International Journal of Innovative Computing, Information and Control
Volume **19**, Number **1**, February **2023**
 ISSN 1349-4198
pp. **279–287**

A KIND OF INFORMATION GRANULAR FUZZY BROAD LEARNING SYSTEM BASED ON THE TAKAGI-SUGENO MODEL

Degang Wang[1], Menghang Li[1] and Wenyan Song[2,*]

[1]School of Control Science and Engineering
Dalian University of Technology
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China
wangdg@dlut.edu.cn; 964136862@qq.com

[2]School of Economics
Dongbei University of Finance and Economics
No. 217, Jianshan Street, Shahekou District, Dalian 116025, P. R. China
*Corresponding author: songwydufe@dufe.edu.cn

Received June 2022; revised September 2022

Abstract. *Based on the Takagi-Sugeno fuzzy model, a broad learning system, whose output has the form of information granule, is designed in this paper. The feature nodes and the enhancement nodes are constructed by integrating the membership degree of categories and the information of inference rules. This model mines and utilizes the distribution characteristics of data set itself and the correlation among attributes, and then it generates the information granular output with information extraction ability in virtue of the structure of broad learning system. As the results of experiments carried out show, the proposed method achieves satisfactory results in some public data set and a real data set.*
Keywords: Information granule, TS fuzzy model, Broad learning system, Clustering method

1. **Introduction.** The broad learning system (BLS) is a kind of model with strong feature expression ability and high learning efficiency [1,2]. It has attracted the attention of many researchers since it was proposed. A variety of broad learning models have been developed in many application scenarios, and many satisfactory results have been obtained [3-6]. Among them, it is worth mentioning that as the fuzzy system has a simple structure and certain semantic characteristics, a class of fuzzy broad learning systems (FBLS) has also been designed [7,8].

On the other hand, with the development of modeling technology and the enrichment of model forms, an increasing number of scholars pay attention to the feature extraction capability of objects and the comprehensibility in application for the modeling process. Therefore, the knowledge representation method, typified by the information granular theory, has gradually received more and more attention in the research of modeling methods [9-12]. Besides, in order to evaluate the quality of information granules, many scholars have proposed some evaluation criteria from different perspectives. In recent years, the principle of justifiable granularity has been widely accepted as a common evaluation criterion when designing an information granule [13,14].

It can be noticed that most of the current fuzzy broad learning systems are mainly used to handle numerical data [1-3], and the more enhancement nodes there are, the more complex the transformation of the original data, which will lead to neglect of the

DOI: 10.24507/ijicic.19.01.279

concentration degree of features. In addition, although by some transformation functions adopted in enhancement nodes, some nonlinear characteristics of the data have been extracted to a certain extent, the correlation between attributes of different dimensions in the data has not been utilized. In consideration of the fact that the information granule can better describe the attributes or objects with variable characteristics [9,10,13], based on the Takagi-Sugeno (TS) model, this paper proposes a fuzzy broad learning system, which combines the data distribution and the feature concentration degree together, and designs the output of the model as an information granule, which makes it more applicable and has certain semantic interpretation ability.

The rest of this paper is organized as follows. In Section 2, the granulating technique of data is proposed. In Section 3, the construction method of the feature node and the enhancement node in the fuzzy broad learning system based on the TS model is introduced. Besides, the modeling process of the FBLS is carried out. Section 4 shows the experiment results of five public data sets and a real-world data set. Section 5 presents conclusions.

2. **Information Granule Design.** It is known that information granule has been effectively used for knowledge representation. Generally speaking, the constructed information granules are required to have ability of information extraction and characterization, that is, to have high specificity and coverage degree at the same time [13,14]. To evaluate the quality of an information granule, the principle of justifiable granularity proposed in [13] is adopted.

Drawing on ideas from [15], by improving the granularity distribution module according to the aggregation of data set, a series of granules are designed in the following process. For a data set $\boldsymbol{Y} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\}$, in which $\boldsymbol{y}_k \in \boldsymbol{R}^m$ $(k = 1, \ldots, N)$, a generation method of hyper-cubic type information granule is considered in this paper.

Firstly, the data set is clustered into M classes. And the center module is obtained based on the center and the category membership function of each class. For the center module, let the numeric feature $\boldsymbol{m}_j$ $(j = 1, \ldots, M)$ represent every cluster, where $\boldsymbol{m}_j$ can be chosen as the cluster center itself like in this paper, or the weighting average points of all the data belonging to the j-th cluster. Then for the granularity distribution module, the level of granularity $\tilde{\boldsymbol{\varepsilon}}_j$ is introduced for current information granule design. Here, a rule-based fuzzy system is adopted to construct the granularity distribution module, from which we can achieve a nonlinear mapping from the data set to a granularity level. The rules for the granularity module are as follows:

$$\text{Rule } j\text{: if } \boldsymbol{y}_k \text{ is } B_j(\boldsymbol{y}_k) \text{ then } \tilde{\boldsymbol{\varepsilon}}_j \text{ is } \left(\varepsilon_j^-, \varepsilon_j^+\right), \tag{1}$$

where $B_j(\boldsymbol{y}_k)$ is the membership function of data $\boldsymbol{y}_k$ in the j-th rule, which measures the extent to that data $\boldsymbol{y}_k$ belongs to the j-th class. The vector $\left(\varepsilon_j^-, \varepsilon_j^+\right)$ controls the lower and upper boundaries of the granularity level respectively, and each component of it has a value between 0 and 1. In this paper, the membership degree of data in each rule is calculated by considering the amount of spatial bias to each hyper-cubic information granule.

The detailed implementation method is as follows. Consider a hyper-cubic $\boldsymbol{\Omega}$ in the space $\boldsymbol{R}^m$. If $m = 1$, then $\boldsymbol{\Omega}$ is an interval. If $m = 2$, then $\boldsymbol{\Omega}$ is a rectangle. And if $m = 3$, then $\boldsymbol{\Omega}$ is a cube. As shown in Figure 1, a two-dimensional hyper-cubic information granule is a rectangle, which can be uniquely determined by the kernel $\boldsymbol{v} = (v^1, v^2)$ and the extension vector $\boldsymbol{d} = \left(\boldsymbol{d}_1^-, \boldsymbol{d}_1^+, \boldsymbol{d}_2^-, \boldsymbol{d}_2^+\right)$ in two directions. Obviously, the two sides of the rectangle are $\boldsymbol{d}_1 = \boldsymbol{d}_1^- + \boldsymbol{d}_1^+$ and $\boldsymbol{d}_2 = \boldsymbol{d}_2^- + \boldsymbol{d}_2^+$, respectively.

After clustering the data set into M classes, we can get all the cluster centers $\boldsymbol{m}_j$, and take them as the kernels $\boldsymbol{v}_j$. Accordingly, the data set is divided into M subsets according

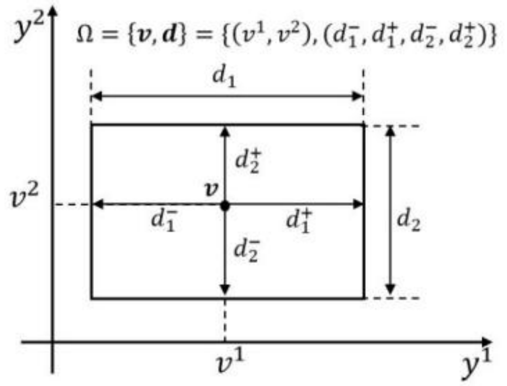

FIGURE 1. Two-dimensional hyper-cubic information granule

to the clustering labels as D_j, $j = 1, \ldots, M$. Then several interval information granules can be generated by using the principle of justifiable granularity in each dimension to determine the extension vector of the corresponding hyper-cubic.

To evaluate the equality of hyper-cubic information granule as constructed above, the specificity of the p-th dimension on the j-th hyper-cubic is measured by the following formula:

$$spec_j^p = 1 - \frac{\left|b_j^p - a_j^p\right|}{y_{\max}^p - y_{\min}^p}, \quad p = 1, \ldots, m, \;\; j = 1, \ldots, M, \tag{2}$$

where a_j^p and b_j^p are the left and right boundaries of the p-th dimension of the j-th hyper-cubic respectively, and the extension vector on this dimension is $\left(d_{jp}^-, d_{jp}^+\right) = \left(v_j^p - a_j^p,\right.$ $\left.b_j^p - v_j^p\right)$. $y_{\max}^p$ and $y_{\min}^p$ are the maximum and minimum values of the p-th dimension of the data which are included D_j.

In addition, the coverage of the information granule on this dimension can be expressed as

$$cov_j^p = \frac{1}{N} \sum_{k=1}^{N} incl\left(y_k^p \middle| \boldsymbol{y}_k \in D_j\right), \tag{3}$$

in which

$$incl\left(y_k^p \middle| \boldsymbol{y}_k \in D_j\right) = \begin{cases} 1, & y_k^p \in \left[a_j^p, b_j^p\right] \text{ and } \boldsymbol{y}_k \in D_j, \\ 0, & \text{otherwise.} \end{cases}$$

Subsequently, define the quality evaluation function of the information granule as the product of the specificity and the coverage, i.e.,

$$V_j^p = spec_j^p * cov_j^p. \tag{4}$$

Then the parameters d_{jp}^- and d_{jp}^+ can be determined by maximizing the above function V_j^p. So far, we can obtain M hyper-cubic shape information granules $\boldsymbol{\Omega}_j = \left\{\left(\boldsymbol{v}_j^p, d_{jp}^-, d_{jp}^+\right)\right\}$, $p = 1, \ldots, m$, $j = 1, \ldots, M$.

As mentioned above, the traditional membership degree calculation method is to calculate the distance between each data and the cluster center, and then we can normalize these distances to get the membership degree belonging to each rule. In this paper, since the data space is constructed into several hyper-cubic information granules, the offset of information granule δ_k^j is considered to evaluate the distance from data to kernel of a granule, i.e., $\delta_k^j = \frac{\sum_{l=1}^{2^m} \|\boldsymbol{r}_l - \boldsymbol{y}_k\|}{2^m} - \|\boldsymbol{r}_0 - \boldsymbol{g}_j\|$, where $\boldsymbol{g}_j$ is the geometric center of the hyper-cubic, which can be determined by $\boldsymbol{d}_j$, and $\boldsymbol{r}_l$ is the l-th vertex of the hyper-cubic. Because any vertex has the same distance from the center, $\boldsymbol{r}_0$ can be any vertex of all the 2^m vertices. In this way, the membership function $B_j(\boldsymbol{y}_k)$ is determined as $B_j(\boldsymbol{y}_k) = 1\Big/\sum_{t=1}^{M}\left(\frac{\delta_k^j}{\delta_k^t}\right)^2$.

Hereby, the level of granularity for current information granule is $\varepsilon^+ = \sum_{j=1}^{M} B_j(\boldsymbol{y}_k)\varepsilon_j^+$ and $\varepsilon^- = \sum_{j=1}^{M} B_j(\boldsymbol{y}_k)\varepsilon_j^-$. When $m = 2$, the rectangle type information granule can be described as $G_j = \left[v_j^1 - \varepsilon^- * range_{y^1}, v_j^1 + \varepsilon^+ * range_{y^1}\right] \times \left[v_j^2 - \varepsilon^- * range_{y^2}, v_j^2 + \varepsilon^+ * range_{y^2}\right]$, in which $range_{y^1}$ is the interval length of the value for the first dimension of the data, and $range_{y^2}$ is the interval length of the value for the second dimension of the data. Given a certain level of constraint like $\frac{1}{M}\sum_{j=1}^{M}\left(\varepsilon_j^- + \varepsilon_j^+\right) \leq \varepsilon_0$, where $\varepsilon_0 \in [0, 1]$, the consequent part $\left(\varepsilon_j^-, \varepsilon_j^+\right)$ in the fuzzy rule (1) can be determined by solving a series of optimization problems, in which the goals are to maximize V_j^p in Equation (2) to Equation (4). Since these optimization problems are nonlinear and have no analytical solutions, in this paper the particle swarm optimization (PSO) algorithm is adopted to deal with them. Similarly, other intelligent optimization methods can also be chosen to solve these problems. Moreover, it should be pointed out that the closer ε_0 is to 1, the longer the sides of the rectangle are.

3. **Fuzzy Broad Learning System Based on Information Granule.** In a typical broad learning system, the original inputs are mapped into random feature nodes by some feature mappings. Then by means of nonlinear transformation, the random features are put into the enhancement nodes. Next, the random feature nodes and the enhancement nodes are connected to the output layer of the learning system [1,2]. In order to make full use of the distribution characteristics of the data set itself, the TS fuzzy model, which can be used to handle complex data modeling [16], is adopted to design the feature nodes and the enhancement nodes of the broad learning system in this paper.

Firstly, by a selected clustering method, a TS fuzzy system can be represented, in which the premises of the fuzzy rules are determined by the clustering centers and the membership functions. For simplicity, the consequent of the fuzzy rule is chosen as the linear function of the input vectors, and the least square method is used to train the parameters. After the TS model has been constructed, the output of each fuzzy rule is taken as the feature node in the broad learning system. Meanwhile, the product of the membership function in the antecedent part of the rule and the output of the same rule is taken as the transitional information, which is transferred into the enhancement node. Considering the influence of clustering algorithm on the class center and the membership function, different clustering methods can be used together to design multi-group of feature nodes and transitional input of enhancement nodes in parallel.

3.1. **Feature node.** Given an input data set $\boldsymbol{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subseteq \boldsymbol{R}^n$, where $\boldsymbol{x}_k = (x_{k1} \ldots, x_{kn})^T$, $k = 1, \ldots, N$, and a corresponding output data set $\boldsymbol{Y} = \{y_1, \ldots, y_N\} \subseteq \boldsymbol{R}$, by some clustering method, c clustering centers and corresponding membership functions can be obtained. Accordingly, the fuzzy rules in the TS model can be described as follows:

$$\text{Rule } i\text{: if } \boldsymbol{x}_k \text{ is } A_i(\boldsymbol{x}_k) \text{ then } \tilde{y}_i \text{ is } y_i(\boldsymbol{x}_k), \quad i = 1, \ldots, c.$$

Here $A_i(\boldsymbol{x}_k)$ is the membership function, defined as $A_i(\boldsymbol{x}_k) = 1\Big/\sum_{j=1}^{c}\left(\frac{\|\boldsymbol{x}_k - \boldsymbol{v}_i\|}{\|\boldsymbol{x}_k - \boldsymbol{v}_j\|}\right)$, which satisfies $\sum_{i=1}^{c} A_i(\boldsymbol{x}_k) = 1$. And the local consequent $\tilde{y}_i$ is a linear function of all the input variables, i.e., $\tilde{y}_i(\boldsymbol{x}_k) = w_i + \boldsymbol{p}_i^T(\boldsymbol{x}_k - \boldsymbol{v}_i)$, where $(\boldsymbol{v}_i, w_i) \in \boldsymbol{R}^{n+1}$ is the clustering center prototypes of input-output data set, and $\boldsymbol{p}_i$ is the coefficient vector in the local linear function. For this TS model, the output is the combination of c rules:

$$F(\boldsymbol{x}_k) = \sum_{i=1}^{c} A_i(\boldsymbol{x}_k)\left[w_i + \boldsymbol{p}_i^T(\boldsymbol{x}_k - \boldsymbol{v}_i)\right].$$

Furthermore, if N_f clustering methods are used to design N_f TS fuzzy models, then we can get N_f group of feature nodes. In each group, the feature nodes are the outputs of every input data, i.e., $\{F_j(\boldsymbol{x}_k)\}$, $k = 1, \ldots, N$, $j = 1, \ldots, N_f$. Hence, $\boldsymbol{F}^{N_f} = \left(F_1, \ldots, F_{N_f}\right)$ is the output of the feature layer.

3.2. **Enhancement node.** In this paper, different from the usual way in an FBLS model, we do not choose $\boldsymbol{F}^{N_f}$ as the input of the enhancement node. We take the combination of every consequent in the fuzzy rule and the membership function value together as the input of the enhancement node, which can preserve more of the distribution characteristics from the original data.

Specifically, set the intermediate variables of the j-th TS model as $\boldsymbol{Z}_j = \left(A_{j1}(\boldsymbol{x}_k)\tilde{y}_{j1}(\boldsymbol{x}_k), \ldots, A_{jc_j}(\boldsymbol{x}_k)\tilde{y}_{jc_j}(\boldsymbol{x}_k)\right)$, $j = 1, \ldots, N_f$, where c_j is the number of clusters for the j-th clustering method, i.e., the number of rules. Then $\boldsymbol{Z}^{N_f} = \left(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{N_f}\right)$ is transferred to the enhancement nodes as $\boldsymbol{H}^{N_e} = (\boldsymbol{H}_1, \ldots, \boldsymbol{H}_{N_e})$, where N_e is the number of enhancement nodes in the FBLS. And $\boldsymbol{H}_l = \xi\left(\boldsymbol{Z}^{N_f}\boldsymbol{W}_{h_l} + \boldsymbol{\beta}_{h_l}\right)$, $l = 1, \ldots, N_e$, is the output matrix of the l-th enhancement node, where $\boldsymbol{W}_{h_l}$ and $\boldsymbol{\beta}_{h_l}$ are the weight and offset of the feature layer with the l-th enhancement node group. $\xi(\cdot)$ is the activation function of the enhancement node. Here, $\boldsymbol{W}_{h_l}$ and $\boldsymbol{\beta}_{h_l}$ are assumed to obey $[0, 1]$ uniform distribution, and $\xi(\boldsymbol{u}) = \frac{2}{1+e^{-2u}} - 1$.

3.3. **FBLS with granular output.** At last, the output of feature layer $\boldsymbol{F}^{N_f}$ and the output of enhancement layer $\boldsymbol{H}^{N_e}$ are imported to the top output layer at the same time. Then the final output of the FBLS can be determined as $\hat{y}(\boldsymbol{x}_k) = \left(\boldsymbol{F}^{N_f}, \boldsymbol{H}^{N_e}\right)\boldsymbol{W}_O$, in which $\boldsymbol{W}_O$ is the weight between the whole input nodes formed by the feature layer, the enhancement layer, and the top output layer. When the target value of the training data is given as $\boldsymbol{Y}$, $\boldsymbol{W}_O$ can be calculated by the ridge regression algorithm or other optimization regression algorithm. In this paper, considering that the amount of data processed is not huge, we use the Moore-Penrose pseudo inverse method to calculate it, i.e.,

$$\boldsymbol{W}_O = \left(\boldsymbol{F}^{N_f}, \boldsymbol{H}^{N_e}\right)^{\dagger}\boldsymbol{Y}.$$

Further, in order to obtain the outputs of the learning system with the shape of information granule, we use the granular technique proposed in Section 2 to convert $\hat{\boldsymbol{Y}} = \{\hat{y}_1, \ldots, \hat{y}_N\}$ into several information granules $\boldsymbol{G} = \{G_1, \ldots, G_M\}$. At last, the FBLS with granular output (FBLS-IG) is obtained, whose structure is shown in Figure 2.

4. **Experiment Results.** In order to investigate the effect of the designed model, two experiments are considered in this section. As to the PSO algorithm, without loss of generality, the number of initial particle population is set as 100, the learning factor is set as 1.49, and the maximum iteration number is set as 100. In the proposed information granule design process, the granularity level parameter ε_0 is initially set to 0.1 and grows to 1 in steps of 0.1.

Example 4.1. *Five public data sets, including Boston housing, Auto-MPG, Stock, Red wine and Airfoil Self-Noise, from the UCI machine learning repository and the StaLib repository are discussed*[1]*. The brief information about these data sets is listed in Table 1.*

The K-means clustering method, the fuzzy C-means (FCM) clustering method and the Ward's hierarchical clustering method are used to get three groups of feature nodes and

[1] http://archive.ics.uci.edu/ml/; http://www.dcc.fc.up.pt/∼Itorgo/Regression.

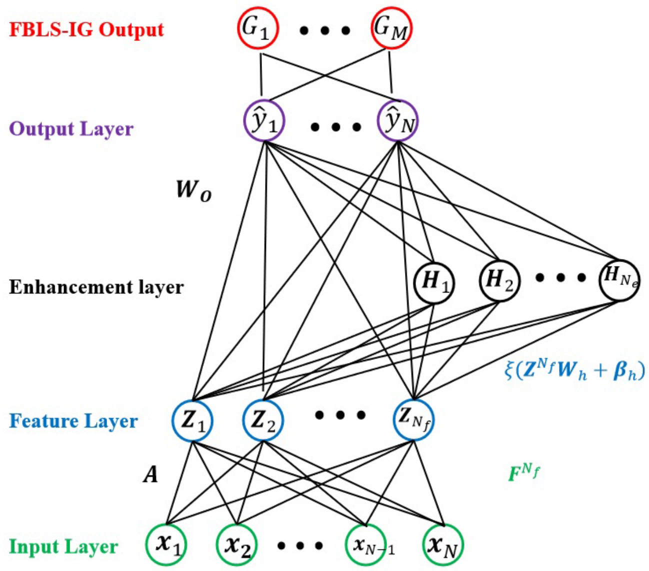

FIGURE 2. The structure of the proposed FBLS-IG

TABLE 1. Public data sets used in experiment: A brief description

Data set	Number of data	Input dimensions
Boston housing	506	13
Auto-MPG	392	7
Stock	950	9
Red wine	1599	11
Airfoil Self-Noise	1503	5

enhancement nodes, i.e., $N_f = 3$. Since the identification process of model parameters is stochastic, the average statistical results after 50 experiments are recorded. In this experiment, for each data set, 70% of the data are randomly selected as the training set and the remaining 30% as the test set. And consider that the dimension of the data set does not exceed 13 and as few nodes as possible should be included, the number of enhancement nodes N_e starts at 1 and increases by 1 to 40.

In this experiment, for comparison with models in literature, the following indicator is used [17]:

$$Q = \frac{1}{c}\left(\sum_{i=1}^{c}\frac{(v_i - \hat{v}_i)}{\sigma_v^2} + \sum_{i=1}^{c}\frac{\left(d_i - \hat{d}_i\right)}{\sigma_d^2}\right),$$

in which $\hat{v}_i$ and $\hat{d}_i$ are the center and radius of the output information granule, and σ_v and σ_d are the standard deviations. This indicator Q reflects the degree to which the constructed information granule can describe the data distribution and feature concentration. The method presented in this paper (FBLS-IG) is compared with three representative granule models discussed in [15], including the model whose output is an information granule (OutIG), the model whose parameters are information granules (ParaIG), and the model whose clustering centers are information granules (ClusIG). The mean values of Q are shown in Table 2. As we can see for four data sets, the FBLS-IG model has better result.

Example 4.2. *A medical data set is discussed in this case, which contains information of 148 patients with heart disease. This data set records the relationship between blood*

TABLE 2. The comparison of average granularity index under different models

Data set	Model	$c = 3$	$c = 5$	$c = 7$	Mean Q
Boston housing	OutIG	0.5058	0.5095	0.5700	0.5453
	ParaIG	0.2815	0.2867	0.2912	0.2839
	ClusIG	0.5934	0.6019	0.6203	0.6119
	FBLS-IG	0.6931	0.6977	0.6881	**0.6926**
Auto-MPG	OutIG	0.6695	0.6645	0.6733	0.6692
	ParaIG	0.2054	0.2068	0.1926	0.1975
	ClusIG	0.6469	0.6424	0.7413	0.6724
	FBLS-IG	0.6993	0.7031	0.7055	**0.7010**
Stock	OutIG	0.7612	0.7466	0.7423	0.7582
	ParaIG	0.5214	0.5403	0.5039	0.5158
	ClusIG	0.7780	0.7985	0.7537	0.7706
	FBLS-IG	0.8142	0.8180	0.8188	**0.8165**
Red wine	OutIG	0.5934	0.5905	0.5806	**0.5864**
	ParaIG	0.2764	0.2687	0.2650	0.2631
	ClusIG	0.5357	0.5651	0.5827	0.5684
	FBLS-IG	0.5291	0.5248	0.5255	0.5258
Airfoil Self-Noise	OutIG	0.5367	0.5383	0.5424	0.5411
	ParaIG	0.2651	0.2720	0.2796	0.2609
	ClusIG	0.5843	0.4823	0.4537	0.4660
	FBLS-IG	0.5616	0.5668	0.5626	**0.5635**

sodium, potassium, calcium, phosphorus, magnesium and heart rate (HR), QRS time, PR cycle, QT interval, QT correction time (QTC) and other important ECG indicators within 5 hours. Based on the proposed model (FBLS-IG), a quantitative model is established to analyze the changes of electrolytes in the blood within five hours, so as to predict five most focused physical indicators, including HR, QRS time, PR cycle, QT interval and QT correction time, and to facilitate the subsequent treatment of the patients.

In this experiment, since the constructed information granules are intervals, the following three indicators are used for comparison:

$$ARV^I = \frac{\sum_{k=1}^{N}(a_k - \hat{a}_k)^2 + \sum_{k=1}^{N}\left(b_k - \hat{b}_k\right)^2}{\sum_{k=1}^{N}(a_k - \bar{a})^2 + \sum_{k=1}^{N}\left(b_k - \bar{b}\right)^2},$$

$$COV^I = \frac{\sum_{k=1}^{N}\sum_{t=1}^{T} card\left\{\boldsymbol{x}_k^t \middle| \boldsymbol{x}_k^t \in \boldsymbol{G}_k^y\right\}}{N * T},$$

$$SP^I = \frac{\sum_{k=1}^{N}\left(1 - \max\left(0, \frac{\hat{b}_k - \hat{a}_k}{range_y}\right)\right)}{N},$$

where a_k and b_k are the lower and upper bounds of the k-th target interval. $\bar{a}$ and $\bar{b}$ are the mean values of them. Besides, $\hat{a}_k$ and $\hat{b}_k$ are the endpoints of the predicted interval granule. $\boldsymbol{x}_k^t$ is the record value of the k-th patient when he or she took the t-th test. $card\{\cdot\}$ is the counting function. N is the number of all the patients, and T is the length of testing time.

The FCM clustering method is used to cluster the data set, and the number of the enhancement nodes is set to be 5, i.e., $N_f = 1$, $N_e = 5$. The number of clusters is set to be 2. The comparison results with the back propagation (BP) network and the support

Table 3. The comparison between the proposed model, BP and SVM

ECG indicator	FBLS-IG			BP			SVM		
	$\overline{ARV^I}$	$\overline{COV^I}$	$\overline{SP^I}$	$\overline{ARV^I}$	$\overline{COV^I}$	$\overline{SP^I}$	$\overline{ARV^I}$	$\overline{COV^I}$	$\overline{SP^I}$
HR	**1.5981**	**0.6036**	0.6879	1.7235	0.5673	**0.7427**	2.3087	0.5305	0.7413
QRS	**1.1765**	**0.5032**	0.7809	1.7319	0.3982	0.8511	3.1688	0.5009	**0.8539**
PR	**1.8961**	**0.7110**	0.7081	2.1382	0.6286	0.4634	3.7672	0.4833	**0.7692**
QT	**1.4909**	**0.5900**	**0.7117**	1.7890	0.5195	0.6790	3.2642	0.5468	0.7109
QTC	**1.5286**	**0.6682**	0.6750	1.8362	0.5686	**0.7227**	3.8336	0.5809	0.7028

vector machine (SVM) are listed in Table 3. It can be seen that though the proposed model is not as good as the BP network and the SVM model in terms of the specificity, it is better than them in terms of the coverage and the overall accuracy of the interval granules.

5. **Conclusions.** In this paper, based on the fuzzy TS system, some clustering algorithms are used to generate feature mapping nodes and enhancement nodes in the FBLS model. Besides, the output of the proposed system is transferred into the form of hypercubic shape information granule. Benefited from the fuzzy system construction method and the information granule generation technology, the data distribution and feature agglomeration are exploited fully in the construction process of the FBLS-IG model. And in the following studies, we will further investigate the optimization method and the role of different forms of information granules in the construction and application of the broad learning system.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China under Grant (12071056, 61773088), the Fundamental Research Funds for the Central Universities (DUT22YG227), the Humanities and Social Science project for Ministry of Education of China (21YJA630079) and the National Key R&D Program of China (2018AAA0100300). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] C. L. Philip Chen and Z. Liu, Broad learning system: An effective and efficient incremental learning system without the need for deep architecture, *IEEE Transactions on Neural Networks and Learning Systems*, vol.29, no.2, pp.10-24, 2018.

[2] C. L. Philip Chen, Z. Liu and S. Feng, Universal approximation capability of broad learning system and its structural variations, *IEEE Transactions on Neural Networks and Learning Systems*, vol.30, no.4, pp.1191-1204, 2019.

[3] S. Feng, W. Ren, M. Han and Y. W. Chen, Robust manifold broad learning system for large-scale noisy chaotic time series prediction: A perturbation perspective, *Neural Networks*, vol.117, pp.179-190, 2019.

[4] J. Jin, Y. Li, T. Yang, L. Zhao, J. Duan and C. L. Philip Chen, Discriminative group-sparsity constrained broad learning system for visual recognition, *Information Sciences*, vol.576, pp.800-818, 2021.

[5] Z. Liu, S. Huang, W. Jin and Y. Mu, Graph-based broad learning system for classification, *Neurocomputing*, vol.463, pp.535-544, 2021.

[6] L. Liu, L. Cai, T. Liu, C. L. Philip Chen and X. Tang, Cauchy regularized broad learning system for noisy data regression, *Information Sciences*, vol.603, pp.210-221, 2022.

[7] S. Feng and C. L. Philip Chen, Nonlinear system identification using a simplified fuzzy broad learning system: Stability analysis and a comparative study, *Neurocomputing*, vol.337, pp.274-286, 2019.

[8] S. Feng and C. L. Philip Chen, Fuzzy broad learning system: A novel neuro-fuzzy model for regression and classification, *IEEE Transactions on Cybernetics*, vol.50, no.2, pp.414-424, 2020.
[9] L. A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, vol.90, no.2, pp.111-127, 1997.
[10] M. M. Ahmed and N. A. Mat Isa, Knowledge base to fuzzy information granule: A review from the interpretability-accuracy perspective, *Applied Soft Computing*, vol.54, pp.121-140, 2017.
[11] D. Wang, H. Liu, W. Pedrycz, W. Song and H. Li, Design Gaussian information granule based on the principle of justifiable granularity: A multi-dimensional perspective, *Expert Systems with Applications*, vol.197, 116763, 2022.
[12] T. M. Hossain, J. Watada, I. A. Aziz, M. Hermana, S. T. Meraj and H. Sakai, Lithology prediction using well logs: A granular computing approach, *International Journal of Innovative Computing, Information and Control*, vol.17, no.1, pp.225-244, 2021.
[13] W. Pedrycz and W. Homenda, Building the fundamentals of granular computing: A principle of justifiable granularity, *Applied Soft Computing*, vol.13, no.10, pp.4209-4218, 2013.
[14] W. Pedrycz and X. Wang, Designing fuzzy sets with the use of the parametric principle of justifiable granularity, *IEEE Transactions on Fuzzy Systems*, vol.24, no.2, pp.489-496, 2016.
[15] X. Hu, W. Pedrycz and X. Wang, Development of granular models through the design of a granular output space, *Knowledge-Based Systems*, vol.134, pp.159-171, 2017.
[16] Y. Xie, J. Liu and L. Wang, H_∞ filtering design for Takagi-Sugeno fuzzy model with immeasurable premise variables by applying a switching method, *ICIC Express Letters*, vol.14, no.3, pp.257-264, 2020.
[17] X. Zhu, W. Pedrycz and Z. Li, Development and analysis of neural networks realized in the presence of granular data, *IEEE Transactions on Neural Networks and Learning Systems*, vol.31, no.9, pp.3606-3619, 2020.

Author Biography

Degang Wang received the B.Sc. degree in mathematics from Liaoning Normal University, China, in 2002; the M.Sc. degree in mathematics from Liaoning Normal University, China, in 2005; the Ph.D. degree in applied mathematics from Beijing Normal University, China, in 2008.

Dr. Wang is currently a full-time associate professor at the School of Control Science and Engineering, Dalian University of Technology, China. His research interests include fuzzy system modelling and applications, machine learning and intelligent control.

Menghang Li received the B.Sc. degree in automation from Dalian Maritime University, China, in 2018; the M.Sc. degree in control science and engineering from Dalian University of Technology, in 2021. His research interests include machine learning and knowledge representation.

Wenyan Song received the B.Sc. degree in mathematics from Beijing Normal University, China, in 2002; the Ph.D. degree in applied mathematics from Beijing Normal University, China, in 2007.

Dr. Song is currently a full-time associate professor at the School of Economics, Dongbei University of Finance and Economics, China. Her research interests include fuzzy system modelling, neural network and management optimization.