Knowledge-Based Systems xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A two-grade approach to ranking interval data

Peng Song [a,b], Jiye Liang [b,c,*], Yuhua Qian [b,c]

[a] School of Management, Shanxi University, Taiyuan, 030006 Shanxi, China
[b] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Taiyuan, 030006 Shanxi, China
[c] School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China

A R T I C L E I N F O

Article history:
Received 31 January 2011
Received in revised form 26 May 2011
Accepted 14 October 2011
Available online xxxx

Keywords:
Ranking decision
Rough sets
Dominance relation
Information entropy
Interval ordered information systems

A B S T R A C T

Ranking decision for interval data is a very important issue in decision making analysis. In recent years, several ranking approaches based on dominance relations have been developed. In these approaches, a dominance degree and an entire dominance degree are employed. However, one cannot obtain the complete rank of objects. To address this problem, this work will propose a two-grade approach to ranking interval data. In this approach, we keep the ranking result induced by the entire dominance degree in the first grade, and then refine the objects that cannot be ranked through introducing a so-called entire directional distance index. An example and a real case are employed to verify the effectivity of the two-grade ranking approach proposed in this paper.

1. Introduction

In reality, one often encounters a number of alternatives which need to be evaluated on the basis of a set of criteria in investment decision [28,31,39,41], universities ranking [10,34], road safety risk evaluation [9], and so on. In these cases, the alternatives and the related criteria are often combined to a data table. In decision making, one needs to rank these alternatives through using some criteria that are characterized by attributes in the data table according to an increasing or a decreasing preference. This kind of decision making tasks are called ranking decision, which is becoming an important research point in decision making analysis. At present, ranking decision has been widely used in economy, management, engineering and other broad areas.

For effective and rational ranking decision, many decision making methods have been developed, which include TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [13,51], AHP (Analytic Hierarchy Process) [16,32,35], ELECTRE (ELimination Et Choix Traduisant la REalité) [33,37], and the methods based on fuzzy set theory [3,10–12,52,39], etc.

In the past twenty years, rough set theory introduced by Pawlak [24–27], has increasingly played an important role in the field of decision making analysis. One of its prominent advantages is to effectively deal with vagueness and uncertainty information without requiring any prior knowledge. As we know, Pawlak's rough set theory does not consider attributes with preference-ordered domains, that is, criteria. To solve the problem of ranking decision, several extended rough set models have been developed in the literature. Greco et al. [5–8] proposed an extension of rough set theory induced by a dominance relation, called a dominance-based rough set approach (DRSA), in which the ordering properties of criteria are taken into account. In what follows, we briefly review several works related to dominance-based decision making. By adding order relations on attribute values, Yao et al. [34,50] studied ordered information tables, and raised a convenient model to mine ordering rules through transforming an ordered information table into a binary information table. Yang et al. [47] introduced a similarity dominance relation, and developed a new dominance-based rough set model in incomplete ordered information systems. Hu et al. [11,12] presented a fuzzy preference rough set model by integrating fuzzy preference relations with a fuzzy rough set model. Moreover, evaluation on decision performance is also an important task in rough set theory [30,40,42,44]. In Ref. [44], concepts of knowledge granulation, knowledge entropy and knowledge uncertainty have been given to measure the discernibility ability of different knowledge in ordered information systems. Under the condition of homomorphism, Wang et al. [40] researched data compression in ordered information systems to perform equivalent attribute reductions and rule extraction in the smaller compressed image database for improving efficiency and saving decision-making costs. In a word, the dominance-based rough set theory has contributed a basic theoretical framework for ranking decision.

* Corresponding author at: School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China. Tel./fax: +86 0351 7018176.
E-mail addresses: songpeng@sxu.edu.cn (P. Song), ljy@sxu.edu.cn (J. Liang), jinchengqyh@sxu.edu.cn (Y. Qian).

0950-7051/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2011.10.007

In decision making analysis, we often need to deal with various types of data sets, in which objects may be characterized by single value, set value, null value, or interval value [2,14,15,17,18,28–31,36,45–49]. Among these kinds of data sets, interval data is an important class of data and a generalized form of single-valued data. Hence, how to rank objects with interval values is a very desirable issue.

As mentioned above, although many results have been developed in the context of interval ordered information systems, how to rank objects using a dominance relation has not been reported. To address this problem, Qian et al. [28] proposed a ranking approach for all objects based on dominance classes and the entire dominance degree. This is the first attempt to rank objects with interval values. This method adopts a cautious decision strategy, in which we say that one object is superior to the other object if and only if the value of one object is dominant than that of the other under each attribute. Obviously, this approach is credible because it meets with practical decision situations, in which the risk aversion is one of major characteristics for decision makers.

However, it can be seen, from Qian's work [28], that the final rank obtained is not a complete rank, in which there may exist several objects being put into the same place. To overcome this drawback, we will further develop a new version of Qian's ranking approach according to the following two motivations.

(1) In practical issues, decision makers often want to get a complete rank of objects according to a user's requirement. A complete rank of objects will be helpful for obtaining a more satisfactory decision scheme. This opinion can be illustrated by using ranking decision of investment projects [31]. Generally, a complete rank of investment projects is necessary since decision makers only have limited capital. If several investment projects lie in the same place in final ranking result, that will confuse decision makers. Therefore, how to obtain a complete rank needs to be further addressed in the context of interval ordered information systems.
(2) In the process of looking for a complete rank, the rank induced by the entire dominance degree should be remained. As we know, a cautious ranking project is often desirable for the vast majority of decision makers. Hence, we argue that looking for a complete rank should be based on the cautious rank. For this reason, we will establish a two-grade ranking approach considering the property of rank preservation.

Therefore, in this paper, our objective is to develop an approach to obtaining a complete rank of objects with interval values. In this study, we first introduce the concept of a directional distance index and give some of its nice properties. Then, we define an ordered mutual information to calculate the weight of each criterion and the directional distance index with weights. Based on this consideration, we propose a two-grade approach to ranking objects with interval values. Finally, we also employ a real case about stock selection for verifying the effectivity of the proposed approach in this paper.

The remainder of this paper is organized as follows. Section 2 reviews some preliminary concepts and important properties of interval ordered information systems. Section 3 establishes a two-grade approach to ranking completely objects with interval values by combining a directional distance index with a dominance degree. In Section 4, through introducing weights of criteria based on an ordered mutual information, we propose a more rational two-grade approach to ranking completely interval data. In Section 5, we use a stock selection case to illustrate how to make a decision by using the ranking approach proposed in this paper. Finally, Section 6 concludes this paper with a remark.

2. Preliminaries

In this section, we briefly review some basic concepts and important properties of interval ordered information systems.

An *information system* (IS) is a quadruple $S = (U,AT,V,f)$, where U is a finite non-empty set of objects and AT is a finite non-empty set of attributes, $V = \bigcup_{a\in AT} V_a$ and V_a is a domain of attribute a, $f: U \times AT \to V$ is a total function such that $f(x,a) \in V_a$ for every $a \in AT$, $x \in U$, called an information function [31]. An information system is called an *interval information system* (IIS) if V_a is a set of interval numbers. We denote $f(x,a) \in V_a$ by

$$f(x,a) = [a^L(x), a^U(x)] = \{p \mid a^L(x) \leqslant p \leqslant a^U(x), a^L(x), a^U(x) \in \mathbf{R}\}.$$

It is the interval number of x under the attribute a.

Here, single-valued information systems, in which $f(x,a) = a^L(x) = a^U(x)$, can be seen as a special form of interval information systems. Example 2.1 shows an interval information system.

Example 2.1 [28]. An interval information system is listed in Table 1, where $U = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}$, $AT = \{a_1,a_2,a_3,a_4,a_5\}$.

Definition 2.1. An interval information system is called an *interval ordered information system* (IOIS) if all attributes are criteria, that is, the domain of each attribute is ordered according to an increasing or a decreasing preference.

It is assumed that the domain of a criterion $a \in AT$ is completely pre-ordered by an outranking relation $\succcurlyeq_a$; $y \succcurlyeq_a x$ means that y is at least as good as (outranks) x with respect to the criterion a. Furthermore, we define $y \succcurlyeq_A x \iff \forall a \in A (A \subseteq AT), y \succcurlyeq_a x$.

Based on the above illustration, we review the dominance relation that identifies dominance classes especially in an interval ordered information system. In a given IOIS, we say that y dominates x with respect to $A \subseteq AT$ if $y \succcurlyeq_A x$, and denoted by $yR_A^{\geqslant}x$. That is

$$\begin{aligned} R_A^{\geqslant} &= \{(y,x) \in U \times U \mid a_1^L(y) \geqslant a_1^L(x),\ a_1^U(y) \geqslant a_1^U(x) \\ &\quad (\forall a_1 \in A_1);\ a_2^L(y) \leqslant a_2^L(x),\ a_2^U(y) \leqslant a_2^U(x)\ (\forall a_2 \in A_2)\} \\ &= \{(y,x) \in U \times U \mid (y,x) \in R_A^{\geqslant}\}, \end{aligned}$$

where the attributes set A_1 according to increasing preference and A_2 according to decreasing preference, and $A = A_1 \cup A_2$.

According to the definition of $R_A^{\geqslant}$, the dominance class $[x]_A^{\geqslant}$ which is the set of objects dominating x can be induced as follows

$$\begin{aligned} [x]_A^{\geqslant} &= \{y \in U \mid a_1^L(y) \geqslant a_1^L(x),\ a_1^U(y) \geqslant a_1^U(x) (\forall a_1 \in A_1); a_2^L(y) \\ &\quad \leqslant a_2^L(x),\ a_2^U(y) \leqslant a_2^U(x) (\forall a_2 \in A_2)\} = \{y \in U \mid (y,x) \in R_A^{\geqslant}\}. \end{aligned}$$

Analogously, $R_A^{\leqslant}$ and $[x]_A^{\leqslant}$ can be defined too.

From the definitions of $R_A^{\geqslant}$ and $[x]_A^{\geqslant}$, a partial order can be defined on the attribute set. Let $S = (U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$. We say that A is coarser than B (or B is finer than A) if and only if $[x_i]_B^{\geqslant} \subseteq [x_i]_A^{\geqslant}$, $\forall i \in \{1,2,\ldots,|U|\}$, just $B \preccurlyeq A$.

Using the above definitions, one has the following properties [28].

Table 1
An interval information system.

U	a_1	a_2	a_3	a_4	a_5
x_1	1	[0,1]	2	1	[1,2]
x_2	[0,1]	0	[1,2]	0	1
x_3	[0,1]	0	[1,2]	1	1
x_4	0	0	1	0	1
x_5	2	[1,2]	3	[1,2]	[2,3]
x_6	[0,2]	[1,2]	[1,3]	[1,2]	[2,3]
x_7	1	1	2	1	2
x_8	[1,2]	[1,2]	[2,3]	2	[2,3]
x_9	[1,2]	2	[2,3]	[0,2]	3
x_{10}	2	2	3	[0,1]	3

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007

Property 2.1. *Let S = (U,AT,V,f) be an IOIS and $A \subseteq AT$, we have that*

(1) $R_A^{\geqslant} = \bigcap_{a \in A} R_a^{\geqslant}$;
(2) $R_A^{\leqslant} = \bigcap_{a \in A} R_a^{\leqslant}$.

Property 2.2. *Let S = (U,AT,V,f) be an IOIS and $A \subseteq AT$, we have that*

(1) $R_A^{\geqslant}$, $R_A^{\leqslant}$ *are reflexive;*
(2) $R_A^{\geqslant}$, $R_A^{\leqslant}$ *are unsymmetric; and*
(3) $R_A^{\geqslant}$, $R_A^{\leqslant}$ *are transitive.*

Property 2.3. *Let S = (U,AT,V,f) be an IOIS and A, $B \subseteq AT$, we have that*

(1) *if* $B \subseteq A \subseteq AT$, *then* $R_B^{\geqslant} \supseteq R_A^{\geqslant} \supseteq R_{AT}^{\geqslant}$;
(2) *if* $B \subseteq A \subseteq AT$, *then* $[x]_B^{\geqslant} \supseteq [x]_A^{\geqslant} \supseteq [x]_{AT}^{\geqslant}$;
(3) *if* $B \subseteq A \subseteq AT$, *then* $AT \preccurlyeq A \preccurlyeq B$;
(4) *if* $x_j \in [x_i]_A^{\geqslant}$, *then* $[x_j]_A^{\geqslant} \subseteq [x_i]_A^{\geqslant}$ *and* $[x_i]_A^{\geqslant} = \bigcup\{[x_j]_A^{\geqslant} : x_j \in [x_i]_A^{\geqslant}\}$; *and*
(5) $[x_i]_A^{\geqslant} = [x_j]_A^{\geqslant}$ *iff* $f(x_i,a) = f(x_j,a)$ $(\forall a \in A)$.

In the rest part of this paper, without any loss of generality, we only consider all attributes with an increasing preference.

3. A two-grade approach to ranking interval data

Ranking objects is an important problem in many practical decision making fields. Under rough set theory framework, Zhang and Qiu [53] proposed a ranking method in classical ordered information systems by defining a concept of dominance degree. But this approach only deals with the ranking problem aiming at single-valued information systems. Through extending the definition of dominance relation, Qian et al. [28] established an approach to ranking objects in interval ordered information systems. However, this method still needs to be further improved because it cannot completely rank objects with interval values. In this section, we will construct a two-grade ranking approach to obtaining a complete rank of interval objects.

To further improve the ranking approach, we review the concept of dominance degree and some of its properties in the following.

Definition 3.1 [28]. Let S = (U,AT,V,f) be an IOIS and $A \subseteq AT$. *Dominance degree* between two objects with respect to the dominance relation $R_A^{\geqslant}$ is defined as

$$D_A(x_i,x_j) = \frac{|[x_i]_A^{\geqslant c} \cup [x_j]_A^{\geqslant}|}{|U|},$$

where $|\cdot|$ denotes the cardinality of a set, $[x_i]_A^{\geqslant c} = U - [x_i]_A^{\geqslant}$, $x_i, x_j \in U$.

Property 3.1 [28]. *$D_A(x_i,x_j)$ has the following properties*

(1) $\frac{1}{|U|} \leqslant D_A(x_i,x_j) \leqslant 1$;
(2) *if* $(x_j,x_k) \in R_A^{\geqslant}$, *then* $D_A(x_i,x_j) \leqslant D_A(x_i,x_k)$; *and*
(3) *if* $(x_j,x_k) \in R_A^{\geqslant}$, *then* $D_A(x_j,x_i) \geqslant D_A(x_k,x_i)$.

Definition 3.2 [28]. Let S = (U,AT,V,f) be an IOIS and $A \subseteq AT$. *Entire dominance degree* of each object is defined as

$$D_A(x_i) = \frac{1}{|U|-1}\sum_{j \neq i} D_A(x_i,x_j), \quad x_i, \quad x_j \in U.$$

From Definitions 3.1 and 3.2, one can easily find that the entire dominance degree of an object can be used to judge its ranking place. The bigger the value of $D_A(x_i)$, and the better this object. This is illustrated by Example 3.1.

Example 3.1 (*Continued from Example 2.1*). Rank all objects in U according to the entire dominance degree $D_{AT}(x_i)$.

First, according to the definition of $[x_i]_{AT}^{\geqslant}$, we have that

$$[x_1]_{AT}^{\geqslant} = \{x_1,x_5,x_7,x_8\}, \quad [x_2]_{AT}^{\geqslant} = \{x_1,x_2,x_3,x_5,x_6,x_7,x_8,x_9,x_{10}\},$$

$$[x_3]_{AT}^{\geqslant} = \{x_1,x_3,x_5,x_6,x_7,x_8\},$$
$$[x_4]_{AT}^{\geqslant} = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\},$$

$$[x_5]_{AT}^{\geqslant} = \{x_5\}, \quad [x_6]_{AT}^{\geqslant} = \{x_5,x_6,x_8\}, \quad [x_7]_{AT}^{\geqslant} = \{x_5,x_7,x_8\},$$

$$[x_8]_{AT}^{\geqslant} = \{x_8\}, \quad [x_9]_{AT}^{\geqslant} = \{x_9\}, \quad [x_{10}]_{AT}^{\geqslant} = \{x_{10}\}.$$

Therefore, one can obtain the dominance degree relation matrix as

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0.7 & 0.8 & 0.9 & 0.7 & 0.6 & 0.6 \\ 0.5 & 1 & 0.7 & 1 & 0.2 & 0.4 & 0.4 & 0.2 & 0.2 & 0.2 \\ 0.8 & 1 & 1 & 1 & 0.5 & 0.7 & 0.7 & 0.5 & 0.4 & 0.4 \\ 0.4 & 0.9 & 0.6 & 1 & 0.1 & 0.3 & 0.3 & 0.1 & 0.1 & 0.1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0.9 & 0.9 & 0.9 \\ 0.9 & 1 & 1 & 1 & 0.8 & 1 & 0.9 & 0.8 & 0.7 & 0.7 \\ 1 & 1 & 1 & 1 & 0.8 & 0.9 & 1 & 0.9 & 0.7 & 0.7 \\ 1 & 1 & 1 & 1 & 0.9 & 1 & 1 & 1 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 & 1 & 0.9 & 0.9 & 0.9 & 0.9 & 1 & 0.9 \\ 0.9 & 1 & 0.9 & 1 & 0.9 & 0.9 & 0.9 & 0.9 & 0.9 & 1 \end{pmatrix}.$$

Then, we get that

$$D_{AT}(x_1) = 0.81, \quad D_{AT}(x_2) = 0.42, \quad D_{AT}(x_3) = 0.67, \quad D_{AT}(x_4) = 0.32, \quad D_{AT}(x_5) = 0.97,$$

$$D_{AT}(x_6) = 0.87, \quad D_{AT}(x_7) = 0.89, \quad D_{AT}(x_8) = 0.97, \quad D_{AT}(x_9) = 0.92, \quad D_{AT}(x_{10}) = 0.92.$$

Thus, according to the values of $D_{AT}(x_i)$, the ranking result can be listed as follows

$$\begin{pmatrix} x_5 \\ x_8 \end{pmatrix} \succ \begin{pmatrix} x_9 \\ x_{10} \end{pmatrix} \succ x_7 \succ x_6 \succ x_1 \succ x_3 \succ x_2 \succ x_4.$$

Remark 3.1. From Example 3.1, we obtained a ranking result by the dominance degree and the entire dominance degree. However, decision makers often want to get a complete rank of objects according to a user's requirement. Therefore, it is not satisfactory because some objects cannot be ranked completely in present rank (seeing x_5 and x_8, x_9 and x_{10}). In fact, the entire dominance degree $D_A(x_i)$ is constructed based on the dominance degree $D_A(x_i,x_j)$, and the value is determined by the comparison of dominance classes of two objects. Obviously, $D_A(x_i,x_j)$ mainly investigates the relative ranking position of objects in entire universe from the viewpoint of rank, while ignoring more detailed difference of two objects under their attributes values. Hence, it is inevitable that some objects may be put into the same place.

To get a complete rank of objects, from the numerical point of view, we first define a directional distance index to measure the preferability degree of the object x_i over the object x_j.

Definition 3.3. Given two interval numbers $f(x_i,a) = [a^L(x_i), a^U(x_i)]$ and $f(x_j,a) = [a^L(x_j), a^U(x_j)]$. *Directional distance index* between two objects under the attribute a is defined as

$$DDI_a(x_i,x_j) = \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))},$$

where $max(a^U(x)) = max\{a^U(x_1), a^U(x_2), \ldots, a^U(x_{|U|})\}$, $min(a^L(x)) = min\{a^L(x_1), a^L(x_2), \ldots, a^L(x_{|U|})\}$, and $x_i, x_j \in U$. In particular, $DDI_a(x_i,x_j) = \frac{1}{2}$ if $max(a^U(x)) = min(a^L(x))$.

Property 3.2. *$DDI_a(x_i,x_j)$ has the following properties*

(1) *$0 \leqslant DDI_a(x_i,x_j) \leqslant 1$;*
(2) *if $(x_j,x_k) \in R_a^{\geqslant}$, then $DDI_a(x_i,x_j) \leqslant DDI_a(x_i,x_k)$;*
(3) *if $(x_i,x_j) \in R_a^{\geqslant}$, then $DDI_a(x_i,x_j) \geqslant \frac{1}{2}$;*
(4) *if $DDI_a(x_i,x_j) \geqslant \frac{1}{2}$ and $DDI_a(x_j,x_k) \geqslant \frac{1}{2}$, then $DDI_a(x_i,x_k) \geqslant \frac{1}{2}$; and*
(5) *$DDI_a(x_i,x_j) + DDI_a(x_j,x_i) = 1$.*

Proof. Let $f(x_i,a) = [a^L(x_i), a^U(x_i)]$, $f(x_j,a) = [a^L(x_j), a^U(x_j)]$, and $f(x_k, a) = [a^L(x_k), a^U(x_k)]$.

(1) According to the definitions of $max(a^U(x))$ and $min(a^L(x))$, one has that

$$min(a^L(x)) - max(a^U(x)) \leqslant a^U(x_i) - a^U(x_j) \leqslant max(a^U(x)) - min(a^L(x)) \Longleftrightarrow -1 \leqslant \frac{a^U(x_i) - a^U(x_j)}{max(a^U(x)) - min(a^L(x))} \leqslant 1.$$

Similarly, we have that

$$-1 \leqslant \frac{a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \leqslant 1.$$

Thus,

$$-2 \leqslant \frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \leqslant 2 \Longleftrightarrow 0 \leqslant \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \leqslant 1 \Longleftrightarrow 0 \leqslant DDI_a(x_i,x_j) \leqslant 1.$$

(2) If $(x_j,x_k) \in R_a^{\geqslant}$, it follows from the dominance relation $R_a^{\geqslant}$ that the interval number $f(x_j,a)$ is bigger than the interval number $f(x_k,a)$, i.e., $a^U(x_j) \geqslant a^U(x_k)$ and $a^L(x_j) \geqslant a^L(x_k)$. Then, we have that

$$\begin{aligned} DDI_a(x_i,x_j) - DDI_a(x_i,x_k) &= \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} - \frac{1}{4}\frac{a^U(x_i) - a^U(x_k) + a^L(x_i) - a^L(x_k)}{max(a^U(x)) - min(a^L(x))} \\ &= \frac{1}{4}\frac{a^U(x_k) - a^U(x_j) + a^L(x_k) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \\ &\leqslant \frac{1}{4}\frac{a^U(x_j) - a^U(x_j) + a^L(x_j) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} = 0, \end{aligned}$$

that is $DDI_a(x_i,x_j) \leqslant DDI_a(x_i,x_k)$.

(3) If $(x_i,x_j) \in R_a^{\geqslant}$, then the interval number $f(x_i,a)$ is bigger than the interval number $f(x_j,a)$, i.e., $a^U(x_i) \geqslant a^U(x_j)$ and $a^L(x_i) \geqslant a^L(x_j)$. Therefore,

$$\begin{aligned} DDI_a(x_i,x_j) &= \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \\ &\geqslant \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_i) + a^L(x_i) - a^L(x_i)}{max(a^U(x)) - min(a^L(x))} \\ &= \frac{1}{2} + 0 = \frac{1}{2}. \end{aligned}$$

(4) If $DDI_a(x_i,x_j) \geqslant \frac{1}{2}$ and $DDI_a(x_j,x_k) \geqslant \frac{1}{2}$, then

$$\begin{aligned} DDI_a(x_i,x_k) &= \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_k) + a^L(x_i) - a^L(x_k)}{max(a^U(x)) - min(a^L(x))} \\ &= \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_k) + a^L(x_i) - a^L(x_k) - a^U(x_j) + a^U(x_j) - a^L(x_j) + a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \\ &= \frac{1}{2} + \frac{1}{4}\frac{[a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)] + [a^U(x_j) - a^U(x_k) + a^L(x_j) - a^L(x_k)]}{max(a^U(x)) - min(a^L(x))} \\ &= \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} + \frac{1}{2} + \frac{1}{4}\frac{a^U(x_j) - a^U(x_k) + a^L(x_j) - a^L(x_k)}{max(a^U(x)) - min(a^L(x))} \\ &- \frac{1}{2} = DDI_a(x_i,x_j) + DDI_a(x_j,x_k) - \frac{1}{2} \geqslant \frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

(5) From Definition 3.3, it follows that

$$\begin{aligned} DDI_a(x_i,x_j) + DDI_a(x_j,x_i) &= \frac{1}{2} + \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j)}{max(a^U(x)) - min(a^L(x))} \\ &+ \frac{1}{2} + \frac{1}{4}\frac{a^U(x_j) - a^U(x_i) + a^L(x_j) - a^L(x_i)}{max(a^U(x)) - min(a^L(x))} \\ &= 1 + \frac{1}{4}\frac{a^U(x_i) - a^U(x_j) + a^L(x_i) - a^L(x_j) + a^U(x_j) - a^U(x_i) + a^L(x_j) - a^L(x_i)}{max(a^U(x)) - min(a^L(x))} \\ &= 1 + 0 = 1. \end{aligned}$$

This completes the proof. □

Remark 3.2. From the numerical viewpoint, the index $DDI_a(x_i,x_j)$ can measure the preferability degree of the object x_i over the object x_j. From Definition 3.3, it is easy to see that $DDI_a(x_i,x_j)$ can characterize both the difference of two interval numbers and the partial directional property. In other words, when the object x_i is preferable to the object x_j,one has that $DDI_a(x_i,x_j) \geqslant \frac{1}{2}$, and vice versa. In this paper, we hence call the index a directional distance index.

In fact, from Definition 3.3, we can make a comparison between two objects under an attribute in IOIS. Furthermore, it is no doubt that two objects can be compared under all considered attributes according to the following formula

$$DDI_A(x_i,x_j) = \frac{1}{|A|}\sum_{\forall a \in A} DDI_a(x_i,x_j),$$

where $A \subseteq AT$, and $|A|$ denotes the cardinality of a considered attribute set.

Analogously, $DDI_A(x_i,x_j)$ has the same properties as $DDI_a(x_i,x_j)$.

Based on the definition of $DDI_A(x_i,x_j)$, let $(x_i,x_j) \in U \times U$, one can construct a directional distance index relation matrix with respect to A. Furthermore, the entire directional distance index of each object can be calculated through the constructed matrix, which is as follows.

Definition 3.4. Let $S = (U, AT, V, f)$ be an IOIS and $A \subseteq AT$. *Entire directional distance index* of each object is defined as

$$DDI_A(x_i) = \frac{1}{|U| - 1}\sum_{j \neq i} DDI_A(x_i,x_j), \quad x_i, \quad x_j \in U.$$

According to the entire directional distance index of each object, one can easily rank all objects on the basis of the values of $DDI_A(x_i)$.

Remark 3.3. Essentially, the entire directional distance index $DDI_A(x_i)$ is mainly defined for ranking objects from the perspective of distance under attributes values, which can help us to obtain a much finer ranking result.

However, from the viewpoint of decision making, the entire dominance degree $D_A(x_i)$ only gives a cautious ranking result, which is built on the basis of the decision strategy that one object is superior to the other object under all considered attributes. As we know, for practical decision making issues, a cautious ranking result is often desirable since risk aversion is one of major characteristics for decision makers. Therefore, we argue that $D_A(x_i)$ should be a prior grade and $DDI_A(x_i)$ is a second grade, and the latter can give a much finer ranking result. Based on this idea, we establish a two-grade ranking approach in the following.

Definition 3.5. Let $S=(U,AT,V,f)$ be an IOIS and $A\subseteq AT$, and x_i, $x_j\in U$.

If $D_A(x_i)\geqslant D_A(x_j)$, then x_i dominates x_j, denoted by $x_i \succcurlyeq x_j$;
If $D_A(x_i)\leqslant D_A(x_j)$, then x_i is dominated by x_j, denoted by $x_i \preccurlyeq x_j$;
If $D_A(x_i)=D_A(x_j)$, then

(1) If $DDI_A(x_i)\geqslant DDI_A(x_j)$, then x_i dominates x_j, denoted by $x_i \succcurlyeq x_j$;
(2) If $DDI_A(x_i)\leqslant DDI_A(x_j)$, then x_i is dominated by x_j, denoted by $x_i \preccurlyeq x_j$;
(3) If $DDI_A(x_i)=DDI_A(x_j)$, then x_i is the same as x_j, denoted by $x_i=x_j$.

From Remark 3.1, it is known that some objects may be put into the same place in the ranking result by using the entire dominance degree $D_A(x_i)$. According to the motivations of this paper, we try to establish a complete ranking approach keeping the rank induced by the entire dominance degree unchanged. From Definition 3.5, the entire dominance degree $D_A(x_i)$ is made as the first grade of this ranking approach, and the entire directional distance index $DDI_A(x_i)$ as the second grade. Clearly, the first grade remains the rank preservation, while the second grade is used to obtain a much finer ranking result. Thus, we can get a complete rank of objects by the two-grade ranking approach. In the following, the Example 3.2 will be employed for showing the effectivity of this approach.

Example 3.2 (*Continued from Example 3.1*). Rank all objects in U according to the two-grade ranking approach.

Firstly, one can get the following ranking result according to the first grade.

$$\begin{pmatrix} x_5 \\ x_8 \end{pmatrix} \succcurlyeq \begin{pmatrix} x_9 \\ x_{10} \end{pmatrix} \succcurlyeq x_7 \succcurlyeq x_6 \succcurlyeq x_1 \succcurlyeq x_3 \succcurlyeq x_2 \succcurlyeq x_4.$$

Secondly, one can utilize the second grade to rank those objects, x_5 and x_8, x_9 and x_{10}, which cannot be ranked by the first grade. Here, the directional distance index relation matrix dose not need to be calculated because there are only two objects in each comparison. Thus, we have that

$$\begin{aligned} DDI_{AT}(x_5) &= DDI_{AT}(x_5,x_8)=\frac{1}{5}\sum_{m=1}^{5}DDI_{a_m}(x_5,x_8) \\ &=\frac{1}{5}\left(\frac{5}{8}+\frac{1}{2}+\frac{5}{8}+\frac{3}{8}+\frac{1}{2}\right)=\frac{21}{40}, \end{aligned}$$

$$\begin{aligned} DDI_{AT}(x_8) &= DDI_{AT}(x_8,x_5)=1-DDI_{AT}(x_5,x_8)=\frac{19}{40}<\frac{21}{40} \\ &=DDI_{AT}(x_5). \end{aligned}$$

Therefore, we can conclude the result $x_5 \succcurlyeq x_8$.

Similarly,

$$\begin{aligned} DDI_{AT}(x_9) &= DDI_{AT}(x_9,x_{10})=\frac{1}{5}\sum_{m=1}^{5}DDI_{a_m}(x_9,x_{10}) \\ &=\frac{1}{5}\left(\frac{3}{8}+\frac{1}{2}+\frac{3}{8}+\frac{5}{8}+\frac{1}{2}\right)=\frac{19}{40}, \end{aligned}$$

$$\begin{aligned} DDI_{AT}(x_{10}) &= DDI_{AT}(x_{10},x_9)=1-DDI_{AT}(x_9,x_{10})=\frac{21}{40}>\frac{19}{40} \\ &=DDI_{AT}(x_9). \end{aligned}$$

That is $x_{10} \succcurlyeq x_9$.

Finally, we can easily obtain the complete ranking result, that is $x_5 \succcurlyeq x_8 \succcurlyeq x_{10} \succcurlyeq x_9 \succcurlyeq x_7 \succcurlyeq x_6 \succcurlyeq x_1 \succcurlyeq x_3 \succcurlyeq x_2 \succcurlyeq x_4$.

From Example 3.2, a complete rank of objects can be obtained by using the two-grade ranking approach. In the complete ranking result, several objects (i.e., x_5 and x_8, x_9 and x_{10}) which are put into the same place according to the entire dominance degree have been further ranked. Simultaneously, the property of rank preservation is still remained, that is to say, each of objects x_5 and x_8 also dominates each of objects x_9 and x_{10} in the complete rank. Obviously, the effectivity of the two-grade ranking approach has been verified by Example 3.2. In fact, the directional distance index $DDI_A(x_i,\ x_j)$ assumes that every attribute has the same weight of $w_a=\frac{1}{|A|},\forall a\in A$. However, this assumption is only presented under special circumstances in practical decision issues. This problem will be addressed in next section.

4. The two-grade ranking approach with weights

In many situations, the significance of every attribute is often not equal to each other. This implies an important problem of how to determine the weight of each attribute for more rational decision making. Under this consideration, the directional distance index $DDI_A(x_i,x_j)$ mentioned by Section 3 can be seen as the index with equal weights. In this section, we continue to develop the version of the directional distance index considering weights.

For information systems (IS), entropy of a system is a useful mechanism for characterizing the information content, which is defined by Shannon [38], and gives a measure of uncertainty about its actual structure. Several authors [1,4,10,19–23,43] have used Shannon's concept and its extension to measure uncertainty in rough set theory. On this basis, mutual information induced by information entropy has also been proved that is an effective tool for measuring attributes importance [10]. In the following, through substituting a dominance relation for an indiscernibility relation, we will propose a new concept of ordered mutual information, which considers complement behavior of information gain. It can be used to modify the directional distance index $DDI_A(x_i,x_j)$ for more rationally ranking interval data.

In what follows, we briefly review several related concepts of information entropy in rough set theory.

Definition 4.1 [19]. Let $S=(U,AT,V,f)$ be an IS and $A\subseteq AT$, $U/IND(A)=\{X_1,X_2,\ldots,X_m\}$, information entropy of A in rough set theory is defined as

$$E(A)=\sum_{i=1}^{m}\frac{|X_i|}{|U|}\frac{|X_i^c|}{|U|}=\sum_{i=1}^{m}\frac{|X_i|}{|U|}\left(1-\frac{|X_i|}{|U|}\right),$$

where X_i^c is the complement set of X_i, i.e., $X_i^c=U-X_i$, $|X_i|/|U|$ denotes the probability of X_i within the universe U, and $|X_i^c|/|U|$ is the probability of the complement set of X_i within the universe U.

Definition 4.2 [19]. Let $S=(U,AT,V,f)$ be an IS and A, $B\subseteq AT$, $U/IND(A)=\{X_1,X_2,\ldots,X_m\}$, and $U/IND(B)=\{Y_1,Y_2,\ldots,Y_n\}$. Then, conditional entropy $E(B|A)$ is defined by

$$E(B|A)=\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{|Y_j\cap X_i|}{|U|}\frac{\left|Y_j^c-X_i^c\right|}{|U|}$$

and mutual information $E(B;A)$ is defined as

$$E(B;A)=\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{|Y_j\cap X_i|}{|U|}\frac{|Y_j^c\cap X_i^c|}{|U|}.$$

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007

Remark 4.1. From Definition 4.2, it is clear that the mutual information based on an indiscernibility relation can measure the consistency of two partitions in the universe U. It is because that it depicts both the overlap degree of two equivalent classes (X_i and Y_j) and that of their complement sets (X_i^c and Y_j^c). Nevertheless, it is also obvious that the mutual information in Definition 4.2 cannot characterize ordered consistency. Therefore, we need to extend several definitions listed above to measure the uncertainty in ordered information systems effectively.

In fact, from Definition 4.1, we can directly induce a new information entropy in interval ordered information systems, which has been also mentioned in Xu's work [44].

Definition 4.3. Let $S=(U,AT,V,f)$ be an IOIS, $A \subseteq AT$, and $U/R_A^{\geqslant} = \{[x_1]_A^{\geqslant}, [x_2]_A^{\geqslant}, \ldots, [x_{|U|}]_A^{\geqslant}\}$. Then, interval ordered information entropy of A is defined as

$$E(A^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_A^{\geqslant}|}{|U|}\right).$$

In the following, we put forward the new definition of mutual information in interval ordered information systems.

Definition 4.4. Let $S=(U,AT,V,f)$ be an IOIS, A, $B \subseteq AT$, and $U/R_A^{\geqslant} = \{[x_1]_A^{\geqslant}, [x_2]_A^{\geqslant}, \ldots, [x_{|U|}]_A^{\geqslant}\}$, $U/R_B^{\geqslant} = \{[x_1]_B^{\geqslant}, [x_2]_B^{\geqslant}, \ldots, [x_{|U|}]_B^{\geqslant}\}$. Then, the joint entropy of $A^{\geqslant} \cup B^{\geqslant}$ is defined as

$$E(A^{\geqslant} \cup B^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right).$$

Property 4.1. *$E(A^{\geqslant} \cup B^{\geqslant})$ has the following properties:*

(1) $E(A^{\geqslant} \cup B^{\geqslant}) \geqslant E(A^{\geqslant})$;
(2) $E(A^{\geqslant} \cup B^{\geqslant}) \geqslant E(B^{\geqslant})$.

Proof. They are straightforward. □

Corollary 4.1. *Let $S=(U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$. If $B^{\geqslant} \preccurlyeq A^{\geqslant}$, then $E(A^{\geqslant} \cup B^{\geqslant}) = E(B^{\geqslant})$.*

Definition 4.5. Let $S=(U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$, the conditional entropy of $A^{\geqslant}$ with respect to $B^{\geqslant}$ is defined as

$$E(A^{\geqslant}|B^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(\frac{|[x_i]_B^{\geqslant}|}{|U|} - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right).$$

Property 4.2. *Let $S=(U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$. Then, $E(A^{\geqslant}|B^{\geqslant}) = E(A^{\geqslant} \cup B^{\geqslant}) - E(B^{\geqslant})$.*

Proof. From Definition 4.5, we have that

$$\begin{aligned}
E(A^{\geqslant}|B^{\geqslant}) &= \sum_{i=1}^{|U|} \frac{1}{|U|}\left(\frac{|[x_i]_B^{\geqslant}|}{|U|} - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right) \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}\left(\frac{|[x_i]_B^{\geqslant}|}{|U|} - 1 + 1 - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right) \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}\left[\left(1 - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right) - \left(1 - \frac{|[x_i]_B^{\geqslant}|}{|U|}\right)\right] \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right) - \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_B^{\geqslant}|}{|U|}\right) \\
&= E(A^{\geqslant} \cup B^{\geqslant}) - E(B^{\geqslant}).
\end{aligned}$$

This completes the proof. □

Corollary 4.2. *Let $S=(U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$. If $B^{\geqslant} \preccurlyeq A^{\geqslant}$, then $E(A^{\geqslant}|B^{\geqslant}) = 0$.*

Corollary 4.3. *Let $S=(U,AT,V,f)$ be an IOIS and A,B, $D \subseteq AT$. If $D^{\geqslant} \preccurlyeq B^{\geqslant} \preccurlyeq A^{\geqslant}$, then $E(D^{\geqslant}|B^{\geqslant}) \leqslant E(D^{\geqslant}|A^{\geqslant})$.*

Definition 4.6. Let $S=(U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$. Then, *ordered mutual information* between $A^{\geqslant}$ and $B^{\geqslant}$ is defined as

$$E(A^{\geqslant}; B^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]_A^{\geqslant c} \cap [x_i]_B^{\geqslant c}|}{|U|}.$$

Property 4.3. *Let $S=(U,AT,V,f)$ be an IOIS and A, $B \subseteq AT$. Then, $E(A^{\geqslant}; B^{\geqslant}) = E(A^{\geqslant}) - E(A^{\geqslant}|B^{\geqslant}) = E(B^{\geqslant}) - E(B^{\geqslant}|A^{\geqslant})$.*

Proof. From Definition 4.6, we have that

$$\begin{aligned}
E(A^{\geqslant}; B^{\geqslant}) &= \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]_A^{\geqslant c} \cap [x_i]_B^{\geqslant c}|}{|U|} = \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|U - [x_i]_A^{\geqslant} \cup [x_i]_B^{\geqslant}|}{|U|} \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|U| - |[x_i]_A^{\geqslant}| - |[x_i]_B^{\geqslant}| + |[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|} \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}\left[\left(1 - \frac{|[x_i]_A^{\geqslant}|}{|U|}\right) - \left(\frac{|[x_i]_B^{\geqslant}|}{|U|} - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right)\right] \\
&= \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_A^{\geqslant}|}{|U|}\right) - \sum_{i=1}^{|U|} \frac{1}{|U|}\left(\frac{|[x_i]_B^{\geqslant}|}{|U|} - \frac{|[x_i]_A^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|U|}\right) \\
&= E(A^{\geqslant}) - E(A^{\geqslant}|B^{\geqslant}).
\end{aligned}$$

Similarly, $E(A^{\geqslant}; B^{\geqslant}) = E(B^{\geqslant}) - E(B^{\geqslant}|A^{\geqslant})$ also can be proved. This completes the proof. □

Property 4.4. *Let $S=(U,AT,V,f)$ be an IOIS and A,$B \subseteq D \subseteq AT$. If $|[x_i]_D^{\geqslant c} \cap [x_i]_B^{\geqslant c}| \geqslant |[x_i]_D^{\geqslant c} \cap [x_i]_A^{\geqslant c}|$, then $\frac{|[x_i]_D^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|[x_i]_B^{\geqslant}|} \geqslant \frac{|[x_i]_D^{\geqslant} \cap [x_i]_A^{\geqslant}|}{|[x_i]_A^{\geqslant}|}$.*

Proof. From the term (2) in Property 2.3, we have that $[x_i]_D^{\geqslant} \subseteq [x_i]_A^{\geqslant}$, $[x_i]_D^{\geqslant} \subseteq [x_i]_B^{\geqslant}$ and $[x_i]_D^{\geqslant c} \supseteq [x_i]_A^{\geqslant c}$, $[x_i]_D^{\geqslant c} \supseteq [x_i]_B^{\geqslant c}$. Thus,

$$\begin{aligned}
\frac{|[x_i]_D^{\geqslant} \cap [x_i]_B^{\geqslant}|}{|[x_i]_B^{\geqslant}|} &= \frac{|[x_i]_D^{\geqslant}|}{|[x_i]_B^{\geqslant}|} = \frac{|[x_i]_D^{\geqslant}|}{|U| - |[x_i]_B^{\geqslant c}|} = \frac{|[x_i]_D^{\geqslant}|}{|U| - |[x_i]_D^{\geqslant c} \cap [x_i]_B^{\geqslant c}|} \\
&\geqslant \frac{|[x_i]_D^{\geqslant}|}{|U| - |[x_i]_D^{\geqslant c} \cap [x_i]_A^{\geqslant c}|} = \frac{|[x_i]_D^{\geqslant}|}{|U| - |[x_i]_A^{\geqslant c}|} = \frac{|[x_i]_D^{\geqslant}|}{|[x_i]_A^{\geqslant}|} \\
&= \frac{|[x_i]_D^{\geqslant} \cap [x_i]_A^{\geqslant}|}{|[x_i]_A^{\geqslant}|}.
\end{aligned}$$

This completes the proof. □

Corollary 4.4. *Let $S=(U,AT,V,f)$ be an IOIS and $A \subseteq D \subseteq AT$. Then,*

$$\frac{|[x_i]_D^{\geqslant}|}{|U|} \leqslant \frac{|[x_i]_D^{\geqslant} \cap [x_i]_A^{\geqslant}|}{|[x_i]_A^{\geqslant}|} \leqslant 1,$$

where $\frac{|[x_i]_D^{\geqslant} \cap [x_i]_A^{\geqslant}|}{|[x_i]_A^{\geqslant}|} = \frac{|[x_i]_D^{\geqslant}|}{|U|}$ *if* $|[x_i]_D^{\geqslant c} \cap [x_i]_A^{\geqslant c}| = 0$, *and* $\frac{|[x_i]_D^{\geqslant} \cap [x_i]_A^{\geqslant}|}{|[x_i]_A^{\geqslant}|} = 1$ *if* $[x_i]_D^{\geqslant c} = [x_i]_A^{\geqslant c}$.

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007

Remark 4.2. In this section, we aim to depict importance of criteria through the ordered mutual information that considers complement behavior of information gain in interval ordered information systems. From Definition 3.5 and Example 3.2 , the entire dominance degree $D_A(x_i)$ is the first grade, which can be used to induce a cautious ranking result. Here, we call the cautious ranking result as the prior rank, which is ensured by the dominance relation $R_A^{\geqslant}$. Therefore, we argue that the criterion a ($a \in A$) is more important when the rank induced by $R_a^{\geqslant}$ is more consistent with the prior rank. In essence, this can be characterized by the consistency $\frac{|[x_i]_A^{\geqslant} \cap [x_i]_a^{\geqslant}|}{|[x_i]_a^{\geqslant}|}$ between $[x_i]_a^{\geqslant}$ and $[x_i]_A^{\geqslant}$. According to Property 4.4, the bigger $|[x_i]_A^{\geqslant c} \cap [x_i]_a^{\geqslant c}|$, the larger $\frac{|[x_i]_A^{\geqslant} \cap [x_i]_a^{\geqslant}|}{|[x_i]_a^{\geqslant}|}$. From Corollary 4.4, one knows that the consistency of $[x_i]_A^{\geqslant}$ and $[x_i]_a^{\geqslant}$ reaches the maximum when $[x_i]_A^{\geqslant c}$ equals $[x_i]_a^{\geqslant c}$, and vice versa. So, $E(a^{\geqslant}; A^{\geqslant})$ can measure the consistency between $[x_i]_a^{\geqslant}$ and $[x_i]_A^{\geqslant}$. From the analysis, we draw the conclusion that the ordered mutual information with complement behavior of information gain is suitable to measure the importance of criterion a, $\forall a \in A$.

Next, on the basis of Definition 4.6, we will give a new definition of $DDI_A^*(x_i, x_j)$ and propose a two-grade ranking approach with criterion weights in interval ordered information systems.

Definition 4.7. *Directional distance index* $DDI_A^*(x_i, x_j)$ which considered weights of criteria between two objects under attributes set $A(A \subseteq AT)$ is defined as

$$DDI_A^*(x_i, x_j) = \sum_{\forall a \in A} \frac{E(a^{\geqslant}; A^{\geqslant})}{\sum\limits_{\forall a \in A} E(a^{\geqslant}; A^{\geqslant})} DDI_a(x_i, x_j).$$

Here, $DDI_A^*(x_i, x_j)$ has the same properties with $DDI_A(x_i, x_j)$.

Definition 4.8. Let $S = (U, AT, V, f)$ be an IOIS and $A \subseteq AT$. *Entire directional distance index* with criterion weights is defined as

$$DDI_A^*(x_i) = \frac{1}{|U| - 1} \sum_{j \neq i} DDI_A^*(x_i, x_j), \quad x_i, \quad x_j \in U.$$

Analogously, a two-grade ranking approach considering weights of criteria can be established.

Definition 4.9. Let $S = (U, AT, V, f)$ be an IOIS and $A \subseteq AT$, x_i, $x_j \in U$.

If $D_A(x_i) \geqslant D_A(x_j)$, then x_i dominates x_j, denoted by $x_i \succcurlyeq x_j$;
If $D_A(x_i) \leqslant D_A(x_j)$, then x_i is dominated by x_j, denoted by $x_i \preccurlyeq x_j$;
If $D_A(x_i) = D_A(x_j)$, then

(1) If $DDI_A^*(x_i) \geqslant DDI_A^*(x_j)$, then x_i dominates x_j, denoted by $x_i \succcurlyeq x_j$;
(2) If $DDI_A^*(x_i) \leqslant DDI_A^*(x_j)$, then x_i is dominated by x_j, denoted by $x_i \preccurlyeq x_j$;
(3) If $DDI_A^*(x_i) = DDI_A^*(x_j)$, then x_i is the same as x_j, denoted by $x_i = x_j$.

Definition 4.9 gives the two-grade ranking approach with weights, in which the first grade is the same as the two-grade ranking approach in Section 3, while the second grade is modified. The modified entire directional distance index $DDI_A^*(x_i)$ is constructed by the directional distance index $DDI_A^*(x_i, x_j)$ considering the importance of criteria. It is obvious that this approach not only meets with the motivations of this paper but also takes the practical decision situations into account. Therefore, a more reasonable rank regarding decision behaviors of decision makers can be acquired.

The modified ranking approach can be designed as the following algorithm.

Algorithm 1. A modified two-grade ranking approach for interval data (TGRA).

Input: Decision table $S = (U, AT, V, f)$;
Output: The ranked array I.
Step 1: $I \leftarrow [x_1, x_2, \ldots, x_{|U|}]$; // Initializing the array I, $I(1) = x_1$, $I(2) = x_2, \ldots, I(|U|) = x_{|U|}$
Step 2: For $i = 1$ to $|U|$
{for $j = 1$ to $|U|$
{compute $D_A(I(i)) = \frac{1}{|U|-1} \sum_{j \neq i} D_A(I(i), I(j))$}};
Step 3: For $i = 1$ to $|U|$
{for $j = 1$ to $|U|$
{if $D_A(I(i)) < D_A(I(j))$
$x = I(i)$, $I(i) = I(j)$, $I(j) = x$; //x is a temporary variable
if $D_A(I(i)) = D_A(I(j))$
compute $DDI_A^*(x_i) = \frac{1}{|U|-1} \sum_{j \neq i} DDI_A^*(x_i, x_j)$}};
Step 4: For $i = 1$ to $|U|$
{for $j = 1$ to $|U|$
{if $D_A(I(i)) = D_A(I(j)) \& DDI_A^*(I(i)) < DDI_A^*(I(j))$
$x = I(i)$, $I(i) = I(j)$, $I(j) = x$}}; //x is a temporary variable
Step 5: Return I and end.

Example 4.1 (*Continued from Example 3.1*). Rank all objects in U according to the two-grade ranking approach with weights.

Based on Example 3.1, we rank those objects in U which cannot be ranked by the entire dominance degree.

Firstly, according to the definition of $E(a_m^{\geqslant}; AT^{\geqslant})$, we have that

$$\begin{aligned} E(a_1^{\geqslant}; AT^{\geqslant}) &= \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]_{a_1}^{\geqslant c} \cap [x_i]_{AT}^{\geqslant c}|}{|U|} \\ &= \frac{1}{10}\left(\frac{5}{10} + \frac{1}{10} + \frac{1}{10} + 0 + \frac{8}{10} + \frac{5}{10} + \frac{4}{10} + \frac{6}{10} + \frac{6}{10} + \frac{8}{10}\right) \\ &= \frac{44}{100} = \frac{11}{25}, \end{aligned}$$

$$\begin{aligned} E(a_2^{\geqslant}; AT^{\geqslant}) &= \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]_{a_2}^{\geqslant c} \cap [x_i]_{AT}^{\geqslant c}|}{|U|} \\ &= \frac{1}{10}\left(\frac{3}{10} + 0 + 0 + 0 + \frac{5}{10} + \frac{5}{10} + \frac{4}{10} + \frac{5}{10} + \frac{8}{10} + \frac{8}{10}\right) \\ &= \frac{38}{100} = \frac{19}{50}, \end{aligned}$$

$$\begin{aligned} E(a_3^{\geqslant}; AT^{\geqslant}) &= \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]_{a_3}^{\geqslant c} \cap [x_i]_{AT}^{\geqslant c}|}{|U|} \\ &= \frac{1}{10}\left(\frac{4}{10} + \frac{1}{10} + \frac{1}{10} + 0 + \frac{8}{10} + \frac{5}{10} + \frac{4}{10} + \frac{6}{10} + \frac{6}{10} + \frac{8}{10}\right) \\ &= \frac{43}{100}, \end{aligned}$$

$$\begin{aligned} E(a_4^{\geqslant}; AT^{\geqslant}) &= \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]_{a_4}^{\geqslant c} \cap [x_i]_{AT}^{\geqslant c}|}{|U|} \\ &= \frac{1}{10}\left(\frac{4}{10} + 0 + \frac{4}{10} + 0 + \frac{7}{10} + \frac{7}{10} + \frac{4}{10} + \frac{9}{10} + \frac{6}{10} + \frac{2}{10}\right) \\ &= \frac{43}{100}, \end{aligned}$$

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007

ARTICLE IN PRESS

$$
\begin{aligned}
E(a_5^{\geqslant};AT^{\geqslant}) &= \sum_{i=1}^{|U|}\frac{1}{|U|}\frac{|[x_i]_{a_5}^{\geqslant c}\cap[x_i]_{AT}^{\geqslant c}|}{|U|}\\
&=\frac{1}{10}\left(\frac{3}{10}+0+0+0+\frac{5}{10}+\frac{5}{10}+\frac{4}{10}+\frac{5}{10}+\frac{8}{10}+\frac{8}{10}\right)\\
&=\frac{38}{100}=\frac{19}{50}
\end{aligned}
$$

Then, $DDI^*_{AT}(x_i)$ can be calculated

$$
\begin{aligned}
DDI^*_{AT}(x_5) = DDI^*_{AT}(x_5,x_8) &= \sum_{m=1}^{5}\frac{E(a_m^{\geqslant};AT^{\geqslant})}{\sum_{m=1}^{5}E(a_m^{\geqslant};AT^{\geqslant})}DDI_{a_m}(x_5,x_8)\\
&=\frac{0.44}{2.06}\times\frac{5}{8}+\frac{0.38}{2.06}\times\frac{1}{2}+\frac{0.43}{2.06}\times\frac{5}{8}+\frac{0.43}{2.06}\times\frac{3}{8}+\frac{0.38}{2.06}\times\frac{1}{2}\\
&=0.53,
\end{aligned}
$$

$$
\begin{aligned}
DDI^*_{AT}(x_8) = DDI^*_{AT}(x_8,x_5) &= 1-DDI^*_{AT}(x_5,x_8)=0.47<0.53\\
&=DDI^*_{AT}(x_5),
\end{aligned}
$$

$$
\begin{aligned}
DDI^*_{AT}(x_9) = DDI^*_{AT}(x_9,x_{10}) &= \sum_{m=1}^{5}\frac{E(a_m^{\geqslant};AT^{\geqslant})}{\sum\limits_{m=1}^{5}E(a_m^{\geqslant};AT^{\geqslant})}DDI_{a_m}(x_9,x_{10})\\
&=\frac{0.44}{2.06}\times\frac{3}{8}+\frac{0.38}{2.06}\times\frac{1}{2}+\frac{0.43}{2.06}\times\frac{3}{8}+\frac{0.43}{2.06}\times\frac{5}{8}+\frac{0.38}{2.06}\times\frac{1}{2}\\
&=0.47,
\end{aligned}
$$

$$
\begin{aligned}
DDI^*_{AT}(x_{10}) = DDI^*_{AT}(x_{10},x_9) &= 1-DDI^*_{AT}(x_9,x_{10})=0.53>0.47\\
&=DDI^*_{AT}(x_9).
\end{aligned}
$$

Therefore, the ranking result can be obtained as follows

$x_5 \succcurlyeq x_8 \succcurlyeq x_{10} \succcurlyeq x_9 \succcurlyeq x_7 \succcurlyeq x_6 \succcurlyeq x_1 \succcurlyeq x_3 \succcurlyeq x_2 \succcurlyeq x_4.$

From Example 4.1, a complete rank of objects can also be obtained. Obviously, according to the two-grade ranking approach with weights, object x_5 and object x_8 in the original rank have been distinguished, and objects x_9 and x_{10} have the similar result. Moreover, the property of rank preservation is also kept. In fact, from the perspective of decision-making situations, determining the weight of each criterion is a crucial step for more rational ranking decision. Hence, the order mutual information $E(a_m^{\geqslant};AT^{\geqslant})$ with complement behavior of information gain is introduced to measure the importance of criterion a_m. This approach to determining weights can reflect both the characteristics of data sets and decision behaviors of decision makers, which seems more rational and comprehensive.

5. Case study

Stock selection is an important research issue in financial field. In essence, the problem of stock selection is to rank alternatives on the basis of their values of some indicators. The purpose of this section is, through an actual issue of stock selection, to verify the effectivity of the proposed two-grade ranking approach.

In the case study, we will employ twenty stocks x_i ($i = 1,2,\ldots,20$) from the tourism sector in Chinese A-share stock markets. We select three attributes for evaluating these stocks, which are earnings per share, book-to-market equity, and total assets turnover, denoted by a_m ($m = 1,2,3$). Among these three indicators, earnings per share and total assets turnover reflect the profitability and operational capacity of enterprises, while book-to-market equity is also recognized as an important factor which is positive correlation with stock returns. Therefore, we will use these indicators for this case study.

Generally, the attribute values of each stock are depicted by a numerical number. However, it is difficult for us to analyze the range of the value of objects under each attribute. In order to better reveal the entirety of a data set, we can adopt such a strategy to transform a single-valued data set into an interval-valued data set based on the idea of data-packaging. Here, we will make use of quarterly data of each stock from 2008 to 2009, where the minimum value is denoted by $a_m^L(x_i)$ and the maximum value is wrote by $a_m^U(x_i)$ under every attribute. Through data processing, the interval ordered information system about stock selection is established, which is shown in Table 2, where $U = \{x_1,x_2,\ldots,x_{20}\}$ and $AT = \{a_1,a_2,a_3\}$.

Firstly, according to the definition of $[x_i]_{AT}^{\geqslant}$, we have that

$$
\begin{aligned}
&[x_1]_{AT}^{\geqslant}=\{x_1,x_4,x_{16}\},\\
&[x_2]_{AT}^{\geqslant}=\{x_2,x_4,x_7,x_{11},x_{13},x_{14},x_{15},x_{16},x_{17},x_{20}\},\\
&[x_3]_{AT}^{\geqslant}=\{x_3\},\ [x_4]_{AT}^{\geqslant}=\{x_4\},\quad [x_5]_{AT}^{\geqslant}=\{x_4,x_5,x_9,x_{14},x_{15},x_{16}\},\\
&[x_6]_{AT}^{\geqslant}=\{x_6,x_{14},x_{15},x_{20}\},\\
&[x_7]_{AT}^{\geqslant}=\{x_7,x_{14},x_{15}\},\quad [x_8]_{AT}^{\geqslant}=\{x_8,x_9,x_{14},x_{15},x_{16},x_{20}\},\\
&[x_9]_{AT}^{\geqslant}=\{x_9\},\quad [x_{10}]_{AT}^{\geqslant}=\{x_4,x_{10},x_{14},x_{15},x_{16},x_{20}\},\\
&[x_{11}]_{AT}^{\geqslant}=\{x_4,x_{11},x_{14},x_{15},x_{16}\},\\
&[x_{12}]_{AT}^{\geqslant}=\{x_{12}\},\quad [x_{13}]_{AT}^{\geqslant}=\{x_4,x_{13},x_{15}\},\\
&[x_{14}]_{AT}^{\geqslant}=\{x_{14}\},\quad [x_{15}]_{AT}^{\geqslant}=\{x_{15}\},\\
&[x_{16}]_{AT}^{\geqslant}=\{x_{16}\},\quad [x_{17}]_{AT}^{\geqslant}=\{x_{14},x_{16},x_{17},x_{20}\},\quad [x_{18}]_{AT}^{\geqslant}=\{x_{18}\},\\
&[x_{19}]_{AT}^{\geqslant}=\{x_4,x_9,x_{10},x_{13},x_{14},x_{15},x_{16},x_{19},x_{20}\},\quad [x_{20}]_{AT}^{\geqslant}=\{x_{20}\}.
\end{aligned}
$$

Therefore, one can obtain the dominance relation matrix as

$$
\begin{pmatrix}
1 & 0.95 & 0.85 & 0.9 & 0.95 & 0.85 & \cdots & 0.85 & 0.9 & 0.9 & 0.85 & 0.95 & 0.85\\
0.6 & 1 & 0.55 & 0.55 & 0.7 & 0.65 & \cdots & 0.55 & 0.55 & 0.7 & 0.5 & 0.8 & 0.55\\
0.95 & 0.95 & 1 & 0.95 & 0.95 & 0.95 & \cdots & 0.95 & 0.95 & 0.95 & 0.95 & 0.95 & 0.95\\
1 & 1 & 0.95 & 1 & 1 & 0.95 & \cdots & 0.95 & 0.95 & 0.95 & 0.95 & 1 & 0.95\\
0.8 & 0.9 & 0.7 & 0.75 & 1 & 0.8 & \cdots & 0.75 & 0.75 & 0.8 & 0.7 & 0.95 & 0.7\\
0.8 & 0.95 & 0.8 & 0.8 & 0.9 & 1 & \cdots & 0.85 & 0.8 & 0.9 & 0.8 & 0.95 & 0.85\\
0.85 & 1 & 0.85 & 0.85 & 0.95 & 0.95 & \cdots & 0.9 & 0.85 & 0.9 & 0.85 & 0.95 & 0.85\\
0.75 & 0.9 & 0.7 & 0.7 & 0.9 & 0.85 & \cdots & 0.75 & 0.75 & 0.85 & 0.7 & 0.95 & 0.75\\
0.95 & 0.95 & 0.95 & 0.95 & 1 & 0.95 & \cdots & 0.95 & 0.95 & 0.95 & 0.95 & 1 & 0.95\\
0.8 & 0.95 & 0.7 & 0.75 & 0.9 & 0.85 & \cdots & 0.75 & 0.75 & 0.85 & 0.7 & 1 & 0.75\\
0.85 & 1 & 0.75 & 0.8 & 0.95 & 0.85 & \cdots & 0.8 & 0.8 & 0.85 & 0.75 & 0.95 & 0.75\\
0.95 & 0.95 & 0.95 & 0.95 & 0.95 & 0.95 & \cdots & 0.95 & 0.95 & 0.95 & 0.95 & 0.95 & 0.95\\
0.9 & 1 & 0.85 & 0.9 & 0.95 & 0.9 & \cdots & 0.9 & 0.85 & 0.85 & 0.85 & 1 & 0.85\\
0.95 & 1 & 0.95 & 0.95 & 1 & 1 & \cdots & 0.95 & 0.95 & 1 & 0.95 & 1 & 0.95\\
0.95 & 1 & 0.95 & 0.95 & 1 & 1 & \cdots & 1 & 0.95 & 0.95 & 0.95 & 1 & 0.95\\
1 & 1 & 0.95 & 0.95 & 1 & 0.95 & \cdots & 0.95 & 1 & 1 & 0.95 & 1 & 0.95\\
0.85 & 1 & 0.8 & 0.8 & 0.9 & 0.9 & \cdots & 0.8 & 0.85 & 1 & 0.8 & 0.95 & 0.85\\
0.95 & 0.95 & 0.95 & 0.95 & 0.95 & 0.95 & \cdots & 0.95 & 0.95 & 0.95 & 1 & 0.95 & 0.95\\
0.65 & 0.85 & 0.55 & 0.6 & 0.8 & 0.7 & \cdots & 0.6 & 0.6 & 0.7 & 0.55 & 1 & 0.6\\
0.95 & 1 & 0.95 & 0.95 & 0.95 & 1 & \cdots & 0.95 & 0.95 & 1 & 0.95 & 1 & 1
\end{pmatrix}.
$$

Then, we can get that

$D_{AT}(x_1) = 0.8895,\quad D_{AT}(x_2) = 0.6237,\quad D_{AT}(x_3) = 0.95,$
$D_{AT}(x_4) = 0.9684,$

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007

$D_{AT}(x_5) = 0.7947, \quad D_{AT}(x_6) = 0.8632, \quad D_{AT}(x_7) = 0.9,$
$D_{AT}(x_8) = 0.7921,$

$D_{AT}(x_9) = 0.9579, \quad D_{AT}(x_{10}) = 0.8053,$
$D_{AT}(x_{11}) = 0.8395, \quad D_{AT}(x_{12}) = 0.95,$

$D_{AT}(x_{13}) = 0.8974, \quad D_{AT}(x_{14}) = 0.9737,$
$D_{AT}(x_{15}) = 0.9737, \quad D_{AT}(x_{16}) = 0.9711,$

$D_{AT}(x_{17}) = 0.8632, \quad D_{AT}(x_{18}) = 0.95,$
$D_{AT}(x_{19}) = 0.6684, \quad D_{AT}(x_{20}) = 0.9658.$

According to the definition of $D_{AT}(x_i)$, the prior rank can be obtained as follows

$$\begin{pmatrix} x_{14} \\ x_{15} \end{pmatrix} \succcurlyeq x_{16} \succcurlyeq x_4 \succcurlyeq x_{20} \succcurlyeq x_9 \succcurlyeq \begin{pmatrix} x_3 \\ x_{12} \\ x_{18} \end{pmatrix} \succcurlyeq x_7 \succcurlyeq x_{13} \succcurlyeq x_1 \succcurlyeq$$
$$\begin{pmatrix} x_6 \\ x_{17} \end{pmatrix} \succcurlyeq x_{11} \succcurlyeq x_{10} \succcurlyeq x_5 \succcurlyeq x_8 \succcurlyeq x_{19} \succcurlyeq x_2.$$

From the prior rank, one sees that several stocks cannot be ranked (seeing x_{14} and x_{15}, x_3, x_{12} and x_{18}, x_6 and x_{17}). As we know, stock selection is a typical issue of investment decision, in which limited capital is a major constraint. So a complete rank of alternatives is desirable for investors. However, the dominance degree $D_A(x_i,x_j)$ only gives the preferability degree of the object x_i over x_j through investigating the relative ranking position of objects, which is difficult to obtain a complete rank. Next, we will adopt the second grade $DDI^*_A(x_i)$ to rank these objects for a complete rank.

Here, we need to calculate the ordered mutual information $E(a^{\geqslant}_m;AT^{\geqslant})$ as weights of criteria.

$$E(a^{\geqslant}_1;AT^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]^{\geqslant c}_{a_1} \cap [x_i]^{\geqslant c}_{AT}|}{|U|} = \frac{1}{20}\left(\frac{15}{20}+\frac{8}{20}+\frac{18}{20}+\frac{17}{20}+\frac{7}{20} + \frac{8}{20}+\frac{10}{20}+\frac{5}{20}+\frac{16}{20}+\frac{11}{20}+\frac{14}{20}+\frac{12}{20}+\frac{12}{20}+\frac{18}{20}+\frac{14}{20} + \frac{19}{20}+\frac{15}{20}+\frac{17}{20}+\frac{9}{20}+\frac{17}{20}\right) = \frac{262}{400} = \frac{131}{200},$$

Table 2
An interval ordered information system about stock selection.

U	Stock code	a_1	a_2	a_3
x_1	000007	[−0.1577, 0.2254]	[0.2279, 0.6261]	[0.0488, 0.0854]
x_2	000033	[−0.0382, 0.0212]	[0.1971, 0.5813]	[0.0309, 0.0382]
x_3	000069	[0.0180, 0.3270]	[0.3758, 0.7145]	[0.0249, 0.3628]
x_4	000428	[0.0000, 0.2585]	[0.5779, 1.0334]	[0.0580, 0.1142]
x_5	000430	[−0.1600, 0.1000]	[0.3270, 0.6165]	[0.0254, 0.1140]
x_6	000524	[−0.0900, 0.0630]	[0.4382, 0.9639]	[0.0582, 0.0947]
x_7	000610	[−0.0321, 0.0285]	[0.4760, 0.9029]	[0.0614, 0.1285]
x_8	000802	[−0.1527, 0.0481]	[0.3513, 0.7635]	[0.0595, 0.0783]
x_9	000888	[−0.0539, 0.2142]	[0.4422, 0.7668]	[0.0757, 0.1813]
x_{10}	000978	[−0.0440, 0.1190]	[0.3856, 0.8101]	[0.0421, 0.0701]
x_{11}	002033	[0.0000, 0.1700]	[0.3411, 0.7464]	[0.0391, 0.0692]
x_{12}	002059	[−0.0300, 0.0861]	[0.7967, 1.1671]	[0.0216, 0.1968]
x_{13}	002159	[−0.0330, 0.1230]	[0.5325, 0.9263]	[0.0390, 0.1118]
x_{14}	600,138	[0.0403, 0.1772]	[0.4922, 0.9721]	[0.1830, 0.4405]
x_{15}	600,175	[0.0000, 0.1700]	[0.6648, 0.9762]	[0.1151, 0.2449]
x_{16}	600,258	[0.1044, 0.3731]	[0.3867, 0.9343]	[0.1563, 0.3060]
x_{17}	600,358	[0.0010, 0.0410]	[0.2819, 0.6104]	[0.0397, 0.0623]
x_{18}	600,593	[−0.1407, 0.2700]	[0.4161, 1.0364]	[0.0177, 0.1035]
x_{19}	600,749	[−0.0941, 0.0918]	[0.3569, 0.7586]	[0.0066, 0.0635]
x_{20}	601,007	[0.0290, 0.1390]	[0.6873, 1.0706]	[0.0644, 0.0978]

$$E(a^{\geqslant}_2;AT^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]^{\geqslant c}_{a_2} \cap [x_i]^{\geqslant c}_{AT}|}{|U|} = \frac{1}{20}\left(\frac{3}{20}+0+\frac{7}{20}+\frac{17}{20}+\frac{3}{20} + \frac{14}{20}+\frac{13}{20}+\frac{7}{20}+\frac{12}{20}+\frac{9}{20}+\frac{5}{20}+\frac{19}{20}+\frac{15}{20}+\frac{15}{20}+\frac{17}{20} + \frac{12}{20}+\frac{2}{20}+\frac{17}{20}+\frac{7}{20}+\frac{18}{20}\right) = \frac{212}{400} = \frac{53}{100},$$

$$E(a^{\geqslant}_3;AT^{\geqslant}) = \sum_{i=1}^{|U|} \frac{1}{|U|} \frac{|[x_i]^{\geqslant c}_{a_3} \cap [x_i]^{\geqslant c}_{AT}|}{|U|} = \frac{1}{20}\left(\frac{11}{20}+\frac{5}{20}+\frac{18}{20}+\frac{14}{20}+\frac{13}{20} + \frac{13}{20}+\frac{15}{20}+\frac{13}{20}+\frac{16}{20}+\frac{9}{20}+\frac{8}{20}+\frac{15}{20}+\frac{13}{20}+\frac{19}{20}+\frac{17}{20} + \frac{18}{20}+\frac{8}{20}+\frac{9}{20}+\frac{2}{20}+\frac{15}{20}\right) = \frac{251}{400}.$$

Then, we will rank such stocks that cannot be ranked by $D_{AT}(x_i)$.

(1) Let us rank the objects x_{14} and x_{15}.

$$\begin{aligned} DDI^*_{AT}(x_{14}) = DDI^*_{AT}(x_{14},x_{15}) &= \sum_{m=1}^{3} \frac{E(a^{\geqslant}_m;AT^{\geqslant})}{\sum_{m=1}^{3} E(a^{\geqslant}_m;AT^{\geqslant})} DDI_{a_m}(x_{14},x_{15}) \\ &= \frac{0.655}{1.8125}\left(\frac{1}{2}+\frac{1}{4}\frac{0.1772-0.17+0.0403-0}{0.3731-(-0.16)}\right) \\ &\quad + \frac{0.53}{1.8125}\left(\frac{1}{2}+\frac{1}{4}\frac{0.9721-0.9762+0.4922-0.6648}{1.1671-0.1971}\right) \\ &\quad + \frac{0.6275}{1.8125}\left(\frac{1}{2}+\frac{1}{4}\frac{0.4405-0.2449+0.183-0.1151}{0.4405-0.0066}\right) \\ &= 0.5473, \end{aligned}$$

$$\begin{aligned} DDI^*_{AT}(x_{15}) = DDI^*_{AT}(x_{15},x_{14}) &= 1 - DDI^*_{AT}(x_{14},x_{15}) \\ &= 1 - 0.5473 = 0.4527. \end{aligned}$$

Obviously, we can conclude $x_{14} \succcurlyeq x_{15}$.

(2) Let us rank the objects x_3, x_{12} and x_{18}. Through the definition of $DDI^*_{AT}(x_i,x_j)$, one can obtain the directional distance index relation matrix as

$$\begin{pmatrix} 0.5 & 0.5169 & 0.5624 \\ 0.4831 & 0.5 & 0.5455 \\ 0.4376 & 0.4545 & 0.5 \end{pmatrix}.$$

From the above matrix, one can easily obtain the following result

$$DDI^*_{AT}(x_3) = \frac{1}{2}\left[DDI^*_{AT}(x_3,x_{12}) + DDI^*_{AT}(x_3,x_{18})\right] = 0.5397,$$

$$DDI^*_{AT}(x_{12}) = \frac{1}{2}\left[DDI^*_{AT}(x_{12},x_3) + DDI^*_{AT}(x_{12},x_{18})\right] = 0.5143,$$

$$DDI^*_{AT}(x_{18}) = \frac{1}{2}\left[DDI^*_{AT}(x_{18},x_3) + DDI^*_{AT}(x_{18},x_{12})\right] = 0.4461,$$

Thus we can conclude $x_3 \succcurlyeq x_{12} \succcurlyeq x_{18}$.

(3) Let us rank the objects x_6 and x_{17}.

$$DDI^*_{AT}(x_6) = DDI^*_{AT}(x_6,x_{17}) = 0.5369,$$

$$\begin{aligned} DDI^*_{AT}(x_{17}) = DDI^*_{AT}(x_{17},x_6) = 1 - DDI^*_{AT}(x_6,x_{17}) &= 1 - 0.5369 \\ &= 0.4631. \end{aligned}$$

Thus we can conclude $x_6 \succcurlyeq x_{17}$.

Therefore, according to Definition 3.5 and Algorithm 1, we can get a complete ranking result as follows

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007

$x_{14} \succcurlyeq x_{15} \succcurlyeq x_{16} \succcurlyeq x_4 \succcurlyeq x_{20} \succcurlyeq x_9 \succcurlyeq x_3 \succcurlyeq x_{12} \succcurlyeq x_{18} \succcurlyeq x_7 \succcurlyeq x_{13} \succcurlyeq$
$x_1 \succcurlyeq x_6 \succcurlyeq x_{17} \succcurlyeq x_{11} \succcurlyeq x_{10} \succcurlyeq x_5 \succcurlyeq x_8 \succcurlyeq x_{19} \succcurlyeq x_2.$

In this case, we try to completely rank all stocks for a more satisfactory investment decision. However, there are seven stocks cannot be ranked through the first grade $D_A(x_i)$ in the entire stocks. Then, the second grade $DDI_A^*(x_i)$ is employed for measuring more detailed difference of objects, where each weight of criteria is attained by the presented ordered mutual information. Clearly, using the developed ranking approach, we have obtained a complete ranking result of the stocks.

6. Conclusions

As to ranking decision, existing approaches cannot obtain the complete rank of objects with interval values. We want to overcome this shortcoming based on the three viewpoints: (1) one keeps the rank induced by original entire dominance degree unchanged; (2) the objects lying in the same place in this rank should be further ranked; and (3) the difference in-between objects in the same place can be characterized by their values under every attribute. Taking these three viewpoints into account, in this paper, we first have proposed an entire directional distance index, which can be used to further distinguish the difference among objects with the same place in the rank induced by Qian's approach. Based on this proposed index, we have then developed a two-grade approach to ranking objects with interval values, which can obtain a complete rank without destroying the rank induced by Qian's approach. In order to rank objects more rationally, we have also presented a concept of ordered mutual information, which can be used to calculate the weight of each criterion. The two-grade ranking approach with weights is a much better choice to ranking objects. Finally, we have employed a case about stock selection for verifying the effecitivity of the proposed two-grade ranking approach. Results show that the proposed two-grade approach is much better than the original one to ranking objects with interval values.

Acknowledgements

The authors wish to thank the anonymous reviewers and Editor-in-Chief for their constructive comments and suggestions.

This work was partially supported by the National Natural Science Foundation of China (Nos. 71031006, 70971080, and 60903110), Soft Science Research Project of Shanxi Province of China (No. 2010041054-01) and research project supported by Shanxi Scholarship Council of China (No. 201008).

References

[1] T. Beaubouef, F.E. Petry, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109 (1–4) (1998) 185–195.
[2] P. Bermejo, Luis de la Ossa, José A. Gámez, José M. Puerta, Fast wrapper feature subset selection in high-dimensional datasets by means of filter re-ranking, Knowledge-Based Systems (2011), doi:10.1016/j.knosys.2011.01.015.
[3] C.T. Chen, A fuzzy approach to select the location of the distribution center, Fuzzy Sets and Systems 118 (1) (2001) 65–73.
[4] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1) (1998) 109–137.
[5] S. Greco, B. Matarazzo, R. Slowinski, A new rough set approach to multicriteria and multiattribute classification, Lecture Notes in Artificial Intelligence 1424 (1998) 60–67.
[6] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129 (1) (2001) 1–47.
[7] S. Greco, B. Matarazzo, R. Slowinski, J. Stefanowski, An algorithm for induction of decision rules consistent with dominance principle, Lecture Notes in Artificial Intelligence 2005 (2001) 304–313.
[8] S. Greco, B. Matarazzo, R. Slowinski, Rough sets methodology for sorting problems in presence of multiple attributes and criteria, European Journal of Operational Research 138 (2) (2002) 247–259.
[9] E. Hermans, D. Ruan, T. Brijs, G. Wets, K. Vanhoof, Road safety risk evaluation by means of ordered weighted averaging operators and expert knowledge, Knowledge-Based Systems 23 (1) (2010) 48–52.
[10] Q.H. Hu, M.Z. Guo, D.R. Yu, J.F. Liu, Information entropy for ordinal classification, Science in China F: Information Sciences 53 (6) (2010) 1188–1200.
[11] Q.H. Hu, D.R. Yu, M.Z. Guo, Fuzzy preference based rough sets, Information Sciences 180 (10) (2010) 2003–2022.
[12] Q.H. Hu, D.R. Yu, C.X. Wu, Fuzzy preference relation rough sets, in: IEEE Conference on Granular Computing, 2008, pp. 300–305.
[13] C.L. Hwang, K. Yoon, Multiple Attribute Decision Making: Methods and Applications, Springer-Verlag, Berlin, 1981.
[14] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112 (1–4) (1998) 39–49.
[15] M. Kryszkiewicz, Rules in incomplete information systems, Information Sciences 113 (3–4) (1999) 271–292.
[16] L.C. Leung, D. Cao, On consistency and ranking of alternatives in fuzzy AHP, European Journal of Operational Research 124 (1) (2000) 102–113.
[17] Y. Leung, D.Y. Li, Maximal consistent block technique for rule acquisition in incomplete information systems, Information Sciences 153 (2003) 85–106.
[18] Y. Leung, M.M. Fischer, W.Z. Wu, J.S. Mi, A rough set approach for the discovery of classification rules in interval-valued information systems, International Journal of Approximate Reasoning 47 (2) (2008) 233–246.
[19] J.Y. Liang, K.S. Chin, C.Y. Dang, Richard C.M. YAM, A new method for measuring uncertainty and fuzziness in rough set theory, International Journal of General Systems 31 (4) (2002) 331–342.
[20] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (1) (2004) 37–46.
[21] J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wireman, The information entropy, rough entropy and knowledge granulation in incomplete information systems, International Journal of General Systems 34 (1) (2006) 641–654.
[22] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Science in China F: Information Sciences 38 (12) (2008) 2048–2065.
[23] J.Y. Liang, J.H. Wang, Y.H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Information Sciences 17 (9) (2009) 458–470.
[24] Z. Pawlak, Rough sets: theoretical aspects of reasoning about data, system theory, Knowledge Engineering and Problem Solving, vol. 9, Kluwer, Dordrecht, 1991.
[25] Z. Pawlak, Rough set theory and its applications to data analysis, Cybernetics and Systems 29 (7) (1998) 661–688.
[26] Z. Pawlak, Rough sets and intelligent data analysis, Information Sciences 147 (1–4) (2002) 1–12.
[27] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (1) (2007) 3–27.
[28] Y.H. Qian, J.Y. Liang, C.Y. Dang, Interval ordered information systems, Computers and Mathematics with Applications 56 (8) (2008) 1994–2009.
[29] Y.H. Qian, J.Y. Liang, D.Y. Li, F. Wang, N.N. Ma, Approximation reduction in inconsistent incomplete decision tables, Knowledge-Based Systems 23 (5) (2010) 427–433.
[30] Y.H. Qian, J.Y. Liang, D.Y. Li, H.Y. Zhang, C.Y. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, Information Sciences 178 (1) (2008) 181–202.
[31] Y.H. Qian, J.Y. Liang, P. Song, C.Y. Dang, On dominance relations in disjunctive set-valued ordered information systems, International Journal of Information Technology & Decision Making 9 (1) (2010) 9–33.
[32] A. Rad, B. Naderi, M. Soltani, Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran, Expert Systems with Applications 38 (1) (2011) 755–763.
[33] B. Roy, Classement et choix en présence de points de vue multiples: la méthode ELECTRE, Revue Francaise d'Informatique et de Recherche Opérationnelle 8 (1968) 57–75.
[34] Y. Sai, Y.Y. Yao, N. Zhong, Data analysis and mining in ordered information tables, in: Proceedings of the 2001 IEEE International Conference on Data Mining, America, 2001, pp. 497–504.
[35] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, New York, 1980.
[36] S. Senthamarai Kannan, N. Ramaraj, A novel hybrid feature selection via symmetrical uncertainty ranking based local memetic search algorithm, Knowledge-Based Systems 23 (6) (2010) 580–585.
[37] A. Shanian, A.S. Milani, C. Carson, R.C. Abeyaratne, A new application of ELECTRE III and revised Simos' procedure for group material selection under weighting uncertainty, Knowledge-Based Systems 21 (7) (2008) 709–720.
[38] C.E. Shannon, The mathematical theory of communication, The Bell System Technical Journal 27 (3–4) (1948) 373–423. See also pp. 623–656.
[39] F. Tiryaki, M. Ahlatcioglu, Fuzzy stock selection using a new fuzzy ranking and weighting algorithm, Applied Mathematics and Computation 170 (1) (2005) 144–157.
[40] C.Z. Wang, D.G. Chen, Q.H. Hu, Some invariant properties of ordered information systems under homomorphism, Science in China F: Information Sciences 53 (9) (2010) 1816–1825.
[41] G.W. Wei, Maximizing deviation method for multiple attribute decision making in intuitionistic fuzzy setting, Knowledge-Based Systems 21 (8) (2008) 833–836.

[42] W. Wei, J.Y. Liang, Y.H. Qian, F. Wang, C.Y. Dang, Comparative study of decision performance of decision tables induced by attribute reductions, International Journal of General Systems 39 (8) (2010) 813–838.
[43] M.J. Wierman, Measuring uncertainty in rough set theory, International Journal of General Systems 28 (4) (1999) 283–297.
[44] W.H. Xu, X.Y. Zhang, W.X. Zhang, Knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems, Applied Soft Computing 9 (4) (2009) 1244–1251.
[45] Z.S. Xu, Uncertain Multiple Attribute Decision Making: Methods and Applications, TsingHua Press, Beijing, China, 2004.
[46] Z.S. Xu, Q.L. Da, The uncertain OWA operator, International Journal of Intelligent Systems 17 (6) (2002) 569–575.
[47] X.B. Yang, J.Y. Yang, C. Wu, D.J. Yu, Dominance-based rough set approach and knowledge reductions in incomplete ordered information system, Information Sciences 178 (4) (2008) 1219–1234.
[48] X.B. Yang, J. Xie, X.N. Song, J.Y. Yang, Credible rules in incomplete decision system based on descriptors, Knowledge-Based Systems 22 (1) (2009) 8–17.
[49] X.B. Yang, D.J. Yu, J.Y. Yang, L.H. Wei, Dominance-based rough set approach to incomplete interval-valued information system, Data and Knowledge Engineering 68 (11) (2009) 1331–1347.
[50] Y.Y. Yao, Y. Sai, Mining ordering rules using rough set theory, Bulletin of International Rough Set Society 5 (1–2) (2001) 99–106.
[51] Z.L. Yue, An extended TOPSIS for determining weights of decision makers with interval numbers, Knowledge-Based Systems 24 (1) (2011) 146–153.
[52] M. Zerafat Angiz L., A. Emrouznejad, A. Mustafa, A.S. Al-Eraqi. Aggregating preference ranking with fuzzy Data Envelopment Analysis, Knowledge-Based Systems 23 (6) (2010) 512–519.
[53] W.X. Zhang, G.F. Qiu, Uncertain Decision Making Based on Rough Sets, TsingHua Press, Beijing, China, 2005.

Please cite this article in press as: P. Song et al., A two-grade approach to ranking interval data, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.10.007