Research Article

Multilevel Information Granule Construction Model Based on Large Data Environment and Its Application in Time Series

Dongxian Yu[1] and Jian Xiao [2,3]

¹*College of Modern Information Technology, Henan Polytechnic, Zhengzhou 450046, Henan, China*
²*School of Business, Zhejiang Fashion Institute of Technology, Ningbo 315211, China*
³*School of Business, Beijing International Studies University, Beijing 100024, China*

Correspondence should be addressed to Jian Xiao; xiaojian@bisu.edu.cn

Considering the uneven distribution of time series data in time windows, this paper designs an information granulation method based on the idea of "multigranularity," and uses multilayer information granularity to construct the prediction model of time series. Firstly, a multigranularity time series model is established by applying binary relation on the time axis of time series data, and an evaluation effect function of time series data mining is introduced for the purpose of forecasting tasks. Based on this function, the optimal time granularity of prediction is found. In view of the uncertainty of time series, a multidimensional matrix model of attributes of sequence data is constructed, and the clustering category information of each attribute is obtained by clustering operation on the matrix model. Finally, the clustered data are classified by machine learning method to obtain classification knowledge. Prediction is achieved by classifying knowledge. Finally, the ICU data were used to predict the life and death of ICU patients to test the predictive effect of granular computing on time series data. The experimental results show the following: (1) the proposed method is superior to the traditional time series analysis and modeling methods. (2) When EM clustering algorithm, optimal time granularity and ID3 decision tree algorithm are used, the prediction rate of ICU mortality can reach 92.13% by using ICU time series data.

1. Introduction

Time series data exist in many fields such as nature, finance, medicine, and so on. It objectively records the important information of the observed system at every time point. As the output of the observed system, time series often implies some specific laws and potential characteristics of the system. Therefore, time series analysis and modeling has been widely studied. Early researchers used the linear system theory, black box methodology, and fuzzy set theory to analyze and model time series and obtained many classical time series models, such as AR series model [1], artificial neural network model [2], Bayesian model [3], and fuzzy time series model [4]. These models have been applied in various fields and achieved good results. However, in the context of today's big data era, time series are characterized by high-dimensional data representation, massive data, and

complexity of data structure, which makes some traditional time series analysis and modeling methods cannot fully meet the practical application needs. The main manifestations are as follows: (1) the traditional time series analysis and modeling methods are essentially "numerical" centered, when the data sample size is huge, and the "interpretability" of the time series model constructed by the traditional time series modeling methods will be greatly reduced. (2) Traditional models cannot deal with the relationship between the dimensions of complex time series, which leads to poor results in dealing with complex time series data. (3) The models constructed by traditional time series analysis and modeling methods are all output in the form of specific "digital values." Such output is difficult to understand and recognize in large data environments. Therefore, many researchers have tried other methods to deal with time series in large data environments. In reference [5], the convolution

neural network is introduced to the mining of time series data. Reference [6] proposes the use of recurrent neural network (RNN) to solve the prediction and repair of complex time series. Bettini et al. [7] proposed for the first time that the size of time interval in time series should be regarded as time granularity. Time granularity function is used to express multiple description levels in time series and then mining frequent association patterns at different levels. Ruan et al., combined the granulation of fuzzy information with support vector machine, proposed a time series model of fuzzy granular support vector machine [8]. In these studies, the introduction of granular computing into time series analysis is the main research.

Granular computing is a natural model to simulate human thinking and solving large-scale complex problems. It improves computing efficiency by granulating complex data and replacing samples with information granules. In granular computing, the transformation of deterministic information and uncertain information can be achieved by changing granularity; the data scale of original data can be reduced by abstract summary representation of the finest granular raw data by granular structure; and the time efficiency of solution can be improved by establishing domain-oriented granular structure. At present, there are many research results on time series analysis using granular computing. In [9], Hryniewicz and Kaczmarek integrated information granules into the Bayesian posterior estimation process, proposed a Bayesian probability prediction model of time series based on granular computing. In [10], M. Y. Chen and B. T. Chen used the discrete iteration method based on entropy to degranulate the domain of fuzzy time series, and based on this, we constructed a hybrid fuzzy time prediction model to predict stock prices. Documents [11] Al-Hmouz et al. construct information particles in space by collecting amplitude and amplitude changes through time windows, and experiment with a series of time series data sets. Lu et al. transformed the original time series into granular time series, used different window widths to control the size of granulation, and analyzed the influence of granulation window size and granulation degree on the prediction results through experiments. In [13], Wang et al. applied the improved fuzzy C-means clustering algorithm and information granulation to the long-term prediction of time series. In [8], Ruan et al. combined the granulation of fuzzy information with support vector machines (SVM) to predict large-scale data rapidly by using FGSVM (fuzzy granular support vector machines). In [14], Vairavan et al. proposed the method of granulation of multigranularity information to analyze the trend of data transformation. Geometric models were constructed by using the topological grouping structure of moving objects in two-dimensional or high-dimensional space. The transformation trend of time-varying data was detected by adjusting granularity, support size and duration. Although there are many achievements in the research of time series analysis and modeling methods based on granular computing, some of these studies reduce the scale of time series data through information granules and some simplify the calculation through information granules. They only take advantage of some characteristics of

granular computing and lack systematic methods for time series analysis and modeling based on granular computing. In addition, many studies use "single information granule," that is, when granulating time series, only one granulation level is considered. This cannot reasonably reflect the different viewing angles of users. The information granulation of time series data at different granularity levels will help us to analyze and observe time series data from different levels according to the needs of the problem.

Therefore, this paper proposes to improve the processing of time series data through granular computing in three aspects. Firstly, according to the high dimension of time series data, the ICU time series are granulated at different granularity levels, and the multigranularity time series model is established to reduce the original data scale. Aiming at the uncertainty of time series, the multidimensional matrix model is constructed on the basis of the multi-granularity time series model. In order to reduce the difficulty of time series data processing, the multidimensional matrix model processing is divided into two steps. Firstly, the matrix model is clustered to get the clustering category information of each attribute. Then, the clustered data are classified by the machine learning method. Finally, this paper tests the predictive effect of the proposed method by using ICU data to predict the life and death of ICU patients. The experimental results show that when EM clustering algorithm, optimal time granularity, and ID3 decision tree algorithm are used, the prediction rate of ICU mortality can reach 92.13% by using ICU time series data.

2. Time Series Prediction Model Based on Multilayer Information Granules

2.1. Granule Computation

2.1.1. Granular Computing Framework. Granular computing is a very active research field in computer science in recent ten years. By abstracting and dividing complex problems into several simpler ones, it helps to better analyze and solve problems. In the framework of granular computing, according to the actual problem itself, the problem information is granulated into information granules, which are calculated, processed, and returned to the results of processing. Figure 1 shows the basic framework of granular computing.

In Figure 1, information granule is the basic element of granular computing, which is a set of elements aggregated according to indistinguishability, similarity, similarity or functionality. In granular computing, the main task is to construct, represent, and process information granules. Information granules can be divided into different sizes, that is, the "information granularity" of information granules. For example, people can use different time intervals to granulate information and deduce information granules of different sizes (year, month, day, etc.). The size of information granules deduced implies the level of "information granularity" used in information granulation. Therefore, information granules are hierarchical, and the information granules generated at different granularity levels can be transformed into each other, as shown in Figure 2.
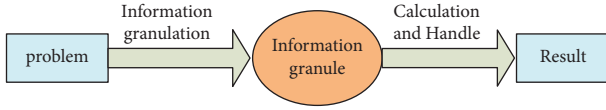
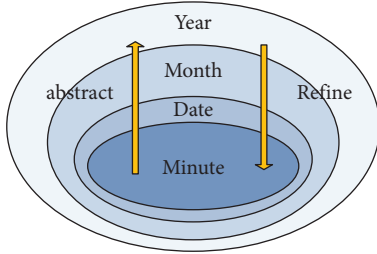FIGURE 1: Basic framework for granular computing.

FIGURE 2: Information granules and information granularity.

In the process of granulation, there are two directions: construction and decomposition. Construction refers to how finer or lower grains are merged into coarser or upper grains. Convert "thinner" and "more special" information granules into "coarser" and "more general" information granules; on the contrary, decompose the coarser or upper granules into finer lower granules. That is to say, "coarser" and "more general" information granules are further refined into "thinner" and "more special" information granules. At present, the specific methods of granulation include fuzzy information granulation [15], rough set approximation [16], quotient space method [17], clustering-based granulation [18] and cloud model method [19–21].

2.1.2. Basic Terms in Granular Computing

(1) Representation of Information Granules. Referring to the representation of topological space in quotient space, the information granules are formally described by triples, namely,

$$IG = (KVS, GM, VM), \tag{1}$$

where KVS (key value pair set) represents the feature subvector describing the information granule, i.e., $KVS = \{\langle key_1, value_1 >, \ldots, < key_n, value_n \rangle\}$ $(i = 1, 2, \ldots, n)$ and $Value_i$ represents the value taken by a feature named Key_i in the information granule.

GM (granularity measure) denotes the granularity measure of the information granule, i.e., the degree of granularity of the information granule. Its formula is [22]

$$GM(\pi) = \sum_{i=1}^{m} \frac{|X_i|}{|U|} \log|X_i| \tag{2}$$

Among them, $\pi = \{X_1, X_2, \ldots, X_m\}$ is a division of the universe U, and X_i is a subset of U. When the granularity is the finest, that is, each granule is a single point set, there is $GM(\pi) = 0$; when the granularity is the thickest, that is, the whole universe is a granule. $GM(\pi) = \log|U|$ VM (value measure) represents the value measure of the information

granule. Its determination is mainly from three aspects: granularity measurement, uncertainty, and domain knowledge.

(2) Representation of Granular Layer. The granular layer (layer) consists of all information granules based on a certain granulation criterion and the relationship between them. The granular layer can be formally expressed as a binary group, that is, a binary group.

$$Layer = (IGS, Intra - LR), \tag{3}$$

where IGS (information granule set) represents a set of information granules in the granular layer, which can be expressed as $IGS = \{IG_1, IG_2, \ldots, IG_M\}$; intra-LR (intra-layer relationships) represents possible relationships between information granules in the granular layer. If the information granule IG_p and IG_q is related to the existence, then $Intra - LR = \{E/E = (IG_p, IG_q), IG_p, IG_q \in IGS\}$.

(3) Representation of Granular Structure. The granular structure in MGrIKR is a topological structure consisting of multiple granular layers obtained by different granulation criteria, the relationship between information granules in different granular layers and the relationship between information granules in the same granular layer. Therefore, the granular structure GS (granular structure) can be expressed in tuple form:

$$GS = (LS, Inter - LR) \tag{4}$$

Among them, $LS = \{Layer_1, \ldots, Layer_{m-1}, Layer_m\}$ represents a set of m granular layers (Layer Set, LS), in which the granular layer $Layer_j$ is a granular layer in the granular structure. Inter-LR (interlayer relationships) denotes the set of transformation relations between two layers of $Layer_j$ and $Layer_k$ information granules, $Inter - LR$ which can be expressed as

$$Inter - LR = \{rp|rp(Layer_j, Layer_k)\} \cdot (j, k = 1, \ldots, m). \tag{5}$$

Among them, $r <$ denotes the satisfied partial order relationship between $Layer_j$ and $Layer_k$ information granules, which can be the relationship between information granules in adjacent two granular layers, or the relationship between information granules across layers.

2.2. Framework for Time Series Analysis, Processing, and Modeling Based on Granular Computing. Sequential data in the era of large data have the characteristics of huge scale, dynamic and instability. Therefore, researchers have proposed a processing framework for time series data based on granular computing to simulate human granulation cognitive mechanism as shown in Figure 3.

2.2.1. Rational Information Granulation of Time Series. In order to reduce the scale of time series data, it replaces the original time series in a simple and clear form. Firstly, it divides the time series into several time windows and then granulates the data on each time window to transform the
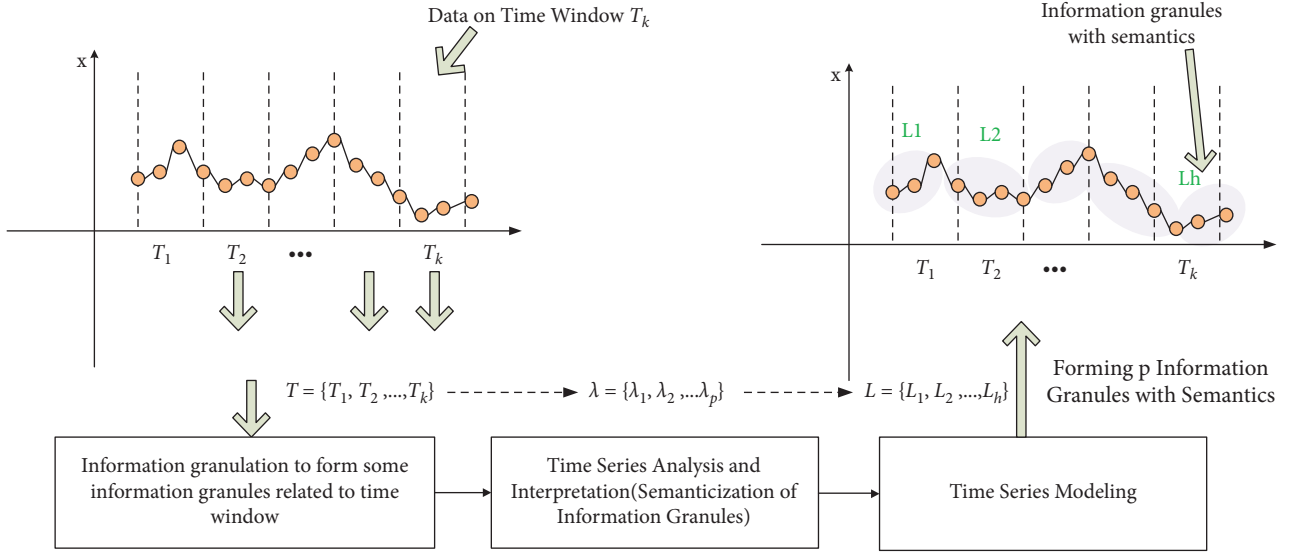
FIGURE 3: Data sequence processing framework based on granular computing.

original time series into a time-related "information granule" sequence (granular time series). Suppose for a time series $x = \{x_1, x_2, \ldots, x_n\}$, firstly, the time series is divided into a P-time window $\{T_1, T_2, \ldots, T_k\}$, and then the data on each time window $T_i (i = 1, \ldots, k)$ are granulated to form information granules $\lambda_i (i = 1, 2, \ldots p)$. Thus, the original time series X is transformed into corresponding granular time series $\lambda = \{\lambda_1, \lambda_2, \ldots \lambda_p\}$ by granulation of information.

2.2.2. Time Series Analysis and Interpretation Based on Information Granulation.
Time series analysis and interpretation takes "granular time series" as the research object. Semantic labels describing the dynamic characteristics of information granules are used as basic elements to describe each element in granular time series. It is assumed that there exists a set of granular semantic tags $L = \{L_1, L_2, \ldots, L_h\}$, in which $\leq p$. This stage is to realize how to describe granular information λ_i with elements in the set L of semantic tags. Researchers have used fuzzy sets to achieve this, that is, to assign corresponding semantics to information granules λ_i: "Yes, λ_i is L_j, its matching degree is ϕ_{ij}." Some researchers also use clustering algorithm to divide P information granules into granular semantic tag set L according to their similarity.

2.2.3. Time Series Modeling Based on Information Granulation.
Time series modeling is to transfer domain knowledge by means of information granules and capture the main relational features of time series by logical association language. Because the granular model of time series is user-oriented, the output of the model can be understood and recognized by users, thus helping users to make reasonable decisions. Its input and output are information granules on the corresponding time window of time series. It is assumed that time series granular model describes the dynamic relationship between information granules

$\lambda_j, \ldots \lambda_{i-2}, \lambda_{i-1}, \lambda_i$ formed by time series in time windows $T_j, \ldots T_{i-2}, T_{i-1}, T_i$.

2.3. Time Series Prediction Model Based on Multilayer Information Granules.
Considering the data presented by time series in time window, the distribution of time series is mostly uneven. The information granules constructed by single information granularity may not capture the essential characteristics of data. Therefore, in order to reasonably granulate the information of time series data, this paper designs an information granulation method based on the idea of "multi-granularity." Different information granularity levels are used to granulate the time series data on a time window to obtain information granularity at different information granularity levels. Finally, a time series prediction model based on multilevel information granules is proposed to predict ICU patient data. The basic flow of the prediction process is shown in Figure 4.

Specific steps are as follows: firstly, this paper establishes a multigranularity time series model by applying binary relationship on the time axis of time series data and introduces the evaluation effect function of time series data mining (the function of judging the accuracy of prediction) as the goal of prediction task. Based on this function, the optimal time granularity of prediction is found. In view of the uncertainty of time series, a multidimensional matrix model is constructed, and then the clustering category information of each attribute is obtained by clustering operation. Finally, the clustered data are classified by machine learning method to obtain classification knowledge. Forecasting can be achieved by classifying knowledge.

2.3.1. Establishment of Multigranularity Time Series Model.
In the research of time series prediction, the time domain U_T is generally considered as a line segment. Therefore, this paper uses the form of the time axis to represent the time domain and assumes that the starting point of the time
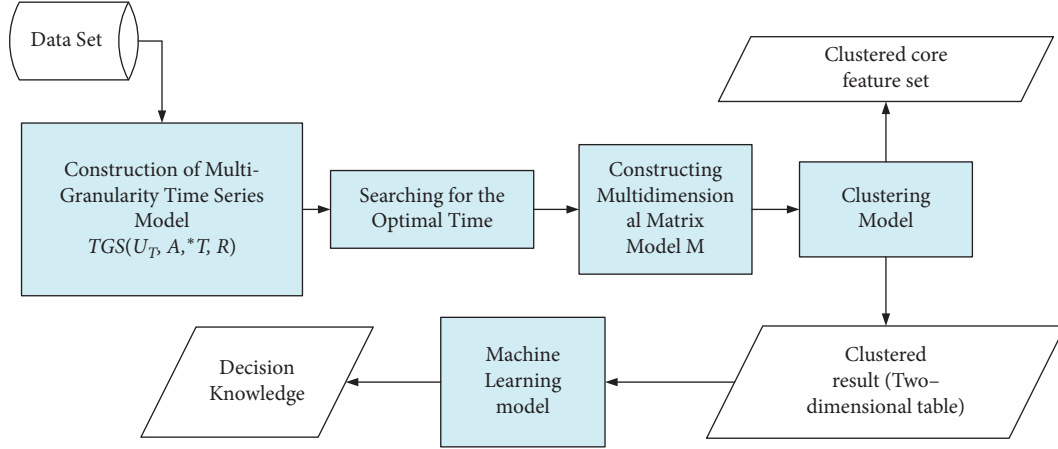
FIGURE 4: Forecasting process of the forecasting model.

domain is a origin of the time axis, which is recorded as t_0. The multigranularity time series $\sum((U_T, A_T), TS)$ is defined as TS with multiple optional time granularities (indicating that AT values can be more than one) in a given time domain U_T.

General analysis is based on the different attributes of objects in the universe to analyze their values and describe them in the form of two-dimensional tables. In this paper, we slice the multidimensional time series and analyze one of its atomic points. The time series are divided into $t_0, t_1, \ldots,$ T_n, and the analysis of tI time points is studied. From this, a multigranularity time series model $TGS(U_T, A, *T, R)$ of ICU is constructed, U_T which is a time domain, $A = A_N \cup A_T$, A_N is a set of related attributes; A_T is a set of time granularity attributes, and $*T$ is a set of time granularity $a_{tx} \in A_T$ attributes obtained by equivalence partition $[U_T]a_{tx}$ of R to U_T formed according to different time granularity A_T.

2.3.2. Finding the Optimal Time Granularity for Prediction.
Because the granularity is variable in constructing multigranularity time series model, the model can be used to predict different prediction effects. Therefore, in order to evaluate the performance of prediction based on different time grains in a given time domain, this paper uses the time series data mining evaluation function $F(a_{ti} \in A_T)$ to try to find the best time granularity in the model. The process of solving the optimal time granularity is shown in Figure 5.

There are two stages:

(i) Optimal time granular search based on time domain.

Assuming the time domain U_T, the alternative I is based on the equivalence partition $I = \{k_1, k_2, k_3\}$ of time granules, the evaluation effect function $F(a_{ti} \in A_T)$ of time series data mining, the diagnostic accuracy λ and the minimum time granularity a_{min}. Firstly, according to the evaluation function, a partition is selected from the candidate I partition to get the granularity a_{ti} as the next time granularity set TG_i.
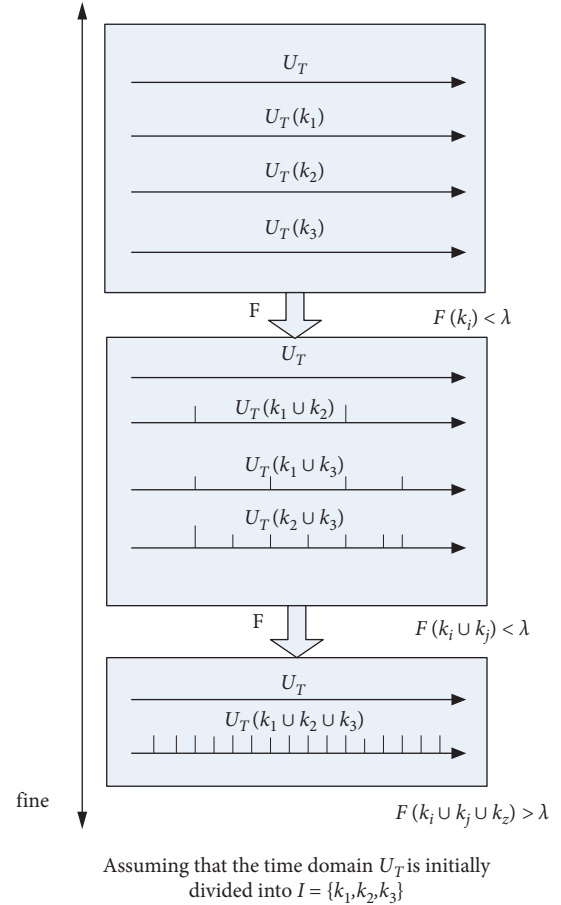

Assuming that the time domain U_T is initially divided into $I = \{k_1, k_2, k_3\}$

FIGURE 5: The process of finding the optimal time granularity.

(iii) Optimal time granular fine search based on time domain.

The prediction results under each TG_i condition are evaluated in turn, and the algorithm ends if the accuracy is less than λ or less than the prescribed minimum time granularity a_{min}. Otherwise, according to the current TG_i time granule, the best time granule expression sequence is selected to further classify the time granule under the current

granularity into I categories, and the results are calculated according to the precursor-successor relationship of the time granule. Finally, the optimal time granularity and its time granularity are recorded as OPT.

2.3.3. Constructing Multidimensional Matrix Model of Indicator Attributes.

Time series data include data of many different indicators, which have been collected throughout the time domain. These data have high latitude and time series. In order to obtain an effective data source, it is necessary to sample the original time series data. TS, which is commonly used in practice, is also a limited set. It is expressed as $TS = \{(t_1, v_1), (t_2, v_2), \ldots\}$, where t_i denotes sampling time, v_i denotes sampling value, and there is a certain functional relationship between T_i and v_i, $v_i = f(t_i)$. From a large number of time series data records, it is found that these records are some interval sampling uncertain time series. Therefore, in order to transform it into deterministic time series, this paper intends to use the blank strategy and data interpolation to determine the uncertain time series, respectively. Then, the matrix model is used to describe the time series data of time series data indicators. Matrix M, which consists of time series corresponding to all indexes of time series $(TS)_{d_j}^{a^i}$ data, is defined as a multidimensional matrix model of indexes as follows:

$$M = \begin{bmatrix} (TS)_{d_1}^{a_1} & (TS)_{d_1}^{a_2} & \ldots & (TS)_{d_1}^{a_n} \\ (TS)_{d_2}^{a_1} & (TS)_{d_2}^{a_2} & \ldots & (TS)_{d_2}^{a_n} \\ \ldots & \ldots & \ldots & \ldots \\ (TS)_{d_m}^{a_1} & (TS)_{d_m}^{a_2} & \ldots & (TS)_{d_m}^{a_n} \end{bmatrix}. \tag{6}$$

Among them, $(TS)_{d_j}^{a^i} = \,< f(t_1), f(t_2), \ldots f(t_h) >_{d_j}^{a^i}$ the first index representing the time series of Article J has time series. M_{ij} represents the time series corresponding to the jth index of the first time series. Meeting conditions $\exists rst. f(t_r)_{d_i}^{a_i} = \varphi \cap f(t_r)_{d_i}^{a_j} or f(t_r)_{d_i}^{a_i} \neq \varphi \cap f(t_r)_{d_i}^{a_j} = \varphi$.

It can also be seen from formula (1) that the index multimatrix model is a set of super-high dimensional time series.

2.3.4. Classification by Clustering.

The index multidimensional matrix model is clustered to obtain the clustering core sequence based on the index data and the two-dimensional table corresponding to the physiological index and the clustering result. The specific steps are as follows:

Input: Data training set DS, initial clustering center $H = \{h_1, h_2, \ldots, h_{|h|}\}$, similarity function $F(x, y)$. Convergence coefficient is θ.

Step 1: calculate the distance between X and y of any time series according to $F(x, y)$. Where $F(x, y)$ is expressed as

$$F(x, y) = \frac{\sum_{i=1}^{48|lw|} \sqrt{\left(d_{x\_i\_\max} - d_{y\_i\_\max}\right)^2 + \left(d_{x\_i\_\min} - d_{y\_i\_\min}\right)^2}}{\sum_{i=1}^{48|lw|} \sqrt{\left(\left(d_{x\_i\_\min} + d_{y\_i\_\min}\right)/2\right)^2 + \left(\left(d_{x\_i\_\min} - d_{y\_i\_\min}\right)/2\right)^2}}. \tag{7}$$

$d_{x\_i\_\max}$ It represents the maximum value in the i time granule of sequence x and $d_{x\_i\_\min}$ represents the minimum value in the i time granule of sequence X.

Step 2: update the clustering center. The newly generated K clustering centers are not more than θ according to $F(x, y)$. Then output the clustering result of DS, otherwise continue iterating.

In this paper, K-means clustering algorithm and x-means clustering algorithm are used.

(i) K-means clustering algorithm

The idea of K-means algorithm is simple. For a given sample set, the sample set is divided into K clusters according to the distance between samples. Let the points in the cluster be as close as possible, and let the distance between the clusters be as large as

possible. Mathematical expressions are used to minimize the mean variance E if the cluster is divided into $(C_1, C_2, \ldots, C_k)$.

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} ||x - \mu_i||_2^2, \tag{8}$$

where μ_i is the mean vector of C_i, the expression is

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x. \tag{9}$$

(ii) X-means algorithm

X-means algorithm is an improvement of K-means algorithm. The main aspects of improvement are as follows:

(a) Using kd-tree to accelerate each iteration of the original K-means

(b) Users specify the scope of K and select the optimal K according to BIC score

(c) Each iteration only takes 2-means.

X-means algorithm steps: user input K_{min}, K_{max}, data set D.

(1) Running K_{min}-means.

(2) Run 2-means on each cluster.

(3) According to BIC score (calculated only on this cluster, i.e. the BIC score when only the clustering data are divided into one and two categories), it is decided whether or not to divide the clustering into two categories.

(4) If $K < K_{max}$, proceed to Step 2, otherwise the result will be returned.

(5) Clustering data classification using machine learning method.

The two-dimensional information table of clustering results is used to obtain classification knowledge by corresponding machine learning method. Then, the prediction is realized by matching the learned classification knowledge in the test set data. In this paper, SVM algorithm, ID3 algorithm and PNN (probability neural networks) algorithm are used for classification.

(iii) SVM algorithm

Support vector machine (SVM) is a classification algorithm. It can improve the generalization ability of learning machine by seeking the smallest structured risk, and minimize the empirical risk and confidence range, so as to achieve the goal of obtaining good statistical rules in the case of fewer statistical samples. Generally speaking, it is a two-class classification model. Its basic model is defined as a linear classifier with the largest spacing in the feature space. That is to say, the learning strategy of support vector machine is to maximize the spacing, which can eventually be transformed into a solution of a convex quadratic programming problem.

(iv) ID3 algorithm

The core of ID3 algorithm is that when selecting attributes at all levels of decision tree nodes, information gain is used as the criterion to select attributes, so that when testing at each nonleaf node, the maximum category information about the tested record can be obtained. The specific method is to detect all attributes, select the attributes with the greatest information gain to generate decision tree nodes, establish branches from different values of the attributes, and then recursively call the method to establish branches of decision tree nodes for each subset until all subsets contain only the same category of data. Finally, a decision tree is obtained, which can be used to classify new samples.

(v) PNN algorithm

PNN is a simple and widely used neural network. It is a radial neural network. On the basis of RBF network, PNN combines density function estimation and Bayesian decision theory. It consists of input layer, hidden layer, summation layer and output layer. The input layer receives the values from the training samples and transfers the data to the hidden layer. Implicit layer mainly accepts input samples from input layer, calculates the distance between input vector and center, and finally returns a scalar value. Vectors are input into the hidden layer. The input/output relationship determined by the j neuron of type I mode in the hidden layer is defined as follows:

$$\varphi_{ij}(x) = \frac{1}{(2\pi)^{1/2}\sigma^d}e^{\left(x - x_{ij}\right)\left(x - x_{ij}\right)^T/\sigma^2}. \tag{10}$$

$I = 1, 2, \ldots, M$, M is the total number of classes in the training sample. d is the dimension of the sample space data and x_{ij} is the jth center of the class i sample. The summation layer weights and averages the output of the hidden neurons belonging to the same class in the hidden layer:

$$v_i = \frac{\sum_{j=1}^{L} \varphi_{ij}}{L}. \tag{11}$$

v_i represents the output of Class i, and L represents the number of neurons in Class i. The number of neurons in summation layer is the same as that in category M.

Output layer selection and the largest class in the layer as output category $y = \arg\max(v_i)$.

3. Experiment

3.1. Experimental Data. In this paper, the data provided by the PhysicoNet website of the National Institutes of Health are used. The data included 575 ICU data from Coronary Care, 871 ICU data from Cardiac Surgery, 1482 ICU data from Recovery Unit and 4000 ICU data from Medical and Surgical. Various physiological indicators were used in all the data. Different physiological indicators also have a certain percentage of missing values in these case records. Of these cases, 13.86% died and 86.14% survived. Because of the unbalanced distribution of data sets, the focus of this classifier is whether death samples can be correctly identified.

3.2. Physiological Indicators and Setting of Experimental Parameters. According to the frequency and importance of physiological indicators in different ICUs, and the short average sampling interval of these indicators. Ten candidate physiological attributes (HR, urine, WBC, temp, PH, PaO_2, $PaCO_2$, NIDias ABP, dias ABP, and sys ABP) were selected from many indicators. In this paper, the data ratio of training set and test set is set to 7 : 3. The minimum time granularity is set to 0.25 hours, and the minimum prediction accuracy is 75%.
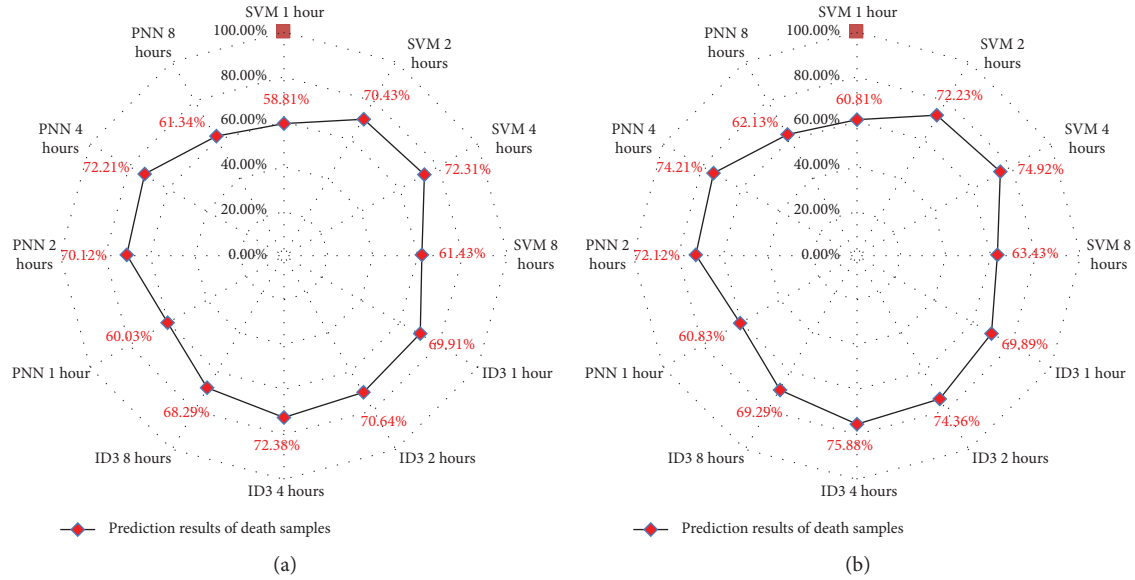
FIGURE 6: Prediction results of schemes 1 and 2. (a) Comparison of the prediction effect of scheme 1. (b) Comparison of the prediction effect of scheme 2.

3.3. Experimental Scheme. In the early stage of this paper, the missing data is filled by simple blank interpolation. After determining the time series of ICU data, two different clustering methods (K-means algorithm and GMM maximum expectation EM algorithm) are used to cluster the sequence. Finally, different machine methods (SVM algorithm, ID3, and PNN) are used to classify and predict the sequence. Therefore, three experimental schemes are adopted in this paper.

Scheme 1: clustering algorithm uses K-means algorithm, and then uses three different machine classification algorithms, using different time granularity for prediction.

Scheme 2: clustering algorithm adopts EM algorithm, and then three different machine classification algorithms are used to predict with different time granularity.

Scheme 3: clustering algorithm adopts EM algorithm, then three different machine classification algorithms are used, respectively, and the optimal time granularity is used to predict.

4. Result Analysis

4.1. Analysis of Experimental Results of Schemes 1 and 2. The experimental results of schemes 1 and 2 are shown in Figure 6.

In Figure 6, three different machine classification methods (SVM, ID3, PNN) are used to predict the results at four different time granularities (1 hour, 2 hours, 4 hours, 8 hours). As can be seen from Figure 6, both scheme 1 and scheme 2 have the best prediction effect when using ID3 algorithm, and PNN algorithm has the worst prediction effect. When using the same machine learning method, different time granularity is selected, and the prediction effect is also quite different. When choosing the time

granularity of 4 hours, the prediction effect of different machine learning methods is generally the best. The prediction rates of SVM method, ID3 method and PNN method in scheme 1 were 72.31%, 72.38% and 72.21%, respectively. The prediction rates of SVM method, ID3 method and PNN method in scheme 2 were 74.92%, 75.88% and 74.21%, respectively. The comparison of the forecasting effect between scheme 1 and scheme 2 is shown in Figure 7.

As can be seen from Figure 7, the blue line represents the prediction effect of Scheme 2 and the red line represents the prediction effect of Scheme 1. In scheme 2, the predictive effect of all machine learning methods has increased in general. It is shown that the clustering algorithm EM can not only reduce the dimension, but also improve the distinguishability of various physiological index attributes for life and death results by comparing K-means algorithm.

4.2. Experimental Results of Scheme 3. In experiment 3, the given time domain was assumed to be January $(30 * 24$ hours). For the initial equivalence of the time domain, it is divided into $I = \{1, 2, 4\}$, where the meaning of 1, 2, 4 represents the equivalence of 1, 2 and 4 for a given time domain. Only the predictive effect of death samples is considered here. The threshold of prediction accuracy of death samples is 75% and the minimum time granularity is 0.25 hours. Then the initial equivalence partition is refined according to the value of mining evaluation effect function. The experimental results of Experiment 3 are shown in Table 1 and Figure 8.

As can be seen from Table 1, when the time granularity is equal to 5.625 hours, i.e. 128 equal parts in the time domain, the value of the evaluation function in the prediction model is more than 75% regardless of the machine classification method used. Therefore, 5.625 hours is the optimal time granule. From Figure 8, we can use the optimal time granule
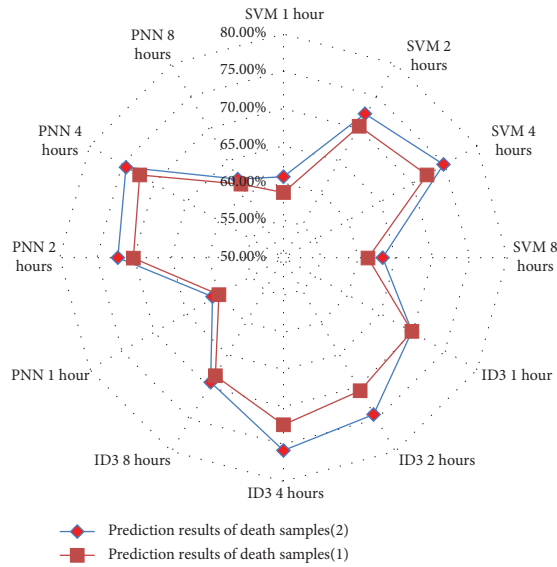
FIGURE 7: Comparison of the predicted results of schemes 1 and 2.

TABLE 1: Predicted results of Scheme 3.

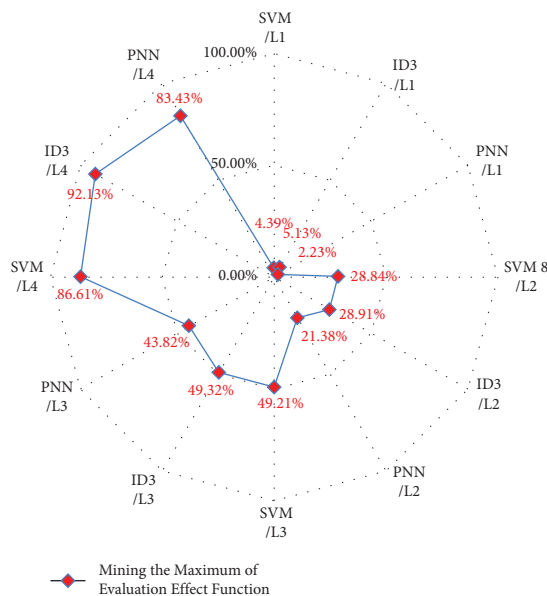| Equivalent partition set	Machine method	Mining the maximum value of evaluation effect function (%)	Function maximum corresponds to time granule
L1 = {1, 2, 4}	SVM	4.39	
	ID3	5.13	180 hours
	PNN	2.23	
L2 = {1, 2, 4, 8}	SVM	28.84	
	ID3	28.91	90 hours
	PNN	21.38	
L3 = {1, 2, 4, 8, 16, 32}	SVM	49.21	
	ID3	49.32	45 hours
	PNN	43.82	
L4 = {1, 2, 4, 8, 16, 32, 64, 128, 256}	SVM	86.61	
	ID3	92.13	5.625 hours
	PNN	83.43	

FIGURE 8: Comparisons of prediction results of different machine classifications with optimal time granule.

TABLE 2: Predictive effect of different mining algorithms for traditional time series analysis.

| Fourier change				Wavelet transform			
Data set source/number (example)	SVM	ID3	PNN	Data set source/number (example)	SVM	ID3	PNN
Coronary care/575	54.83	57.23	54.23	Coronary care/575	60.83	64.23	54.23
Cardiac surgery/871	58.23	63.87	53.22	Cardiac surgery/871	62.23	63.87	62.22
Recovery unit/1482	67.25	68.26	66.53	Recovery unit/1482	61.25	62.26	59.53
Medical and surgical/4000	54.87	56.91	52.19	Medical and surgical/4000	69.87	73.21	65.13
Average value	58.795	61.5675	56.5425	Average value	63.545	65.8925	60.2775
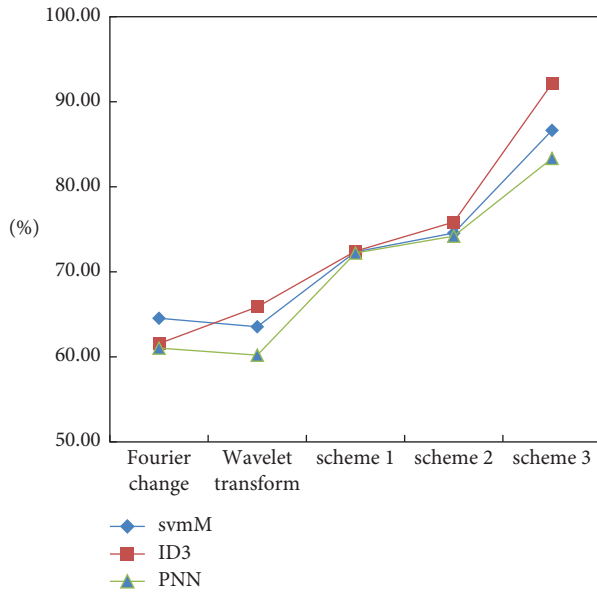
FIGURE 9: Comparison of the predictive effect between the traditional time series analysis modeling method and the proposed method.

prediction algorithm, which basically reaches about 85%. The highest prediction algorithm is ID3 algorithm, which has a prediction rate of 92.13%. It is obvious that the prediction effect of scheme 1 is better than that of scheme 1 and scheme 2. Secondly, as can be seen from Figure 8, the speed of finding the best time granule based on the time domain proposed in this paper is also faster. Basically, the best time granule was found after four cycles.

4.3. Comparisons of Prediction Results between Traditional Time Series Analysis Processing Algorithms and Granular Computing. In order to judge the validity of the proposed method, this paper compares the predictive effect of the traditional time series analysis on the same data set. Fourier transform is used here. Firstly, the original time series is filtered and reduced by discrete Fourier transform and wavelet transform, respectively. Then the data mining algorithm is used to mine the preprocessed time series. The data mining algorithms used here are SVM, ID3, and PNN. The data contains 575 ICU data from Coronary Care and 871 ICU data from Cardiac Surgery. There were 1482 ICU data from Recovery Unit and 4000 ICU data from Medical and Surgical. The prediction results of traditional time series analysis are shown in Table 2.

From Table 2, we can see that the effect of using Fourier transform or wavelet transform to analyze time series data is basically below 65%. Their prediction results are compared with those of the first three schemes as shown in Figure 9. It is obvious that the prediction effect of time series data based on multilayer information granule construction model is better than that of Fourier transform and wavelet transform.

5. Conclusion

With the rapid development of network, the era of big data has come. Among them, time series data is very common in various fields. This kind of data mining and knowledge discovery is of great significance. Time series data is different from ordinary data. In addition to time-related, it also has high-dimensional, uncertain and dynamic characteristics. These characteristics lead to the use of traditional mining methods for data mining is not very good. Therefore, this paper uses granular computing to deal with it according to its characteristics, and uses ICU data to predict the life and death of ICU patients to test the predictive effect of granular computing on time series data. The main work of this paper is as follows:

(1) In this paper, a multigranularity time series model is established by applying binary relation on the time axis of time series data, and an evaluation effect function of time series data mining is introduced for the purpose of forecasting tasks. Based on this function, the optimal time granularity of prediction is found. According to the uncertainties of time series, an index multidimensional matrix model is constructed, and then the clustering category information of each attribute is obtained by clustering operation. Finally, the clustered data are classified by machine learning method to obtain classification knowledge. Prediction is achieved by classifying knowledge.

(2) The experimental results show that EM clustering algorithm is better than K-means in improving the discriminability of all kinds of physiological indicators for life and death results; the optimal time granularity can improve the prediction rate; ID3 decision-making algorithm is better than SVM and PNN algorithm in data mining for time series data. In addition, the prediction of time series data based on the multilayer information granule construction model is better than the traditional time series analysis method.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declare no potential conflicts of interest with respect to the research, author-ship, and/or publication of this article.

Acknowledgments

References

[1] G. E. P. Box and G. M. Jenkins, "Time series analysis: forecasting and control," *Journal of Time*, vol. 31, no. 4, p. 303, 2010.

[2] M. Khashei and M. Bijari, "An artificial neural network model for timeseries forecasting," *Expert Systems with Applications*, vol. 37, no. 1, pp. 479–489, 2010.

[3] V. Petridis, A. Kehagias, L. Petrou et al., "A bayesian multiple models combination method for time series prediction," *Journal of Intelligent and Robotic Systems*, vol. 31, no. 1-3, pp. 69–89, 2001.

[4] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets and Systems*, vol. 54, no. 3, pp. 269–277, 1993.

[5] J. Gehring, M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin, *Convolutional Sequence to Sequence Learning*, pp. 1243–1252, Computation and Language, Sydney, Australia, 2017.

[6] Y. Yang and D. Wan, "Time series prediction model based on deep cycle neural network," *Computer technology and development*, vol. 27, no. 3, pp. 35–38, 2017.

[7] C. Bettini, X. S. Wang, S. Jajodia, and J.-L. Lin, "Discovering frequent event patterns with multiple granularities in time sequences," *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, no. 2, pp. 222–237, 1998.

[8] J. Ruan, X. Wang, and Y. Shi, "Developing fast predictors for large-scale time series using fuzzy granular support vector machines," *Applied Soft Computing*, vol. 13, no. 9, pp. 3981–4000, 2013.

[9] O. Hryniewicz and K. Kaczmarek, "Bayesian analysis of time series using granular computing approach," *Applied Soft Computing*, vol. 47, Article ID S1568494614005869, 2014.

[10] M.-Y. Chen and B.-T. Chen, "A hybrid fuzzy time series model based on granular computing for stock price forecasting," *Information Sciences*, vol. 294, no. 2, pp. 227–241, 2015.

[11] R. Al-Hmouz, W. Pedrycz, and A. Balamash, "Description and prediction of time series: a general framework of Granular Computing," *Expert Systems with Applications*, vol. 42, no. 10, pp. 4830–4839, 2015.

[12] W. Lu, W. Pedrycz, X. Liu, J. Yang, and P. Li, "The modeling of time series based on fuzzy information granules," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3799–3808, 2014.

[13] W. Wang, W. Pedrycz, and X. Liu, "Time series long-term forecasting model based on information granules and fuzzy clustering," *Engineering Applications of Artificial Intelligence*, vol. 41, no. C, pp. 17–24, 2015.

[14] S. Vairavan, L. Eshelman, S. Haider, and A. Flower, "Prediction of mortality in an intensive care unit using logistic regression and a hidden Markov model," *Computers in Cardiology*, vol. 39, pp. 393–396, 2012.

[15] 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. 90, pp. 111–127, 1997.

[16] Y. Y. Yao, "Information granulation and rough set approximation," *International Journal of Intelligent Systems*, vol. 16, no. 1, pp. 87–104, 2001.

[17] Y. P. Zhang, L. Zhang, and T. Wu, "The representation of different granular worlds: a quotient space," *Chinese Journal of Computers*, vol. 27, no. 3, pp. 328–333, 2004.

[18] Y.-H. Qian, "Granulation mechanism and data modeling for complex data," Ph.D.Dissertation in Chinese, School of Computer and Information Techonlogy, Shanxi University, Taiyuan, 2011.

[19] M. A. Hongyao, W. Guoyin, Z. Qinghua, and X. U. Ning, "Multi-granularity color image segmentation based on cloud model," *Computer Engineering*, vol. 38, no. 20, pp. 184–187, 2012.

[20] J. Gao and Z. Huang, "Image segmentation based on cloud model," *Journal of Heilongjiang Bayi Agricultural Reclamation University*, pp. 88–91, 2015.

[21] Y. Yiyu, "Notes on rough set approximations and associated measures," *Journal of Zhejiang Ocean University*, vol. 29, no. 5, pp. 399–410, 2010.

[22] J. Xu, Y. Zhang, L. D. ZhouDuanning, and L. Yu, "Modeling and clustering of uncertain time series based on granular computing," *Journal of Nanjing University (Natural Science)*, vol. 50, no. 1, pp. 87–94, 2014.