Knowledge-Based Systems 91 (2016) 126–142

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Fuzzy probabilistic rough sets and their corresponding three-way decisions

Xue Rong Zhao [a,b], Bao Qing Hu [a,b,*]

[a] School of Mathematics and Statistics, Wuhan University, Wuhan 430072, PR China
[b] Computational Science Hubei Key Laboratory, Wuhan University, Wuhan 430072, PR China

ARTICLE INFO

Article history:
Received 20 October 2014
Revised 16 September 2015
Accepted 16 September 2015
Available online 30 September 2015

Keywords:
Rough set
Fuzzy event
Interval-valued fuzzy event
Fuzzy probabilistic rough set
Interval-valued fuzzy probabilistic rough set
Three-way decisions

ABSTRACT

The concept of probabilistic rough sets, as a main subject of this paper, is intimately connected with the concept of decision-theoretic rough sets. This paper investigates fuzzy and interval-valued fuzzy probabilistic rough sets within frameworks of fuzzy and interval-valued fuzzy probabilistic approximation spaces, respectively. Four types of fuzzy probabilistic rough sets as well as interval-valued fuzzy probabilistic rough sets are established in terms of different constraints on parameters. To find a suitable way of explaining and determining these parameters in each model, three-way decisions are studied based on Bayesian minimum risk decision procedure, i.e., the decision-theoretic rough set approach. The proposed models in this paper broaden applications of probabilistic rough sets due to their abilities of directly dealing with real-valued and interval-valued data.

1. Introduction

Probabilistic rough sets (short for PRSs), as a combination of rough set theory [35,36] and probability theory [9], have been studied at length in literatures [5,31,32,37,41,49,67,68]. The pair of probabilistic approximation operators are built in terms of conditional probabilities and parameter(s) (representing to what degree we can bare the uncertainty or misclassification). When applying PRSs to some concrete situation, parameters, playing a key role in establishing probabilistic approximation operators, are usually provided by experts who are familiar with that situation. It cannot make up for the lack of mathematical foundation even with the help of experts when deciding values of parameters; and it is not confirmed which way is most reasonable when choosing these values. However, from the mathematical viewpoint, we prefer such methods with solid mathematical foundations when there is no concordant criterion. The decision-theoretic rough set (DTRS, for short) approach, proposed by Yao [46–52,55,56], studies rough sets in terms of the Bayesian decision procedure. This approach, just like we said before that it is closely connected with PRSs, provides a mathematic and systematic way to explain and calculate the parameters on the basis of losses or costs of various decisions. The DTRS approach approximates a given concept or a set by three regions (positive, negative and boundary regions) which correspond to positive, negative and boundary rules in three-way decisions (3WDs), respectively [11,50,51]. It has been applied to text classification [20,23], web-based support systems [54], cluster analysis [27,58], investment decisions [30], multi-classification [6,26,27], email filtering [17,63], government decisions [28], face recognition [22] etc. In the following, we first review developments of PRSs, DTRSs and 3WDs; then we present the main work of this paper.

1.1. Reviews of PRSs, DTRSs and 3WDs

Let U be a finite universe of discourse and R be an equivalence relation on U. Two subsets of U are denoted by X and Y.

(1) Probabilistic approaches to rough sets were first studied by Wong and Ziarko [44]. The PRS model was proposed in 1988 by Pawlak et al. [37]. Their model was built based on a fixed precision 0.5, called 0.5-PRS [45]. The variable precision rough sets (VPRSs) [18,64] were formulated by a graded set-inclusion relation, i.e. $c(X,Y) = 1 - \frac{|X\cap Y|}{|X|}$. They can be regarded as special kinds of PRSs if the conditional probabilities are estimated by cardinalities of sets, i.e., $P(X|Y) = \frac{|X\cap Y|}{|Y|}$ [68]. In order to avoid the parameter appeared in VPRSs, Ślęzak and Ziarko [39,41] proposed the non-parametric Bayesian rough set (BRS) model, where the set approximations are defined by adopting the prior probability as references. (Further studies of BRS can be found in [34,40,61,66].)

* Corresponding author at: School of Mathematics and Statistics, Wuhan University, Wuhan 430072, PR China. Tel.: +86 2768775350.
E-mail address: bqhu@whu.edu.cn, bqhujjhe@hotmail.com (B.Q. Hu).

http://dx.doi.org/10.1016/j.knosys.2015.09.018
0950-7051/

(2) The DTRS approach was first introduced by Yao in 1990 [45]. The lower and upper approximations of a concept are derived from the Bayesian minimum risk decision procedure, where the universe of objects are partitioned into three disjoint regions—positive, negative and boundary regions. It finds out that the Pawlak rough set can be obtained based on a special restriction on the loss function. Also, the 0.5-PRS can be obtained as a special case of DTRS by setting another restriction on the loss function. Yao further studied about DTRS approach in [46–48,53] where he did detailed discussions on different restrictions of loss function and thus obtained the basic (α, β)-PRS model $(0 \leq \beta < \alpha \leq 1)$, the α-PRS model, etc.

(3) Li and Yang studied the axiomatic characterization of PRSs [21]. They derived two sets of axioms using the probabilistic approximation operators. The proposed approach helps to understand PRSs from an axiomatic way. Probabilistic rough set model on two universes was first discussed by Ma and Sun [31,32]. They have studied the interrelationship between the Bayesian risk decision and probabilistic approximations on two universes. The rough entropy for this generalized PRS model was proposed based on Shannon entropy. Based on the local rough set and the dynamic granulation principle Sang et al. proposed a new DTRS model under dynamic granulation which satisfies the monotonicity of positive regions [38]. The two parameters α and β dynamically update for each granulation. Game-theoretic rough set was obtained by combining the DTRS approach with game theory [1,10] and it has been applied into recommender system [3].

(4) The probabilistic approximations of fuzzy sets have appeared in [7,42,57]. In literature [7], Deng and Yao deal with fuzzy sets actually based on cut sets of fuzzy sets. An element whose membership grade is greater than or equal to α is put into the positive region; an element whose membership grade is less than or equal to β is put into the negative region; and an element whose membership is between α and β is put into the boundary region. They gave two ways to determine parameters α and β: one is to minimize the total error caused by aforementioned operations on all elements in U; the other is based on Bayesian risk decision procedure like that in [45–48]. The probabilistic rough fuzzy set model for a fuzzy set was defined based on an equivalence relation on U [42]. It is also studied the decision-theoretic rough fuzzy set from the viewpoint of Bayesian decision theory. The limitation of the model was that it depends on equivalence relation. Yang et al. proposed fuzzy probabilistic rough set model based on fuzzy relations [57]. Even though the fuzzy relation was adopted in their model, it is the λ-cut sets of fuzzy relation that really work. That means it is still based on classical relations.

(5) Interval-valued decision-theoretic rough set model has been studied by Liang and Liu [24]. However, their model is built within probabilistic approximation spaces (i.e., still based on classical equivalence relations). It is only the loss function that is interval-valued. In another paper [25], Liang et al. proposed the triangular fuzzy decision-theoretic rough set model in the framework of a probabilistic approximation space. Likewise, only the loss function is made up of triangular numbers. A new model for incomplete information system was studied by Liu et al. in reference [29] where the conditional probabilities are computed based on a new similarity relation and the loss function is represented by interval-numbers.

(6) The concept of three-way decisions actually coexists with rough set theory in which it is interpreted as positive, negative and boundary regions. It was first clearly proposed by Yao in [51,52]. Then, Hu studied 3WDs from a mathematical viewpoint and proposed three-way decision spaces based on fuzzy lattices and partially ordered sets, respectively [11,12].

1.2. What do we do?

The notion of an event in probability theory [9] is a precisely specified collection of elements in the sample space. However, in everyday experience one frequently encounters situations in which an "event" is fuzzy rather than crisp [60]. For example, the temperature is *around* 21°C; a student is *most probable* to pass the exam, etc. These events are fuzzy because of the ill-defined description "around" and "most probable". Besides, some properties (such as, the using temperature range of different bolts; the best quantum of water sprinkling for a certain kind of plants) cannot be described by exact values for which reason interval values are more desirable. On the other hand, since measurement errors are unavoidable in principle, the measuring result is often accompanied by an error range. Instead of using a single value to represent the measuring result, an interval number is, sometimes, more reasonable. Considering these, there is a need to generalize PRSs and DTRS approach for fuzzy events and interval-valued fuzzy events within the frameworks of fuzzy and interval-valued fuzzy probabilistic approximation spaces, respectively, which is the main work of this paper.

In the fuzzy probabilistic approximation space, we first propose four types of fuzzy probabilistic approximation operators (defined for fuzzy events); then, applying Bayesian decision theory, we study three-way decisions for fuzzy events and figure out the relationship between DTRS approach and fuzzy probabilistic rough sets; finally, we study fuzzy probabilistic rough sets with two different universes of discourse. Compared with existing results [7,42,57], our model has several advantages listed below:

(1) the fuzzy probabilistic rough set models presented in our paper can deal with fuzzy set directly instead of its λ-cut set;
(2) the models are established within the framework of fuzzy probabilistic approximation space which ensures fuzzy relations to directly take a part in computing conditional probabilities instead of using their λ-cut sets;
(3) the models are constructed in terms of fuzzy probability instead of the cardinality-based estimation in most literatures [32,46,48,51,52,54].

Within the framework of interval-valued fuzzy probabilistic approximation space, we first provide two different ways to define probabilistic approximations for interval-valued fuzzy events. One is based on the interval-valued fuzzy probability; the other is based on fuzzy probability. Then, we study three-way decisions for interval-valued fuzzy events by employing Bayesian decision procedure. The case of two different universal sets is considered at last. Comparing with reference [24], our model is more flexible since it can directly deal with interval-valued fuzzy information systems.

It is worth reminding that investigations of DTRS approach for fuzzy events and interval-valued fuzzy events presented in this paper are actually generalizations of those introduced in reference [62] which studied DTRS approach for classical events in fuzzy and interval-valued fuzzy approximation spaces, respectively. Besides, the main purpose of this paper is to construct various (interval-valued) fuzzy probabilistic rough sets of (interval-valued) fuzzy sets for different application demands.

The remainder of this paper is organized as follows. Section 2 reviews basic notions of fuzzy set, fuzzy event, interval-valued fuzzy set, interval-valued fuzzy event, etc. Section 3 discusses fuzzy probabilistic rough sets and their three-way decisions. Section 4 studies interval-valued fuzzy probabilistic rough sets as well as associated three-way decisions in the framework of interval-valued fuzzy probabilistic approximation space. Brief examples are attached to