$

and let

$$M_0^{\alpha} = \{D_{AT}^{\alpha}(x_i, x_k): \forall D_{AT}^{\alpha}(x_i, x_k) \in M^{\alpha}, D_{AT}^{\alpha}(x_i, x_k) \neq \emptyset\}, \quad (15)$$

then to obtain a α-dominance consistent set according to α-dominance matrix, the following theorem can be used

Theorem 1. Let S be a stochastic information system, $\forall A \subseteq AT$, we have

$$DOM_A^{\alpha} = DOM_{AT}^{\alpha} \Leftrightarrow A \cap D_{AT}^{\alpha}(x_i, x_k) \neq \emptyset, \forall D_{AT}^{\alpha}(x_i, x_k) \in M_0^{\alpha}.$$

Theorem 1 ensures that adequate and essential conditions to evaluate whether a subset of attributes is a α-dominance consistent set in S.

Definition 7. Let S be a stochastic information system, let us define

$$\Delta = \bigwedge_{D_{AT}^{\alpha}(x_i, x_k) \in M_0^{\alpha}} \bigvee D_{AT}^{\alpha}(x_i, x_k), \quad (16)$$

Δ is referred to as the α-dominance discernibility function in S.

Theorem 2. Let S be a stochastic information system, $\forall A \subseteq AT$, A is a α-dominance reduct in S if and only if $\bigvee A$ is a prime implicant of Δ.

According to those described above, we can see that it ensures a practical framework for attribute reduction, which can be expressed with the following example.

Table 2 0.6-discernibility set for pairs of objects given in Table 1.

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1		$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	a_1a_4	a_4	$a_1a_2a_3a_4$	$a_1a_2a_4$
x_2			$a_1a_2a_3a_4$	$a_1a_3a_4$	a_1a_3	a_1a_3	a_1		a_1	a_2
x_3		a_2a_4		a_1a_4	a_1a_4	a_1	a_1a_4		a_1	a_2
x_4		a_2a_3	a_2a_3		a_2a_3	a_1a_3				a_2
x_5		$a_2a_3a_4$	$a_2a_3a_4$	$a_1a_2a_3a_4$		$a_1a_3a_4$	a_4		a_1	a_2
x_6	a_2	$a_2a_3a_4$	$a_2a_3a_4$	$a_1a_2a_3a_4$	$a_2a_3a_4$		a_4		a_2	a_2a_4
x_7	a_2a_3	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$			$a_1a_2a_3$	$a_2a_3a_4$
x_8	$a_1a_2a_3a_4$	$a_1a_2a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_3a_4$	$a_1a_2a_4$	$a_1a_2a_3a_4$		$a_1a_2a_3a_4$	$a_1a_2a_3a_4$
x_9	a_3a_4	$a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	a_4			$a_2a_3a_4$
x_{10}	a_1a_3	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_3a_4$	a_1a_4		a_1a_3	

Example 1.3 (Continued from Example 1.2) For stochastic information system given in Table 1, following Example 1.2, the 0.6-discernibility sets for all pairs of objects are shown in Table 2. By the definition of discernibility matrix, one can

obtain the discernibility sets for all pairs of objects are shown in Table 2.

By Theorems 1 and 2, it is not difficult to obtain that $Red^{0.6}(S) = \{\{a_1, a_2, a_4\}\}$, i.e., $\{a_1, a_2, a_4\}$ is the only one 0.6-dominance reduct in Table 1.

4. ILLUSTRATIVE EXAMPLES

In this part, illustrative examples for the three cases are evaluated to express the potential implementation of the proposed approach.

Example 2

This is a modified example of selecting the most desirable strategy for an electricity retailer, which considered [35, 41]. In the example, there are nine objects $(A_1, A_2, \dots, A_9)$ and four attribute: long-term profit (C_1), shortterm profit (C_2), market share (C_3) and green market share (C_4). Assume that the result of object A_i with respect to criterion C_j is a random variable with normal probability distribution $f_{ij}(x) = \frac{1}{\sigma_{ij}\sqrt{2\pi}} e^{-(x-\mu_{ij})^2/2\sigma_{ij}^2}$, which is given in Table 3. To determine the attribute reduct(s), computation processes and results using the proposed method are summarized as follows

Table 3 Means and standard deviations for consequences of objects with respect to attribute.

Objects	(μ_{ij}, σ_{ij})			
	C_1	C_2	C_3	C_4
A_1	(439,143)	(163,36)	(12.1, 0.5)	(9.3, 6.5)
A_2	(426,125)	(159,31)	(12.1, 0.5)	(14.8, 6.5)
A_3	(264,135)	(104,32)	(13.1, 0.5)	(9.3,6.5)
A_4	(444,125)	(163,31)	(12.1, 0.5)	(9.3, 6.5)
A_5	(605,115)	(220,31)	(11.0, 0.5)	(9.3, 6.5)
A_6	(449,126)	(166,32)	(12.1, 0.5)	(4.3, 6.5)
A_7	(449,107)	(164,27)	(12.1, 0.5)	(9.3, 6.5)
A_8	(457,126)	(165,32)	(12.1, 0.5)	(9.3, 6.5)
A_9	(453,107)	(163,27)	(12.1, 0.5)	(14.8, 6.5)

Firstly, using Eq. (7), four dominance degree matrices D_1, D_2, D_3 *and* D_4 are built as follows:

$$D_1 = \begin{bmatrix} 0.5 & 0.5273 & 0.8132 & 0.4895 & 0.1828 & 0.4791 & 0.4777 & 0.4624 & 0.4688 \\ 0.4727 & 0.5 & 0.8107 & 0.4594 & 0.1460 & 0.4484 & 0.4444 & 0.4307 & 0.4348 \\ 0.1868 & 0.1893 & 0.5 & 0.1640 & 0.0273 & 0.1582 & 0.1414 & 0.1480 & 0.1363 \\ 0.5105 & 0.5406 & 0.8360 & 0.5 & 0.1716 & 0.4888 & 0.4879 & 0.4708 & 0.4782 \\ 0.8172 & 0.8540 & 0.9728 & 0.8284 & 0.5 & 0.8198 & 0.8397 & 0.8072 & 0.8334 \\ 0.5209 & 0.5516 & 0.8418 & 0.5112 & 0.1802 & 0.5 & 0.5 & 0.4821 & 0.4903 \\ 0.5223 & 0.5556 & 0.8586 & 0.5121 & 0.1603 & 0.5 & 0.5 & 0.4807 & 0.4895 \\ 0.5376 & 0.5693 & 0.8520 & 0.5292 & 0.1928 & 0.5179 & 0.5193 & 0.5 & 0.5097 \\ 0.5312 & 0.5652 & 0.8637 & 0.5218 & 0.1666 & 0.5097 & 0.5105 & 0.4903 & 0.5 \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 0.5 & 0.5336 & 0.8897 & 0.5 & 0.1151 & 0.4752 & 0.4911 & 0.4834 & 0.5 \\ 0.4665 & 0.5 & 0.8915 & 0.4637 & 0.0821 & 0.4376 & 0.4516 & 0.4464 & 0.4612 \\ 0.1103 & 0.1085 & 0.5 & 0.0927 & 0.0046 & 0.0853 & 0.0759 & 0.0888 & 0.0794 \\ 0.5 & 0.5363 & 0.9073 & 0.5 & 0.0968 & 0.4732 & 0.4903 & 0.4821 & 0.5 \\ 0.8849 & 0.9179 & 0.9954 & 0.9032 & 0.5 & 0.8873 & 0.9134 & 0.8915 & 0.9172 \\ 0.5248 & 0.5624 & 0.9147 & 0.5268 & 0.1128 & 0.5 & 0.5191 & 0.5088 & 0.5286 \\ 0.5089 & 0.5484 & 0.9241 & 0.5097 & 0.0866 & 0.4810 & 0.5 & 0.4905 & 0.5104 \\ 0.5166 & 0.5536 & 0.9112 & 0.5179 & 0.1085 & 0.4912 & 0.5095 & 0.5 & 0.5191 \\ 0.5 & 0.5388 & 0.9206 & 0.5 & 0.0828 & 0.4714 & 0.4896 & 0.4810 & 0.5 \end{bmatrix},$$

$$D_3 = \begin{bmatrix} 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.9214 & 0.9214 & 0.5000 & 0.9214 & 0.9985 & 0.9214 & 0.9214 & 0.9214 & 0.9214 \\ 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.0599 & 0.0599 & 0.0015 & 0.0599 & 0.5000 & 0.0599 & 0.0599 & 0.0599 & 0.0599 \\ 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.5000 & 0.0787 & 0.5000 & 0.9401 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \end{bmatrix},$$

$$D_4 = \begin{bmatrix} 0.5000 & 0.2748 & 0.5000 & 0.5000 & 0.5000 & 0.7068 & 0.5000 & 0.5000 & 0.2748 \\ 0.7252 & 0.5000 & 0.7252 & 0.7252 & 0.7252 & 0.8733 & 0.7252 & 0.7252 & 0.5000 \\ 0.5000 & 0.2748 & 0.5000 & 0.5000 & 0.5000 & 0.7068 & 0.5000 & 0.5000 & 0.2748 \\ 0.5000 & 0.2748 & 0.5000 & 0.5000 & 0.5000 & 0.7068 & 0.5000 & 0.5000 & 0.2748 \\ 0.5000 & 0.2748 & 0.5000 & 0.5000 & 0.5000 & 0.7068 & 0.5000 & 0.5000 & 0.2748 \\ 0.2932 & 0.1267 & 0.2932 & 0.2932 & 0.2932 & 0.5000 & 0.2932 & 0.2932 & 0.1267 \\ 0.5000 & 0.2748 & 0.5000 & 0.5000 & 0.5000 & 0.7068 & 0.5000 & 0.5000 & 0.2748 \\ 0.5000 & 0.2748 & 0.5000 & 0.5000 & 0.5000 & 0.7068 & 0.5000 & 0.5000 & 0.2748 \\ 0.7252 & 0.5000 & 0.7252 & 0.7252 & 0.7252 & 0.8733 & 0.7252 & 0.7252 & 0.5000 \end{bmatrix}.$$

Then, using Eq. (8), overall dominance degree matrix D can be obtained, i.e.,

$$D = \begin{bmatrix} 0.5000 & 0.4589 & 0.5704 & 0.4974 & 0.4345 & 0.5403 & 0.4922 & 0.4856 & 0.4359 \\ 0.5411 & 0.5000 & 0.6265 & 0.5371 & 0.4734 & 0.5648 & 0.5303 & 0.5256 & 0.4740 \\ 0.4296 & 0.3735 & 0.5000 & 0.4195 & 0.3826 & 0.4679 & 0.4097 & 0.4146 & 0.3530 \\ 0.5026 & 0.4629 & 0.5805 & 0.5000 & 0.4271 & 0.5422 & 0.4946 & 0.4882 & 0.4383 \\ 0.5655 & 0.5267 & 0.6174 & 0.5729 & 0.5000 & 0.6185 & 0.5783 & 0.5647 & 0.5213 \\ 0.4597 & 0.4352 & 0.5321 & 0.4478 & 0.3816 & 0.5000 & 0.4531 & 0.4460 & 0.4114 \\ 0.5078 & 0.4697 & 0.5904 & 0.5055 & 0.4218 & 0.5470 & 0.5000 & 0.4928 & 0.4437 \\ 0.5136 & 0.4744 & 0.5855 & 0.5118 & 0.4354 & 0.5540 & 0.5072 & 0.5000 & 0.4509 \\ 0.5641 & 0.5260 & 0.6471 & 0.5618 & 0.4787 & 0.5886 & 0.5563 & 0.5491 & 0.5000 \end{bmatrix}.$$

D_1=0.52, D_2=0.47, D_3=0.60, D_4=0.51, D_5=0.43, D_6=0.56, D_7=0.51, D_8=0.50, D_9=0.45.

In the following, we rank all objects according to the number of $D(i)$.

$$X_3 \succcurlyeq X_6 \succcurlyeq X_1 \succcurlyeq \binom{X_4}{X_7} \succcurlyeq X_8 \succcurlyeq X_2 \succcurlyeq X_9 \succcurlyeq X_5.$$

If we consider the stochastic information system given in Table 3 and assume that a threshold of dominance degree 0.6, then we can obtain a Boolean matrix, which is corresponding to $D_{AT}^{0.6}$ such that:

$$D_{AT}^{0.6} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Take for instance x_2, under the given dominance degree $\alpha = 0.6$, objects x_2 *and* x_3 are regarded as dominated by x_2 with respect to set of attributes AT. By the definition of discernibility matrix, one can obtain the discernibility sets for all pairs of objects are shown in Table 4.

Table 4 0.6-discernibility set for pairs of objects given in Table 3.

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1		$a_1a_2a_3a_4$	a_3a_4	$a_1a_2a_3a_4$	$a_1a_2a_4$	$a_1a_2a_3$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$
x_2	$a_1a_2a_3$		a_3	$a_1a_2a_3$	a_1a_2	$a_1a_2a_3$	$a_1a_2a_3$	$a_1a_2a_3$	$a_1a_2a_3a_4$
x_3	$a_1a_2a_3$	$a_1a_2a_4$		$a_1a_2a_4$	$a_1a_2a_4$	a_1a_2	$a_1a_2a_4$	$a_1a_2a_4$	$a_1a_2a_4$
x_4	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	a_3a_4		$a_1a_2a_4$	$a_1a_2a_3$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$
x_5	a_3a_4	a_3a_4	a_3a_4	a_3a_4		a_3	a_3a_4	a_3a_4	a_3a_4
x_6	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	a_3a_4	$a_1a_2a_3a_4$	$a_1a_2a_4$		$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$
x_7	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	a_3a_4	$a_2a_3a_4$	$a_1a_2a_4$	$a_1a_2a_3$		$a_1a_2a_3a_4$	$a_1a_2a_3a_4$
x_8	$a_1a_2a_3a_4$	$a_1a_2a_3a_4$	a_3a_4	$a_1a_2a_3a_4$	$a_1a_2a_4$	$a_1a_2a_3$	$a_1a_2a_3a_4$		$a_1a_2a_3a_4$
x_9	$a_1a_2a_3$	$a_1a_2a_3a_4$	a_3	a_2a_3	a_1a_2	$a_1a_2a_3$	$a_1a_2a_3$	$a_1a_2a_3$	

By Theorems 1 and 2, it is not difficult to obtain that $Red^{0.6}(S) = \{\{a_1, a_3\}, \{a_2, a_3\}\}$, i.e., $\{a_1, a_3\}$ and $\{a_2, a_3\}$ are the 0.6-dominance reduct and $\{a_3\}$is the core in Table 3.

Table 5. Discrete probability distributions and normal probability distributions for consequences of objects with respect to attribute.

Attribute	Objects	Scores										(μ_{ij}, σ_{ij})
		1	2	3	4	5	6	7	8	9	10	
C1	A1	0	0.1	0.1	0.3	0.2	0.2	0.1	0	0	0	
	A2	0	0	0.1	0.3	0.2	0.2	0.1	0.1	0	0	
	A3	0.1	0	0.2	0.2	0.3	0.1	0.1	0	0	0	
	A4	0	0	0	0.2	0.3	0.4	0.1	0	0	0	
C2	A1											(5.5,1.3)
	A2											(5,1.1)
	A3											(6,1.5)
	A4											(4.5,1)
C3	A1	0	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0	0	
	A2											(4.5,1.2)
	A3	0	0	0.2	0.3	0.2	0.2	0.1	0	0	0	
	A4											(4.6,1.1)
C4	A1											(6,1.2)
	A2	0	0	0	0.1	0.2	0.3	0.2	0.2	0	0	
	A3											(6.3,1)
	A4	0	0	0	0	0.1	0.2	0.3	0.3	0.1	0	

Example 3

This is a modified example of (Lui et al.) choosing the most desirable vendor(s). There are four alternatives $(A_1, A_2, A_3\ and\ A_4)$ and four attribute: responsiveness to customer needs (C_1), price (C_2), on-time delivery (C_3) and product quality (C_4). Considerations of the alternatives according to the attribute are shown in Table 5. To determine the attribute reduct(s), computation processes and results using the proposed method are summarized as follows

Firstly, using Equations (3), (4), (5), (6) and (7), four dominance degree matrices $D_1, D_2, D_3\ and\ D_4$ are built as follows:

$$D_1 = \begin{bmatrix} 0.5 & 0.405 & 0.545 & 0.335 \\ 0.595 & 0.5 & 0.635 & 0.435 \\ 0.455 & 0.365 & 0.5 & 0.29 \\ 0.665 & 0.565 & 0.71 & 0.5 \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 0.5 & 0.6155 & 0.4006 & 0.7290 \\ 0.3845 & 0.5 & 0.2954 & 0.6317 \\ 0.5994 & 0.7046 & 0.5 & 0.7973 \\ 0.2710 & 0.3683 & 0.2027 & 0.5 \end{bmatrix},$$

$$D_3 = \begin{bmatrix} 0.5 & 0.6125 & 0.575 & 0.5989 \\ 0.3875 & 0.5 & 0.4680 & 0.4755 \\ 0.425 & 0.5320 & 0.5 & 0.5090 \\ 0.4011 & 0.5245 & 0.4910 & 0.5 \end{bmatrix},$$

$$D_4 = \begin{bmatrix} 0.5 & 0.4548 & 0.42386 & 0.2554 \\ 0.5452 & 0.5 & 0.4778 & 0.31 \\ 0.5762 & 0.5222 & 0.5 & 0.3002 \\ 0.7446 & 0.69 & 0.6998 & 0.5 \end{bmatrix},$$

Then, using Eq. (8), overall dominance degree matrix D can be set up, i.e.

$$D = \begin{bmatrix} 0.5000 & 0.5220 & 0.4861 & 0.4796 \\ 0.4781 & 0.5000 & 0.4691 & 0.4631 \\ 0.5139 & 0.5310 & 0.5000 & 0.4741 \\ 0.5204 & 0.5370 & 0.5259 & 0.5000 \end{bmatrix}.$$

D_1=0.51 , D_2=0.53, D_3=0.50, D_4=0.48.

In the following, we rank all objects according to the number of $D(i)$.

$$X_2 \succcurlyeq X_1 \succcurlyeq X_3 \succcurlyeq X_4.$$

If we consider the stochastic information system given in Table 5 and assume that a threshold of dominance degree 0.6, then we can obtain a Boolean matrix, which is corresponding to $D_{AT}^{0.6}$ such that:

$$D_{AT}^{0.6} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Take for instance x_2, under the given dominance degree $\alpha = 0.6$, objects x_2 is regarded as dominated by x_2 with respect to set of attributes AT.

By the definition of discernibility matrix, one can obtain the discernibility sets for all pairs of objects are shown in Table 6.

Table 6 0.6-discernibility set for pairs of objects given in Table 5.

U	x_1	x_2	x_3	x_4
x_1		a_1a_4	$a_1a_2a_3a_4$	$a_1a_3a_4$
x_2	$a_1a_2a_3a_4$		$a_2a_3a_4$	$a_1a_3a_4$
x_3	$a_1a_2a_3a_4$	$a_1a_3a_4$		$a_1a_3a_4$
x_4	a_2a_3	$a_1a_2a_3$	a_2a_3	

By Theorems 1 and 2, it is not difficult to obtain that $Red^{0.6}(S) = \{\{a_1, a_2\}, \{a_1, a_3\}, \{a_2, a_4\}, \{a_3, a_4\}\}$, i.e., $\{a_1, a_2\}, \{a_1, a_3\}, \{a_2, a_4\}$ and $\{a_3, a_4\}$ are the 0.6-dominance reduct in Table 5.

5. CONCLUSIONS

In this study, a general method is proposed to model the dominance relation in stochastic information systems. To check the dominance degree between two stochastic data, a threshold is utilized such α-dominance relation. This is useful in adapting information granulation according to varied levels.

In the framework, firstly, the dominance degree matrix of objects according to each attribute is established based on comparisons of probability distributions. The dominance degree provides that identifying process of stochastic dominance relations to compare objects. According to dominance relation, a rough set approach in stochastic information system is established as a substitution of the indiscernibility relation. Also we have discussed stochastic information systems in order to extract attribute reduct(s) based on the discernibility matrices. The proposed approach can be computed basic and easy procedures, beside the method provides to find much simpler attribute reduct(s) in a simple stochastic information system.

In terms of future study, the proposed model can be extended with multiple information forms such as crisp numbers, interval numbers, and stochastic information and so on.

REFERENCES

[1]. Pawlak, Z. (1982). Rough sets. International Journal of Computer & Information Sciences, 11(5), 341-356.

[23]. Nowak, M. (2007). Aspiration level approach in stochastic MCDM problems. European Journal of Operational Research, 177(3), 1626-1640.

[24]. Zaras, K., & Martel, J. M. (1994). Multiattribute analysis based on stochastic dominance (pp. 225-248). Springer Netherlands.

[25]. Zaras, K. (2001). Rough approximation of a preference relation by a multi-attribute stochastic dominance for determinist and stochastic evaluation problems. European Journal of Operational Research, 130(2), 305-314.

[2]. Skowron, A., & Pawlak, Z. (2007). Rough sets: Some extensions. Information Sciences, 177(1), 28-40.

[3]. Pawlak, Z. (1991). Rough sets: theoretical aspects of reasoning about data, system theory, Knowledge Engineering and Problem Solving, vol. 9.

[4]. Hu, Q., Yu, D., Liu, J., & Wu, C. (2008). Neighborhood rough set based heterogeneous feature subset selection. Information sciences, 178(18), 3577-3594.

[5]. Li, T., Ruan, D., Geert, W., Song, J., & Xu, Y. (2007). A rough sets based characteristic relation approach for dynamic attribute generalization in data mining. Knowledge-Based Systems, 20(5), 485-494.

[6]. Qian, Y., Liang, J., Pedrycz, W., & Dang, C. (2010). Positive approximation: an accelerator for attribute reduction in rough set theory. Artificial Intelligence, 174(9), 597-618.

[7]. Zhang, J., Li, T., Ruan, D., Gao, Z., & Zhao, C. (2012). A parallel method for computing rough set approximations. Information Sciences, 194, 209-223.

[8]. Zhang, J., Li, T., Ruan, D., & Liu, D. (2012). Neighborhood rough sets for dynamic data mining. International Journal of Intelligent Systems, 27(4), 317-342.

[9]. Guan, J. W., & Bell, D. A. (1998). Rough computational methods for information systems. Artificial intelligence, 105(1), 77-103.

[10]. Qian, Y., Liang, J., & Dang, C. (2008). Converse approximation and rule extraction from decision tables in rough set theory. Computers & Mathematics with Applications, 55(8), 1754-1765.

[11]. Greco, S., Matarazzo, B., & Slowinski, R. (1998, June). A new rough set approach to multicriteria and multiattribute classification. In Rough sets and current trends in computing (pp. 60-67). Springer Berlin Heidelberg.

[12]. Greco, S., Matarazzo, B., & Slowinski, R. (2001). Rough sets theory for multicriteria decision analysis. European journal of operational research, 129(1), 1-47.

[13]. Greco, S., Matarazzo, B., & Slowinski, R. (2002). Rough sets methodology for sorting problems in presence of multiple attributes and criteria. European journal of operational research, 138(2), 247-259.

[14]. Greco, S., Matarazzo, B., Slowinski, R., & Stefanowski, J. (2000, October). An algorithm for induction of decision rules consistent with the dominance principle. In Rough sets and current trends in computing (pp. 304-313). Springer Berlin Heidelberg.

[15]. Dembczyński, K., Pindur, R., & Susmaga, R. (2003). Generation of exhaustive set of rules within dominance-based rough set approach. Electronic Notes in Theoretical Computer Science, 82(4), 96-107.

[16]. Dembczyński, K., Pindur, R., & Susmaga, R. (2003). Generation of exhaustive set of rules within dominance-based rough set approach. Electronic Notes in Theoretical Computer Science, 82(4), 96-107.

[17]. Sai, Y., Yao, Y. Y., & Zhong, N. (2001). Data analysis and mining in ordered information tables. In Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on (pp. 497-504). IEEE.

[18]. Kotłowski, W., Dembczyński, K., Greco, S., & Słowiński, R. (2008). Stochastic dominance-based rough set model for ordinal classification. Information Sciences, 178(21), 4019-4037.

[19]. J. Martel, K. Zaras, Dominance stochastique en analyse multicritére face au risque, Cahier No. 100. Laboratoire d'Analyseet Modélisation de Systèmes pour l'Aide à la Décision. Université Paris-Dauphine, Paris, 1990.

[20]. J. Martel, S.H. Azondékon, K. Zaras, Preference relations in multicriterion analysis under risk, Belgian Journal of Operations Research, Statistics and Computer Science 31 (1992) 55–83.

[21]. Nowak, M. (2004). Preference and veto thresholds in multicriteria analysis based on stochastic dominance. European Journal of Operational Research, 158(2), 339-350.

[22]. Nowak, M. (2006). INSDECM—an interactive procedure for stochastic multicriteria decision problems. European Journal of Operational Research, 175(3), 1413-1430.

[26]. Zaras, K. (2004). Rough approximation of a preference relation by a multi-attribute dominance for deterministic, stochastic and fuzzy decision problems. European Journal of Operational Research, 159(1), 196-206.

[27]. Zawisza, M., & Trzpiot, G. (2002). Multicriteria Analysis Based On Stochastic and Probabilistic Dominance in Measuring Quality of Life. In Multiple Objective and Goal Programming (pp. 412-423). Physica-Verlag HD.

[28]. Zhang, Y., Fan, Z. P., & Liu, Y. (2010). A method based on stochastic dominance degrees for stochastic multiple criteria decision making. Computers & Industrial Engineering, 58(4), 544-552.
[29]. Lin, G., Qian, Y., & Li, J. (2012). NMGRS: Neighborhood-based multigranulation rough sets. International Journal of Approximate Reasoning, 53(7), 1080-1093.
[30]. Qian, Y., Zhang, H., Sang, Y., & Liang, J. (2014). Multigranulation decision-theoretic rough sets. International Journal of Approximate Reasoning, 55(1), 225-237.
[31]. Qian, Y., Li, S., Liang, J., Shi, Z., & Wang, F. (2014). Pessimistic rough set based decisions: a multigranulation fusion strategy. Information Sciences, 264, 196-210.
[32]. Xu, W., Sun, W., Zhang, X., & Zhang, W. (2012). Multiple granulation rough set approach to ordered information systems. International Journal of General Systems, 41(5), 475-501.
[33]. Yang, X., Qi, Y., Song, X., & Yang, J. (2013). Test cost sensitive multigranulation rough set: model and minimal cost selection. Information Sciences, 250, 184-199.
[34]. Yang, X., Song, X., She, Y., & Yang, J. (2013). Hierarchy on multigranulation structures: a knowledge distance approach. International Journal of General Systems, 42(7), 754-773.
[35]. Liu, Y., Fan, Z. P., & Zhang, Y. (2011). A method for stochastic multiple criteria decision making based on dominance degrees. Information Sciences, 181(19), 4139-4153.
[36]. Graves, S. B., & Ringuest, J. L. (2009). Probabilistic dominance criteria for comparing uncertain alternatives: A tutorial. Omega, 37(2), 346-357.
[37]. Zhang, W. X., & Qiu, G. F. (2005). Uncertain decision making based on rough sets. Publishing of Tsinghua University, Beijing.
[38]. D. Bertsekas, J. Tsitsiklis, Introduction to Probability, Athena Scientific, 2002.
[39]. W. Feller, An Introduction to Probability Theory and Its applications, Wiley, New York, 1971.
[40]. Fouss, F., Achbany, Y., & Saerens, M. (2010). A probabilistic reputation model based on transaction ratings. Information Sciences, 180(11), 2095-2123.
[41]. Lahdelma, R., Makkonen, S., & Salminen, P. (2006). Multivariate Gaussian criteria in SMAA. European Journal of Operational Research, 170(3), 957-970.
[42]. Yager, R. R., Detyniecki, M., & Bouchon-Meunier, B. (2001). A context-dependent method for ordering fuzzy numbers using probabilities. Information Sciences, 138(1), 237-255.
[43]. Qian, Y., Liang, J., & Dang, C. (2008). Interval ordered information systems. Computers & Mathematics with Applications, 56(8), 1994-2009.
[44]. Yang, X., Qi, Y., Yu, D. J., Yu, H., & Yang, J. (2015). α-Dominance relation and rough sets in interval-valued information systems. Information Sciences, 294, 334-347.
[45]. Shao, M. W., & Zhang, W. X. (2005). Dominance relation and rules in an incomplete ordered information system. International journal of intelligent systems, 20(1), 13-27.
[46]. X.B. Yang, J.Y. Yang, C. Wu, D.J. Yu, Dominance-based rough set approach and knowledge reductions in incomplete ordered information system, Inform. Sci. 178 (2008) 1219–1234.