Multigranulation Super-Trust Model for Attribute Reduction

Weiping Ding, *Senior Member*, IEEE, Witold Pedrycz, *Fellow,* IEEE, Isaac Triguero, *Member*, IEEE,
Zehong Cao, *Member*, IEEE, and Chin-Teng Lin, *Fellow*, IEEE

Abstract—As big data often contains a significant amount of uncertain, unstructured and imprecise data that are structurally complex and incomplete, traditional attribute reduction methods are less effective when applied to large-scale incomplete information systems to extract knowledge. Multigranular computing provides a powerful tool for use in big data analysis conducted at different levels of information granularity. In this paper, we present a novel multigranulation super-trust fuzzy-rough set-based attribute reduction (MSFAR) algorithm to support the formation of hierarchies of information granules of higher types and higher orders, which addresses newly emerging data mining problems in big data analysis. First, a multigranulation super-trust model based on the valued tolerance relation is constructed to identify the fuzzy similarity of the changing knowledge granularity with multimodality attributes. Second, an ensemble consensus compensatory scheme is adopted to calculate the multigranular trust degree based on the reputation at different granularities to create reasonable subproblems with different granulation levels. Third, an equilibrium method of multigranular-coevolution is employed to ensure a wide range of balancing of exploration and exploitation and can classify super elitists' preferences and detect noncooperative behaviors with a global convergence ability and high search accuracy. The experimental results demonstrate that the MSFAR algorithm achieves a high performance in addressing uncertain and fuzzy attribute reduction problems with a large number of multigranularity variables.

Index Terms—Multigranulation super-trust model, fuzzy-rough attribute reduction, valued tolerance relation, ensemble consensus compensatory scheme

---

W. Ding is with School of Information Science and Technology, Nantong University, Nantong 226019, China (e-mail: dwp9988@163.com).
W. Pedrycz is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6R 2V4, Canada (e-mail: wpedrycz@ualberta.ca).
I. Triguero is with the School of Computer Science, University of Nottingham, Nottingham NG7 2RD, United Kingdom (e-mail: Isaac.Triguero@nottingham.ac.uk).
Z. Cao is with the Discipline of ICT, School of Technology, Environment and Design, College of Science and Engineering, University of Tasmania, Australia (e-mail: zhcaonctu@gmail.com).
C.-T. Lin is with the Centre for Artificial Intelligence, FEIT, University of Technology Sydney, Ultimo, NSW 2007, Australia (e-mail: chin-teng.lin@uts.edu.au).

I. INTRODUCTION

Recently, there has been explosive growth in the amount of data generated, and the term 'big data' is being used to refer to the challenges of handling data with a high volume, variety, velocity, intrinsic value and uncertain veracity. These 'five V's' are the key features defining the essence of big data [1][2]. Big data has attracted much attention from a variety of circles of scientific research, marketing, business management, and government decision making, leading to an upsurge of research [3][4][5][6][7][8]. Although a large candidate set of attributes is provided in big data problems, most may be redundant or irrelevant, which highly diminishes the learning performance of decision-making algorithms. The complexity of the big data problem mainly arises from the very large number of decision variables and various types of constraints. Thus, it has become highly desirable to develop some effective attribute reduction (feature selection) methods to extract useful knowledge hidden in large-scale data repositories. Since big data can often be incomplete, uncertain and vague in reality, conventional knowledge discovery techniques, ranging from models, algorithms, and systems to applications, have been challenged in terms of how to store, manage, process, and analyze the complex attribute sets of big data [9][10][11].

With the increase of big data, researchers have started realizing the existence of data space alongside natural and social spaces and shown remarkable interest in its exploration. Structuralized knowledge organization and reasoning is considered an effective paradigm for handling large-scale tasks. In the recent past, a considerable amount of work has focused on a new research area—granular computing (GrC)—which is situated against the background of other human-centered information processing paradigms. GrC, which is a term coined by Zadeh [12], refers to a new knowledge representation and reasoning paradigm with information granules. Fuzzy sets and rough sets are the two main formal frameworks among active branches of granular computing [13], which are of vital importance for the understanding of big data analysis completed at different granularity levels. They provide two powerful conceptual and algorithmic vehicles for multiple-view data analysis. Fuzzy set theory introduced by Zadeh in 1965 [14] is a formal mechanism by which to represent and manipulate concepts with ambiguous boundaries and to understand and apply the processes employed in human reasoning. However, a fuzzy set is characterized by only a membership function, which ignores uncertain information and thus degrades its performance in big data analysis. Rough set theory proposed by Pawlak in 1982 [15] has been applied to contend with uncertainty caused by indiscernibility and incompleteness [16][17][18][19][20][21].

Since rough set theory is complementary to fuzzy set theory, fuzzy-rough sets have appeared as a newly emerging combination delivering the advantages of both complementary areas and are considered to provide a more powerful model for analyzing uncertainty in big data [22][23][24]. A fuzzy-rough set is defined by two fuzzy sets, fuzzy upper and lower approximations, which are obtained by extending the corresponding notions of a rough set.

In the fuzzy-rough framework, elements have membership grades located within some range, which allows for greater flexibility in handling uncertain information [25]. In the Boolean case, elements that belong to the lower approximation are represented as belonging to the approximated set with absolute certainty. Therefore, it has become timely and strongly justified to develop effective fuzzy-rough set algorithms with multigranulation to enhance understanding and reasoning in big data analytics. Type-2 fuzzy sets, as a higher type with a higher order of information granules, extend the expressive capabilities of Type-1 fuzzy sets, and they are able to represent the imprecision of the membership function of fuzzy sets. A Type-2 interval number (IN) is a mathematical object that can be interpreted either probabilistically or possibilistically. The use of an IN is particularly appropriate when modeling linguistic concepts [26][27]. Type-2 fuzzy sets have the potential to model uncertainties despite the large number of associated computations, especially when applied to non-real-time applications [28][29]. However, Type-2 fuzzy sets do not place any constraints on the continuity or other properties of their embedded sets.

Fuzzy-rough set models provide a method by which discrete or a real-valued noisy data or a mixture of both can be greatly reduced so that it can be effectively applied to both regression and classification of large-scale datasets. Fuzzy-rough set research has attracted considerable attention in recent years. Some approaches have been proposed to improve the performance of traditional fuzzy sets and rough sets as follows. Wang *et al.* [30] presented a new nearest-neighbor clustering classification algorithm based on fuzzy-rough set theory, in which every training sample was made according to fuzzy roughness, and then training sample points in a class boundary or overlapping regions were removed. Hassanien [31] introduced a hybrid scheme in conjunction with statistical feature extraction techniques by combining the advantages of both rough sets and fuzzy sets, wherein rough sets were employed for the reduction generation of the minimal number of features, and fuzzy sets were considered as an image preprocessing technique to enhance the contrast of the whole image. It was reported, however, that fuzzy-rough sets are sensitive to noisy samples. To alleviate this shortcoming, Hu *et al.* [32] discussed why the models of rough sets are sensitive to noise and developed some robust fuzzy-rough set models based on fuzzy lower approximations. Petrosino and Salvi [33] presented a multiscale algorithm based on rough fuzzy sets in which rough sets handled the vagueness and fuzzy sets handled the coarseness. Sarkar [34] generalized the concept of rough membership functions to rough–fuzzy membership functions, wherein the value signified the rough uncertainty as well as the fuzzy uncertainty associated with the pattern. An *et*

al. [35] proposed a novel robust data-distribution-aware fuzzy-rough set model by computing lower and upper approximations. However, the proposed models cannot be used to handle multimodal big data in real-world applications. Xu *et al.* [36] put forward a novel data redundancy reduction approach based on both fuzzy-rough set theory and information theory. Salama [37] provided fuzzy-rough attribute reduction software to facilitate the reduction of high-dimensional data. Zeng *et al.* [38] proposed the fuzzy-rough set approach for incremental feature selection in hybrid information systems. Maji and Garai [39] presented an interval type-2 (IT2) fuzzy-rough feature selection method, judiciously integrating the merits of the IT2 fuzzy set and rough sets to effectively reduce real-valued noisy features. Zhao *et al.* [40] analyzed the nested topological structure of fuzzy-rough sets with incremental parameters and designed a novel algorithm to compute a nested classifier by reflecting all possible parameters. Feng and Mi [41] considered the multigranulation fuzzy-rough sets of an information system by the minimal and maximal membership degrees based on multifuzzy tolerance relations. Yang *et al.* [42] presented two incremental algorithms for attribute reduction with fuzzy-rough sets for one and multiple incoming samples. Wang *et al.* [43] introduced a fitting fuzzy-rough set model to guarantee the maximal membership degree of a sample to its own category. Because fuzziness is employed in rough set theory, more reduction information relevant to continuous attributes can be successfully acquired. Hu *et al.* [50] proposed a model of multikernel fuzzy-rough sets and described a parallel strategy to handle large-scale multimodality fuzzy data attribute reduction.

Although fuzzy-rough attribute reduction methods have shown promising performance, they cannot cope well with the multimodality of big data and a large variety of real-world applications that involve the challenging complexity of big data. In practice, most attribute reduction and classification tasks are associated with mixtures of numerical and categorical attribute features. The size of multimodal big data is usually very large, resulting in extensive time consumption and the use of massive parallel processing databases in performing attribute reduction. Obviously, this can be greatly detrimental to the traditional attribute reduction performance for analyses of incomplete large-scale information systems. Furthermore, noisy attributes are also one of the main sources of uncertainty in big data applications. Although a few fuzzy-rough attribute reduction methods are robust toward complex noisy attributes, they require more user-supplied information, and there is a lack of continuity and inheritance in their internal relationships, which results in unsatisfactory performances. Meanwhile, when the volume of massive data objects increases in the database, much more computing time and space overhead are necessary to address the new rendering decision attributes. Currently, few works have considered the multigranulation algorithm for big data analysis at different granularity levels, and there has been a shortage of fuzzy-rough set analytical research.

The focus of this paper is to devise a fuzzy-rough attribute reduction approach capable of addressing structurally complex and granular large-scale attributes. The efficiency of fuzzy-rough attribute reduction algorithms in a very

large-scale dataset is an important research topic for the future. The multigranulation-fuzzy-rough model is an appropriate solution by which to accelerate the process of finding attribute reduction sets. In this paper, we propose a novel multigranulation super-trust fuzzy-rough attribute reduction (MSFAR) algorithm to support the formation of hierarchies of information granules of higher types and higher orders. MSFAR is suitable not only to address newly emerging attribute reduction problems associated with an irregular distribution of changing large-scale datasets but also to satisfy scenarios with complex noisy attributes. Furthermore, the multigranulation super-trust model offers a new way to classify data with different degrees of overlap, resulting in the creation of reasonable subproblems with different levels of granulation. Its main advantages are its high efficiency and robustness. Therefore, the main contributions of this paper are as follows:

(1) A multigranulation super-trust model based on valued tolerance relations is constructed to identify the fuzzy similarity of the changing knowledge granularity for fuzzy classification with multimodality attributes, which effectively solves the attribute reduction problem of missing data in a large-scale information system.

(2) An ensemble consensus compensatory scheme is adopted to calculate the multigranular trust degree in different granularities, resulting in the creation of reasonable subproblems with different granulation levels, from coarsened to refined.

(3) An equilibrium method of multigranular-coevolution is employed to ensure a wide range of balancing of exploration and exploitation. This strategy can classify super elitists' preferences and detect noncooperative behaviors with a global convergence ability and high search accuracy.

The remainder of the paper is organized as follows. We provide background information about the fuzzy-rough set model based on the valued tolerance relation in Section II. Section III introduces a novel multigranulation super-trust model with a self-evolving compensatory scheme, wherein the multigranulation super-trust model, the ensemble consensus compensatory scheme, and an equilibrium method of multigranular-coevolution are described in detail. In Section IV, the primary steps of MSFAR are detailed. Extensive experimental evaluations are described in Section V. Finally, conclusions are drawn in Section VI.

II. FUZZY-ROUGH SET MODEL BASED ON VALUED TOLERANCE RELATIONS

This section provides the relevant definitions for the fuzzy-rough set model based on valued tolerance relations.

Definition 1 [15] In the rough set theory, the universe is divided into a set of equivalence classes according to the attribute values of objects. An information system can be defined as a decision table by $T = (U, C, D, V, f)$, where C is the set of condition attributes, D is the set of decision attributes, V is the value set of all attributes, and $f : U \times \{C \bigcup D\} \to V$ is the information function, such that $f(x, a) \in V_a$ for each $a \in \{C \bigcup D\}, x \in U$.

Definition 2 [16] [17] For $P \subseteq \{C \bigcap D\}$, an equivalence relation $IND(P)$ is defined as follows.

$$IND(P) = \{(x, y) \in U \times U \mid \forall a \in P, f(x, a) = f(y, a)\} \quad (1)$$

$IND(P)$ partitions U into disjoint subsets. Let U/P denote the family of all equivalence classes of relation $IND(P)$, i.e., $U/P = \{P_1, P_2, ..., P_i ..., P_n\}$, where P_i is an equivalence class of P, which is denoted as $[x_i]_P$. Note that equivalence classes are defined with respect to their own attribute sets. Equivalence classes U/C and U/D will be called condition and decision classes, respectively.

Definition 3 [14] Let $\{\tilde{D}_1, \tilde{D}_2, ..., \tilde{D}_r\}$ be a family of fuzzy sets on U if $\sum_{i=1}^{r} \tilde{D}_i(x) = 1 \ (\forall x \in U)$, where $\{\tilde{D}_1, \tilde{D}_2, ..., \tilde{D}_r\}$ is considered a fuzzy partition. For $\forall x \in U$, the fuzzy decision of x is defined by

$$\tilde{D}_i(x) = \frac{\left| [x]_B \bigcap D_i \right|}{\left| [x]_B \right|}, i = 1, 2, ..., r, \text{and } x \in U, \quad (2)$$

where $\tilde{D}_i(x)$ denotes the fuzzy decision of x to D_i.

Definition 4 For an object $x \in U$, the fuzzy positive region of D relative to B is defined as

$$POS_B(D)(x) = \bigcup_{i=1}^{r} \underline{B}D_i(x), \quad (3)$$

where $D_i(x)$ is a set of decision attributes.

Definition 5 [42] For each condition attribute $a \in A$, one can define a fuzzy binary relation R_a, which is called a fuzzy equivalence relation if R_a is reflexive $(R(x, x) = 1)$, symmetric $(R(x, y) = R(y, x))$, and sup-min transitive $(R(x, y) \geq sup_{z \in U} min\{R(x, z), R(z, y)\})$ for $\forall x, y \in U$. A subset $B \subseteq A$ can also define a fuzzy equivalence relation, denoted by $R_B = \bigcap_{a \in B} R_a$. Based on the fuzzy equivalence relation, the concept of fuzzy rough set is defined as follows.

Let $F(U)$ be the fuzzy power set of U and $B \subseteq A$. For each $x \in U$, a pair of lower and upper approximation operationer of $X \subseteq F(U)$ based on R_B is defined by

$$\underline{R_B}(X)(x) = inf_{u \in U} max\{1 - R_B(x, u), X(u)\}, \quad (4)$$

$$\overline{R_B}(X)(x) = sup_{u \in U} min\{R_B(x, u), X(u)\}. \quad (5)$$

where the lower approximation $\underline{R_B}(X)(x)$ is considered as the degree of x certainly belonging to X, while the upper approximation $\overline{R_B}(X)(x)$ is the degree of x possibly belonging to X.

$(\underline{R_B}(X), \overline{R_B}(X))$ is referred to as the fuzzy rough set of X with respect to B, which is defined based on Max t-conorm and Min t-norm.

Definition 6 For the fuzzy-rough attribute reduction process, it is necessary to determine the dependency degree of the decision features. The dependency function of D relative to B is formally described by

$$\partial_B(D) = \frac{\sum_{x_i \in U} POS_B(D)(x_i)}{|U|}, \qquad (6)$$

where $0 \le \partial_B(D) \le 1$, and it is defined as the ratio of the sizes of the positive region relative to all samples in the feature space.

Definition 7 The main thought behind the attribute reduction process using fuzzy rough sets is to find a minimal subset of attributes that keeps the positive region unchanged, so that those wiped features will not affect the decision making. Considering a decision table $(U, A \bigcup D)$, a subset $Red \subseteq A$ is called a reduct of A relative to D if the following conditions are satisfied.

$$i) \quad \text{for } \forall x \in U, Pos_A(D)(x) = Pos_{Red}(D)(x); \qquad (7)$$

$$ii) \forall a \in P, \exists y \in U, Pos_{P-\{a\}}(D)(y) = Pos_{Red}(D)(y). \qquad (8)$$

Notice that the reduct is usually not unique. Let $Red_D(A)$ denote the set of all reducts with respect to $(U, A \bigcup D)$, and then $Core_D(A) = \bigcap Red_D(A)$ is called the core of $(U, A \bigcup D)$. It is easier to obtain the core first and then find a reduct based on the core.

Definition 8 [44][45] In the improved quantitative tolerance model, for an incomplete information system $DIIS = \{U, AT, V, f\}, a \in B \subseteq AT$,

$V_a' = \{(k_a^1, v_1), (k_a^2, v_2), \ldots, (k_a^m, v_m)\}$, where $v_1, v_2, \ldots, v_m$ are all the possible values of the attribute a. $k_a^i = |\{x \in U | a(x) = v_i\}|$, where $/ \bullet /$ denotes the cardinality of a set. $\forall x \in U, a(x) = v_i$, and its probability is $k_a^i / (k_a^1 + k_a^2 + \ldots + k_a^k)$. $\forall x, y$. Their similarity degree based on the value of $a \in A$ is defined by

$$VS_{\{a\}}(x, y) = \begin{cases} 1, & a(x) = a(y) \wedge a(x) \ne * \wedge a(y) \ne * \\ 0, & a(x) \ne a(y) \wedge a(x) \ne * \wedge a(y) \ne * \\ k_a^i / (\sum_{j=1}^{m} k_a^j), & a(x) = v_i \wedge a(y) \ne (a(x) \ne * \wedge a(y) = v_i) \\ \sum_{j=1}^{m} (k_a^j / (\sum_{j=1}^{m} k_a^j)), & a(x) = * \wedge a(x) = * \end{cases}$$

$$(9)$$

Therefore, the similarity between x and y in $A \subseteq AT$ is defined as

$$VT_A(x, y) = \prod_{v_i \in A} VT_{v_i}(x, y), \qquad (10)$$

and the quantitative tolerance class of x in $A \subseteq AT$ is denoted by

$$VT_A^w(x) = \{y = U : VT_A(x, y) \ge w\} \bigcup \{x\}. \qquad (11)$$

Definition 9 A multigranular valued tolerance relation is considered a good decomposition fuzzy-rough set framework for addressing large-scale problems with dynamically increasing complexity. For the information system $DIS = \{U, A, V, f\}, A_1, A_2, \ldots, A_m \subseteq A$, which corresponds to the classification threshold with a sequence of m sequence of attribute sets. $\forall X \subseteq U$, where the upper approximation

and lower approximation based on the grade multigranulations are defined as follows.

$$\sum_{i=1}^{m} A_i^0 w_i X = \{x \in U : VT_{A_1}^{w_1} \subseteq X \vee VT_{A_2}^{w_2} \subseteq X \vee \ldots \vee VT_{A_m}^{w_m}\}, (12)$$

$$\overline{\sum_{i=1}^{m} A_i^0 w_i X} = \square \underline{\sum_{i=1}^{m} A_i^0 w_i (\square X)}. \qquad (13)$$

Therefore, $(\underline{\sum_{i=1}^{m} A_i^0 w_i X}, \overline{\sum_{i=1}^{m} A_i^0 w_i X})$ is represented as the optimistic multigranularity fuzzy-rough set model based on quantitative tolerance.

III. MULTIGRANULATION SUPER-TRUST MODEL WITH A SELF-EVOLVING COMPENSATORY SCHEME

Traditional attribute reduction methods are satisfactory to a certain extent, but they are not capable of addressing massive amounts of complex large-scale data. Thus, there is a need for devising an effective trust method to efficiently handle the inherent multimodality attributes characteristics of big data. A multigranulation super-trust fuzzy-rough set model based on the valued tolerance relation is constructed to extract the fuzzy similarity of changing knowledge granularity for fuzzy classification. This model effectively solves the problem of missing data in an incomplete large-scale information system. With the increasing dimensionality of multigranulation space, most approaches to extract the fuzzy similarity of knowledge granularity are easily trapped in local optima due to overexploitation; therefore, their performance deteriorates. To achieve a better balance between the exploration and exploitation of knowledge granularity for solving complex large-scale data sets, we propose a novel multigranulation super-trust model with a self-evolving compensatory scheme to calculate the multigranular trust degree according to Eq. (17) at different granularities, thereby splitting the large dataset into reasonable subdatasets. In addition, this model can explore the search space and locate the global best region during the fuzzy-rough attribute reduction process, as well as accelerate the premature convergence speed.

A. Multilgranulation Super-Trust Model

Since there are various methods for calculating the credibility of a population, a practical method aims to process the multigranulation super-trust framework to adjust trust relationships based on subpopulations' interactions in different granularity spaces. We construct a granu-population architecture according to the trust degree of evolutionary elitists, including super elitists, denoted by "●", and ordinary elitists, denoted by "◉". The credibility of subpopulations in the same granu-population is calculated according to the trust calculation mechanism based on their respective reputations, which has been proven to be a good reflection of trust relationships between subpopulations of different granularities. Since the interaction between elitists in the same granu-population will be more frequent than those between granu-subpopulations in the common topology, the trust degree can be quickly established, and the granu-subpopulation can be evaluated effectively. The dynamic trust execution process is described in Fig. 1. Two types of super-trust relationships are employed to play two roles within different granu-subpopulations for fuzzy-rough attribute reduction. As specified in Fig. 2, the super-trust

relationships between the super elitist and the ordinary elitist and between ordinary elitists are both direct trust relationships, and those between the super elitists within different granu-subpopulations are recommendation trust relationships. The main steps of the process are described in **Algorithm 1**.
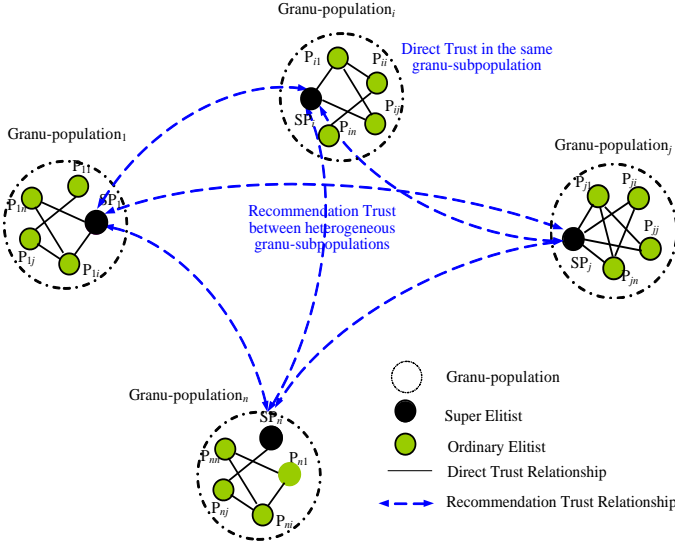

Fig. 1. Dynamic trust execution process of the multigranularity super-trust model

---

Algorithm 1: Multigranulation Super-Trust Model (MSTM)

1. Set the number of granularity subpopulations N $(N \geq 2)$ and initialize N granularity subpopulations GS_h, where $h \in \{1, ..., N\}$.

2. Initialize the first granularity subpopulation by assigning the collective preference in round P_t to granularity subpopulation center GS_1^t. Then, initialize the second granularity subpopulation center GS_2^t as the elitist preference E_i^t, which is the farthest from E_c^t.

3. For $GS_h^t (h \geq 3)$, compute the minimum distance between each of the remaining ordinary elitists' preferences E_i^t and all current initial granularity subpopulation centers and find the super elitist preference whose minimum distance is the largest by

$$MS(GS_h^t) = \max_i \left(\min_{u<h} d(E_i^t, GS_u^t) \right). \quad (14)$$

Assign it to GS_h^t. Repeat this step until all N granularity subpopulations are initialized.

4. In the granu-population$i$, to compute the distances between preference relations (both elitists' preferences and cluster centers indistinctly), the trust degree of different elitists in the same granularity subpopulation is defined by

$$dev(D_i) = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \| P_{ij} - SP_i \|}, \quad (15)$$

where n is the total number of elitists, SP_i is the super

elitist, and P_{ij} is the ordinary elitist in the granu-population$i$.

5. Compute the trust degree of each super elitist SP_i toward each granularity subpopulation center $GS_h^t(h \geq 2)$, $_{uGS_h}(P_i) \in [0,1]$, as follows:

$$_{uC_h}(P_i) = \frac{(1/d(P_i, C_h))^{1/(b-1)}}{\sum_{u=1}^{N} (1/d(P_i, C_u))^{1/(b-1)}}. \quad (16)$$

6. The granularity subpopulation centers $GS_h^t(h \geq 2)$ and cluster trust degrees $_{uGS_h}(P_i)$ are updated iteratively. Reputations represent the trust degrees of different granu-populations, which can be expressed as follows:

$$GS_h = \frac{\sum_{i=1}^{m} {}_{uGS_h}(P_i)P_i}{\sum_{i=1}^{m} {}_{uGS_h}(P_i)}. \quad (17)$$

7. The similarity between granularity subpopulation $GS_h^t(h \geq 2)$ is determined in the current round t, $t \in \{2, ..., Maxrounds - 1\}$, and each granularity subpopulation $C_u^{t-1}(u \geq 2)$ is computed in the previous round $(t-1)$. Thus, the scale of the subdatasets is dynamically updated iteratively by the trust degree relationships based on the subpopulation interactions in different granularity spaces.

8. A granularity subpopulation similarity measure $sim\left(GS_h^t, GS_u^{t-1}\right)$ is defined by

$$sim\left(GS_h^t, GS_u^{t-1}\right) = 1 - \frac{\sum_{i=1}^{m} \Delta_{hu}^t(P_i)}{m}, \quad (18)$$

where $\Delta_{hu}^t(P_i) \in [0,1]$ is the variation in the P_i membership of both granularity subpopulations, which is computed by

$$\Delta_{hu}^t(P_i) = \left| \mu_{GS_h}^t(P_i^t) - \mu_{GS_u}^{t-1}(P_i^{t-1}) \right|. \quad (19)$$

For the granularity subpopulation GS_h^t, if $\exists GS_u^{t-1} : sim(GS_h^t, GS_u^{t-1}) \geq k$, $k \in [0,1]$ is a similarity threshold. Then, GS_h^t and GS_u^{t-1} are assumed to represent the same granu-population.

B. Ensemble Consensus Compensatory Scheme

In this section, an ensemble consensus compensatory scheme is presented to evaluate the trust information of recommended granu-subpopulations. Its main novelty lies in calculating the multigranular trust degrees in different granularities based on the fuzzy granularity compensatory scheme, as well as addressing reasonable subproblem splitting. By analyzing the change in the knowledge granularity produced by coarsening and refining in the process of attribute reduction, the overall performance is greatly improved. Fig. 2 shows the framework of the ensemble

consensus compensatory scheme, and its main steps are described in **Algorithm 2**.
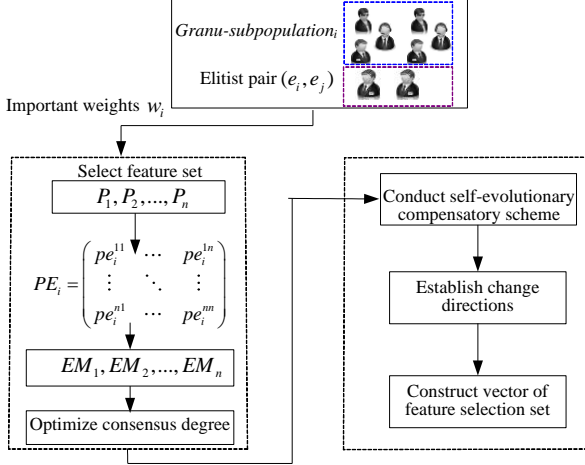

Fig. 2. Ensemble consensus compensatory scheme

Algorithm 2: Ensemble Consensus Compensatory Scheme (ECCS)

1. Dedicate each granu-population to its corresponding attribute set. The fitness evaluation is distributed equally among the participating *Granu-subpopulationi* with the importance weight w_i of the ith elitist. *Granu-subpopulationi* is evaluated in its own domain.

2. Compute the weights by the fuzzy measures over each *Granu-subpopulationi*. $\left(d_{i\sigma(1)} \geq d_{i\sigma(2)} \geq \ldots \geq d_{i\sigma(n)} \right)$ is the permutation of the evaluation values of the alternative A_i, and $C_{i\sigma(j)}$ is the criterion corresponding to $d_{i\sigma(j)}$. The weight criterion is calculated by
$$w_{i\sigma(j)} = \mu\left(A_{i\sigma(j)} \right) - \mu\left(A_{i\sigma(j-1)} \right), \quad j = 1,2,\ldots,n, \quad (20)$$
where $A_{i\sigma(j)} = \left\{ C_{\sigma(t)} \middle| t \leq j \right\}$ for $j \geq 1$, $A_{i\sigma(0)} = \phi$. $w_{i\sigma(j)}$ is the weight of criterion $C_{i\sigma(j)}$, and $\mu\left(A_{i\sigma(j)} \right)$ is the fuzzy measure.

3. Based on the given weight vector $w = (w_1, w_2, \cdots, w_m)^{\mathrm{T}}$, the intuitionistic fuzzy matrix is a matrix of pairs of nonnegative weight numbers, which is constructed by
$$I = (a_{ij})_{n\times m}(i=1,2,\cdots,n; j=1,2,\cdots,m). \quad (21)$$

4. Each super elitist $SP_i \in E$ provides its preference for alternatives according to the fuzzy preference relation $P_i = (p_i^{lk})^{n\times n}$, which consists of a matrix of assessments p_i^{lk} for each pair of super elitists $(x_i, x_j), l, k \in \{1,2,\ldots,n\}$. The ensemble consensus of preferences can be improved if the super elitists provide reciprocal assessments. If $p_i^{lk} = p, p \in [0,1], l \neq k$, then $p_i^{lk} = 1 - p$.

5. Compute the similarity degree of each pair of super elitists (SP_i, SP_j), and then the similarity matrix

$SM_{ij} = (sm_{ij}^{lk})^{n\times m}$ is defined by
$$SM_{ij} = \begin{pmatrix} sm_{ij}^{11} & \cdots & sm_{ij}^{1n} \\ \vdots & \ddots & \vdots \\ sm_{ij}^{n1} & \cdots & sm_{ij}^{nn} \end{pmatrix}_{n\times n}, \quad (22)$$
where $sm_{ij}^{lk} \in [0,1]$ is the similarity degree between super elitists SP_i and SP_j in their assessments p_i^{lk} and p_j^{lk}, as obtained by the similarity function
$$sm_{ij}^{lk} = 1 - \sqrt{\left| p_i^{lk} - p_j^{lk} \right|}. \quad (23)$$

6. A consensus matrix $\mathbf{CM} = (cm^{lk})^{n\times m}$ is computed by aggregating similarity matrices, including the importance weights $w_{ij} \in [0,1]$ of each pair of super elitists (SP_i, SP_j). The weighted average of the similarity degrees of each elitist $cm^{lk} \in [0,1], (l \neq k)$ is computed by
$$cm^{lk} = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} w_{ij} sm_{ij}^{lk}}{\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} w_{ij}}. \quad (24)$$

7. If all super elitists are given equal importance weights, cm^{lk} is redefined by
$$cm^{lk} = \frac{\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} w_{ij} sm_{ij}^{lk}}{m(m-1)/2}, \quad (25)$$
where $m(m-1)/2$ is the number of different pairs of elitists (e_i, e_j) in the *granu-subpopulationi*.

8. Design the average weight $\overline{cm_{ij}^{lk}}$ associated with each pair of super elitists (e_i, e_j) to select each pair of features (x_l, x_k) as
$$\overline{cm_{ij}^{lk}} = \frac{w_{ij} sm_{ij}^{lk}}{\sum_{j=i+1}^{m} w_{ij}}, \quad (26)$$
where $w_{ij} \in [0,1]$, and the value is computed based on a single super elitist's weights w_i, w_j.

9. An ensemble consensus degree with a compensatory scheme is computed by three different levels:

(*i*) The level of pairs of super elitists (cp^{lk}): This value is obtained from **CM** as $cp^{lk} = cm^{lk}$, with $l, k \in \{1,\ldots,n\}, l \neq k$.

(*ii*) The level of super elitists (ca^l): The level of agreement of each super elitist $x_l = X$ is computed by
$$ca^l = \frac{\sum_{k=1, k\neq l}^{n} cp^{lk}}{n-1}. \quad (27)$$

(*iii*) The level of preference relation of the overall consensus

degree cr is obtained by

$$cr = \frac{\sum_{l=1}^{n} ca^l}{n}. \qquad (28)$$

10. Compare the consensus degree EM_i with a consensus threshold $\mu \in [0,1]$.

(*i*) If $EM_i \geq \mu$, the *ganu-subpopulation_i* moves to the attribute selection process.

(*ii*) If $EM_i < \mu$, the moderator advises the elitists to modif their preferences to increase the level of agreement in the following rounds.

C. Equilibrium Adjustment Strategy of Multigranular Coevolution

Information granularities are not completely independent, and they usually overlap and overlay each other. Therefore, a dynamic-approximation equilibrium adjustment strategy is needed in the multigranular-coevolution space to avoid the super elitists running into optima, thus providing good guidance for all elitists. While super elitists are in the local optima state according to the premature judgment mechanism, the mutative-scale equilibrium adjustment strategy is used to ensure a wide range of balancing of exploration and exploitation. This strategy can classify super elitists' preferences and detect noncooperative behaviors with a global convergence ability and high search accuracy. Fig. 3 shows the updating approximations of the equilibrium adjustment strategy when using different multigranularities for super elitists. Our projection produces different isosceles right triangles, and the arrow shows the direction of sorting super elitists in each triangle area. If two super elitists start with the lower granularity of a_{N_3}, the attribute values of the updating approximations converge to the equilibrium pair (a_{N_3}, a_{N_3}). Similarly, if both super elitists start with very a high granularity of a_{N_1}, the attribute values of the updating approximations convergence to the equilibrium pair (a_{N_1}, a_{N_1}). Thus, the results show that using the equilibrium adjustment strategy of multigranular-coevolution for super elitists based on isosceles right triangles can lead to an increase in the size of the basin of multigranularity attraction. Its main steps are described in **Algorithm 3**.

Algorithm 3: Equilibrium Adjustment Strategy of Multigranular Coevolution (EASMC)

1. Assuming that C_h^t and C_u^{t-1} are considered to be in the same elitist cluster, let $S_h^t = \sum_{i=1}^{m} \mu_{C_h^t}(P_i^t)$ and $S_u^{t-1} = \sum_{i=1}^{m} \mu_{C_u^{t-1}}(P_i^{t-1})$ be the sums of the elitists' membership degrees to cluster $C_h^t (h \geq 2)$, respectively. Analogously, let $S_1^t = \sum_{i=1}^{m} \mu_{C_1^t}(P_i^t)$ be the sum of the elitists' membership degrees to the collective preference.

2. If $S_h^t > S_u^{t-1}$, the super elitists become more assembled around C_h^t. On the other hand, if $S_h^t < S_u^{t-1}$, the super elitists become less assembled around $C_1^t = P_c^t$.

3. Perform a mixed competitive and cooperative convolution among the multigranularity subpopulations and suppose S_i as the ith super elitist. For $i=1$ to $|S_i|$, do the following:

(*i*) Insert the representative of S_i as $S_{i,rep}$ into P_i^C.

(*ii*) If $n_x > |S_i|$, select the competing super elitist randomly and insert competitors from the selected *Granu-subpopulation_i* into P_i^C. Otherwise, insert competitors from other granu-populations into P_i^C and insert random super elitists from S_i into P_i^C.

(*iii*) Assemble the complete solution with $S_{i,j}$ and the representatives from the other granu-populations, assign Pareto rank to $S_{i,j}$, and calculate niche count of $S_{i,j}$.

(*iv*) Update the representative super elitist of S_i and archive the nondominated solution.

(*v*) Determine the winning granu-population S_k and update $S_i = S_k$.

4. Super elitists' fuzzy membership degrees $_{uC_h}(P_i)$ to *granu-subpopulation_i* are computed using similarity measures, where the distance between preference P_i and the super elitist center C_h is represented as $d(P_i, C_h)$.

5. Calculate the consistency ratio *CR* for each of the super elitists, $C^t = (c_{ii}^t)$ and $D^t = (d_{ii}^t)$, with $t \in \{1, 2, ..., s\}$.
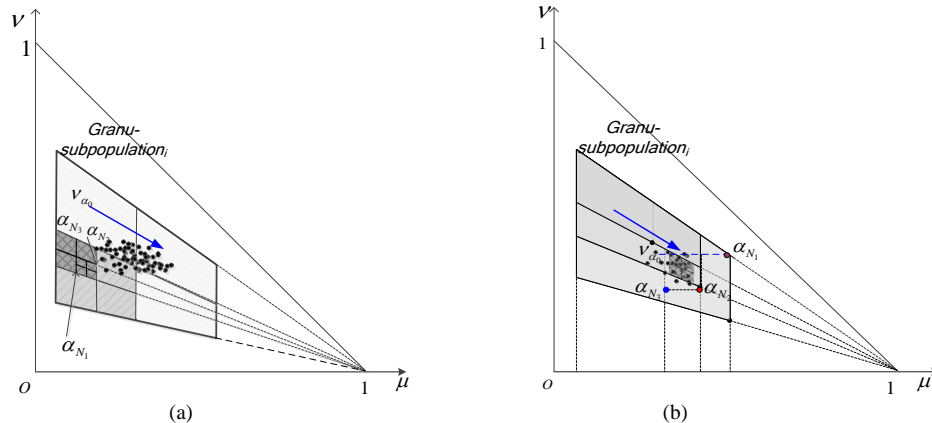

Fig. 3. Coarsening and refining attribute values for the updating approximations in multigranulation space.
(a) Dynamic adjustment, and (b) Equilibrium.

IV. PROPOSED MSFAR ALGORITHM

We propose a novel multigranulation super-trust fuzzy-rough attribute reduction (MSFAR) algorithm to support the formation of hierarchies of information granules of higher types and higher orders to support big data analysis. To accomplish this, we implement the described multigranulation fuzzy-rough sets and super-trust model with a self-evolving compensatory scheme to calculate the multigranular trust degree. It explicitly permits identifying the interdependent variables and adaptively decomposes them in the multigranulation space, so that the complexity and nonseparability of interdependent variables can be minimized among different fuzzy attribute subsets. It also extracts the fuzzy similarity of the changing knowledge granularity for fuzzy classification with multimodality attributes, effectively solving the problem of missing data in incomplete large-scale information systems. The proposed MSFAR algorithm can split a large dataset into reasonable subdatasets with the multigranulation super-trust model. It incorporates an additional multigranulation module and super-trust-coevolution to achieve the desirable goal of detecting complex interdependent variables, which can serve as a guide to carry out fuzzy-rough attribute reduction tasks with multigranulation flexible classification thresholds in big data. Its main steps are detailed as follows.

First, the fuzzy attribute sets are mapped into the evolutionary population space, and the fuzzy reduction model is completed as the optimization objective model.

Second, a multigranulation fuzzy-rough model based on the valued tolerance relation is constructed to identify the fuzzy similarity of the changing knowledge granularity for fuzzy classification with multimodality attributes, which effectively solves the problem of missing data in an incomplete large-scale information system. Then, an equilibrium method of multigranular coevolution is employed to classify super elitists' preferences and detect noncooperative behaviors.

Third, the multigranulation super-trust-coevolution model with a self-evolving compensatory scheme is adopted to calculate the multigranular trust degrees of different granularities, which represent their reputations in the group, resulting in the creation of reasonable subproblems with different levels from coarsened to refined granulation. It can self-adapt among different multigranulation layers and capture interdependent fuzzy-rough attribute subsets.

The pseudocodes of MSFAR are listed as follows:

Algorithm 5: Proposed MSFAR algorithm

---

1. Initialize the searching space of the fuzzy attribute sets and the granularity subpopulations by assigning the collective preference in round P_t, initialize the granularity subpopulations GS_h, and generate a list of candidate fuzzy-rough attribute subsets $(A_1, A_2, ..., A_n)$.

2. Decompose the fuzzy-rough attribute sets, compute equivalence class of the decision table, and classify super elitists' preferences using **Algorithm 1** (MSTM). Then, obtain $S = \{E_{r1}, E_{r2}, ..., E_{rm}\}$.

3. Conduct the ensemble consensus compensatory scheme by

---

using **Algorithm 2** (ECCS) to determine whether a granularity subpopulation is composed of the same super elitists as follows.

Suppose that GS_h^t and GS_u^{t-1} are considered to represent the two granularity subpopulations of two super elitists and ordinary elitists and that their super elitists' trust degrees $\mu_{GS_h}^t(P_i^t)$ and $\mu_{GS_h}^{t-1}(P_i^{t-1})$ have values close to each other for all $e_i \in E$. Then, conduct two similar granularity subpopulation compositions.

4. Perform the equilibrium adjustment strategy of multigranular coevolution using **Algorithm 3** (EASMC) and then obtain the corresponding perfect consistency equilibrium degree *CED* based on isosceles right triangles as $\tilde{C}^t = \left(\tilde{c}_{ij}^t\right)_{n \times n}$ $t \in \{1, 2, ..., s\}$ for any inconsistent *CED* $C^t = \left(c_{ij}^t\right)_{n \times n}$, where $\tilde{c}_{ij}^t = \dfrac{\tau_i}{\tau_j}$ for $i, j = 1, 2, ..., n$.

5. Obtain the perfect *CED* as $\tilde{D}^t = \left(\tilde{d}_{ij}^t\right)_{n \times n}, t \in \{1, 2, ..., s\}$.

Thus, a perfect consistency pair $\tilde{T}^t = <\tilde{C}^t, \tilde{D}^t>$ can be constituted by

$$\tilde{C}^t = \left(\tilde{c}_{ij}^t\right)_{n \times n} \text{ and } \tilde{D}^t = \left(\tilde{d}_{ij}^t\right)_{n \times n} \text{ for } t \in \{1, 2, ..., s\}. \quad (29)$$

6. The energy function is reconstructed by

$$EF(e_k \mid p_{N_i}^{(m)}) = \sum_{j=1}^{K} \mathbf{G}_{kj} v_j. \quad (30)$$

where $\mathbf{G}$ denotes a $K \times K$ matrix defining the connectivity between elitist populations k and j and v_j denotes the accumulated class probabilities in the neighborhood information N_i of elitist population i.

7. The definition of the multigranulation flexible threshold $FT_R(X)$ of R is formulated as

$$FT_R(X) = \frac{|B_{nR}(X)|}{|U|} \log_2 |U| - \left[\sum_{i=1}^{n} \frac{|X_i|}{|U|} \log_2 \frac{1}{|X_i|}\right], \quad (31)$$

where $B_{nR}(X)$ is the threshold region of the boundaries.

8. Conduct classification and learning of fuzzy-rough attribute sets with a parallel accelerating framework and $FT_R(X)$ as follows:

(i) Calculate the degree of dependency as the criterion for attribute selection and obtain the similarity degree $VS_{\{a\}}(x, y)$ based on quantitative tolerance.

(ii) Select feature subsets {subset₁,...,subsetᵢ,...,subsetₙ }.

(iii) Select the best feature subset FS_i^{best} for each granu-subpopulation and achieve a cascade feature set of fuzzy-rough attribute subsets as follows:

$$FS = FS_1^{best} \bigcup FS_2^{best}, ..., \bigcup FS_n^{best}. \quad (32)$$

9. Evaluate whether or not the accuracy of the fuzzy-rough attribute reduction is satisfied with respect to the predefined accuracy. If satisfied, then output the optimal se

$$\mathbf{FS}_{Opt} = \bigcup_{i=1}^{n} FS_i^{best} \text{ ; otherwise, go to Step } \mathbf{6}.$$

V. EXPERIMENTAL EVALUATION AND DISCUSSIONS

To validate the efficiency and effectiveness of the proposed MSFAR algorithm, we carry out a thorough series of experiments in this section, including a comparison of the computational times and accuracies of different algorithms and robustness comparisons for big datasets with consideration of the attribute noise. We provide performance comparisons between MSFAR and other representative algorithms for large-scale datasets and conduct robustness comparisons using attribute noise datasets for MSFAR versus representative algorithms.

A. Experimental Setup

All experiments are performed on computers with a Windows 10 operating system, an i7 3770k Intel CPU and 64 GB RAM using Java 11 programming language. We select five publicly available large-scale datasets from the UCI repository with heterogeneous attributes [46]. In addition, we produce three synthetic large-scale WEKA datasets with the WEKA data mining software [47], with these datasets involving a large number of samples with different statistical characteristics. Descriptive information about the eight datasets is given in Table I, where "KddCup99" is the network connectivity data set from USA air force simulation over nine weeks and "RLCP" means "Record Linkage Comparison Patterns". The datasets Susy and PokerHand are duplicated several times. We employ a stratified 10-fold cross-validation for data validation. The original dataset is equally partitioned into 10 parts, wherein two parts are used for testing and the remaining eight parts are used as the training set for attribute reduction. A classifier is then learned with the reduced training set and the classification accuracy is obtained on the reduced testing data. The cross-validation process is then repeated 10 times. The average values are calculated for the final performance. We compare the experimental results of MSFAR with the results achieved by other representative algorithms.

TABLE I
DESCRIPTION OF EXPERIMENTAL DATASETS

No	Datasets	Samples	Attributes
1	PokerHand	1,025,010	10
2	KddCup99	5,000,000	41
3	Susy	5,000,000	18
4	RLCP	5,749,132	14
5	Higgs	11,000,000	28
6	Weka-1.8G	32,000,000	10
7	Weka-3.2G	40,000,000	15
8	Weka-6.4G	80,000,000	15

B. Computational Times and Accuracies of Different Algorithms

To test the computational feasibility of MSFAR for use with large-scale datasets, we quantitatively compare its classification accuracy with those of some representative fuzzy attribute reduction algorithms, such as B-FRFS [48], MIBARFRAR [36], UFRFS [49], MFDAR [50] and MLFRS

[51]. To avoid the influence of random selection, independent runs for each dataset are repeated 50 times, and the averaged computational time and accuracy are presented as the final results. As described in Table II, we employ '$-$' to indicate no acceptable solution and a bold number to represent the best result of the computational time (time/s $\times 10_2$). As indicated in Table II, MSFAR clearly exhibits the highest speed of fuzzy-rough attribute reduction for most datasets compared with those of B-FRFS, MIBARFRAR, UFRFS, MLFRS and MFDAR. The attribute reduction time denotes the running time required for different fuzzy attribute reduction algorithms select features. For example, in the Weka-6.4G dataset, the attribute reduction time of MSFAR is 2,728 s, whereas the corresponding times for B-FRFS, MIBARFRAR, UFRFS, MLFRS, and MFDAR are 4,916 s, 6,321 s, 5,129 s, 4,723 s and 3,643 s, respectively. The computational time of MSFAR is only 55.49%, 43.16%, 53.19%, 57.76% and 74.88% of those of B-FRFS, MIBARFRAR, UFRFS, MLFRS and MFDAR, respectively. Similar results are evident for most datasets. Table III shows the numerical comparison of the selected features of different fuzzy attribute reduction algorithms. As described in Tables II and III, the proposed MSFAR algorithm can reduce the amount of redundant uncertain, unstructured and imprecise data and significantly improve the computational time. Therefore, it acquires the optimal number of selected features of the fuzzy-rough attribute sets, which is consistently a much better performance than those of its rivals for seven large-scale datasets. The main reasons behind these results are that the multigranulation super-trust-coevolution model with a self-evolving compensatory scheme employed in MSFAR is constructed to calculate the multigranular trust degree at different granularities and to split large datasets into reasonable subdatasets. MSFAR can consider both strongly relevant features and their corresponding correlated features simultaneously, and it selects important correlated features from a set of attribute features for classification.

In the following experiment, we further evaluate the classification accuracy of MSFAR for selected dynamically increasing sample sizes in the large-scale Higgs and Weka-1.8G datasets compared with those of representative algorithms. We employ only two classifiers, namely, SVM [52] and C4.5 [53], to process datasets whose attributes have been selected by six different methods, namely, B-FRFS, MIBARFRAR, UFRFS, MLFRS, MFDAR and MSFAR. Tables IV and V report the classification accuracy versus the dynamically increasing sample size of large datasets with the SVM classifier and C4.5 classifier, respectively. The six different attribute reduction algorithms result in the different sets of attributes for the large-scale increasing datasets. It is obvious that MSFAR significantly surpasses most of the representative algorithms. As an extreme case, both B-FRFS and MIBARFRAR fail with increasing sample sizes of the Higgs dataset because they are easily overwhelmed when processing high-dimensional large datasets. However, MSFAR achieves much better classification accuracy because it can benefit from the advantages of that the multigranulation fuzzy-rough set model can accurately capture interdependent variables associated with structurally complex and incomplete attribute sets and can greatly eliminate most irrelevant attribute sets without lowering the classification performance. For example, for the $80 \times 10_6$ sample size of Weka-1.8G, MSFAR can reach 97.81% classification accuracy, whereas for B-FRFS, UFRFS and MLFRS, these values are 92.21%, 93.43%, 94.62%, respectively. For the ever-growing

large-scale datasets, MSFAR performs significantly better. Therefore, the performance of MSFAR improves for large datasets—the larger is the dataset, the higher is the classification accuracy.

to a value indicates that the performance is statistically better than MSFAR, and a '*w*' symbol shows that the performance is worse statistically. The final line in Tables VI summarizes the count of the number of statistically better, equivalent and worse results for each representative fuzzy attribute reduction algorithm in comparison to MSFAR. The statistical

TABLE II
COMPUTATIONAL TIME (IN SECONDS $\times 10_2$) COMPARISON OF DIFFERENT FUZZY ATTRIBUTE REDUCTION ALGORITHMS

	Datasets	B-FRFS	MIBARFRAR	UFRFS	MLFRS	MFDAR	MSFAR
1.	PokerHand	29.61±0.46	27.89±1.41	32.45±0.52	33.09±1.29	26.34±0.20	**21. 35**±0.23
2.	KddCup99	45.18±1.28	40.98±0.35	36.89±0.38	41.51±1.41	34.90±0.62	**24.23** ±0.24
3.	Susy	34.61±1.41	38.02±1.48	39.36±2.51	34.07±0.32	30.12±0.23	**27.09** ±0.20
4.	RLCP	35.92±0.37	—	29.35±1.32	31.10±0.28	32.24±0.45	**23.97** ±0.31
5.	Higgs	—	30.91±0.42	20.51±0.34	19.25±0.41	23.21±0.32	**15.33** ±0.38
6.	Weka-1.8G	41.19±0.53	—	47.29±1.37	34.22±0.31	35.32±0.87	**32.23** ±0.28
7.	Weka-3.2G	31.68±2.47	38.11±0.68	23.90±1.51	24.19±0.48	24.32±0.39	**19.56** ±0.39
8.	Weka-6.4G	49.16±1.54	63.21±2.59	51.29±1.48	47.23±0.37	36.43±0.69	**27.28** ±0.40

TABLE III
NUMERICAL COMPARISON OF THE SELECTED FEATURES OF DIFFERENT FUZZY ATTRIBUTE REDUCTION ALGORITHMS

	Datasets	B-FRFS	MIBARFRAR	UFRFS	MLFRS	MFDAR	MSFAR
1.	PokerHand	6	6	7	8	8	9
2.	KddCup99	32	34	31	36	37	40
3.	Susy	13	14	14	15	16	18
4.	RLCP	12	—	11	12	13	14
5.	Higgs	—	18	22	24	28	25
6.	Weka-1.8G	7	—	6	7	8	8
7.	Weka-3.2G	12	12	13	12	14	15
8.	Weka-6.4G	12	11	13	11	12	15

As indicated in experimental results, MSFAR is suitable for use in fuzzy-rough attribute reduction and classification for large-scale increasing datasets, thereby overcoming the limitations of the representative algorithms.

The experimental results clearly indicate that the classification system employing MSFAR as the fuzzy-rough attribute reduction algorithm can acquire the optimal reduction results of structurally complex and incomplete attribute sets, and lead to an appealing performance in classification accuracy, irrespective of different classifiers.

C. Statistical Analysis

In order to sufficiently report classification accuracies, an appropriate statistical test need to be applied to evaluate the statistical significance of the results of the fuzzy-rough attribute reduction algorithms. In this paper, a paired t-test was used to determine the statistical significance of the results at the 0.05 level. Since a paired t-test is a parametric test, it always assumes approximate normality. The results are shown in Tables VI for the SVM classifier, where a 'b' symbol next

comparisons results between each compared algorithm and MSFAR indicate that only three datasets (KddCup99, Susy and RLCP) for which MSFAR is bettered by MLFRS and MFDAR, but for the remainder MSFAR achieves statistically equivalent to or better accuracy results than other representative fuzzy attribute reduction algorithms.

It is obvious in Table VI that B-FRFS and MIBARFRAR do not perform as well as does MSFAR in producing classification accuracy results which are consistent across the SVM classifier, as there are four or five cases wherein they produces worse statistical results compared with MSFAR. From the results reported in this table, it can be seen that the representative fuzzy attribute reduction algorithms attain a few of results which are statistically comparable to accuracy for the unreduced data, but MSFAR achieve significantly better classification results than them. Generally speaking, it can be concluded that MSFAR is superior to the other algorithms in the paired t-test. It offers the significance level for the tests whilst retaining the semantics of the data. The better performance of the proposed MSFAR algorithm is achieved due to the fact that it provides an efficient way to select a

reduced attribute set of real valued data sets, having maximum significance and relevance without lowering the classification performance.

D. Discussion

From the results for average classification accuracy presented in Tables IV and V, it can be observed that the multigranulation fuzzy-rough set model can perform better with lower average subset sizes in comparison to both the UFRFS and MLFRS algorithms. In particular, for the proposed MSFAR algorithm, from the abovementioned results, it is clear that the multigranulation model offers a greater reduction of the large-scale size. This reduction is to be expected since

there is much discriminative information contained in the decision features. The main reason behind this result is that the multigranulation fuzzy-rough set model can accurately capture interdependent variables associated with structurally complex and incomplete attribute sets and can greatly eliminate most irrelevant attribute sets without lowering the classification performance. The possible reasons behind this effect are that we construct a multigranulation fuzzy-rough set model based on a valued tolerance relation to extract the fuzzy similarity of

TABLE IV
CLASSIFICATION ACCURACY COMPARISON WITH THE SVM CLASSIFIER FOR THE INCREASING
LARGE-SCALE EVER-GROWING HIGGS DATASETS (TEST±STD /%)

Algorithms	Sample size ($\times 10_8$)				
	2	4	10	20	40
MSFAR	**91.13±0.23**†	**93.21±0.35**	**95.18±0.41**	**96.24±0.29**	**96.93±0.38**
B-FRFS	87.89±0.31	89.29±0.51	91.23±0.43	91.90±0.68	-
MIBARFRAR	82.19±0.28	84.39±0.42	84.23±0.54	85.07±0.31	-
UFRFS	87.12±0.56	91.72±0.21	92.29±0.27	93.10±0.43	93.65±0.12
MLFRS	88.21±0.28	89.36±0.76	90.11±0.65	92.54±0.36	94.65±0.52
MFDAR	89.23±0.23	92.11±0.28	91.47±0.49	92.45±0.67	93.78±0.34

TABLE V
CLASSIFICATION ACCURACY COMPARISON WITH THE C4.5 CLASSIFIER FOR THE INCREASING
LARGE-SCALE EVER-GROWING WEKA-1.8G DATASETS (TEST±STD /%)

Algorithms	Sample size ($\times 10_8$)				
	2	10	30	50	80
MSFAR	**92.38±0.15**	**93.39±0.46**	**95.90±0.28**	**96.32±0.76**	**97.81±0.41**
B-FRFS	88.19±0.28	89.89±0.35	90.29±0.37	91.89±0.25	92.21±0.32
MIBARFRAR	83.78±0.47	84.19±0.12	84.69±0.44	-	-
UFRFS	88.56±0.34	89.19±0.42	91.19±0.43	92.34±0.49	93.43±0.51
MLFRS	91.27±0.37	91.90±0.56	92.43±0.37	93.12±0.34	94.62±0.12
MFDAR	91.56±0.28	92.36±1.01	93.07±0.52	94.21±0.52	95.68±0.35

TABLE VI
STATISTICAL SIGNIFICANCE DETERMINED USING THE PAIRED T−TEST: SVM

Datasets	MSFAR	B-FRFS	MIBARFRAR	UFRFS	MLFRS	MFDAR
1. PokerHand	86.32	84.12	75.13 *w*	69.90 *w*	82.54	83.16
2. KddCup99	78.34	77.25	67.27 *w*	79.90	81.46 *b*	71.42 *w*
3. Susy	92.18	87.43	89.65	91.06	90.65	93.92 *b*
4. RLCP	83.33	67.83 *w*	81.65	82.56	87.43 *b*	80.90
5. Higgs	95.53	85.89 *w*	88.56	82.19 *w*	92.27	92.56
6. Weka-1.8G	92.27	84.45 *w*	87.92 *w*	91.35	88.07 *w*	92.04
7. Weka-3.2G	93.56	88.78	88.76	85.04 *w*	84.37 *w*	85.17 *w*
8. Weka-6.4G	94.96	82.87 *w*	84.27 *w*	91.57	87.65 *w*	90.38 *w*
Summary	*(b/ /w)*	*(0/4/4)*	*(0/4/4)*	*(0/5/3)*	*(2/3/3)*	*(1/4/3)*

the changing knowledge granularity for fuzzy classification with consideration of multimodality attributes, which effectively solves the problem of missing data in the large-scale information system. Meanwhile, the multilayered super-trust-coevolution model with a self-evolving compensatory scheme can calculate the multigranular trust degree at different granularities based on the reputation in the group to split the large problem into reasonable subproblems.

It is necessary for the algorithm to perform efficiently in real time so that it consistently functions without requiring a significant amount of computing resources. In our experiments, the time complexity of the proposed MSFAR algorithm is also analyzed by observing the real-time performance for selected big datasets of varying sizes. The experimental results indicate that the computational complexity of MSFAR is obviously less than those of UFRFS and MLFRS. Hence, if we were to perform the evaluation using larger datasets, the computational time would be unacceptable for the comparison algorithms, but MSFAR requires a smaller amount time to obtain the optimal solution.

Of course, in a few special cases, the results of MSFAR are slightly poorer than those of the representative algorithms. In general, the two main components in MSFAR should work together to allow it to obtain better results. Furthermore, MSFAR can dynamically adapt its main operators to suit various large-scale instances with dynamically increasing percentages of noise.

VI. Conclusions

Recently, big data has been an emerging topic that has attracted the attention of many researchers. The significant amount of unstructured, uncertain and imprecise large-scale data exhibits structurally complex and granular characteristics. Uncertainty data have been widely adopted for attribute reduction, but alone, they may be insufficient for use in batch feature selection. The recent progress in fuzzy-rough set approaches can be helpful for analyses of big data problems. In this paper, we present a novel multigranulation super-trust fuzzy-rough attribute reduction (MSFAR) algorithm for use in big data analysis at different granularity levels. A multigranulation fuzzy-rough set model based on a valued tolerance relation is constructed to identify the fuzzy similarity of the changing knowledge granularity for fuzzy classification with multimodality attributes. Meanwhile, the multigranulation super-trust-coevolution model with a self-evolving compensatory scheme is adopted to calculate the multigranular trust degree at different granularities, and it can be directly applied to a variety of knowledge analytical problems with continuous or numerical large-scale datasets. The experimental results demonstrate that MSFAR produces very good results. It is theoretically and experimentally indicated that the multigranulation super-trust-coevolution model with a self-evolving compensatory scheme can provide a much better performance by the MSFAR compared to those of the representative models. These represent important developments for improving the reasoning and understanding of big data.

In the era of big data, the size of large data usually dynamically increases, including current changing and interconnected datasets. It is time consuming to perform efficient attribute reduction and classification for these uncertain and redundant datasets. In the future, it is expected that using analytical methods to learn from big data can significantly improve the fuzzy-rough attribute reduction process. We will also explore the effective and robust multigranulation mechanisms of fuzzy-rough reduction estimation to achieve improved understanding of large-scale feature selection. We intend to exert great effort in promoting our research to offer a new avenue by which to address the problem of optimum predicting disorder from neonatal brain MRIs.

REFERENCES

[1] C.-H. Fung , Z.T. Ho Tse, and K.-W. Fu , "Converting big data into public health, " *Science* 347 (6222), 620.

[2] C. Min, M. Shiwen, and L. Yunhao, "Big data: A survey," *Mobile Netw. Appl.*, no. 2, pp. 171-209, Apr. 2014.

[3] C. P. Chen and C. Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," *Inf. Sci.*, vol. 275, pp. 314-347, Aug. 2014.

[4] H. V. Jagadish, J. Gehrke, and A. Labrinidis, et al., "Big data and its technical challenges," *Commun. ACM*, vol. 57, no. 7, pp. 86-94, Jul. 2014.

[5] A. Phinyomark, E. Ibáñez-Marcelo, and G. Petr, "Resting-state fMRI functional connectivity: Big data preprocessing pipelines and topological data Aanalysis," IEEE Trans. Big Data, vol. 3, no. 4, pp. 415-428, Oct.-Dec. 2017

[6] L. Dong, Z. Lin, and Y. Liang, et al, "A hierarchical distributed processing framework for big image data," *IEEE Trans. Big Data*, vol. 2, no. 4, pp. 297-309, Oct.-Dec. 2016

[7] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized Deep Learning for EEG Big Data and Seizure Prediction BCI via Internet of Things," *IEEE Trans. Big Data*, vol. 3, no. 4, pp. 392-404, Oct.-Dec., 2017

[8] A. Segatori, A. Bechini, P. Ducange, and F. Marcelloni, "A distributed fuzzy associative classifier for big data," *IEEE Trans. Cybern.*, Doi: 10.1109/TCYB.2017.2748225.

[9] S. Madden, "From databases to big data," *IEEE Internet Comput*. vol. 16, no. 3, pp. 4-6, May-Jun. 2012.

[10] J. Li, H. Liu, "Challenges of feature selection for big data analytics," *IEEE Intell. Syst.* vol. 32, no. 2, pp. 9-15, Mar.--Apr., 2017.

[11] M. Duan, K. Li, and X. Liao, "A parallel multiclassification algorithm for big data using an extreme learning machine," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2337-2351, Jun. 2018.

[12] L.A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets Syst.* vol. 90, no. 2,pp.111-127,1997.

[13] W. Pedrycz, "Granular computing: Analysis and design of intelligent systems," CRC Press/Francis Taylor, Boca Raton, 2013.

[14] L.A.Zadeh, "Fuzzy sets," *Inf. Control*,vol 8, pp.338-353, Jun. 1965.

[15] Z. Pawlak, "Rough sets," *Int. J. Comput. Inf. Sci.* vol. 11, no. 5, pp. 341-356, Oct. 1982.

[16] Z. Bonikowski, and E. Brynarski, "Extensions and intensions in rough set theory," *Inf. Sci.* vol. 107, no. 1-4, pp. 149-167, Jun. 1998.

[17] I. Couso, and D. Dubois, "Rough sets, coverings and incomplete information," *Fundam. Inform.*, vol. 108, pp. 223-247, Jan. 2011.

[18] A. E. Hassanien, A. Abraham, and J. F. Peters, "Rough sets and near sets in medical imaging: A review," *IEEE Transactions on Information Technology in Biomedicine*, vol.13, no.6, pp.955-968, Nov. 2009.

[19] W. Ding, C.-T. Lin, and Z. Cao, "Shared nearest-neighbor quantum game-based attribute reduction with hierarchical coevolutionary Spark and its application in consistent segmentation of neonatal cerebral cortical surfaces," *IEEE Trans. Neural Netw. Learn. Syst.,* vol. 30, no. 7, Jul. 2013-2027, 2019.

[20] A. Albanese, S. K. Pal, and A. Petrosino, "Rough sets, kernel set, and spatiotemporal outlier detection," *IEEE Trans. Knowl. Data Eng.,* vol. 26, no. 1, pp.194-207, Jan. 2014.

[21] W. Ding, C.-T. Lin, M. Prasad, et al, "A layered-coevolution-based attribute-boosted reduction using adaptive quantum behavior PSO and its consistent segmentation for neonates brain tissue," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1177-1191, Jun. 2018.

[22] D. G. Chen and Y. Y. Yang, "Attribute reduction for heterogeneous data based on the combination of classical and fuzzy rough set models," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 5, pp. 1325-1334, Oct. 2014.

[23] A.M. Radzikowska, and E. E. Kerre, "A comparative study of fuzzy rough sets," Fuzzy Sets and Systems, vol. 126, no, 2, pp. 137-155, 2002.

[24] W. Ding, C.-T. Lin, and Z. Cao, "Deep neuro-cognitive co-evolution for

fuzzy attribute reduction by quantum leaping PSO with nearest-neighbor memeplexes," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2744-2757, Jul. 2019.

[25] Y. Qian, H. Cheng, J. Wang, et al. "Grouping granular structures in human granulation intelligence," *Inf. Sci.*, vol. 382–383, pp.150-169, Mar. 2017.

[26] V.G. Kaburlasos, G.A. Papakostas, "Learning distributions of image features by interactive fuzzy lattice reasoning (FLR) in pattern recognition applications", *IEEE Comput. Intell. Mag.*, vol. 10, no. 3, pp. 42-51, Aug. 2015.

[27] V. G. Kaburlasos, S.E. Papadakis, and A. Amanatiadis, "Binary image 2D shape learning and recognition based on lattice-computing (LC) techniques," *J. Math. Imag. & Vis.*, vol. 42, no. 2-3, pp. 118-133, Feb. 2012.

[28] C. Wagner, S. Miller, J. M. Garibaldi, D.T. Anderson, T. C. Havens. "From interval-valued data to general type-2 fuzzy sets", *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 2, 248-269, Apr. 2015.

[29] M. Almaraashi, R. John, A. Hopgood, and S. Ahmadi. "Learning of interval and general type-2 fuzzy logic systems using simulated annealing: Theory and practice", *Inf. Sci.*, vol. 360, pp. 21-42, Sep. 2016.

[30] X. Wang, J. Yang, X. Teng, and N. Peng, "Fuzzy classifiers fuzzy rough set based nearest neighbor clustering classification algorithm," *Lecture Notes Comput. Sci.*, vol. 3613, pp. 370-373, 2005.

[31] A. Hassanien, "Fuzzy-rough sets hybrid scheme for breast cancer detection," *Image Comput. Vision J.*, vol. 25, no. 2, pp. 172-183, Feb. 2006.

[32] Q. H. Hu, L. Zhang, S. An, D. Zhang, and D. R. Yu, "On robust fuzzy rough set models," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 4, pp. 636-651, Aug. 2012.

[33] A. Petrosino and G.Salvi, "Rough fuzzy set based scales pace transforms and their use in image analysis," *Int. J. Approx. Reason.*, vol. 41, no. 2, pp. 212-228, 2006.

[34] M. Sarkar, "Rough-fuzzy functions in classification," *Fuzzy Sets Syst.*,vol. 132, no. 3, pp. 353-369, Dec. 2002.

[35] S. An, Q. Hu, W. Pedrycz, P. Zhu, and E. C. C. Tsang, "Data-distribution-aware fuzzy rough set model and its application to robust classification," *IEEE Trans. Cybern.*, vol. 46, no.12, pp.3073-3085, Dec. 2016.

[36] F. F. Xu, D. Q. Miao, and L. Wei, "Fuzzy-rough attribute reduction via mutual information with an application to cancer classification," *Computers and Mathematics with Applications*, vol. 57, no. 6, pp. 1010-1017, Mar. 2009.

[37] A. S. Salama, "Fuzzy rough data reduction in large datasets with building application software (FRDRS)," *Int. J. Granular Computing, Rough Sets and Intelligent Systems*, vol. 2, no. 4, pp. 292-313, 2012

[38] A. P. Zeng, T. R. Li, and D. Liu, et al, "A fuzzy rough set approach for incremental feature selection on hybrid information systems," *Fuzzy Sets Syst.* vol, 258, pp. 39-60, Jan. 2015.

[39] P. Maji and P. Garai, "IT2 fuzzy-rough sets and max relevance-max significance criterion for attribute selection," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1657-1668, Aug. 2015

[40] S. Zhao, H, Chen, and C. Li, et al, "A novel approach to building a robust fuzzy rough classifier," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 4, pp. 769-786, Aug. 2015

[41] T. Feng, and J. Mi, "Variable precision multigranulation decision-theoretic fuzzy rough sets," *Knowl.-Based Syst.*, vol. 91, pp. 93-101, Jan. 2016.

[42] Y. Yang, D. Chen, and H. Wang, et al, "Fuzzy rough set based incremental attribute reduction from dynamic data with sample arriving," *Fuzzy Set. Syst.*, vol. 312, pp. 66-86, Apr. 2017.

[43] C. Z. Wang, Y. L. Qi, and M. W. Shao, et al, "A fitting model for feature selection with fuzzy rough sets," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 4, pp. 741-753, Aug. 2017.

[44] Wang G,Guan L,Wu W, et al, "Data-driven valued tolerance relation based on the extended rough set," *Fundam. Inform.*, vol. 132 no.3, pp. 349-363, 2014.

[45] Xu Y, Li L S, "Extended rough set model based on known same probability dominant valued tolerance relation," *Int. J. Approx. Reason.*, vol.74, pp. 108-119, Jul. 2016.

[46] K. Bache and M. Lichman. (2013). UCI Machine Learning Repository. [Online]. Available: http://archive.ics.uci.edu/ml.

[47] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann, Fourth Edition, 2016. https://www.cs.waikato.ac.nz/ml/weka/datasets.html

[48] R. Jensen, and Q. Shen, "New Approaches to Fuzzy-Rough Feature Selection," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 4, pp. 824-838, Aug., 2009.

[49] N.M. Parthaláin, and R. Jensen,"Unsupervised fuzzy-rough set-based dimensionality reduction," *Inf. Sci.*,vol.229, pp.106-121, Apr. 2013.

[50] Q.Hu, L. Zhang, Y. Zhou, W. Pedrycz. "Large-scale multi-modality attribute reduction with multi-kernel fuzzy rough sets," *IEEE Trans. Fuzzy Syst.*, vol. 26 , no.1, pp. 226-238, Feb. 2018.

[51] Y. Lin, Y. Li, C. Wang, J. Chen, "Attribute reduction for multi-label learning with fuzzy rough set," *Knowl.-Based Syst.*, vol. 152, pp. 51-61, Jul. 2018.

[52] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no.3, pp. 27:1-27:27, May 2011.

[53] R. Quinlan, C4.5: Programs for Machine Learning. San Mateo, CA, USA: Morgan Kaufmann, 1993.

Weiping Ding (M'16-SM'19) received a Ph.D. in Computation Application, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 2013. He was a Visiting Scholar at University of Lethbridge (UL), Alberta, Canada, in 2011. From 2014 to 2015, he was a Postdoctoral Researcher at the Brain Research Center, National Chiao Tung University (NCTU), Hsinchu, Taiwan. In 2016, he was a Visiting Scholar at National University of Singapore (NUS), Singapore. From 2017 to 2018, he was a Visiting Professor at University of Technology Sydney (UTS), Ultimo, NSW, Australia. Now, Dr. Ding is the Chair of IEEE-CIS Task Force on Granular Data Mining for Big Data. He is a member of Senior IEEE, IEEE-CIS, and Senior CCF. His main research directions involve data mining, granular computing, machine learning and big data analytics. He has co-authored more than 60 research peer-reviewed journal and conference papers in this field, including IEEE TFS, IEEE TNNLS, IEEE TCYB, IEEE TII, IEEETSMC, IEEE TBME, etc, and he has held 15 approved invention patents till now. Dr. Ding was a recipient of Computer Education Excellent Paper Award (First-Prize) from the National Computer Education Committee of China, in 2009. He was an Excellent-Young Teacher (Qing Lan Project) of Jiangsu Province in 2014, a High-Level Talent (Six Talent Peak) of Jiangsu Province in 2016, and a Middle-aged and Young Academic Leaders (Qing Lan Project) of Jiangsu Province in 2019. He was awarded the Best Paper of ICDMA'15. Dr. Ding was a recipient of the Education Teaching and Research Achievement Award (Third Prize) of Jiangsu Province, China in 2018. He was a recipient of Outstanding Associate Editor of 2018 for IEEE Access Journal. Dr. Ding was awarded two Chinese Government Scholarships for Overseas Studies in 2011 and 2016. Dr. Ding currently serves on the Editorial Advisory Board of *Knowledge-based Systems* and Editorial Board of *Information Fusion* and *Applied Soft Computing*. He serves/served as an Associate Editor of *IEEE Transactions on Fuzzy Systems*, *Information Sciences*, *Swarm and Evolutionary Computation, IEEE Access,* and *Journal of Intelligent & Fuzzy Systems*, and a lead guest editor for four international journals. He has delivered more than 10 keynote speeches at international conferences and has co-chaired several international conferences and workshops in the area of fuzzy decision-making, data mining and knowledge engineering.

Witold Pedrycz (F'99) received M.Sc., Ph.D., and D.Sci. degrees from Silesian University of Technology, Gliwice, Poland. He is Professor and Canada Research Chair in Computational Intelligence in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He is also with the Systems Research Institute of the Polish Academy of Sciences, Warsaw, Poland. He holds an appointment of special professorship with the School of Computer Science, University of Nottingham, Nottingham, UK. His current research interests include computational intelligence, fuzzy modeling and granular computing, knowledge discovery and data mining, fuzzy control, and software engineering. He has published numerous papers in the above areas. He has also authored 15 research monographs. Dr. Pedrycz was a recipient of the prestigious Norbert Wiener Award from the IEEE Systems, Man, and Cybernetics Society in 2007, the IEEE Canada Computer Engineering Medal 2008, the Cajastur Prize for Soft Computing from the European

Centre for Soft Computing for "pioneering and multifaceted contributions to granular computing" in 2009, the Killam Prize in 2013, and the Fuzzy Pioneer Award 2013 from the IEEE Computational Intelligence Society. He was elected as a Foreign Member of the Polish Academy of Sciences in 2009 and a fellow of the Royal Society of Canada in 2012. He has been a member of numerous program committees of the IEEE conferences in the area of fuzzy sets and neurocomputing. Dr. Pedrycz is intensively involved in editorial activities. He is an Editor-in-Chief of *Information Sciences*, Editor-in-Chief of *WIREs Data Mining and Knowledge Discovery* (Wiley), and *Int. J. of Granular Computing* (Springer). He serves on an Advisory Board of *IEEE Transactions on Fuzzy Systems* and is a member of a number of editorial boards of other international journals.

Isaac Triguero received his M.Sc. and Ph.D. degrees in Computer Science from the University of Granada, Granada, Spain in 2009 and 2014, respectively. He has been an Assistant Professor of Data Science since June 2016. His work is mostly concerned with the research of novel methodologies for big data analytics. Dr Triguero has published more than 65 international publications in the fields of Big Data, Machine Learning and Optimization (h-index=21 and more than 1700 citations on Google Scholar). He is a Section Editor-in-Chief of the Machine Learning and Knowledge Extraction journal, and an associate editor of the Big Data and Cognitive Computing journal, and the IEEE Access journal. He has acted as Program Co-Chair of the IEEE Conference on Smart Data (2016), the IEEE Conference on Big Data Science and Engineering (2017), and the IEEE International Congress on Big Data (2018). Dr Triguero is currently leading a KTP project funded by Innovative UK and the energy provider E.ON that investigates Smart Metering data.

Zehong (Jimmy) Cao (M'17) is a Lecturer (a.k.a. Assistant Professor) with Discipline of Information and Communication Technology (ICT), School of Technology, Environments and Design, College of Sciences and Engineering, University of Tasmania, Hobart, Australia, and a Adjust Fellow with the Faculty of Engineering and IT, University of Technology Sydney (UTS), Australia. He received a dual Ph.D. degree in Information Technology from UTS, and Electrical and Control Engineering from National Chiao Tung University (NCTU), Taiwan. He received the MS and BS from The Chinese University of Hong Kong and Northeastern University, respectively. He serves as an Associate Editor of IEEE Access and the Journal of Intelligent and Fuzzy Systems and a Guest Editor of Swarm and Evolutionary Computation, Neurocomputing, and International Journal of Distributed Sensor Networks. He has 30+ publications in the most respected journals. He was awarded the UTS Centre for Artificial Intelligence Best Paper (2017), UTS Faculty of Engineering and IT Publication Award (2017), UTS President Scholarship (2015), and NCTU & Songshanhu Scholarship (2014). His research interests cover fuzzy sets and systems, fuzzy neural networks, game artificial intelligence, deep reinforcement learning, brain-computer interface and biosignal processing.

Chin-Teng Lin (S'88-M'91-SM'99-F'05) received a BS degree from National Chiao Tung University (NCTU), Taiwan in 1986, and Master and Ph.D. degrees in electrical engineering from Purdue University, USA in 1989 and 1992, respectively. He is currently the Distinguished Professor of Faculty of Engineering and Information Technology and Co-Director of Center for Artificial Intelligence, University of Technology Sydney, Australia. He is also Honorary Chair Professor of Electrical and Computer Engineering, NCTU and Honorary Professorship of University of Nottingham. Dr. Lin was elevated to be an IEEE Fellow for his contributions to biologically inspired information systems in 2005 and was elevated to International Fuzzy Systems Association (IFSA) Fellow in 2012. Dr. Lin received the IEEE Fuzzy Systems Pioneer Awards in 2017. He served as the Editor-in-Chief of IEEE Transactions on Fuzzy Systems from 2011 to 2016. He also served on the Board of Governors at IEEE Circuits and Systems (CAS) Society in 2005-2008, IEEE Systems, Man, Cybernetics (SMC) Society in 2003-2005, IEEE Computational Intelligence Society in 2008-2010. Dr. Lin was the Distinguished Lecturer of IEEE CAS Society from 2003 to 2005 and CIS Society from 2015-2017. He served as the Chair of IEEE-CIS Distinguished Lecturer Program Committee in 2018. He served as the Deputy Editor-in-Chief of IEEE Transactions on Circuits and Systems-II in 2006-2008. Dr. Lin is the coauthor of Neural Fuzzy Systems (Prentice-Hall) and the author of Neural Fuzzy Control Systems with Structure and Parameter Learning (World Scientific). He has published more than 300 journal papers (total citations: 19,232; h-index: 64; i10-index: 243) in the areas of neural networks, fuzzy systems, brain-computer interface, multimedia information processing, and cognitive neuro-engineering, including more than 120 IEEE journal papers.