IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 23, 2019, accepted May 23, 2019, date of publication May 29, 2019, date of current version July 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919739

A Two-Phase Development of Fuzzy Rule-Based Model and Their Analysis

DAN WANG[1,2,3], WITOLD PEDRYCZ[1,2,3,5], (Fellow, IEEE), AND ZHIWU LI[1,2,4], (Fellow, IEEE)

[1]School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China
[2]Key Laboratory of Electric Equipment Structure Design, Ministry of Education, Xidian University, Xi'an 710071, China
[3]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2R3, Canada
[4]Institute of Systems Engineering, Macau University of Science and Technology, Macau, China
[5]Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Zhiwu Li (zhwli@xidian.edu.cn)

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant JB160401 and Grant JBG160415, in part by the Recruitment Program of Global Experts, in part by the Science and Technology Development Fund, Macau SAR, under Grant 078/2015/A3, and in part by the National Science under Grant 61672400. The work of D. Wang was supported by the China Scholarship Council under Grant 201706960027.

ABSTRACT Fuzzy rule-based models form a commonly encountered category of fuzzy models. As such they have enjoyed a great deal of conceptual and algorithmic developments followed by numerous case studies. This paper contributes to this area by bringing forward a two-phase design of fuzzy rules completed on the basis of experimental data. This design directly reflects upon the nature of the rules vis-à-vis the data used in their construction. First, information granules (fuzzy sets) standing in condition and conclusion parts of the individual rules are formed following a commonly used clustering technique of Fuzzy C-Means (FCM). The results of fuzzy clustering are directly used to build a collection of fuzzy sets of conditions and conclusions forming the individual rules. Some optimization aspects are raised in this context by expressing the performance of the condition and conclusion fuzzy sets in terms of the reconstruction abilities of the data captured by the rules. Second, fuzzy sets present in the rules (which are typically described by membership functions having infinite support) are transformed into interval-valued information granules of finite support that capture the essential (core) relationships between the regions in the input and output spaces strongly supported by the experimental data. In this way, the proposed rule-based model exhibits a two-tier architecture built in two successive phases. Subsequently, the proposed architecture invokes two fundamentally different modes of reasoning: 1) a recall mode in case when a new datum is positioned within the interval-valued information granules and 2) approximation mode where we invoke an aggregation of the individual rules given their activation levels in case a new datum does not belong to the core structure of the rules. These two modes produce granular results (represented as intervals). A way of assessing the quality of the obtained results is provided. Along with these two modes, we offer a characterization of the quality of results as well as the quality of the rules (expressed in terms of coverage, specificity of condition space and specificity of conclusion space). Experimental results are reported to illustrate the design process and the performance of the constructed model.

INDEX TERMS Fuzzy sets, information granules, the principle of justifiable granularity, fuzzy rule-based model.

I. INTRODUCTION

In fuzzy system modeling, rule-based architectures are commonly encountered in the literature and come with a great deal of concepts, algorithms, and applications, cf. [3], [8], [10], [11], [34]. The generic topology of an n-dimensional input-single output structure of a rule-based model (a so-called Mamdani type of model [13], [15]) comprises c rules coming in the following form

$$-\text{if } \boldsymbol{x} \text{ is } A_i, \text{ then } y \text{ is } B_i \tag{1}$$

where A_i and B_i are fuzzy sets [1], [2], [9], [26], [33], [39], [40], [43] defined in the input space and output space, respectively; $x \in R^n$ and $y \in R$. Once the model has been constructed, for any given input $\boldsymbol{x}$, the output is determined by aggregating the levels of activation of the rules $A_1(\boldsymbol{x})$,

The associate editor coordinating the review of this manuscript and approving it for publication was Ananya Sen Gupta.

$A_2(\boldsymbol{x}), \ldots, A_c(\boldsymbol{x})$ along with the fuzzy sets forming the conclusion parts of the rules. In the sequel, the result of aggregation is decoded to form a single numeric output. In this way, the output does not distinguish between a situation when the input matches quite well one of the rules (and in this sense the obtained output is strongly supported by the corresponding fuzzy set located in the output space) and a diametrically different situation where the input is localized in the region where membership grades of A_i are quite low (and hence the confidence associated with the output produced by the fuzzy model becomes quite low as well).

To address this issue, the objective of this study is to develop a *granular* rule-based model [5]–[7], [19], [21], [22], [29] coming in the form of intervals exhibiting a finite support. In this way, we establish a two-tier (layered) model to quantify relevance (confidence) of the results. This makes a clear distinction between the two fundamentally different situations encountered when using the rule-based model:

- *core – data driven model*: when a new datum is contained within information granules [4], [12], [14], [18], [25], [30], [32], [41] built upon a basis of the original fuzzy set of condition. The location of information granule is associated with the region of the input space where there was a high density of data.

- *data – invoked approximation*: a new datum is located outside the regions strongly supported by experimental data. We anticipate that the confidence in the result is naturally weaker than that in the first scenario.

From the design perspective to reflect these two categories of situations, based on the original rules we form the rules with interval-valued information granules of finite support, say $\tilde{A}_i$ and $\tilde{B}_i$, where they are built on a basis of the fuzzy sets A_i and B_i but they capture the essence (core) of these fuzzy sets. The i-th rule reads as follows

$$-\text{if } \boldsymbol{x} \text{ is } \tilde{A}_i \text{ then } y \text{ is } \tilde{B}_i \tag{2}$$

The two categories of situations are invoked with reference to the location of $\boldsymbol{x}$ with respect to the constructed information granules. In contrast to the "standard" fuzzy models which produce numeric outputs, the outcome of the granular fuzzy model is an interval information granule. This implies that the performance of the model is quantified by analyzing the behavior of granular output vis-à-vis numeric data. To address this problem, we consider performance evaluation in terms of indices that are pertinent to information granules, namely a coverage and specificity measures.

In this study, we propose an original two-phase approach to the design and ensuing characterization of fuzzy rule-based models in which by starting with fuzzy sets constructed with the aid of the FCM algorithm [20], [36], [37], the rules are refined in the form given by (2). In contrast to the existing fuzzy models, we delineate between the results that come from the core of the model (being supported by the collection of information granules) and those being results of approximation in case the input data are located outside the regions densely populated by the experimental data and thus quantified by information granules of a usually lower level of specificity.

This study is structured as follows. We start with some prerequisites (section 2) which concerns with reconstruction of fuzzy sets of condition and conclusion. Based on the reconstruction results, we form a collection of interval-valued information granules with the use of the principle of justifiable granularity. In section 3, we concentrate in three items – specificity of condition, specificity of conclusion and a joint coverage, which are considered as a characterization of the quality of the rules. The processing in the granular rule-based model and its evaluation of performance are presented in section 4. In section 5, we present a series of experimental studies in which we present the formation of the fuzzy rule-based model and their analysis. Conclusions parts are provided in section 6.

II. PREREQUISITES

In this section, we elaborate on the formation of fuzzy sets standing in the rules by engaging the clustering mechanism of Fuzzy C-Means (FCM). In parallel, we offer a certain way of evaluating the quality of fuzzy sets with respect to their reconstruction abilities [17], [28], [31]. Some auxiliary optimization aspects are also presented. In the sequel, we discuss a principle of justifiable granularity [23], [24], [44], which offers an ability to construct a single information granule on a basis of available experimental evidence.

A. FORMATION OF FUZZY SETS OF CONDITIONS AND CONCLUSIONS AND THEIR OPTIMIZATION

In what follows, we mainly focus on the construction of fuzzy sets standing in the condition and conclusion parts of the rules. Let us consider pairs of data $(\boldsymbol{x}_k, target_k)$, $k = 1, 2, \ldots, N$, where $\boldsymbol{x}_k \in \boldsymbol{R}^n$, $target_k \in \boldsymbol{R}$. A standard way of building information granules (fuzzy sets of condition and conclusion space) is to invoke fuzzy clustering [42] such as Fuzzy C-Means (FCM), which is one of the most commonly encountered alternatives. The clustering is carried out in the joint input-output space $\boldsymbol{R}^{n+1}$. Denote by t_k — the concatenation of $\boldsymbol{x}_k$ and $target_k$, namely $\boldsymbol{t}_k = [\boldsymbol{x}_k \ target_k]^{\mathrm{T}}$. Clustering these $(n+1)$-dimensional data into c clusters returns a collection of prototypes $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_c$. The distance between the data and the prototypes is expressed as a weighted Euclidean distance [16]

$$||\boldsymbol{t}_k - \boldsymbol{g}_i|| = \sum_{j=1}^{n+1} \frac{(t_{kj} - g_{ij})^2}{\boldsymbol{\sigma}_j^2} \tag{3}$$

where σ_j^2 is the variance of the j-th variable. The FCM method returns a collection of prototypes and a partition matrix. With regard to the prototypes, we distinguish between their part positioned in the input and output spaces; essentially we have them expressed as a concatenation of v_i and w_i, namely $\boldsymbol{g}_i = [\boldsymbol{v}_i \ \ w_i]^{\mathrm{T}}$. For the prototypes positioned in the input space, we have a collection of fuzzy sets forming fuzzy sets of condition

A_i with the following membership functions

$$A_i(\boldsymbol{x}) = \frac{1}{\sum_{l_1=1}^{c_1} \left(\frac{||\boldsymbol{x}-\boldsymbol{v}_i||}{||\boldsymbol{x}-\boldsymbol{v}_{l_1}||}\right)^{2/(m-1)}} \tag{4}$$

where $\boldsymbol{x} \in \boldsymbol{R}^n$. Here c_1 is the number of clusters of the condition space.

For conclusion space, we can obtain fuzzy sets B_i with the corresponding membership functions

$$B_i(target) = \frac{1}{\sum_{l_2=1}^{c_2} \left(\frac{||target-w_i||}{||target-w_{l_2}||}\right)^{2/(m-1)}} \tag{5}$$

where $target \in \boldsymbol{R}$, c_2 is the number of clusters of the conclusion space. In this study, we assume $c_1 = c_2 = c$; m stands for a fuzzification coefficient [27], [35]; this coefficient is usually set equal to 2.0.

It is imperative to assess the quality of fuzzy sets constructed through fuzzy clustering. A suitable performance index comes in the form of a reconstruction index. Using this index, the quality of fuzzy sets is assessed by quantifying their reconstruction abilities. For any given input $\boldsymbol{x}_k$, the reconstruction (degranulation) of this input, we consider the prototypes and the membership grades forming the following weighted sum

$$\hat{\boldsymbol{x}_k} = \frac{\sum_{i=1}^{c_1} A_i(\boldsymbol{x}_k)^m \boldsymbol{v}_i}{\sum_{i=1}^{c_1} A_i(\boldsymbol{x}_k)^m} \tag{6}$$

The reconstruction criterion is then expressed by the sum of error being taken over all the data (N-dimension)

$$\mathrm{E} = \frac{1}{N}\sum_{k=1}^{N} \|\boldsymbol{x}_k - \hat{\boldsymbol{x}_k})\|^2 \tag{7}$$

where, as before, the *weighted Euclidean distance* is involved. Intuitively, the lower the values of the sum error E, the better the reconstruction abilities are delivered by fuzzy sets $A_1, A_2, \ldots, A_c$.

In a similar way, given an output $target_k$, we determine its reconstruction in the following form

$$\hat{target}_k = \frac{B_i(target_k)^m w_i}{\sum_{i=1}^{c_2} (B_i(target_k))^m} \tag{8}$$

In this case, the reconstruction criterion reads as

$$\mathrm{F} = (target_k - \hat{target}_k)^2 \tag{9}$$

A certain stopping criterion ε is proposed by analyzing the decrease of the values of the reconstruction criterion with respect to the increasing values of the number of clusters c_1 and c_2, while selecting such value of this parameter where there is no substantial drop in the values of E and F. In other words, by repeating the reconstruction process, the error values E and F decrease with increasing the number of clusters c_1 and c_2, the repeating process will not stop until it meets a stopping condition, for example, $E = \varepsilon$ and $\varepsilon < 0.05$.

It is worth noting that the FCM algorithm is realized in the $\boldsymbol{R}^{n+1}$ space whereas input and output spaces in the algorithm are not distinguished. Therefore, a modified version of the distance by introducing an auxiliary weight λ to tell apart the two spaces in FCM algorithm. More specifically, we have

$$||\boldsymbol{t}_k - \boldsymbol{g}_i|| = \sum_{j=1}^{n} \frac{(x_{kj} - v_{ij})^2}{\sigma_j^2} + \lambda \frac{(target_k - w_i)^2}{\sigma_{n+1}^2} \tag{10}$$

where $\lambda > 0$. The parameter λ can be optimized by running FCM algorithm for its values and choosing λ_{opt} for which the performance indexes E and F attain their minimal values.

To compensate for the difference between the dimensionality of the input (n-dimension) and output (1-dimension) space, we vary the values of λ in the range $(0, 2n)$. If $\lambda \in (0, n)$, the input space has a stronger impact than the output space while if $\lambda \in (n, 2n)$; the output space has more impact.

B. THE PRINCIPLE OF JUSTIFIABLE GRANULARITY

The principle of justifiable granularity [23], [24] provides a general idea on how to construct a single information granule by achieving a sound compromise between two essential requirements characterizing it, namely coverage and specificity [38]. The level of coverage is related to the number of elements which are embraced (covered) by the information granule, whereas specificity is about how detailed the information granule is in standing for the experimental data and it associates with the semantics of the granule. In light of the essence of these criteria, it is intuitively apparent that the two requirements are in conflict. Thus, the increase in coverage associates with the decrease in specificity.

As an illustrative experiment, we consider a simple example for determining the interval information granules with the use of the principle of justifiable granularity. Given a one-dimensional normal distribution data set $D = \{y_k\}$, $y_k \sim N(0,5^2)$, $k = 1, 2, \ldots, N$. $N = 500$ The histogram of the data set presents in Figure 1.

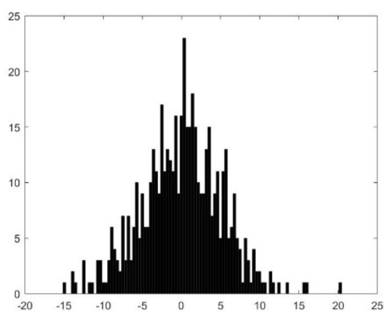


FIGURE 1. Histogram of data set *D*.

Let us start with determining a numeric representative $mean(D)$ — viz. mean value of the data set (in this case study, $mean(D) = 0$). Once the mean value has been determined, the interval information granule $\sum = [a, b]$ is constructed around $mean(D)$. The upper bound b and lower bound a are constructed in a similar way, here we are interested in the formation of upper bound b based on the experimental data which is located on the right hand side of $mean(D)$, coverage and specificity are expressed as follows,

Coverage is determined by the cardinality which concerns about the number of experimental evidence included in the region between the mean value $mean(D)$ and upper bound b.

$$cov = \frac{1}{N} card \left\{ y_k | y_k \in [mean\,(D)\,, b] \right\} \tag{11}$$

It should be noted that the expressions cov and sp are not unique and they depends on the practical problems which is anticipated to be solved.

Specificity reflects the semantics (meaning) of the interval, where a smaller interval information granule implies a higher level of specificity. We introduce parameter $\alpha(\alpha > 0)$ to the expression of specificity which quantifies an impact of specificity in the formation of information granules.

$$sp = e^{-\alpha \frac{|b-mean(D)|}{|y_{\max}-mean(D)|}} \tag{12}$$

The optimization process is realized by maximizing a product of cov and sp, namely,

$$v\,(b) = cov^{*}sp \tag{13}$$

Finally, the optimal upper bound "b_{opt}" is calculated,

$$b_{opt} = \operatorname{argmax}_{y_k \in [mean(D),b]} V(b) \tag{14}$$

In Figure 2, we present the plots of cov, sp and $V(b)$ with b for different parameters $\alpha(= 0.1, 0.3, 0.5$ and $1.0)$.

cov, sp and their product $V(b)$ as function of b are presented in Figure 2. The numeric represents $mean(D) = 0$, considering the experimental evidence positioned in the area of $[mean(D), b]$, the cov increase while sp decrease with an increasing b, which forms a unimodal results of $V(b)$. Therefore, the upper bound of the information granule is determined with the optimal b which is in correspondence to the maximal value of $V(b)$.

To proceed with the construction of the information granules in the context of the problem discussed, two approaches are considered: (i) the standard one encountered in the existing literature and (ii) an augmented version where in the construction of the granules in the input space, we consider an impact implied by the output variable.

(i) generic version of the principle of justifiable granularity

For the i-th cluster, and the j-th variable, $j = 1, 2, \ldots, n$, the corresponding prototype is v_{ij} which is regarded as the modal value of the information granules to be constructed. Considering the data positioned to the right from this modal value, the upper bound b_{ij} can be determined. The coverage is thus expressed as,

$$cov_{ij} = \sum_{x_{ki} \in [V_{ij}, b_{ij}]} \max\{0, [A_i\left(x_{kj}\right) - \max_{\substack{l = 1, 2, \ldots, c \\ l \neq i}} A_i(x_{kj})]\} \tag{15}$$

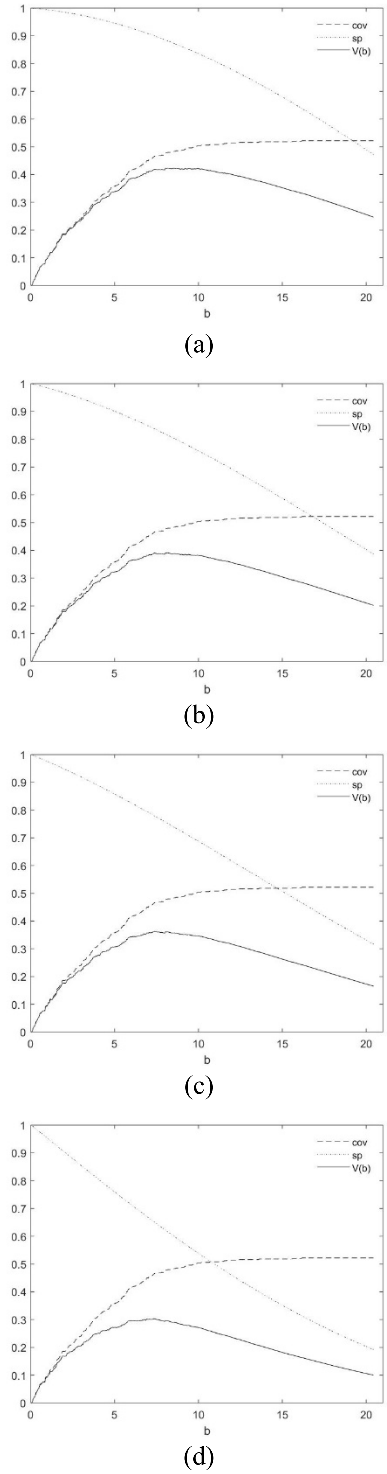


FIGURE 2. Plots of cov, sp and $V(b)$ treated as functions of b: (a) $\alpha = 0.1$. (b) $\alpha = 0.3$. (c) $\alpha = 0.5$. (d) $\alpha = 1.0$.

where $A_i\left(x_{kj}\right) = \frac{1}{\sum_{l_1=1}^{c_1} \left(\frac{(x_{kj}-V_{ij})}{(x_{kj}-V_{l_1 j})}\right)^{2/(m-1)}}$ is the membership grades of the input space associated with the i-th cluster.

Note that we incorporate inhibitory information about other information granules $(A_{l_1}, l_1 = 1, 2, \ldots, c_1, l_1 \neq i)$

so that we penalize the data belonging to other clusters. The specificity is expressed in the following form

$$sp_{ij} = 1 - \frac{|b_{ij} - v_{ij}|}{range_{ij}} \tag{16}$$

where $range_{ij} = |x_{j,max} - v_{ij}|$ and $x_{j,max}$ is the maximal value of the j-th input variable.

The optimized performance index is calculated as the product of these two measures

$$V\left(b_{ij}\right) = cov_{ij}{}^{*}sp_{ij} \tag{17}$$

The optimal value of b_{ij} is found by maximizing $V(b_{ij})$, namely $b_{ij} = arg\max_{b_{ij}} V(b_{ij})$.

Similarly, we can determine an optimal lower bound.

Obviously, there are some differences in the detailed computing related with the data considered in the construction. Now the coverage is taken as the following sum

$$cov_{ij} = \sum_{k:x_{kj}\in[a_{ij},v_{ij}]} A_i(x_{kj}) \tag{18}$$

whereas the specificity is expressed in the form

$$sp_{ij} = 1 - \frac{|v_{ij} - a_{ij}|}{range_{ij}} \tag{19}$$

$range_{ij} = |v_{ij} - x_{j,\min}|$, and $x_{j,\min}$ is the smallest element in the data set for the j-th variable. The optimal a_{ij} comes as a solution to the maximization of the product of coverage and specificity, namely $a_{ij} = arg\max_{a_{ij}} V(a_{ij})$

We repeat the above process of information granulation for each input variable and determine the Cartesian product of information granules obtained in this way

$$\tilde{A}_i = [a_{i1}, b_{i1}] \times [a_{i2}, b_{i2}] \times \ldots \times [a_{in}, b_{in}] \tag{20}$$

(ii) augmented, context-implied principle of justifiable granularity

To consider an impact coming from the output $target_k$, the construction process can be realized similarly stated in (i), only to modify the coverage criteria by including the membership grades of $target_k$.

To determine the upper bounds b_{ij} of the interval, we have coverage expressed in the form

$$cov_{ij} = \sum_{x_{kj}\in[v_{ij},b_{ij}]} \max\{0, A_i\left(x_{kj}\right) - \max_{\substack{l=1,2,\ldots,c\\ l\neq i}} A_l(x_{kj})\} \tag{21}$$

Note that in contrast to (10), we incorporate here membership values of the associated information granules positioned in the output space.

We can follow the same process to determine the upper bound b_{ij} and lower bound a_{ij} as discussed in (i).

In sequel, we talked about two approaches of construction of information granules of input space in (i) and (ii). The determination of the information granules in the output space $\tilde{B}_i = [c_i, d_i]$ is completed following the principle of justifiable granularity as discussed previously only based the data available in the output space.

III. CHARACTERIZATION OF THE QUALITY OF THE RULES

Having the rules composed of the pairs $\tilde{A}_i$ and $\tilde{B}_i$, the quality of the rule is described by investigating the two aspects:

(i) Data coverage provided by the rule. It is expressed by counting the number of data included in the Cartesian product of the intervals. The higher the coverage, the higher the quality of the rule.

$$cov\left(\tilde{A}_i, \tilde{B}_i\right) = \frac{1}{N} card\{(\boldsymbol{x}_k, target_k) | \boldsymbol{x}_k \in \tilde{A}_i, target_k \in \tilde{B}_i\} \tag{22}$$

(ii) Specificity of condition rules $\tilde{A}_i$ versus specificity of conclusion rules $\tilde{B}_i$. Lower the values of specificity of $\tilde{A}_i$ and higher the values of specificity of $\tilde{B}_i$, reflect higher quality of the rules.

There is a sound argument behind this: the rule is applied to many situations (expressed in conditions) and at the same time produces conclusions that are very detailed. The detailed computing of the specificity is outlined in terms of the following expressions.

Specificity of the conclusion part

$$sp_i = 1 - \frac{|d_i - c_i|}{range_{target}} \tag{23}$$

where the bounds c_i and d_i of the interval are determined using the principle of justifiable granularity explained in section 2.2, $\tilde{B}_i = [c_i, d_i]$, while $range_{target}$ is the range of values assumed by the output variable, namely $range_{target} = |target_{max} - target_{min}|$.

Specificity of condition part is the following average

$$sp_i = \frac{1}{n}\sum_{j=1}^{n} sp_{ij} \tag{24}$$

and

$$sp_{ij} = 1 - \frac{|b_{ij} - a_{ij}|}{range_x} \tag{25}$$

where, similarly, the lower and upper bounds of the intervals are constructed by the principle of justifiable granularity, $\tilde{A}_i = [\boldsymbol{a}_i, \boldsymbol{b}_i]$, with $range_x = |x_{\max,j} - x_{\min,j}|$.

The two aspects can be positioned in three-dimensional space where the coordinates are specificity of condition and specificity of conclusion parts as well as the joint coverage. Evidently the quality of rules becomes higher if the specificity of condition part decreases, the specificity of conclusion part increases and the joint coverage increases.

IV. PROCESSING IN GRANULAR RULE-BASED MODEL

A. FORMATION OF GRANULAR RULE-BASED MODEL

Having constructed interval information granules $\tilde{A}_i$ and $\tilde{B}_i$, $i = 1, 2, \ldots, c$ along with the ensuing collection of rules, we look at the mapping of any input $\boldsymbol{x}_k$ through such rules. Two cases are considered depending upon a location of x vis-à-vis the A_i.

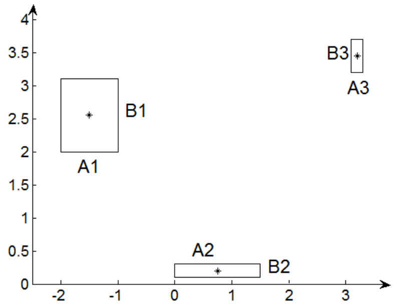


FIGURE 3. Plot of the prototypes and information granules.

(i) $\boldsymbol{x}_k$ is located in $\tilde{A}_i$.

As we are concerned with $\boldsymbol{x}_k$ that is included in the input information granules $\tilde{A}_i$, the output is $\tilde{B}_i$. If there is any overlaps between the two different information granules $\tilde{A}_i$ and $\tilde{A}_l$, $l = 1, 2, \ldots, c, l \neq i$, when $\boldsymbol{x}_k$ is included in either of the information granules, the output can be determined in the form of a union of $\tilde{B}_i$ and $\tilde{B}_l$, namely $\tilde{B}_i \cup \tilde{B}_l$.

(ii) $\boldsymbol{x}_k$ is located outside the core regions of $\tilde{A}_i$, namely $\boldsymbol{x}_k$ union of $\tilde{A}_i$.

In this case, the output is defined by determining a level of matching $\boldsymbol{x}_k$ vis-à-vis individual $\tilde{A}_i$. Here we follow the same formula as commonly encountered in rule-based models by bringing together the levels of activation of the rules and engaging the corresponding interval $\tilde{B}_i$. More specifically, we have

$$\tilde{Y}_k = A_1(\boldsymbol{x}_k) \otimes \tilde{B}_1 \oplus A_2(\boldsymbol{x}_k) \otimes \tilde{B}_2 \oplus \ldots \oplus A_c(\boldsymbol{x}_k) \otimes \tilde{B}_c \tag{26}$$

where the symbols in small circles emphasize that the operations concern intervals. The detailed formulas deal with the multiplication of interval by a non-negative constant and the addition of intervals, namely $\delta \otimes [a, b] = [\delta a, \delta b]$ and $[a, b] \oplus [c, d] = [a + c, b + d]$.

To illustrate the behavior in determining granular output, we assume three information granules defined in one-dimensional input space in form of intervals, $\tilde{A}_1 = [-2, -1]$, $\tilde{A}_2 = [0, 1.5]$ and $\tilde{A}_3 = [3.1, 3.3]$ while the corresponding output intervals $\tilde{B}_1$, $\tilde{B}_2$ and $\tilde{B}_3$ are given as [2.0, 3.1], [0.1, 0.3], and [3.2, 3.7]. The information granules are defined as Cartesian products $\tilde{A}_i \times \tilde{B}_i$, and we determine the prototypes by the center point of the information granules, refer to Figure 3.

For any numeric values x located in the input space, we consider the following two cases: (i) x is located in $\tilde{A}_i$ and (ii) x is not included in the core region of $\tilde{A}_i$. The granular output $Y = [y^-, y^+]$ are determined following the formation process explained above. The results are displayed for selected values $m(m = 1.1, 2.0$ and $3.5)$ as Figure 4.

In Figure 4, we present the upper and lower bound of the granular output Y. As expected, when the input data x is located in the region of $\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$, the upper bound y^+ and lower bound y^- is the upper and lower bounds of $\tilde{B}_i$, $i = 1, 2, 3$. If x is located outside the space of $\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$, the determination of y^+ and y^- is realized following the formation process. In the second case, the granular output behaves differently based on the selected values of m. Consider input space located at the left of $\tilde{A}_1$and right of $\tilde{A}_3$, the bound of the output behaves smoothly when $m = 1.1$, and when $m = 2.0$ and 3.5, the bound of output exhibits more changes. However, when $m = 1.1$, the bounds of granular output have a great reduction and rise to connect the input data outside$\tilde{A}_1$, $\tilde{A}_2$ and $\tilde{A}_3$, while when m is increased to 2.0 and 3.5, the difference between the largest and smallest values of the output become smaller at the junction point.

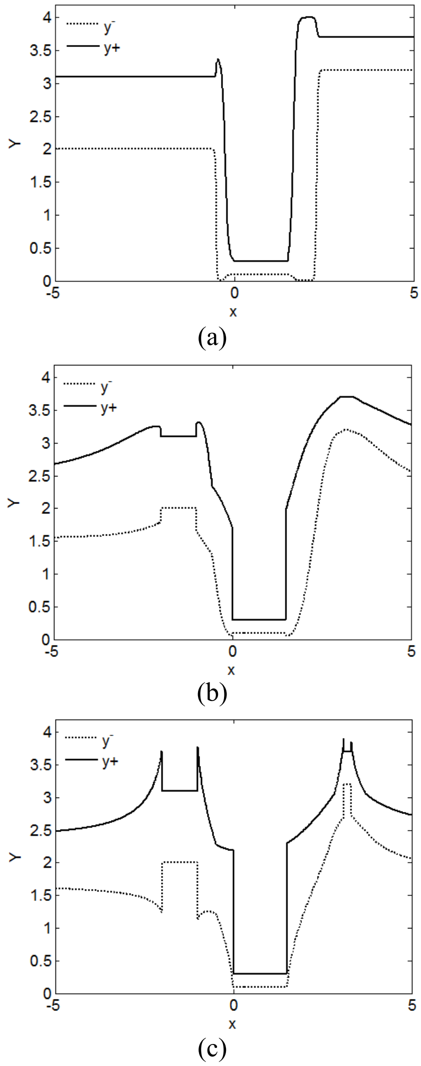


FIGURE 4. Granular output Y with selected m: (a) $m = 1.1$, (b) $m = 2.0$, (c) $m = 3.5$.

B. EVALUATION OF PERFORMANCE OF GRANULAR RULE-BASED MODEL

Taking into account the non-numeric results produced by the granular rule based model, its performance is quantified in terms of the coverage delivered by $\tilde{Y}_k$ and their specificity. Both of them need to be maximized. In light of their conflicting character, the following product comes as a sound performance measure

$$V = [\frac{1}{N}\sum_{k=1}^{N} cov(target_k, \tilde{Y}_k)][\frac{1}{N}\sum_{k=1}^{N} spec(\tilde{Y}_k)] \quad (27)$$

where N is the number of data. In more detail, coverage and specificity are expressed in the following way,

$$cov\left(target_k, \tilde{Y}_k\right) = \begin{cases} 1, & if\, target_k \in [y_k^-, y_k^+] \\ 0, & otherwise \end{cases} \quad (28)$$

$$sp\left(\tilde{Y}_k\right) = 1 - \frac{|y_k^+ - y_k^-|}{range_{target_k}} \quad (29)$$

where $range_{target_k} = |target_{max} - target_{min}|$.

Along with the global index V shown above, it is advantageous to visualize the results for the core and the interpolation part.

-for the core data (the data contained in one of the collection of information granules $\tilde{A}_i$)

$$V_1 = [\frac{1}{N_1}\sum_{k=1}^{N_1} cov(target_k, \tilde{Y}_k)][\frac{1}{N_1}\sum_{k=1}^{N_1} sp(\tilde{Y}_k)] \quad (30)$$

where N_1 is the number of data falling under the core regions.

-for data for which an interpolation mechanism has been invoked (the data not contained in one of the $\tilde{A}_i$)

$$V_2 = [\frac{1}{N_2}\sum_{k=1}^{N_2} cov(target_k, \tilde{Y}_k)][\frac{1}{N_2}\sum_{k=1}^{N_2} sp(\tilde{Y}_k)] \quad (31)$$

where N_2 is the number of data not contained in the core regions.

V. EXPERIMENTS

In this section, we develop a series of experimental studies about how the fuzzy rule-based model is formed with the use of information granules. Synthetic data as well as a collection of UCI data sets are applied to form the fuzzy rules in the input space and output space.

A. SYNTHETIC EXPERIMENTS

Let us consider a one-dimensional synthetic data set, the experimental data coming in pairs $(\boldsymbol{x}_k, target_k)$, $k = 1, 2, \ldots, N$, $N = 1000$. The input space is generated by four groups of Gaussian distribution $N_1\left(0, 0.5^2\right), N_2\left(-5, 0.3^2\right), N_3\left(5, 0.3^2\right), N_4(10, 1)$. The function of $target_k$ with $\boldsymbol{x}_k$ is $target_k = 1 + \sin(x_k) + x_k^{0.5}$. The fuzziness coefficient is $m = 2.0$. Figure 5 shows plotting of the experimental data,

By using FCM algorithm, the experimental data is clustered into c clusters. The number of clusters c is selected as $c = 3, 4, 5, 6, 7$ separately, the prototypes and information granules are displayed in Figure 6.

As shown in Figure 6, stars "*" represent the prototypes produced by the FCM clustering while the rectangles "□" denote for information granules constructed based on the experimental data.

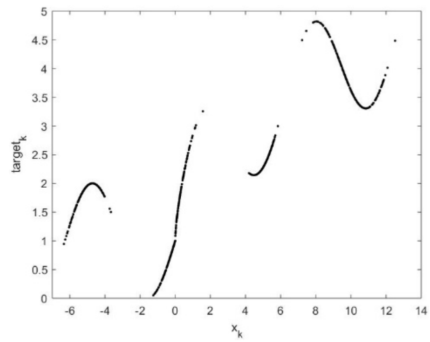


FIGURE 5. Plot of experimental data (x_k, $target_k$).

In Figure 7, the plot of granular output $\tilde{Y}_k$ (with lower bound y_k^- and upper bound y_k^+) are presented.

In Figure 7, we compare experimental output data $target_k$ with the granular output of the fuzzy rule-based model. It indicates that the granular output is generally tend to the curve of $target_k$ with x_k and the main parts of $target_k$ is included in between upper bound and lower bound of the granular output. A good performance of the proposed fuzzy rule-based model can be obtained while as many as possible experimental data are covered by the granular output. In this experiment, the data is clustered to 3-7 clusters, it is concluded that a better performance can be obtained with a larger number of clusters.

B. ANALYSIS ABOUT THE FORMATION OF FUZZY RULE-BASED MODEL

Here we consider the reconstruction of the fuzzy sets of condition and conclusion parts based on the data coming from the UCI repository (http://archive.ics.uci.edu/ml/). For each data set, the data is coming in pairs $(\boldsymbol{x}_k, target_k)$, $k = 1, 2, 3, \ldots, N$, $\boldsymbol{x}_k$ is the first n-th variables while $target_k$ is the $(n+1)$-th variable of the data set, which separately represent the input and output space. The FCM algorithm is applied to transfer numeric data into fuzzy sets and the number of clusters c varies from 2 to 10.

Considering the number of clusters c as well as the fuzzification parameter, both influencing on the reconstruction process. The reconstruction criterion of condition space E and conclusion space F is shown as follows.

Boston housing data set

In Figures 8 and 9, we present the values of the reconstruction criterion – E for the condition space and F for the conclusion space, with regards to different numbers of clusters. We can find that if we get more clusters, the reconstruction criterion decrease for both space. When the number of cluster increases from $c = 2$ to $c = 10$, the reconstruction criterion decrease from 0.7 to 0.5 for condition space while for conclusion the values go from 0.65 to 0.45, both make an improvement of the reconstruction by about 30%.

In Figure 9, the number of clusters is selected as $c = 2, 5$ and 10, we consider the parameter λ changing from 0 to 25,

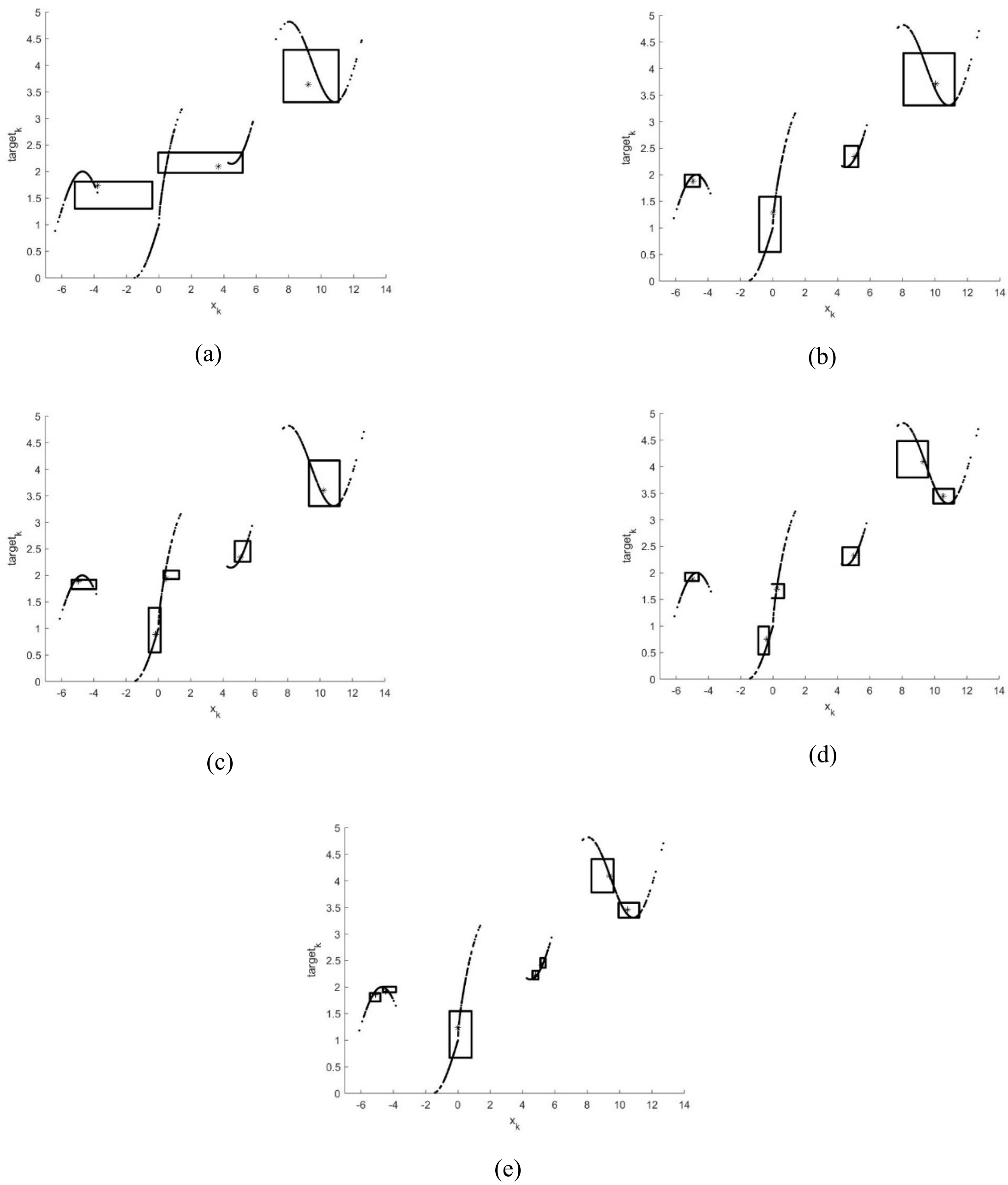


FIGURE 6. **Prototypes and information granules: (a) $c = 3$. (b) $c = 4$. (c) $c = 5$. (d) $c = 6$. (e) $c = 7$.**

which provide from a very weak impact to an extremely strong impact of the condition space in the reconstruction process. Compare the figures in Figure 9, we can find that the reconstruction criterion for both space decrease with an increase value of λ. This implies that for both the condition and conclusion space, greater value of λ can result in better reconstruction results for Boston Housing data.

Forest Fires data set

For the Forest Fires data, the reconstruction criterion of both condition and conclusion space keeps decreasing when the number of clusters increase from 2 to 10, refer to Figure 10. The improvement is about 4.2% and 0.7% for the condition and conclusion space, respectively.

In Figure 11, we observe that when the number of cluster is selected as $c = 2, 5$ and10, the reconstruction criterion E of condition space increase with an increasing value of λ, while in the conclusion space, the reconstruction criterion F decrease with an increasing value of λ. This difference implies that in reconstruction, if we increase the influence of the output part, we can make an positive impact to the reconstruction of the rules of the ouput space but an negative impact to the input space.

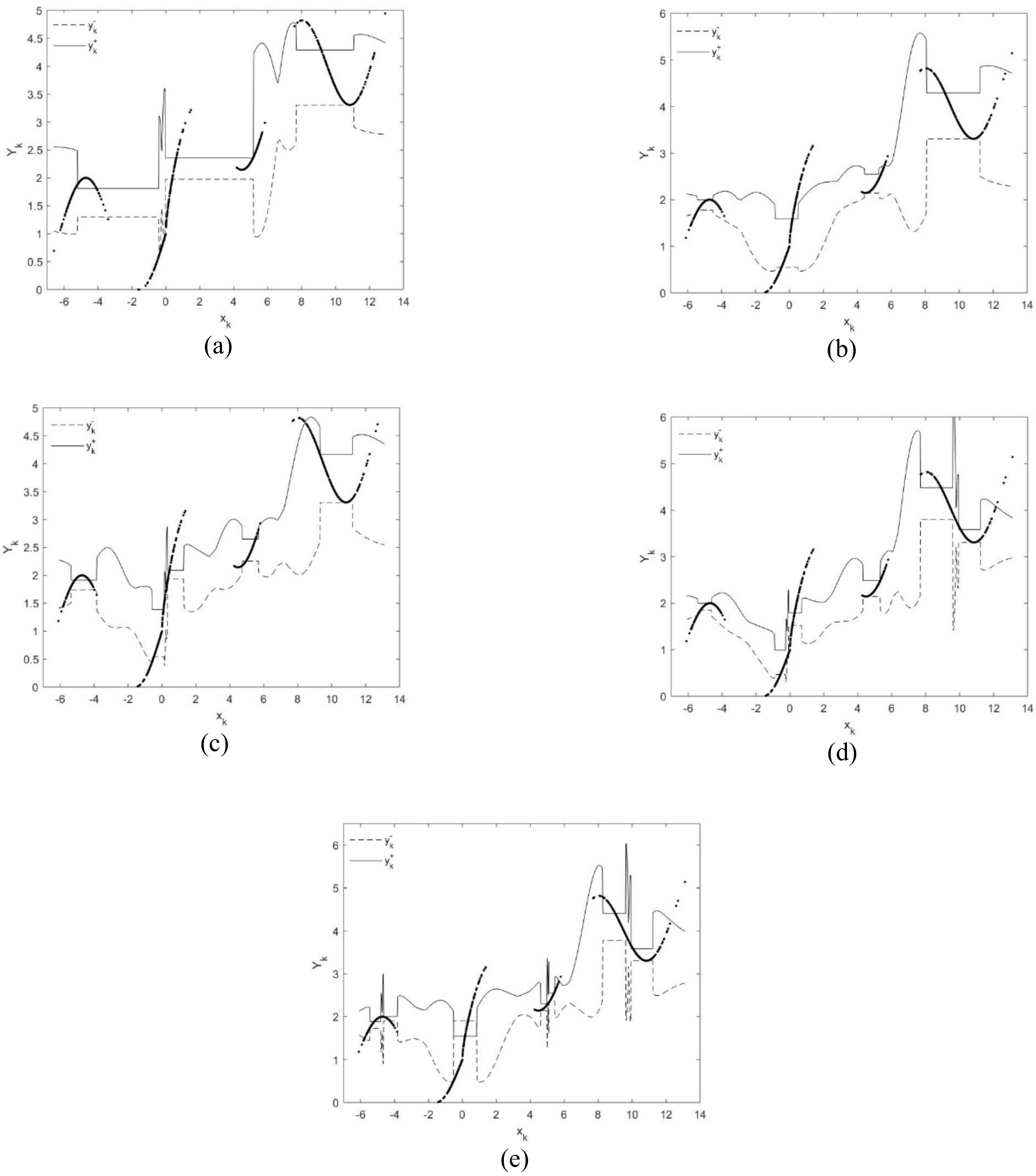


FIGURE 7. Plot of y_k^- and y_k^+ with selected number of clusters: (a) $c = 3$. (b) $c = 4$. (c) $c = 5$. (d) $c = 6$. (e) $c = 7$.

Auto MPG data set

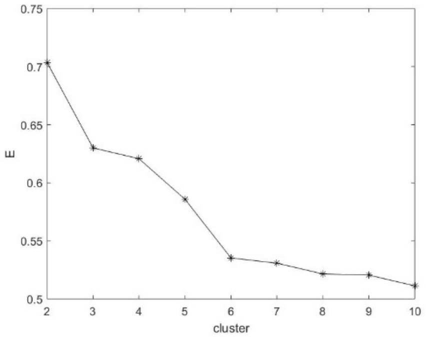
The reconstruction criterion for Auto MPG data decreases smoothly with an increasing number of clusters c from 2 to 10. The improvement of reconstruction with different number of clusters is 52.2% of condition parts and 79.1% of conclusion parts.

As for Auto MPG data, the reconstruction criterion of input space increases with an increasing values of λ while the reconstruction criterion of output space decreases with an increasing values of λ when $c = 2$, 5 and 10.

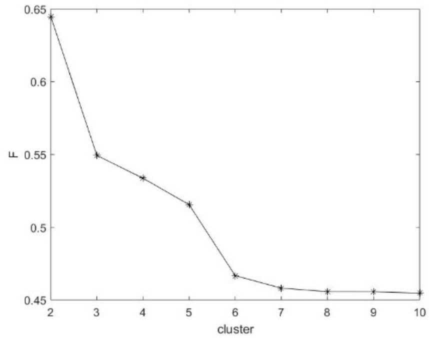
Computer hardware data set

The plots in Figures 14 and 15 shows the same tendency of reconstruction criterion with different number of clusters. The higher the number of cluster, the smaller the reconstruction criterion for both condition and conclusion space, and the improvement is 52.7% and 23.1% for the two spaces separately. In Figure 15, we can find that for some selected number of cluster $c = 2$, 5 and 10, the reconstruction criterion decreases with an increasing value of λ for the condition and conclusion space.

Considering the plots presented in Figures 9 – 15, it is concluded that the reconstruction criterion for both the condition space and conclusion space goes decreasing with an increasing number of clusters. It implies that an increase of clusters can make a better results of the reconstruction of fuzzy sets. However, if we are concerned about the impact

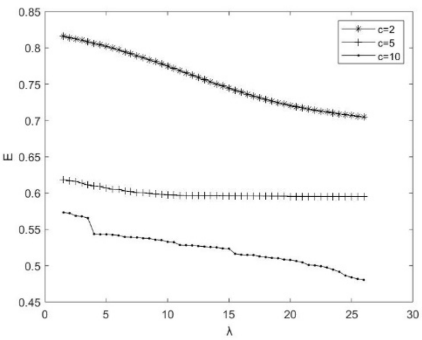


(a)

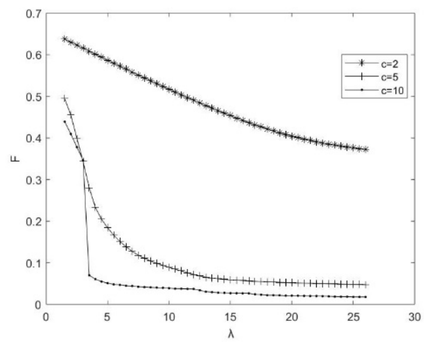


(b)

FIGURE 8. Reconstruction criterion *E* and *F* versus different number of clusters *c* : 2-10 (a) condition space. (b) conclusion space.



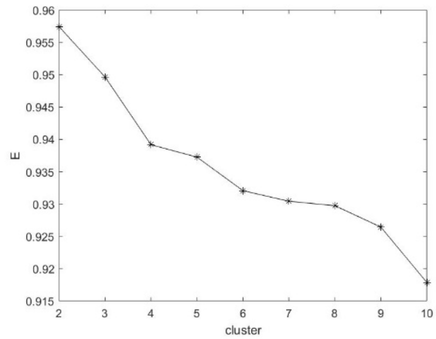
(a)



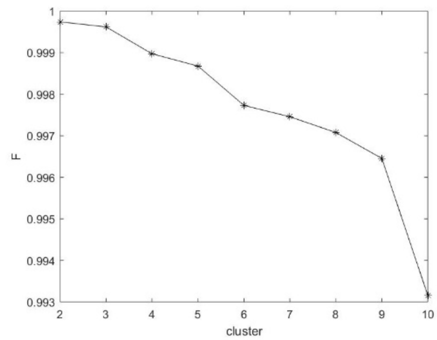
(b)

FIGURE 9. Reconstruction criterion *E* and *F* versus λ. (a) condition space. (b) conclusion space.

of condition and conclusion space at the same time, in other words, if we increase the value λ, the results depend on the data set we used. For example, if we use Boston Housing and Computer Hardware data, the values of *E* and *F* decrease if

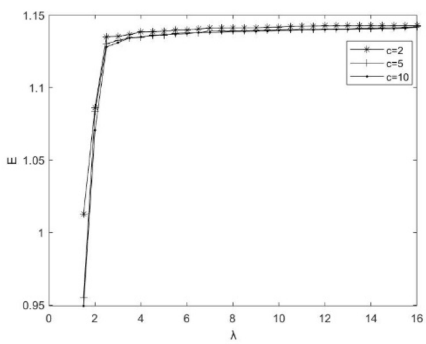


(a)

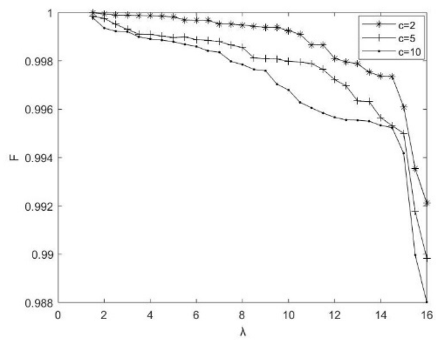


(b)

FIGURE 10. Reconstruction criterion *E* and *F* versus different number of clusters *c* : 2-10 (a) condition space. (b) conclusion space.



(a)



(b)

FIGURE 11. Reconstruction criterion *E* and *F* versus λ. (a) condition space. (b) conclusion space.

λ increases, and if we use Forest Fires data, *E* increase while *F* decrease with an increasing λ. However, in case of Auto MPG data, there is contract situation as that of Forest Fires.

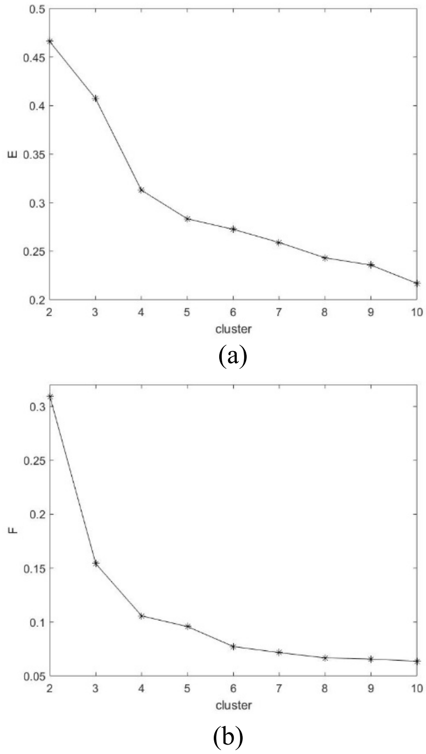


FIGURE 12. Reconstruction criterion *E* and *F* versus different number of clusters *c* : 2-10 (a) condition space. (b) conclusion space.

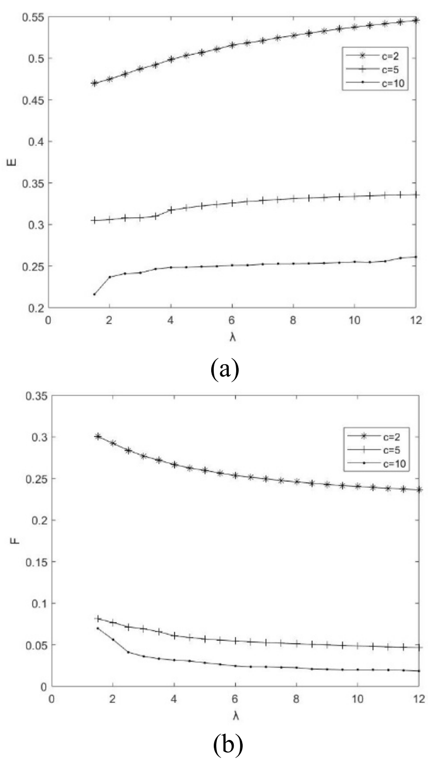


FIGURE 13. Reconstruction criterion *E* and *F* versus λ. (a) condition space. (b) conclusion space.

In this part, we elaborate in detail on the results obtained for Auto MPG data. Fuzzy rules are formed with the use of the FCM algorithm; the fuzzification coefficient is $m = 2.0$.

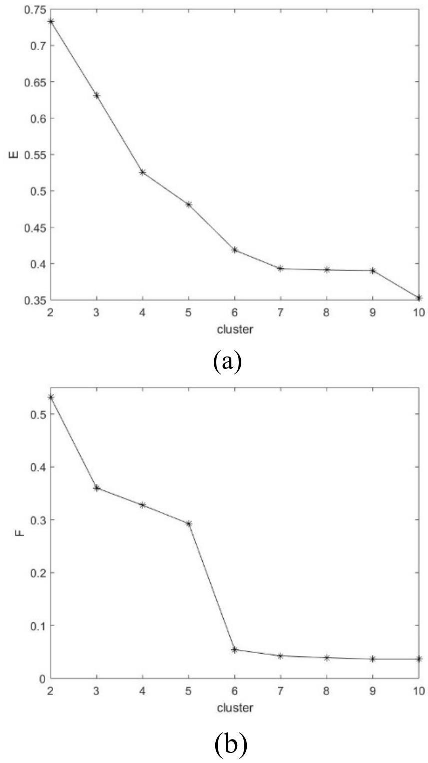


FIGURE 14. Reconstruction criterion *E* and *F* versus different number of clusters *c* : 2-10 (a) condition space. (b) conclusion space.

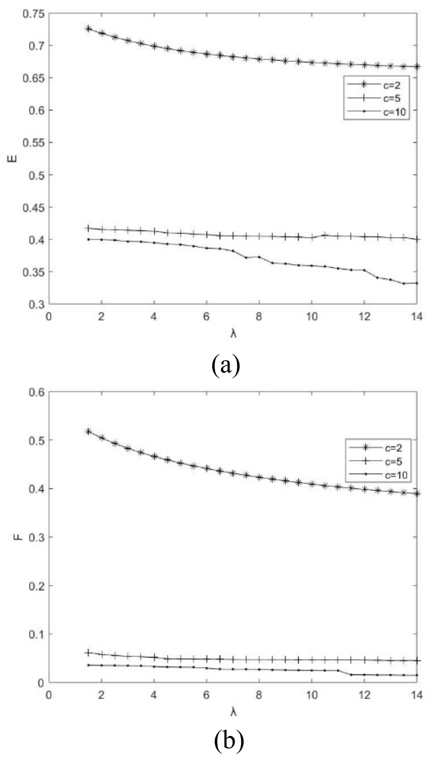


FIGURE 15. Reconstruction criterion *E* and *F* versus λ. (a) condition space (b) conclusion space.

The results are produced for some pairs of condition rules $\tilde{A}_i$ $(i = c_1 = 4)$ and conclusion rules $\tilde{B}_i$ $(i = c_2 = 4)$. We are interested in the three criterion — specificity of condition space, specificity of conclusion space and joint coverage.

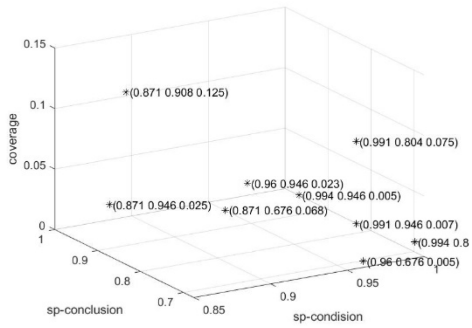


FIGURE 16. Quality of rules plot in 3-dimensional space (*sp*-condition, *sp* -conclusion, and coverage).

The quality of rules is high with regard to high specificity of condition space and joint coverage as well as low specificity of the conclusion space.

The tabular results, shown in Table 1, brings about a summary based on $\tilde{A}_i$ and $\tilde{B}_i$.

TABLE 1. Summary of quality versus rules ($\tilde{A}_i, \tilde{B}_i$).

$\text{cov}(\widetilde{A_I}, \widetilde{B_i})$	$\widetilde{B_1}$	$\widetilde{B_2}$	$\widetilde{B_3}$	$\widetilde{B_4}$
$\widetilde{A_1}$	0.067688	0.000000	0.02513	*0.125126*
$\widetilde{A_2}$	*0.0135025*	0.000000	0.005377	0.000000
$\widetilde{A_3}$	0.000000	*0.075025*	0.0067839	0.000000
$\widetilde{A_4}$	0.00502	0.000000	*0.022613*	0.000000

In Table 1, we present joint coverage based on the fuzzy rules of condition space and conclusion space with the number of rules in both space are four. It is indicated that for each $\tilde{A}_i$, there is different corresponding $\tilde{B}_i$ to attain maximal value of joint coverage. Maximal joint coverage achieve in the following pairs: (A_1, B_4), (A_2, B_1), (A_3, B_2) and (A_4, B_3). A greater joint coverage implies more experimental data can be covered in the information granules which is Cartesian product of condition and conclusion spaces. However, if the value of $cov(\tilde{A}_i, \tilde{B}_i)$ is equal to zero, it means no data is covered by the information granules. In this part, it is concluded that the following information granules — (A_1, B_4), (A_2, B_1), (A_3, B_2) and (A_4, B_3) — play as Cartesian product, perform as the ones to obtain good quality of the rules.

For each of these combinations we report the quality of the obtained results (specificity of condition space, specificity of conclusion space and joint coverage produced by the rules), as shown in Figure 9.

In this study, we use a series of publicly available data coming from the Machine Learning repository (http://archive.ics.uci.edu/ml/), see Table 2.

The data sets are split into the training and testing subset (70%-30% split). FCM algorithm is running for several number of clusters c: 3 — 7, $m = 2.0$. The performance index of the fuzzy rule-based model for both training and testing data are reported in tabular form.

TABLE 2. Characteristic of the data: N — number of data, $(n + 1)$ — dimension of data.

Name	N	n
Housing	506	13
Auto MPG	398	6
Computer Hardware	209	7
Forest Fires	517	8
Concrete Compressive Strength	1030	8
Yacht Hydrodynamics	308	6
Airfoil Self-Noise	1503	5
Wine Quality - red	1599	11

TABLE 3. Performance index of fuzzy rule-based model with traning and testing data.

Data sets		$c=3$	$c=4$	$c=5$	$c=6$	$c=7$
Boston Housing	training	0.112	0.138	0.217	0.228	0.249
	testing	0.097	0.136	0.195	0.203	0.248
Auto MPG	training	0.149	0.174	0.221	0.223	0.280
	testing	0.118	0.141	0.143	0.144	0.255
Computer Hardware	training	0.544	0.552	0.563	0.584	0.606
	testing	0.500	0.533	0.552	0.558	0.559
Forest Fires	training	0.319	0.336	0.342	0.350	0.352
	testing	0.224	0.267	0.277	0.320	0.327
Concrete Compressive Strength	training	0.131	0.197	0.202	0.250	0.263
	testing	0.117	0.182	0.192	0.242	0.256
Yacht Hydrodynamics	training	0.290	0.303	0.314	0.335	0.342
	testing	0.175	0.271	0.275	0.283	0.298
Airfoil Self-Noise	training	0.005	0.010	0.012	0.018	0.024
	testing	0.003	0.004	0.009	0.012	0.020
Wine Quality - red	training	0.033	0.071	0.080	0.083	0.102
	testing	0.023	0.063	0.068	0.074	0.089

In Table 3, we present the performance values of the fuzzy rule-based model based on a collection of data sets (training and testing subset). For each data set, it implies a better performance while the number of cluster c increase from 3 to 7. Considering Computer Hardware data, for both training and testing sets, the performance index can be achieved to around 0.5-0.6 when the number of cluster is $c = 7$, which indicates a good performance of the proposed model. The model also performs well based on Forest Fires data, Yacht Hydrodynamics data, Concrete Compressive Strength data, Auto MPG data and Boston Housing data, all of them achieves a performance values to around 0.25 or higher.

VI. CONCLUSION

In this paper, we proposed a two-phase formation of fuzzy rule-based model. First, as prerequisite, fuzzy sets of condition and conclusion space are reconstructed with the use of Fuzzy C-Means clustering. Two reconstruction criteria E and F are introduced to qualify the reconstruction performance. Intuitively, the two reconstruction criterion both decrease with the increasing number of clusters. Once the fuzzy sets have been reconstructed, they can be applied to develop the interval-valued information granules $\tilde{A}_i$ (input space) and $\tilde{B}_i$ (output space) with the aid of the principle of justifiable granularity. Second, based on the location of input data x_k vis-à-vis information granules $\tilde{A}_i$, two different cases are considered: i) core mode concerns the region where information granules $\tilde{A}_i$ is located in; ii) approximation deals with data x_k that is located outside information granules $\tilde{A}_i$. Based on these two cases, granular output $\tilde{Y}_k$ of the fuzzy rule-based model is determined: i) if x_k is located in the region of input information granules $\tilde{A}_i \cup \tilde{A}_l$, the granular output is the union of output information granules $\tilde{B}_i \cup \tilde{B}_l$, $i = 1, 2, \ldots, c_1,$, $l = 1, 2, \ldots, c_2$, c_1 and c_2 are the number of clusters in the input and output space separately. ii) if x_k is located outside the input information granules $\tilde{A}_i$, the granular output is determined by the output information granules and considering an impact coming from the input memberships, that is $\tilde{Y}_k = A_1(\boldsymbol{x}_k) \otimes \tilde{B}_1 \oplus A_2(\boldsymbol{x}_k) \otimes \tilde{B}_2 \oplus \ldots \oplus A_c(\boldsymbol{x}_k) \otimes \tilde{B}_c$. An illustrative experiment is presented to state that most part of experimental output data can be covered by the granular output of the fuzzy rule-based model.

To evaluate the characterization of the quality of rules, three items are discussed in this study: a joint coverage, specificity of condition part and specificity of conclusion part. It is expected that a high level of the specificity of conclusion space and joint coverage as well as small specificity of condition space lead to high quality of the fuzzy rules of the model.

To analyze the performance, a collection of publicly available data sets coming from the UCI repository is applied to complete the experimental studies. The data are split into training and testing (70%-30% split) for evaluating the performance of the model. It is concluded that by increasing the number of clusters c from 3 to 7, the reconstruction criterion E and F goes decreasing which has a positive impact on reconstruction process. Also when the number of clusters increases from 3 to 7, a better performance can be obtained for the fuzzy rule-based model. In future researches, we expect to design a general model of fuzzy rule-based model using information granules, which is of great interesting in the following aspects: i) collaborative impacts between the condition space and conclusion space in developing the model is worth considering, ii) certify the relationship between specificity of condition space, specificity of conclusion space and the joint coverage based on the fuzzy rule of input and output space.

REFERENCES

[1] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets Syst.*, vol. 20, pp. 87–96, Aug. 1986.

[2] K. Atanassov and G. Gargov, "Interval valued intuitionistic fuzzy sets," *Fuzzy sets Syst.*, vol. 31, no. 3, pp. 343–349, 1989.

[3] P. Angelov, "An approach for fuzzy rule-base adaptation using on-line clustering," *Int. J. Approx. Reasoning*, vol. 35, no. 3, pp. 275–289, 2004.

[4] B. Andrzej and W. Pedrycz, Eds., *Human-Centric Information Processing Through Granular Modelling*, vol. 182. Berlin, Germany: Springer, 2009.

[5] J. C. Bezdek, C. Coray, R. Gunderson, and J. Watson, "Detection and characterization of cluster substructure I. Linear structure: Fuzzy c-lines," *SIAM J. Appl. Math.*, vol. 40, no. 2, pp. 339–357, 1981.

[6] G. Bortolan, "An architecture of fuzzy neural networks for linguistic processing," *Fuzzy Sets Syst.*, vol. 100, nos. 1–3, pp. 197–215, 1998.

[7] B. Carse, T. C. Fogarty, and A. Munro, "Evolving fuzzy rule based controllers using genetic algorithms," *Fuzzy Sets Syst.*, vol. 80, no. 3, pp. 273–293, 1996.

[8] O. Cordón, F. Herrera, and P. Villar, "Analysis and guidelines to obtain a good uniform fuzzy partition granularity for fuzzy rule-based systems using simulated annealing," *Int. J. Approx. Reasoning*, vol. 25, no. 3, pp. 187–215, 2000.

[9] M. J. Bender and S. P. Simonovic, "A fuzzy compromise approach to water resource systems planning under uncertainty," *Fuzzy Sets Syst.*, vol. 115, no. 1, pp. 35–44, 2000.

[10] H. Tzung-Pei and C.-Y. Lee, "Induction of fuzzy rules and membership functions from training examples," *Fuzzy Sets Syst.* vol. 84, no. 1, pp. 33–47, 1996.

[11] X. Hu, W. Pedrycz, O. Castillo, and P. Melin, "Fuzzy rule-based models with interactive rules and their granular generalization," *Fuzzy Sets Syst.*, vol. 307, pp. 1–28, Jan. 2017.

[12] H. Ishibuchi, T. Nakashima, and M. Nii, *Classification and Modeling With Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*. Berlin, Germany: Springer, 2004.

[13] W. J. M. Kickert and E. H. Mamdani, "Analysis of a fuzzy logic controller," *Fuzzy Sets Syst.*, vol. 1, no. 1, pp. 29–44, 1978.

[14] S. Mitra and S. K. Pal, "Fuzzy sets in pattern recognition and machine intelligence," *Fuzzy Sets Syst.*, vol. 156, no. 3, pp. 381–386, 2005.

[15] R. Kothamasu and S. H. Huang, "Adaptive Mamdani fuzzy model for condition-based maintenance," *Fuzzy Sets Syst.*, vol. 158, no. 24, pp. 2715–2733, 2007.

[16] C. R. Maurer, R. Qi, and V. Raghavan, "A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 265–270, Feb. 2003.

[17] Y. Miao and Y. Hua, "Fast subspace tracking and neural network learning by a novel information criterion," *IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 1967–1979, Jul. 1998.

[18] W. Pedrycz, "Fuzzy equalization in the construction of fuzzy sets," *Fuzzy sets Syst.*, vol. 119, no. 2, pp. 329–335, 2001.

[19] W. Pedrycz, *Knowledge-Based Clustering: From Data to Information Granules*. Hoboken, NJ, USA: Wiley, 2005.

[20] W. Pedrycz and P. Rai, "Collaborative clustering with the use of fuzzy c-means and its quantification," *Fuzzy Sets Syst.*, vol. 159, no. 18, pp. 2399–2427, Sep. 2008.

[21] W. Pedrycz, and C. Shyi-Ming, Eds., *Granular Computing and Intelligent Systems: Design with Information Granules of Higher Order and Higher Type*, vol. 13. Berlin, Germany: Springer, 2011.

[22] W. Pedrycz and M. Song, "Granular fuzzy models: A study in knowledge management in fuzzy modeling," *Int. J. Approx. Reasoning*, vol. 53, no. 7, pp. 1061–1079, 2012.

[23] W. Pedrycz and W. Homenda, "Building the fundamentals of granular computing: A principle of justifiable granularity," *Appl. Soft Comput.*, vol. 13, no. 10, pp. 4209–4218, 2013.

[24] W. Pedrycz, R. Al-Hmouz, A. Morfeq, and A. Balamash, "The design of free structure granular mappings: The use of the principle of justifiable granularity," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 2105–2113, Dec. 2013.

[25] W. Pedrycz, "Allocation of information granularity in optimization and decision-making models: Towards building the foundations of granular computing," *Eur. J. Oper. Res.*, vol. 232, no. 1, pp. 137–145, 2014.

[26] D. A. Ralescu, *Applications of Fuzzy Sets to Systems Analysis*. New York, NY, USA: Wiley, 1979.

[27] S.-B. Roh, T.-C. Ahn, and W. Pedrycz, "The design methodology of radial basis function neural networks based on fuzzy K-nearest neighbors approach," *Fuzzy Sets Syst.*, vol. 161, no. 13, pp. 1803–1822, 2010.

[28] O. F. Reyes-Galaviz and W. Pedrycz, "Enhancement of the classification and reconstruction performance of fuzzy C-means with refinements of prototypes," *Fuzzy Sets Syst.*, vol. 318, pp. 80–99, Jul. 2016.

[29] B. P. Shrestha, L. Duckstein, and E. Z. Stakhiv, "Fuzzy rule-based modeling of reservoir operation," *J. Water Resour. Planning Manage.*, vol. 122, no. 4, pp. 262–269, 1996.

[30] A. Skowron and J. Stepaniuk, "Information granules: Towards foundations of granular computing," *Int. J. Intell. Syst.*, vol. 16, no. 1, pp. 57–85, Jan. 2001.

[31] F. E. Thau, "Observing the state of non-linear dynamic systems," *Int. J. Control*, vol. 17, no. 3, pp. 471–479, 1973.

[32] E. Veldhuizen, E. J. Hultink, and A. Griffin, "Modeling market information processing in new product development: An empirical analysis," *J. Eng. Technol. Manage.*, vol. 23, no. 4, pp. 353–373, 2006.

[33] Z. Xu and M. Xia, "Distance and similarity measures for hesitant fuzzy sets," *Inf. Sci.*, vol. 181, no. 11, pp. 2128–2138, 2011.

[34] H. Wang, S. Kwong, Y. Jin, W. Wei, and K. F. Man, "Multi-objective hierarchical genetic algorithm for interpretable fuzzy rule-based knowledge extraction," *Fuzzy sets Syst.*, vol. 149, no. 1, pp. 149–186, 2005.

[35] K. K. Yen, S. Ghoshray, and G. Roig, "A linear regression model using triangular fuzzy number coefficients," *Fuzzy Sets Syst.*, vol. 106, no. 2, pp. 167–177, 1999.

[36] J. Yao, M. Dash, S. T. Tan, and H. Liu, "Entropy-based fuzzy clustering and fuzzy modeling," *Fuzzy Sets Syst.*, vol. 113, no. 3, pp. 381–388, 2000.

[37] M.-S. Yang, P.-Y. Hwang, and D.-H. Chen, "Fuzzy clustering algorithms for mixed feature variables," *Fuzzy Sets Syst.*, vol. 141, no. 2, pp. 301–317, 2004.

[38] F. Yu and W. Pedrycz, "The design of fuzzy information granules: Tradeoffs between specificity and experimental evidence," *Appl. Soft Comput.*, vol. 9, no. 1, pp. 264–273, 2009.

[39] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.

[40] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets Syst.*, vol. 1, no. 1, pp. 3–28, 1978.

[41] 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.

[42] K. Xu, W. Pedrycz, Z. Li, and X. Nie, "Constructing a virtual space for enhancing the classification performance of fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, to be published. doi: 10.1109/TFUZZ.2018.2889020.

[43] K. Xu, W. Pedrycz, W. Nie, and Z. Li, "High-accuracy signal subspace separation algorithm based on Gaussian kernel soft partition," *IEEE Trans. Ind. Electron.*, vol. 66, no. 1, pp. 491–499, Jan. 2019.

[44] D. Wang, W. Pedrycz, and Z. Li, "Granular data aggregation: An adaptive principle of the justifiable granularity approach," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 417–426, Feb. 2019.

DAN WANG received the B.S. degree in the telecommunication engineering from Xidian University, Xi'an, China, in 2013, where she is currently pursuing the Ph.D. degree in control theory and control engineering.

Her current research interests include granular computing, fuzzy modeling, computational intelligence, and algorithm analysis.

WITOLD PEDRYCZ (F'98) received the M.Sc., Ph.D., and D.Sci. degrees from the Silesian University of Technology, Gliwice, Poland.

He is currently a Professor and the Canada Research Chair in computational intelligence with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. He is also with the System Research Institute, Polish Academy of Sciences, Warsaw, Poland, where he is also a Foreign Member. He has authored 15 research monographs covering various aspects of computational intelligence, data mining, and software engineering. His current research interests include the computational intelligence, fuzzy modeling, granular computing, knowledge discovery and data mining, fuzzy control, pattern recognition, knowledge-based neural networks, relational computing, and software engineering.

Dr. Pedrycz currently serves as a member for a number of editorial boards of other international journals. He is a Fellow of the Royal Society of Canada. He was a recipient of the IEEE Canada Computer Engineering Medal, the Cajastur Prize for Soft Computing from the European Center for Soft Computing, the Killam Prize, and the Fuzzy Pioneer Award from the IEEE Computational Intelligence Society. He is intensively involved in editorial activities. He is an Editor-in-Chief of *Information Science*, and Editor-in-Chief of *WIREs Data Mining and Knowledge Discovery* (Wiley) and the *International Journal of Granular Computing* (Springer).

ZHIWU LI (M'06–SM'07–F'16) received the B.S. degree in mechanical engineering, the M.S. degree in automatic control, and the Ph.D. degree in manufacturing engineering from Xidian University, Xi'an, China, in 1989, 1992, and 1995, respectively.

He was a Visiting Professor with the University of Toronto, Toronto, ON, Canada, the Technion—Israel Institute of Technology, Haifa, Israel, the Martin-Luther University of Halle-Wittenburg, Halle, Germany, Conservatoire National des Arts et Métiers, Paris, France, and Meliksah Universitesi, Kayseri, Turkey. He joined Xidian University, in 1992. He is currently with the Institute of System Engineering, Macau University of Science and Technology, Macau, China. His current research interests include Petri net theory and application, supervisor control of discrete-event systems, work-flow modeling and analysis, system reconfiguration, game theory, and data and processing mining.

Dr. Li is a member of Discrete-Event Systems Technical Committee of the IEEE Systems, Man and Cybernetics Society and IFAC Technical Committee on Discrete-Event and Hybrid Systems, from 2011 to 2014. He was a recipient of an Alexander von Humboldt Research Grant, Alexander Von Humboldt Foundation, Germany He is the Founding Chair of Xi'an Chapter of the IEEE Systems, Man and Cybernetics Society. He is listed in Marquis *Who's Who in the World* (27 Ed., 2010). He serves as a Frequent Reviewer for more than 50 international journals, including *Automatica* and a number of the IEEE Transactions and many international conferences.