

## INTERVAL ANALYSIS-BASED HYPERBOX GRANULAR COMPUTING CLASSIFICATION ALGORITHMS

H. LIU, J. LI, H. GUO AND C. LIU

**ABSTRACT.** Representation of a granule, relation and operation between two granules are mainly researched in granular computing. Hyperbox granular computing classification algorithms (HBGrC) are proposed based on interval analysis. Firstly, a granule is represented as the hyperbox which is the Cartesian product of  $N$  intervals for classification in the  $N$ -dimensional space. Secondly, the relation between two hyperbox granules is measured by the novel positive valuation function induced by the two endpoints of an interval, where the operations between two hyperbox granules are designed so as to include granules with different granularity. Thirdly, hyperbox granular computing classification algorithms are designed on the basis of the operations between two hyperbox granules, the fuzzy inclusion relation between two hyperbox granules, and the granularity threshold. We demonstrate the superior performance of the proposed algorithms compared with the traditional classification algorithms, such as, Random Forest (RF), Support Vector Machines (SVMs), and Multilayer Perceptron (MLP).

### 1. Introduction

Classification is a popular application in many fields including computer vision [25], data analysis [5, 6]. As a classification and clustering method, Granular Computing (GrC) involves the computational intelligence transformation between spaces with different granularity [31, 29, 14, 10]. As a fundamental data analysis method, GrC is commonly studied from the theory and application aspects, which include pattern recognition, image processing, and industrial applications [26, 16, 11]. The main issues of GrC are the representation of granule, the operation between two granules, the relation between two granules, and the granularity of a granule.

A granule is a clump or group of objects drawn together by distinguishability, similarity, proximity, or functionality [2]. More specially, a binary granule representation is proposed by a conversion from a set to a binary granule. Correspondingly, set operations are converted to binary granule computing. Besides, from a rough set framework point of view, binary granule representations are investigated by introducing Hamming distance between granules. Moreover, a concept of the granule swarm distance is defined for measuring uncertainty between two granule swarms,

---

Received: May 2016; Revised: December 2016; Accepted: February 2017

*Key words and phrases:* Fuzzy lattice, Granular computing, Hyperbox granule, Fuzzy inclusion relation.

which provides a more comprehensible perspective for measures in rough set theory [3].

The operators of two granules are formally expressed as a certain hypothesis testing where a null hypothesis concerns equality of medians of membership grades produced by the triangular norms [9]. V. G. Kaburlasos defined the join operator and the meet operator to induce the granules with different granularity based on the theory of lattice computing [12]. V. G. Kaburlasos defined the join operation and meet operation between two granules which composed the fuzzy inclusion measure between two granules and form the fuzzy lattice reasoning [13].

The relation between two granules is mainly used to form the association rules between inputs and outputs. The framework is a specialized version of the general framework intended for mining relational data and is defined in granular computing theory [8]. V. G. Kaburlasos formed fuzzy inference system (FIS) based on the lattice theory, and a preliminary industrial application demonstrates the advantages of their proposed schemes [10]. In [19], the author proposes a new concept of granular rule-based models whose rules assume a format "if  $G(A_i)$  then  $G(f_i)$ " where  $G(\cdot)$ s are granular generalizations of the numeric conditions and conclusions of the rules.

Measurement of granularity is one of the foundational issues in granular computing. Yiyu Yao investigates a class of measures of granularity of partitions. The granularity of a set is defined by a strictly monotonic increasing transformation of the cardinality of the set. The granularity of a partition is defined as the expected granularity of all blocks of the partition with respect to the probability distribution defined by the partition [30].

The foundations of interval analysis were established and developed by Moore [17]. The capability of interval analysis to solve a wide variety of real life problems in an efficient manner enabled the extension of its concepts to the probabilistic framework. In this way, the classical concept of random variable was extended to cover the interval random variable concept, which allows not only the modeling of randomness, using the concepts of probability theory, but also the modeling of imprecision and non-specificity, using the concepts of interval analysis. The interval analysis based approach provides for the development of mathematical methods and computational tools that enable modeling data and solving optimization problems under uncertainty.

A general method for predicting future observations from any arbitrary continuous distribution is proposed in [1], and two pivotal statistics are modified to construct prediction intervals for future observations when the sample size is fixed or random.

An upper bound condition is introduced to deal with incomplete interval valued fuzzy preference relations [15].

A. Khalid and I. Beg provides a comprehensive analysis of computational problems concerning calculation of general correlation coefficients for interval data [18]. Exact algorithms solving this task have unacceptable computational complexity for larger samples. General correlation coefficients for interval data are also given by intervals. The bounds are derived as lower and upper endpoints.

By choosing the concept of event interval as the building block of the approach and by considering a variety of such intervals, problems with varying degrees of information in logs (e.g. in relation to time stamps) can be addressed in a systematic manner [28]. Interval analysis is widely used in real-world applications. Interval analysis may increase the diagnostic yield of interictal scalp electroencephalography (EEG) combined with the Fourier analysis [23].

In this work, we present hyperbox granule classification algorithms from the view of interval analysis. Firstly, a granule is represented as the hyperbox with the beginning point and the end point. Secondly, the relation between two hyperbox granules is measured by the novel positive valuation function. Thirdly, the operations between two hyperbox granules are designed so as to form the hyperbox granular computing classification algorithm. Finally, the superior performance of the proposed algorithm is demonstrated by the benchmark data sets.

## 2. Algebraic Lattices for Vector Space

In this section, we discuss the algebraic lattices for vector space induced  $N$ -dimensional space, including the operation between two vectors, the partial order relation between two vectors, and the algebraic lattices induced by vector space and operations between two vectors. For vector space, the theory of algebraic lattices has been presented in literature [20]. In the paper, the main definitions and theorems are shown as follows.

**Definition 2.1.** Suppose  $R$  is a nonempty real number set, operations  $\vee$  and  $\wedge$  between two elements  $a$  and  $b$  are defined as,  $\forall a, b \in R, a \vee b = \max\{a, b\}, a \wedge b = \min\{a, b\}$ .

$\forall a, b, c \in R$ , operations  $\vee$  and  $\wedge$  between two elements have the following laws.

- (1) Commutative law,  $a \vee b = b \vee a, a \wedge b = b \wedge a$ ;
- (2) Associate law,  $a \vee (b \vee c) = (a \vee b) \vee c, a \wedge (b \wedge c) = (a \wedge b) \wedge c$ ;
- (3) Absorptive law,  $a \vee (a \wedge b) = a, a \wedge (a \vee b) = a$ ;
- (4) Distributive law,  $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ .

**Theorem 2.2.** For the nonempty set  $R$ , if the binary relation induced by operations  $\vee$  and  $\wedge$  is  $a \leq b \Leftrightarrow a \vee b = b, a \wedge b = a$ , then the relation  $a \leq b$  between  $a$  and  $b$  is the partial order relation.

**Definition 2.3.** Suppose  $(R, \leq)$  is an algebraic system induced by the partial order relation  $\leq$ , if there is the least upper bound and the greatest lower bound for any two elements in  $R$ , then  $(R, \leq)$  is a lattice.

**Theorem 2.4.**  $(R, \leq)$  and  $(R, \geq)$  are lattices,  $\geq$  is the dual relation of  $\leq$ .

Relation  $\leq$  was defined in  $R$ , the following describes extensions to  $R^N$  space, where  $\mathbf{x}$  is a vector,  $x$  is a scalar,  $\preceq$  is the partial order relation between two vectors which is defined as follows.

**Definition 2.5.** The partial order relation between two vectors  $\mathbf{x}=(x_1, x_2, \dots, x_N)$  and  $\mathbf{y}=(y_1, y_2, \dots, y_N)$  is  $\mathbf{x} \preceq \mathbf{y} \Leftrightarrow (x_1 \leq y_1), (x_2 \leq y_2), \dots, (x_N \leq y_N)$ .

The reflexivity, antisymmetry, and transitive of relation  $\preceq$  are verified easily. So the relation  $\preceq$  between two vectors is the partial order relation.

**Definition 2.6.** The operations between two vectors  $\mathbf{x} = (x_1, x_2, \dots, x_N)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_N)$  are  $\mathbf{x} \vee \mathbf{y} = (x_1 \vee y_1, x_2 \vee y_2, \dots, x_N \vee y_N)$ ,  $\mathbf{x} \wedge \mathbf{y} = (x_1 \wedge y_1, x_2 \wedge y_2, \dots, x_N \wedge y_N)$ .

Furthermore, it follows Theorem 2.3.

**Theorem 2.7.**  $(R^N, \preceq)$  and  $(R^N, \succeq)$  are lattices.

### 3. Algebraic Lattices Induced by Interval for Granule Space

In this section, we detail granular computing from the view of interval analysis, such as the representation of a granule, the operations between two granules, the inclusion relation between two granules.

**3.1. The Representation of Hyperbox Granule.** For granular computing in  $N$ -dimensional space, we represent a granule as a hyperbox with the beginning point  $\mathbf{x}$  and the end point  $\mathbf{y}$  which satisfy the partial order relation  $\mathbf{x} \preceq \mathbf{y}$ . For any two vectors  $\mathbf{x} = (x_1, x_2, \dots, x_N)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_N)$ , the two vectors  $\mathbf{x}^*$  and  $\mathbf{y}^*$  are induced by

$$\begin{aligned}\mathbf{x}^* &= \mathbf{x} \wedge \mathbf{y} \\ \mathbf{y}^* &= \mathbf{x} \vee \mathbf{y}\end{aligned}$$

Obviously,  $\mathbf{x}^* \preceq \mathbf{y}^*$ , therefore we compute the hyperbox granule  $G = [\mathbf{x}^*, \mathbf{y}^*]$ , where  $\mathbf{x}^*$  is the beginning point, and  $\mathbf{y}^*$  is the end point of the granule. In the following sections, we represent hyperbox granule by  $G = [\mathbf{x}, \mathbf{y}]$  for two vectors  $\mathbf{x}$  and  $\mathbf{y}$ . In 2-dimensional space, the granule induced by  $\mathbf{x}$  and  $\mathbf{y}$  is box, and in  $N$ -dimensional space, the granule induced by  $\mathbf{x}$  and  $\mathbf{y}$  is hyperbox.

From the view of interval analysis, the hyperbox granule in  $N$ -dimensional space can be regarded as the Cartesian product. Namely, for  $\mathbf{x}$  and  $\mathbf{y}$ , the hyperbox granule can be represented as

$$G = [\mathbf{x}, \mathbf{y}] = [x_1, y_1] \otimes [x_2, y_2] \otimes \dots \otimes [x_N, y_N]$$

The granularity is computed by the  $L_2$  distance between the beginning point and the end point. For example, in  $R^2$  space,  $(0.1, 0.2) \preceq (0.4, 0.6)$ , the hyperbox granule shown in Figure 1 is  $G = [0.1, 0.2, 0.4, 0.6] = [0.1, 0.4] \otimes [0.2, 0.6]$  with the beginning point  $(0.1, 0.2)$  and the end point  $(0.4, 0.6)$ , its granularity is 0.5.

**3.2. The Operations Between Two Hyperbox Granules.** For two hyperbox granules  $G_1 = [\mathbf{x}_1, \mathbf{y}_1]$  and  $G_2 = [\mathbf{x}_2, \mathbf{y}_2]$ , the join hyperbox granule is the following form by the join operation

$$G_1 \vee G_2 = [\mathbf{x}_1 \wedge \mathbf{x}_2, \mathbf{y}_1 \vee \mathbf{y}_2] \quad (1)$$

where

$$\begin{aligned}\mathbf{x}_1 \wedge \mathbf{x}_2 &= (x_{11} \wedge x_{21}, x_{12} \wedge x_{22}, \dots, x_{1N} \wedge x_{2N}) \\ \mathbf{y}_1 \vee \mathbf{y}_2 &= (y_{11} \vee y_{21}, y_{12} \vee y_{22}, \dots, y_{1N} \vee y_{2N}) \\ \mathbf{x}_i &= (x_{i1}, x_{i2}, \dots, x_{iN}) \\ \mathbf{y}_i &= (y_{i1}, y_{i2}, \dots, y_{iN})\end{aligned}$$

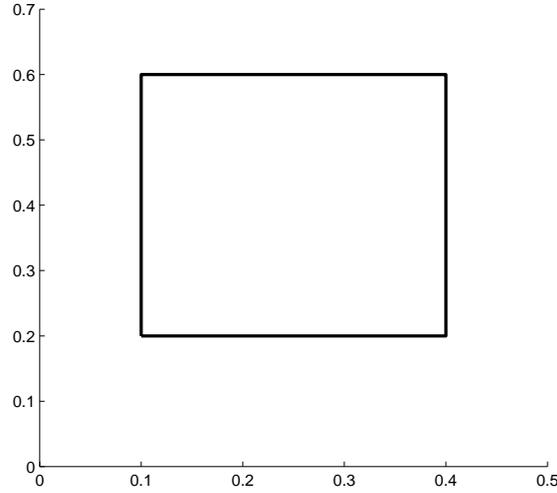


FIGURE 1. Hyperbox Granule in 2-dimensional Space

The join hyperbox granule has greater granularity than the original hyperbox granules. The original hyperbox granules and the join hyperbox granule have the following relations

$$G_1 \subseteq G_1 \vee G_2$$

$$G_2 \subseteq G_1 \vee G_2$$

The meet hyperbox granule has the following form by the meet operation

$$G_1 \wedge G_2 = \begin{cases} [\mathbf{x}_1 \vee \mathbf{x}_2, \mathbf{y}_1 \wedge \mathbf{y}_2] & \text{if } \mathbf{x}_1 \vee \mathbf{x}_2 \preceq \mathbf{y}_1 \wedge \mathbf{y}_2 \\ \emptyset & \text{otherwise} \end{cases} \quad (2)$$

The meet hyperbox granule has smaller granularity than the original hyperbox granules. The meet hyperbox granule and the original hyperbox granules have the following relations,

$$G_1 \wedge G_2 \subseteq G_1$$

$$G_1 \wedge G_2 \subseteq G_2$$

**3.3. The Inclusion Relation Between Two Hyperbox Granules.** For hyperbox granular computing, the inclusion relation between two granules is discussed in literature [21]. In the paper, according to the definitions and theorems, the comments are shown as follows.

**Definition 3.1.** The inclusion relation between two hyperbox granules  $G_1 = [\mathbf{x}_1, \mathbf{y}_1]$  and  $G_2 = [\mathbf{x}_2, \mathbf{y}_2]$  is defined as

$$G_1 \subseteq G_2 \Leftrightarrow \mathbf{x}_2 \preceq \mathbf{x}_1, \mathbf{y}_1 \preceq \mathbf{y}_2$$

**Theorem 3.2.** *The inclusion relation between two hyperbox granules is the partial order relation.*

For the granularity space induced  $N$ -dimensional space, the relation between two hyperbox granules can be measured by the partial order relation between two vectors with  $2N$  length.

For two hyperbox granules  $G_1 = [\mathbf{x}_1, \mathbf{y}_1]$  and  $G_2 = [\mathbf{x}_2, \mathbf{y}_2]$ , if  $G_1 \subseteq G_2$ , then  $\mathbf{x}_2 \preceq \mathbf{x}_1$  and  $\mathbf{y}_1 \preceq \mathbf{y}_2$ , namely

$$G_1 \subseteq G_2 \Leftrightarrow x_{21} \leq x_{11}, \dots, x_{2N} \leq x_{1N}, y_{11} \leq y_{21}, \dots, y_{1N} \leq y_{2N} \quad (3)$$

For two vectors  $(\mathbf{x}_1, \mathbf{y}_1)$  and  $(\mathbf{x}_2, \mathbf{y}_2)$  with  $2N$  length, the partial order relation between these two vectors is

$$(\mathbf{x}_1, \mathbf{y}_1) \preceq (\mathbf{x}_2, \mathbf{y}_2) \Leftrightarrow x_{11} \leq x_{12}, \dots, x_{1N} \leq x_{2N}, y_{11} \leq y_{21}, \dots, y_{1N} \leq y_{2N} \quad (4)$$

Compared with formula (3), formula (4) and formula (3) are not consistent if we consider the hyperbox granule as the vector with  $2N$  length. The order preserving function is introduced to eliminate the inconsistency.

**Definition 3.3.** For hyperbox granule  $G = [\mathbf{x}, \mathbf{y}]$ ,  $\theta(G) = [\theta(\mathbf{x}), \mathbf{y}]$  is the order preserving function which satisfies  $G_1 \subseteq G_2 \Leftrightarrow \theta(G_1) \preceq \theta(G_2)$ .

Owing to the dual property of  $\preceq$  and  $\succeq$ , if  $\mathbf{x}_1 \succeq \mathbf{x}_2$ ,  $\theta(\mathbf{x}_1) \preceq \theta(\mathbf{x}_2)$ . Namely,  $\theta(\cdot)$  is the somorphic mapping between lattice  $(R^N, \preceq)$  and its dual lattice  $(R^N, \succeq)$ .

If  $\theta(\cdot)$  is a decreasing function, then partial order relation (3) and (4) induced by  $\theta(\cdot)$  are consistent. Considering the operations (1) and (2) between two hyperbox granules, the inclusion relation between two hyperbox granules can be transformed into the operation between two hyperbox granules

$$G_1 \subseteq G_2 \Leftrightarrow G_1 = G_1 \wedge G_2, G_2 = G_1 \vee G_2 \quad (5)$$

For the classification problem  $S$  in  $R^N$  space,  $GS$  is the hyperbox granule set, the inclusion relation induces the following lattice.

**Theorem 3.4.**  $(GS, \subseteq)$  is lattice.

For  $N$ -dimensional space, we represent the hyperbox granule as the vector with  $2N$  length induced by the beginning point and the end point, and the hyperbox granule is the Cartesian product of  $N$  intervals, so we discuss the hyperbox granular computing from the view of interval analysis.

#### 4. Fuzzy Lattice for Granule Space

For a classification problem, the inclusion relation between two hyperbox granules is fuzzy, and different from the traditional crisp inclusion relation. We discuss the fuzzy inclusion relation which is compounded of the valuation functions.

**Definition 4.1.** The fuzzy inclusion relation is defined as the following mapping for  $GS$

$$\sigma : GS \times GS \rightarrow [0, 1]$$

In Figure 2, the inclusion relation between two hyperbox granules is fuzzy, the granularities of hyperbox granules do not change between Figure 2(a) and Figure 2(b). Nevertheless, the inclusion measure  $\sigma(G_2, G_1)$  in Figure 2(b) between  $G_1$  and  $G_2$  is greater than the inclusion measure  $\sigma(G_2, G_1)$  in Figure 2(a).

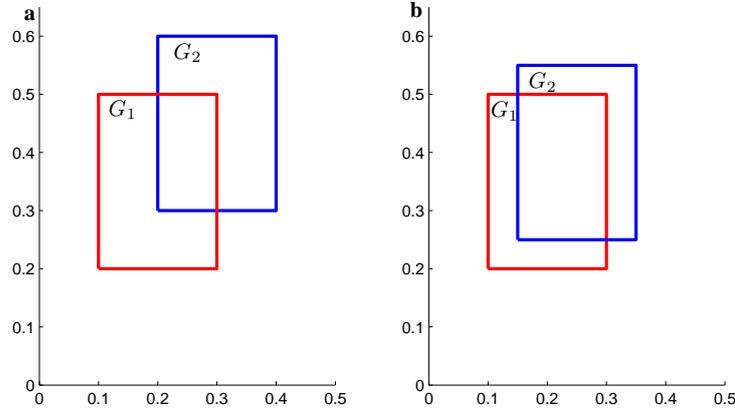


FIGURE 2. Different Fuzzy Inclusion Measures Between Two Hyperbox Granules in 2-dimensional Space

The fuzzy inclusion function between two hyperbox granules satisfies the following properties [13].

$$G_1 \subseteq G_2 \Rightarrow \sigma(G_1, G_2) = 1 \quad (6)$$

$$G_1 \subseteq G_2 \Rightarrow \sigma(G, G_1) \leq \sigma(G, G_2) \quad (7)$$

For  $N$ -dimensional space, the hyperbox granule is considered in  $R^{2N}$  space, and the fuzzy inclusion relation between two hyperbox granules belongs to  $R$  space. So the mapping function between the granule space and the  $R$  space must be used to form the above mentioned fuzzy inclusion relation in terms of the hyperbox granules and their granularities. The linear and nonlinear positive valuation functions are used to realise the mapping between hyperbox granule space and  $R$  space.

A positive valuation function  $f:R \rightarrow R$  has to be a strictly increasing function. The linear positive valuation function is

$$v(\mathbf{x}) = \sum_{i=1}^N x_i \quad (8)$$

For the order preserving function and the nonlinear positive valuation function, the former is used to eliminate the inconsistent of the partial order relation between vectors and the inclusion relation between two hyperbox granules, and the latter transforms  $R$  space and  $R^N$  space into the normalized the suitable interval and preserve the order of elements in the transformed interval, such as the interval  $(0,1)$  for  $R$  space, and  $(0,1)^N$  for  $N$ -dimensional space. Namely, if  $v(x) \in (0,1)$ , then  $\theta(x) \in (0,1)$  for  $R$  space.

According to the monotone properties of the order preserving function and the linear positive valuation function, we can define the following nonlinear positive valuation function.

For  $R$  space, we define the nonlinear positive valuation function as follows

$$v(x) = \frac{1}{2} \left( \frac{e^x - e^{-x}}{e^x + e^{-x}} + 1 \right) \quad (9)$$

The nonlinear positive valuation function maps the  $R$  space into  $(0,1)$ .

**Theorem 4.2.** *The nonlinear positive valuation function*

$$v(x) = \frac{1}{2} \left( \frac{e^x - e^{-x}}{e^x + e^{-x}} + 1 \right)$$

*is the positive valuation for  $R$  space.*

For  $R^N$  space

$$v(\mathbf{x}) = \sum_{i=1}^N v(x_i) \quad (10)$$

Which maps the  $R^N$  space into  $(0,1)^N$  space. Similarly, the following theorem is achieved.

**Theorem 4.3.**  $v(\mathbf{x}) = \sum_{i=1}^N v(x_i)$  *is the positive valuation for  $R^N$  space.*

For one-to-one mapping between  $(0,1)$  and  $R$ , spaces  $R^N$  and  $(0,1)^N$  are isomorphism, and both  $((0,1), \preceq)$  and  $((0,1)^N, \preceq)$  are lattices.

Suppose  $(G((0,1)), \subseteq)$  is the induced hyperbox granule space by the Cartesian product of intervals  $(0,1)$  and  $(0,1)$ , we discuss the preserving function and the positive valuation function in hyperbox granule space from the view of interval analysis.

**Theorem 4.4.** *If  $\theta(x) = 1 - x$  is the isomorphism mapping from lattice  $((0,1), \geq)$  to its dual lattice  $((0,1), \leq)$ , then*

$$v([a, b]) = v(\theta(a)) + v(b) \quad (11)$$

*is the positive valuation function.*

In Theorem, we discuss the partial order relation for the hyperbox granule space induced by 1-dimensional space, such as  $R$  space, and the elimination method of inconsistency of partial order relations for the induced hyperbox granule space and the 2-dimensional space. The partial order relation for the hyperbox granule space and the elimination method can be extended to  $N$ -dimensional space and used to form the classification algorithm and clustering algorithm in  $N$ -dimensional space by the following fuzzy lattice.

**Theorem 4.5.**  $(G(R), k)$  and  $(G(R), s)$  *are fuzzy lattice, where  $R$  is real number, where*

$$k([a, b], [c, d]) = \frac{v([c, d])}{v([a, b] \vee [c, d])} \quad (12)$$

$$s([a, b], [c, d]) = \frac{v([a, b] \wedge [c, d])}{v([a, b])} \quad (13)$$

*are the fuzzy inclusion relation between two hyperbox granules.*

The above discussion highlights the fuzzy lattice induced by  $R$  space from the view of interval analysis which can be extended to  $R^N$  space and achieves the following theorem.

**Theorem 4.6.** *For classification problem  $S$  in  $R^N$  space,  $(G(R^N), k)$  and  $(G(R^N), s)$  are fuzzy lattice, where*

$$k(G_1, G_2) = \frac{v(G_2)}{v(G_1 \vee G_2)}, s(G_1, G_2) = \frac{v(G_1 \wedge G_2)}{v(G_1)} \quad (14)$$

## 5. Hyperbox Granular Computing Classification Algorithms

As mentioned above, the operations between two hyperbox granules and the fuzzy inclusion relations between two hyperbox granules can be used for decision-making by a classification algorithm. We design the classification algorithms by the join operation and the granularity threshold.

### 5.1. The Conditional Join Process Between Two Hyperbox Granules.

Fuzzy inclusion relation measures the inclusion relation between two hyperbox granules and guides the join process of two hyperbox granules. For hyperbox granule  $G_1$ , any one hyperbox granule  $G$  in the hyperbox granule space has the fuzzy inclusion relation with  $G_1$ , the hyperbox granule with the maximal fuzzy inclusion relation is selected to perform the join process, moreover the granularity of the join hyperbox granule also controls the join process.

For example,  $G_1 = [0.05, 0.15, 0.48, 0.68]$ ,  $G_2 = [0.1, 0.2, 0.5, 0.7]$ , and  $G_3 = [0.7, 0.8, 0.7, 0.8]$  shown in Figure 3 are induced by  $R^2$  space. One of  $G_2$  and  $G_3$  is selected to join with  $G_1$ .

$k(G_2, G_1) = \frac{v(1-0.05)+v(0.48)+v(1-0.15)+v(0.68)}{v(1-0.05)+v(0.5)+v(1-0.15)+v(0.7)} = 0.9956$ , for  $G_1 \vee G_2 = [0.05, 0.15, 0.5, 0.7]$  and granularity is 0.7106.

$k(G_3, G_1) = \frac{v(1-0.05)+v(0.48)+v(1-0.15)+v(0.68)}{v(1-0.05)+v(0.7)+v(1-0.15)+v(0.8)} = 0.9656$ , for  $G_1 \vee G_3 = [0.05, 0.15, 0.7, 0.8]$  and granularity is 0.9192. If the granularity threshold is set to 0.8,  $G_1$  is replaced by  $G_1 \vee G_2$  because  $k(G_2, G_1) > k(G_3, G_1)$ .

**5.2. Interval Analysis-based Granular Computing Classification Algorithms.** Based on the above discussion, we propose the following hyperbox granular computing algorithms, the training phase is described as algorithm1, and the testing phase is described as algorithm2.

The task of the test phase is to predict the class of the unknown datum on the basis of the trained hyperbox granule set  $HB$  and the corresponding class labels  $Lab$ . The testing phase is described as follows.

We explain the training phase in 2-dimensional space, suppose the trained temporary hyperbox granule set  $HB$  and the corresponding class label  $Lab$  are

$$HB = \begin{pmatrix} 0.1 & 0.5 & 0.3 & 0.6 \\ 0.2 & 0.3 & 0.3 & 0.4 \\ 0.3 & 0.1 & 0.4 & 0.2 \\ 0.6 & 0.2 & 0.7 & 0.4 \end{pmatrix}, Lab = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

---

**Algorithm 5.1.** Hyperbox granular computing training algorithm

Input: training set  $S$ , the threshold  $\rho$  of granularity

Output: hyperbox granule set  $HB$ , and the corresponding class label  $Lab$

---

- (1) initialize  $HB = \emptyset$ ;
  - (2)  $i = 1$ ;
  - (3) extract the samples with class label  $i$  from data set  $S$ , and form  $X$ ;
  - (4) initialize the temporary hyperbox granule set  $HBt = \{[x_1, x_1]\}$  where  $x_1 \in X$ , and remove  $x_1$  from  $X$ ;
  - (5) if  $X = \emptyset$ ,  $HB = HB \cup HBt$  and  $Lab = Lab \cup \{i\}$ , otherwise;
  - (6) extract the  $j$ th sample from  $X$  at random, and form the atomic hyperbox granule  $HB_0 = [x_j, x_j]$ ;
  - (7)  $s = 1$ ;
  - (8) compute the fuzzy inclusion relation between the atomic hyperbox granule  $HB_0$  and the  $s$ th hyperbox granule  $HBt(s, :)$  by formula  $k(HB_0, HBt(s, :))$ ;
  - (9)  $s = s + 1$ ;
  - (10) find the maximal  $k(\cdot, \cdot)$ , namely  $id = \arg \max_s k(\cdot, \cdot)$ ;
  - (11) join the  $HB_0$  with  $HBt(id, :)$ , namely  $HB_0 \vee HBt(id, :)$ ;
  - (12) if the granularity of join hyperbox granule is less than or equal to  $\rho$ , the  $id$ th hyperbox granule  $HBt(id, :)$  is updated by  $HB^* = HB_0 \vee HBt(id, :)$ , otherwise  $HBt = HBt \cup \{HB_0\}$  and remove  $x_j$  from  $X$ ;
  - (13)  $i = i + 1$ .
- 

**Algorithm 5.2.** Hyperbox granular computing testing algorithm

Input: hyperbox granule set  $HB$ , and the class labels  $Lab$ , the testing datum  $x$ .

Output: the class label of datum  $x$ .

---

- (1) form the atomic hyperbox granule  $HB_x = [x, x]$ ;
  - (2)  $i = 1$ ;
  - (3) computing the inclusion measure  $\sigma_i = k(HB_x, HB(i, :))$  between  $HB_x$  and  $HB(i, :)$ ;
  - (4)  $i = i + 1$ ;
  - (5) find the  $id$  of the hyperbox granule with the maximal inclusion measure in  $HB$ , namely  $id = \arg \max_i \sigma_i$ ;
  - (6) the class label of datum  $x$  is  $Lab(id)$ .
-

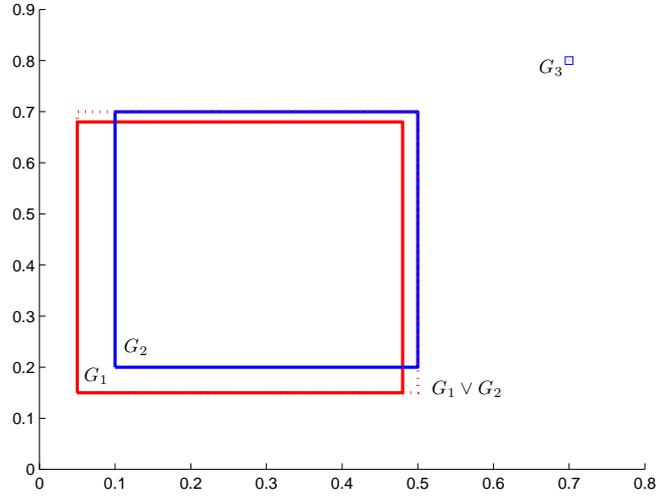


FIGURE 3. Join Between Two Granules

which are shown in Figure 4 marked as  $G_1, G_2, G_3, G_4$ , and the last column is the class labels. For the sample  $(0.45, 0.5)$  with the class label 1 in training set  $S$ . According to algorithm 1, we compute the fuzzy inclusion measure between the atomic hyperbox granule  $G_0 = [0.45, 0.5, 0.45, 0.5]$  induced by the sample and the hyperbox granules with the class labels 1 in  $HB$ ,  $k(G_0, G_1) = 0.9703, k(G_0, G_2) = 0.9516, k(G_0, G_1) > k(G_0, G_2)$ , and the granularity of the join hyperbox granule of  $G_0$  and  $G_1$  is 0.3640, which is less than the user-defined threshold  $\rho = 0.40$ , so  $G_0$  and  $G_1$  are joined, and the hyperbox granule  $G_1$  is replaced by  $G_0 \vee G_1 = [0.1, 0.5, 0.45, 0.6]$ , and the hyperbox granule set  $HB$  and the corresponding class label  $Lab$  are

$$HB = \begin{pmatrix} 0.1 & 0.5 & 0.45 & 0.6 \\ 0.2 & 0.3 & 0.3 & 0.4 \\ 0.3 & 0.1 & 0.4 & 0.2 \\ 0.6 & 0.2 & 0.7 & 0.4 \end{pmatrix}, Lab = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

The join hyperbox granule is shown in Figure 4.

If the threshold of granularity is set to 0.3,  $G_0$  is add to the hyperbox granule set, and the hyperbox granule set  $HB$  and the corresponding class label  $Lab$  are

$$HB = \begin{pmatrix} 0.1 & 0.5 & 0.3 & 0.6 \\ 0.2 & 0.3 & 0.3 & 0.4 \\ 0.3 & 0.1 & 0.4 & 0.2 \\ 0.6 & 0.2 & 0.7 & 0.4 \\ 0.45 & 0.5 & 0.45 & 0.5 \end{pmatrix}, Lab = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}$$

So, the size of hyperbox granule set and the size of corresponding class label  $Lab$  are determined by the user-defined threshold  $\rho$ .

For the testing phase of the user-defined threshold  $\rho = 0.4$  in the training phase, suppose the testing datum is  $x = [0.5, 0.5]$ , the inclusion measures between the hyperbox granule  $HB_x = [0.5, 0.5, 0.5, 0.5]$  formed by datum  $x$  and the hyperbox granules in hyperbox granule set  $HB$  trained by algorithm1 are 0.9615, 0.9430, 0.9235, and 0.9644. The 4th hyperbox granule in the hyperbox granule set includes the hyperbox granule  $HB_x$  maximally, the class label of the testing datum is 3.

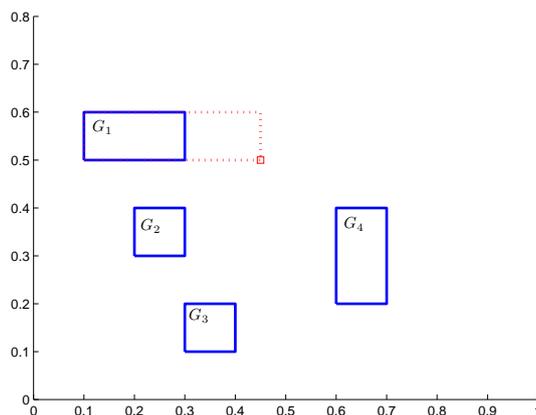


FIGURE 4. Explain of HBGrC in 2-dimensional Space

## 6. Experiments

The effectiveness of the HBGrC is evaluated with a series of empirical studies including the classification problems in 2-dimensional space and classification problems in  $N$ -dimensional space. We compare HBGrC with classification algorithms, such as random forests (RF), support vector machines (SVMs), and evaluate the performance of classification algorithms by training accuracy (TAC) and testing accuracy (AC).

Random forest (RF) is a notion of the general technique of random decision forests that are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of overfitting to their training set [7].

Support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis [4].

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate outputs. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique

called backpropagation for training the network [22]. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable.

For the parameters, HBGrC includes the threshold of granularity  $\rho$ , RF includes the parameter Ntr (nodes of tree), SVMs include the parameter kernel functions, such as poly $d$  (polynomial kernel with the order  $d$ ) and RBF $\sigma$  (Radial basis function with the width  $\sigma$ ), MLP includes the number of hidden layers, iterations, momentum, and learning rate.

**6.1. Classification Problems in 2-dimensional Space.** In the first benchmark study, two spiral curve classification problem [27], Ripley classification problem [27], and sensor2 classification problem (wall-following robot navigation data in website <http://archive.ics.uci.edu/ml/>) which were created in two dimensions, and used to assess efficacy of classification algorithms and visualize the boundary of classification. The details of the data sets and classification performance are summarized in Table 1, and Tr represents the number of training data, Ts represents the number of testing data, TA represents the testing accuracy. Compared with RF and SVMs, the HBGrC has higher or comparable testing accuracies. For data set *spiral*, the testing accuracy of MLP is much lower than that of HBGrC, RF, and SVMs because the classification boundary by MLP is formed by the sigmoid function.

For the data set *Ripley*, Figure 5 shows comparison of the boundary by HBGrC and the boundaries by RF, SVMs, and MLP. In Figure 5(a), each area is composed by the hyperbox granules with the same class labels in the trained hyperbox granule set, in Figure 5(b), each area is determined the data points classified by RF which predicts the class labels from the root node to the leaf nodes, in Figure 5(c), each area is determined by the data which are mapped into high-dimensional space and have the same class labels predicted by SVMs in high-dimensional space, and in Figure 5(d), each area is composed of the data points with the same class labels predicted by the trained sigmoid function.

Data sets	Tr	Ts	Algorithms	Parameter	TA
Spiral	970	194	HBGrC	0.08	100
			RF	ntree=500	100
			SVMs	RBF0.1	100
			MLP	100,500,0.1	51.5464
Ripley	250	1000	HBGrC	0.27	90.1
			RF	ntree=500	89.8
			SVMs	RBF0.1	82.9
			MLP	100,500,0.1	90
Sensor2	4887	569	HBGrC	0.2	99.47
			RF	ntree=500	99.47
			SVMs	RBF0.1	NA
			MLP	100,500,0.1	96.31

TABLE 1. The Classification Problems and Their Performances in 2-dimensional Space

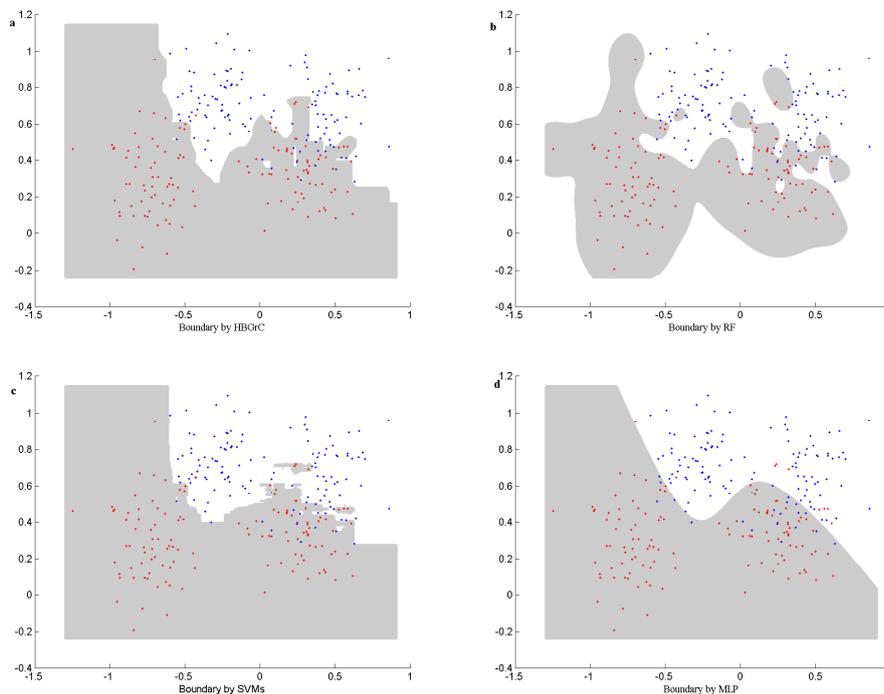


FIGURE 5. Boundaries Performed by HBGrC, RF, SVMs, and MLP for Data Set *Ripley*

**6.2. Classification Problems in  $N$ -dimensional Space.** In this section, we verify the performance of the proposed classification algorithms which are extended to  $N$ -dimensional space compared with the other state-of-art classification algorithms, such as RF, SVMs, and MLP, by the selected benchmark data sets from the website (<http://archive.ics.uci.edu/ml/>). These data sets are the most popular data sets since 2007, and the characters and the performance of the data sets are listed in Table 2. In order to facilitate the selection of parameters of thresholds of granularities, the  $R^N$  space is normalized into the  $[0, 1]^N$  space, the granularity parameters are set to from 0.2 to 0.1 with step 0.005 for the  $n$ -class classification problems performed by HBGrC.

10-fold cross validation is used to evaluate the extension of algorithms. The performances of classification algorithms include the maximal testing accuracies, the mean testing accuracies, the minimal testing accuracies, and the standard deviation of testing accuracies. The superiority of algorithms is evaluated by the mean testing accuracies, the stability of algorithms is verified by the standard deviation which are shown in Table 3.

The testing accuracies are the main evaluation indexes for the classification algorithms, the  $t$ -test values shown in Table 4 are used to verify the testing accuracies by

Data sets	N	outputs	Samples
iris	4	3	150
phoneme	5	2	5404
Sensor4	4	4	5456
Cancer	30	2	532

TABLE 2. The Classification Problems in  $N$ -dimensional Space

Data set	Algorithms	max AC	mean AC	min AC	std AC
iris	HBGrC	100.00	99.3333	93.3333	2.1082
	RF	100.00	96.6667	93.3333	3.5136
	SVMs	100.00	98.00	86.67	4.50
	MLP	100.00	91.3333	66.6667	14.4188
phoneme	HBGrC	92.5788	90.8783	88.3117	1.2052
	RF	94.2486	91.3445	89.4249	1.7909
	SVMs	84.2301	81.5150	79.4063	1.6294
	MLP	89.0538	87.9012	86.6184	0.9242
Sensor4	HBGrC	100.0000	98.3325	95.9484	1.2434
	RF	100.0000	99.7606	98.1584	0.5891
	SVMs	97.6059	95.1968	90.0552	2.5216
	MLP	98.3425	95.6932	89.3186	2.7510
Cancer	HBGrC	100.0000	97.2236	94.2308	1.9334
	RF	100.0000	96.2620	90.3846	3.0849
	SVMs	100.0000	95.6851	92.3077	2.5316
	MLP	67.3077	58.0889	36.5385	11.6020

TABLE 3. The Performances of Classification Problems in  $N$ -dimensional Space

classification algorithms statistically. If  $h = 0$ , then the achieved testing accuracies by classification algorithms have no significant difference statistically, especially, although  $h = 0$ , but  $p$  is relatively small, close to 0.05, we regard the achieved testing accuracies have significant difference. If  $h = 1$ , then the achieved testing accuracies by classification algorithms have significant difference, and we can illustrate the superiority of the algorithm by the mean testing accuracy, especially, although  $h = 1$ , but  $p$  is relatively small, close to 0.05, we regard the achieved testing accuracies have no significant difference.

For data sets *Iris*,  $h = 0$  shown in Table 4. Statistically, the testing accuracies by HBGrC, RF, SVMs, and MLP have no significant difference from the  $h$  values of  $t$ -test listed in Table 4, and the testing accuracies of HBGrC is slightly higher than those of RF, SVMs, and MLP in terms of maximal testing accuracies, mean testing accuracies, and the minimal testing accuracies listed in Table 3.

For the data sets *Phoneme*,  $h = 0$  for HBGrC and RF,  $h = 1$  for HBGrC and SVMs, HBGrC and MLP, HBGrC is superior to SVMs and MLP because the mean

testing accuracies of HBGrC are greater than those of SVMs and MLP (see Table 3), and HBGrC and RF are comparable because  $h = 0$  for HBGrC and RF.

For the data sets *Sensor4*,  $h = 1$  means HBGrC and the traditional classification algorithms have the significant difference from the view of testing accuracies, and we can select the optimal classification algorithm by the mean testing accuracies. RF is superior to HBGrC because the mean testing accuracy of RF is greater than that of HBGrC, HBGrC is superior to SVMs and MLP because the mean testing accuracies of SVMs and MLP is less than those of HBGrC.

For the data sets *Cancer*, HBGrC and RF have no significant difference and HBGrC and SVMs have no significant difference because  $h = 0$ , and HBGrC is superior to MLP obviously for the meaning testing accuracies shown in table 3 and  $h = 1$  shown in Table 4.

Algorithms	Iris		Phoneme		Sensor4		Cancer	
	$h$	$p$	$h$	$p$	$h$	$p$	$h$	$p$
HBGrC-RF	0	0.0544	0	0.5034	1	0.0025	0	0.4146
HBGrC-SVMs	0	0.4073	1	0.0000	1	0.0080	0	0.1441
HBGrC-MLP	0	0.0996	1	0.0000	1	0.0342	1	0.0000

TABLE 4. The  $t$ -test Values of Comparison of HBGrC and Other Algorithms

## 7. Conclusions

In the article, we present the hyperbox granular computing classification algorithms from the view of interval analysis, and demonstrate the feasibility and superiority of HBGrC compared with the traditional classification algorithms, such as RF, SVMs, and MLP, for the benchmark data sets. As an improved granular computing algorithm for classification problems, there are some improvements, such as the adaptive selection of granularity threshold.

**Acknowledgment.** Authors would like to express their sincere thanks to the anonymous referees and the Editor-in-Chief for their valuable comments. This work was supported in part by the Natural Science Foundation of China (Grant Nos. 61170202 and 61202287) and the Teacher Education Curriculum Reform Project of Henan Province in China (2016-JSJYZD-027).

## REFERENCES

- [1] H. M. Barakat, M. E. El-Adll and A. E. Alyb, *Prediction intervals of future observations for a sample of random size from any continuous distribution*, Mathematics and Computers in Simulation, **97** (2014), 1-13.
- [2] G. Bortolan and W. Pedrycz, *Hyperbox classifiers for arrhythmia classification*, Kybernetes, **36(3-4)** (2007), 531-547.
- [3] Y. Chen, Q. Zhu, K. Wu, S. Zhu and Z. Zeng, *A binary granule representation for uncertainty measures in rough set theory*, Journal of Intelligent and Fuzzy Systems, **28(2)** (2015), 867-878.
- [4] C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning, **20(3)** (1995), 273-297.

- [5] Z. Fu, K. Ren, J. Shu, X. Sun and F. Huang, *Enabling personalized search over encrypted outsourced data with efficiency improvement*, IEEE Transactions on Parallel and Distributed Systems, **27(9)** (2016), 2546-2559.
- [6] Z. Fu, X. Sun, N. Linge and L. Zhou, *Achieving effective cloud search services: multi-keyword ranked search over encrypted cloud data supporting synonym query*, IEEE Transactions on Consumer Electronics, **60(1)** (2014), 164-172.
- [7] T. K. Ho, *The random subspace method for constructing decision forests*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **20(8)** (1998), 832-844.
- [8] P. Honko, *Association discovery from relational data via granular computing*, Information Sciences, **234** (2013), 136-149.
- [9] X. Hu, W. Pedrycz and X. Wang, *Comparative analysis of logic operators: A perspective of statistical testing and granular computing*, International Journal of Approximate Reasoning, **66** (2015), 73-90.
- [10] V. G. Kaburlasos and A. Kehagias, *Fuzzy Inference System (FIS) extensions based on the lattice theory*, IEEE Transactions on Fuzzy Systems, **22(3)** (2014), 531-546.
- [11] V. G. Kaburlasos and T. P. Pachidis, *A Lattice-Computing ensemble for reasoning based on formal fusion of disparate data types, and an industrial dispensing application*, Information Fusion, **16** (2014), 68-83.
- [12] V. G. Kaburlasos, S. E. Papadakis and G. A. Papakostas, *Lattice computing extension of the FAM Neural classifier for human facial expression recognition*, IEEE Transactions on Neural Networks and Learning Systems, **24(10)** (2013), 1526-1538.
- [13] V. G. Kaburlasos and G. A. Papakostas, *Learning distributions of image features by interactive fuzzy lattice reasoning in pattern recognition applications*, IEEE Computational Intelligence Magazine, **10(3)** (2015), 42-51.
- [14] J. Kerr-Wilson and W. Pedrycz, *Design of rule-based models through information granulation*, Expert System with Applications, **46** (2016), 274-285.
- [15] A. Khalid and I. Beg, *Incomplete interval valued fuzzy preference relations*, Information Sciences, **348** (2016), 15-24.
- [16] L. Maciura and J. G. Bazan, *Granular computing in mosaicing of images from capsule endoscopy*, Natural Computing, **14(4)** (2015), 569-577.
- [17] R. E. Moore, R. B. Kearfott and M. J. Cloud, *Introduction to Interval Analysis*, SIAM Press, Philadelphia, 2009.
- [18] K. R. Opara and O. Hryniewicz, *Computation of general correlation coefficients for interval data*, International Journal of Approximate Reasoning, **73** (2016), 56-75.
- [19] W. Pedrycz, *Granular fuzzy rule-based architectures: Pursuing analysis and design in the framework of granular computing*, Intelligent Decision Technologies, **9(4)** (2015), 321-330.
- [20] V. Petridis and V. G. Kaburlasos, *Fuzzy lattice neural network (FLNN): a hybrid model for learning*, IEEE Transactions on Neural Networks, **9(5)** (1998), 877-890.
- [21] V. Petridis and V. G. Kaburlasos, *Learning in the framework of fuzzy lattices*, IEEE Transactions on Fuzzy Systems, **7(4)** (1999), 422-440.
- [22] B. Ploj, *Advances in Machine Learning Research*, Nova Press, New York, 2014.
- [23] J. Pyrzowski, M. Siemiski, A. Sarnowska, J. Jedrzejczak and W. M. Nyka, *Interval analysis of interictal EEG: pathology of the alpha rhythm in focal epilepsy*, Scientific Reports 5, **16230** (2015), 1-10.
- [24] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, 1996.
- [25] M. Ristin, M. Guillaumin, J. Gall and L. V. Gool, *Incremental learning of random forests for large-scale image classification*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **38(3)** (2016), 490-503.
- [26] A. V. Savchenko, *Fast multi-class recognition of piecewise regular objects based on sequential three-way decisions and granular computing*, Knowledge-Based Systems, **91** (2016), 252-262.
- [27] H. Sossa and E. Guevara, *Efficient training for dendrite morphological neural networks*, Neurocomputing, **131** (2014), 132-142.

- [28] S. Suriadi, C. Ouyang, W. M. P. van der Aalst and A. H. M. ter Hofstede, *Event interval analysis: Why do processes take time?* Decision Support Systems, **79** (2015), 77-98.
- [29] Y. Yao and Y. She, *Rough set models in multigranulation spaces*, Information Sciences, **327** (2016), 40-56.
- [30] Y. Y. Yao and L. Q. Zhao, *A measurement theory view on the granularity of partitions*, Information Sciences, **213** (2012), 1-13.
- [31] L. A. Zadeh, *Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems*, Soft Computing, **2(1)** (1998), 23-25.

HONGBING LIU\*, CENTER OF COMPUTING, XINYANG NORMAL UNIVERSITY, XINYANG 464000, P. R. CHINA

*E-mail address:* liuhbing@126.com

JIN LI, CENTER OF COMPUTING, XINYANG NORMAL UNIVERSITY, XINYANG 464000, P. R. CHINA

*E-mail address:* lijn@xynu.edu.cn

HUAPING GUO, SCHOOL OF COMPUTER AND INFORMATION TECHNOLOGY, XINYANG NORMAL UNIVERSITY, XINYANG 464000, P. R. CHINA

*E-mail address:* hpguo-cm@163.com

CHUNHUA LIU, SCHOOL OF COMPUTER AND INFORMATION TECHNOLOGY, XINYANG NORMAL UNIVERSITY, XINYANG 464000, P. R. CHINA

*E-mail address:* zzdxliuch@163.com

\*CORRESPONDING AUTHOR