INT J COMPUT COMMUN, ISSN 1841-9836
9(3):313-324, June, 2014.

Numerical Prediction of Time Series Based on FCMs with Information Granules

W. Lu, J. Yang, X. Liu

Wei Lu*, Jianhua Yang, Xiaodong Liu
School of Control Science and Engineering
Dalian University of Technology
Dalian 116023, China
*Corresponding author: luwei@dlut.edu.cn

Abstract: The prediction of time series has been widely applied to many fields such as enrollments, stocks, weather and so on. In this paper, a new prediction method based on fuzzy cognitive map with information granules is proposed, in which fuzzy c-means clustering algorithm is used to automatically abstract information granules and transform the original time series into granular time series, and subsequently fuzzy cognitive map is used to describe these granular time series and perform prediction. two benchmark time series are used to validate feasibility and effectiveness of proposed method. The experimental results show that the proposed prediction method can reach better prediction accuracy. Additionally, the proposed method is also able to precess the modeling and prediction of large-scale time series.
Keywords: Fuzzy Cognitive Maps (FCMs), time series, prediction, , information granules.

1 Introduction

Time series is a sequence of real-data, with each element in this sequence representing a value recorded at some time moment. As a classic issue, time series prediction has been used in diverse fields, which utilizes prediction model describing some useful temporal relationship that is developed by observing a past certain variable or a past family of variables to extrapolate future values. How to construct prediction model of time series is core for prediction of time series. Many researcher early focus on how to construct predictive model of time series with the aid of the linear system theory [1], the stochastic process theory [2] and the black-box methodology [3], and dynamical system analysis [4]. However, the constructed prediction model by using these methods cannot solve prediction problem in which the historical data are missing or uncertain.

Fuzzy sets theory can be used to make semantics and represent the data themselves and fuzzy reasoning offers a viable alternative to ensure robustness to the inherent uncertainty, which has involved into modeling and prediction of time series. Song and Chissom [5–7] defined the concept of fuzzy time series and developed two fuzzy time series prediction models — the time-invariant model [6] and the time-variant model [7]. Following the work of Song and Chissom, many the improved prediction models associated with fuzzy time series are emerged such as the Markov model [8], Chens model [9], Hwangs model [10], the heuristic model [11], the high order model [12], the local trend model [13] and so on. These models have been used to predict enrollment of university, stock index, temperature etc., which also shows better prediction performance. As data, time series is inherently associated with large size, high dimensionality and a stream-like nature. Whereas construction of the fuzzy logic relationships which has important impact on performance of models is a tedious work in the development process of existing fuzzy time series models, which is difficult and complicated for modeling of the large-scale fuzzy time series. To overcome the deficiency, fuzzy cognitive map (FCM for short) with information granules seem to can become an alternative.

Fuzzy cognitive map as a soft computing technology for modeling complex systems was proposed by Kosko in 1986 [14], which are treated as an alternative way for knowledge-based representation and inference process for its easy of usage and numeric matrix operation for complex system. Based on the nature, FCM has capable to handle prediction problem of time series [15, 16]. Further, information granules and information granulation [17, 25] play key roles when dealing with large data. As expression of meaningful the abstract entities, information granules come with information granulation. The role of information granulation is to organize detailed numerical data into some meaningful, semantically sound entities (information granules). In particular, granulate information are able to achieve a high level of interpretability and manage phenomena which are complex and consequently the data are overwhelming. The formalism of information granules includes intervals, fuzzy sets, rough sets, shadow sets or alike. There are many methods supporting design of information granules, one of which is fuzzy c-means clustering. It can provide an ability abstracting fuzzy information granules from data with multivariate attribute.

In this paper, starting with global view of information granules, we attempt to construct fuzzy cognitive map with information granules to realize prediction of time series. Our proposed prediction method of time series includes two stage: the first stage is to construct fuzzy cognitive map prediction model on basis of information granules, and the second stage is to perform prediction by exploiting the fuzzy cognitive map with information granules based model which is constructed by the first stage.

The remainder of this paper is organized as follows. Section 2 briefly introduces some concepts related to FCM, and then focus on learning method of FCM. Section 3 presents a method of constructing FCM with information granules by using fuzzy c-means clustering algorithm.The proposed prediction method of time series based on FCM with information granules is detailed in section 4. In section 5, two benchmark time series are used to validate the feasibility and effectiveness of proposed method, and experimental results is also discussed. Finally, some conclusions is provided in section 6.

2 Fuzzy cognitive map and its learning method

In this section, some concept associated with fuzzy cognitive map is first recalled, and then the learning method based on particle swarms optimization (PSO) technology, an effective learning method of FCM, is presented to learn weights of FCM.

2.1 Fuzzy cognitive map

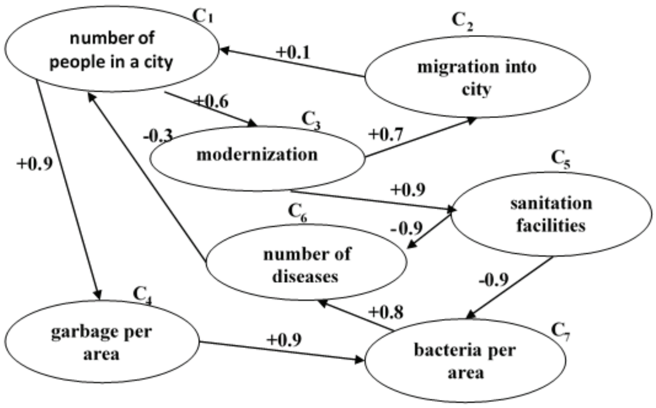
FCM is simple, very powerful tool for representation of human knowledge and performing reasoning. FCM can describe a given system by concepts and mutual relationships among them, which play an important role for time series modeling and prediction in this paper.

FCM is composed of a collection of nodes and directed links (edges) between nodes. In the FCM, nodes is used to represent concepts, say $C_1, C_2, \cdots, C_n$. These concepts can be envisioned as status, variables etc. which is used to describe main dynamic characteristic of problem/process/system. Values of nodes (concepts) are fuzzy and change with time. The directed edges represent casuality between nodes or more precisely a way in which one node affects another one. The connections between nodes could be asymmetric. The strength of connections (it is also called weights of FCM) from node C_j to node C_i, denoted by ω_{ij}, is quantified to be range from -1 to 1. The value of ω_{ij} reflects different casuality between C_i and C_j, viz.,

- $\omega_{ij} > 0$, which indicates positive causality between C_j and C_i, i.e., an increase of value of C_j leads to an increase of value of C_i (and vice versa).

- $\omega_{ij} = 0$, which indicates neutral causality between C_j and C_i, i.e., no relationship between C_j and C_i. In this case, the connection from C_i to C_j can be removed.
- $\omega_{ij} < 0$, which indicates negative causality between C_j and C_i, i.e., an increase of value of C_j leads to an decrease of value of C_i (and vice versa).

The fuzzy cognitive map can be represented as not only the directed graphic but also a square matrix. The square matrix, which is also called relationship matrix, stores all values of weights of FCM. Fig.1 shows an example of FCM model and its relationship matrix that concerns public city health issues.



(a) FCM model of public city health issues

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
C_1	0	0	0.6	0.9	0	0	0
C_2	0.1	0	0	0	0	0	0
C_3	0	0.7	0	0	0.9	0	0
C_4	0	0	0	0	0	0	0.9
C_5	0	0	0	0	0	0.9	0.9
C_6	-0.3	0	0	0	0	0	0
C_7	0	0	0	0	0	0.8	0

(b) The relationship matrix of FCM

Figure 1: FCM model of public city health issues and its relationship matrix.

Behind FCM, there is mathematical mechanism which is described in forms of (1):

$$C_i(t+1) = f\Big(\sum_{j=1}^{n} \omega_{ij} C_j(t) + \omega_{0i}\Big) \tag{1}$$

where $C_j(t)$ is the the active level (value) of jth node at the tth time moment, $\omega_{ij} \in [1, 1]$ is the value of weight from the concept C_j to the concept C_i and $\omega_{0i} \in [0, 1]$ is the bias associated with the ith node. Besides, n is the number of nodes of FCM, and f is the transformation function that is generally selected as sigmoid function with steepness parameter σ — $f(u) = 1/(1 + e^{-\sigma u}))(u \in R, \sigma > 0)$, where the steepness parameter σ is associated with the individual node of the FCM. The role of this parameter is to provide some additional calibration of the value of the node [16].

Let $\mathbf{w}_i = [\omega_{i1}\ \omega_{i2}\ \cdots\ \omega_{in}]^T$ be the weight vector which includes all values of weights from all other nodes to ith node. Likewise the vector of active level of nodes is described as $\mathbf{C} = [C_1\ C_2\ \cdots\ C_n]^T$, which includes all values of nodes of FCM. Thus we can rewrite (1) to obtain a more concise vector notation as follows.

$$C_i(t+1) = f(\mathbf{w}_i^T \mathbf{C}(t) + \omega_{0i}) \tag{2}$$

Once nodes of FCM and weights between these nodes are determined, FCM starts with a given initial state vector $\mathbf{C(t)} = [C_1(t), C_2(t), \cdots, C_n(t)]^T$ to perform iteration computation according to (1). After finite iteration, FCM can reach an equilibrium point or a limit cycle. It is worthy noting that at $t+1$ time moment the active level of ith node depends on all the value of node of FCM at t time moment.

2.2 The learning of fuzzy cognitive map

The sound values of weights can ensure FCM accurately describing dynamic behavior of system to research. Several methods such as the Hebbian learning [14] on basis of the scheme of unsupervised learning, the numeric data-based PSO algorithm [16] and evolutionary optimization [19] have been proposed. As an optimization method, PSO algorithm offers possibilities of global, population-based optimization yet not imposing a very heavy computational overload. The method allows for determining parameters of FCM from original time series data without human input. In this paper, we take PSO algorithm to learn all parameters of FCM.

In brief, PSO involves a population of particles whose dynamics is guided by the mechanisms of social interactions and personal experience. The details of PSO and its enhancement version are documented in the literature [20] and [21].

The objective of utilizing PSO algorithm to learn parameters of FCM is to develop candidate FCM and enable it to mimic the given input data. This optimization problem requires to establish $n \times (n+2)$ parameters. Consequently, the particles structure is defined as

$$\begin{aligned}\hat{\mathbf{W}} = \ [\ & \omega_{11}, \omega_{12}, \cdots, \omega_{1n}, \omega_{21}, \omega_{22}, \cdots, \omega_{2n}, \cdots, \omega_{n1}, \omega_{n2}, \cdots, \omega_{nn}, \\ & \omega_{01}, \omega_{02}, \cdots, \omega_{0n}, \sigma_1, \sigma_2, \cdots, \sigma_n] \end{aligned} \tag{3}$$

where $w_{ij} \in [1, 1]$ is the weight from node j to node it, $w_{0i} \in [0, 1]$ is the bias associated with the ith node, and $\sigma_i > 0$ is the steepness parameters of transformation function f.

Objective function (4) is used to evaluate quality of particles in population, and is defined by taking advantage of an inherent property of FCM model.

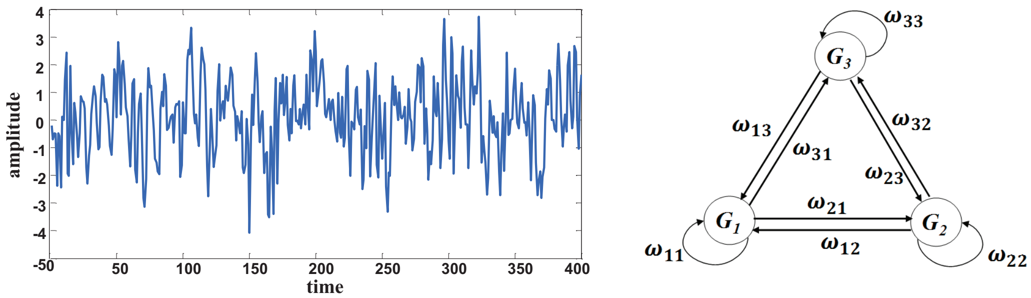
$$min : f = \frac{1}{(l-1)\,n} \sum_{t=1}^{l-1} \sum_{i=1}^{n} \parallel \hat{C}_i(t+1) - C_i(t+1) \parallel^2 \tag{4}$$

where $\mathbf{C}(t+1) = [C_1(t+1)\ C_2(t+1)\ \cdots\ C_n(t+1)]^T$ is a actual response for initial state vector $\mathbf{C(t)} = [C_1(t)\ C_2(t)\ \cdots\ C_n(t)]^T$, $\hat{\mathbf{C}}(t) = [\hat{C}_1(t+1)\ \hat{C}_2(t+1)\ \cdots\ \hat{C}_n(t+1)]$ is a response of the condidate FCM for initial state vector $\mathbf{C(t)}$, l is a number of input data points, n is a number of nodes of FCM. The objection function can realize comparison between the single-step response of the candidate FCM and the actual response for the same initial state vector.

3 The design of fuzzy cognitive map with information granules

One fundamental problem when FCM is used in time series modeling is how to design structure of FCM for a given time series, i.e., how to map the given time series onto the structure of FCM, viz. expressing a meaning of the nodes of the graph and specifying possible causal linkages between pairs of nodes. In general, if the concepts of system to research are easily identifiable, these concepts can be directly cast into FCM. However, the clearly and understandable concepts are hardly obtained from time series including a series of numeric data, which need to be discovered or mined. Here the idea of information granulation is adopted. On a basis of numeric data formed are information granules and they can be treated as nodes of FCM. Fuzzy c-means clustering [22] can serve as a convenient vehicle to construct information granules from data. In what follows, an illustrative example is presented how to design structure of FCM for time series.

Let us consider a certain time series as shown in Fig.2 (a). The numeric data included in the time series is clustered by using the standard version of fuzzy c-means clustering algorithm with Euclid distance, where the number of clusters c is set into 3 and the fuzzification coefficient m is set into 2 as well. The results of clustering is reported in Table 1 in where G_1, G_2, G_3



(a) The original time series plot.

(b) The expression of FCM with information granules

Figure 2: FCM expression of a certain time series.

are prototypes. These prototypes can be viewed as information granulations represented by the following semantics:

- G_1 — the amplitude of time series is negative high where locates nearby -1.9019.
- G_2 — the amplitude of time series is zero where locates nearby 0.
- G_3 — the amplitude of time series is positive high where locates nearby 1.7832.

From perspective of information granules, the dynamic behaviour of timer series can be described through establishing relationships among these granules.

Table 1: The prototypes and granular description of time series showed in Fig.2 (a)

Cluster No.	1	2	3
prototypes	-1.9019	0.025	1.7832
Description of prototypes information granules	G_1	G_2	G_3

From Table 1, if the number of clusters is envisioned as the total number of node of FCM, and the fuzzy semantics associated with prototypes, say G_1, G_2 and G_3, are regarded as the concepts of corresponding nodes of FCM, FCM is constructed to express time series showed in Fig.2 (a), which is shown in Fig.2 (b). FCM with sound weights can represent dynamic characters of the time series, which can become realization by PSO algorithm to minimize objective function (4) (refer to section 2.2).

4 The proposed prediction method of time series based on fuzzy cognitive map with information granules

In this section, the proposed prediction method of time series based on FCM with information granules is detailed. FCM is exploited to realize prediction of time series, which is based on the idea that the structure of it can store the values of weights between nodes that describe dynamic behaviour of time series at each iteration step.

Suppose that $X = \{x_1, x_2, \cdots, x_l\}$ is a time series. The outline of proposed method is presented in Fig.3, which consists of two stages — the first stage (from step 1) to step 5)) is to construct FCM model to express time series and the second stage (from step 6) to step 8)) is to use the constructed FCM model to perform prediction. In what follows, the two stages are detailed respectively.

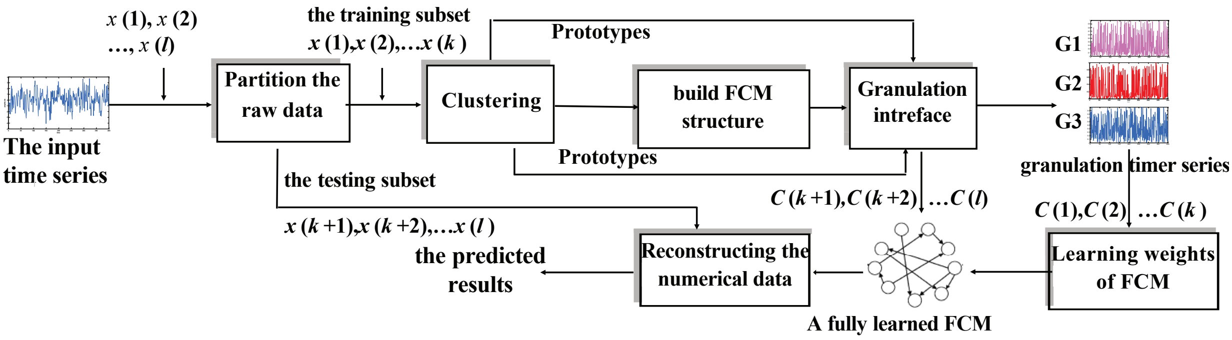


Figure 3: The outline of proposed prediction method.

Step 1) Partition the raw time series $\mathbf{X}$ into the training set consisted of k observations and the testing set consisted of lk observations. The former is used to develop FCM model, whereas the latter one is separate and is used to carry out prediction of the non-observed data.

Step 2) Generate information granules by fuzzy c-means clustering algorithm. Its object is to capture the linguistic and numerical characters of the training set. The stand version of fuzzy c-means clustering algorithm is adopted in this paper, which accepts the all observations from the training set and the predefined number of clustering c, and by the end of clustering algorithm, a prototype vectors $\mathbf{P} = [p_1 \; p_2 \; \cdots \; p_c]$ can be obtained. Each prototype p_i $(i = 1, 2, \cdots, c)$ is assigned into semantics, which result in the formation of information granules, say $G_1, G_2, \cdots, G_c$ (see section 3), i.e., prototypes can be make semantics, which form information granules.

Step 3)) Build the structure of candidate FCM via the information granules generated by step 2). The step plays an important role. As mentioned previously in section 3, the predefined number of clusters c is regarded as the total number of nodes of FCM and information granules are directly taken as the concept of corresponding nodes of FCM. Subsequently, the structure of fuzzy cognitive maps is formed.

Step 4) Granulate each observation in the training set by granular interface, which creates ordered data points by the well-known (5) according to prototypes $\mathbf{P}$ obtained previously and the each observation in the training set. Note that u_i can become the active level of the ith node of FCM related to ith granules G_i which is formed by assigning semantics into prototype p_i.

$$u_i(p_i, x(t)) = \frac{1}{\sum_{j=1}^{c} \| \frac{x(t)-p_i}{x(t)-p_j} \|^{\frac{2}{m-1}}}, i = 1, 2, \cdots, c; t = 1, 2, \cdots, k \tag{5}$$

where $x(t)$ is an observation at t time moment from the training set, c is the number of cluster, p_j $(j = 1, 2, \cdots, c)$ is the jth prototype obtained previously, m is fuzzification coefficient, $u_i(p_i, x(t))$ is the membership value that $x(t)$ belonged to prototype p_i and t is time label.

$$\mathbf{F}=\begin{bmatrix}\mathbf{G}_1\\ \mathbf{G}_2\\ \vdots\\ \mathbf{G}_c\end{bmatrix}=\begin{bmatrix}\mathbf{C}(1)\\ \mathbf{C}(2)\\ \vdots\\ \mathbf{C}(l)\end{bmatrix}^T=\begin{bmatrix}c_1(1) & c_1(2) & \cdots & c_1(l)\\ c_2(1) & c_2(2) & \cdots & c_2(l)\\ \vdots & \vdots & \vdots & \vdots\\ c_c(1) & c_c(2) & \cdots & c_c(l)\end{bmatrix}= \\ \begin{bmatrix}u_1(p_1,x(1)) & u_1(p_1,x(2)) & \cdots & u_1(p_1,x(l))\\ u_2(p_2,x(1)) & u_2(p_2,x(2)) & \cdots & u_2(p_2,x(l))\\ \vdots & \vdots & \vdots & \vdots\\ u_c(p_c,x(1)) & u_c(p_c,x(2)) & \cdots & u_c(p_c,x(l))\end{bmatrix} \tag{6}$$

According to (5), each observations $x_i(t)$ $(i=1,2,\cdots,l-k)$ from the training subset can be transformed along with time variable t, and become available in the form of their membership values corresponding to each granules G_j, which means original time series is converted to granular time series. We unfold (5) along with time label t, (6) can be obtained. Let us observe (6), $G_1,G_2,\cdots,G_c$ are defined as granular time series which includes c subsequence. Note that each row vector of $\mathbf{F}$, say $\mathbf{G}_1,\mathbf{G}_2,\cdots,\mathbf{G}_c$, expresses the level that the given time series can be characterized by corresponding fuzzy semantics p_i, whereas each column vector of $\mathbf{F}$, say $\mathbf{C}(1),\mathbf{C}(2),\cdots,\mathbf{C}(l)$, expresses the level that an observation of time series $x(t)$ at t time moment can be characterized by all semantics.

Step 5) Learn weights of FCM by PSO algorithm according to the transformed granular time series data. Its role is to establish a fully learned FCM model. The PSO algorithm described in Section 2.2 is used to learn all parameters vector $\hat{\mathbf{W}}$ of FCM constructed by step 3) on basis of the granulation time series data $\mathbf{C}(1),\mathbf{C}(2),\cdots,\mathbf{C}(l)$ which is formed by step 4).

Step 6) Granulate datum from the testing set by granulation interface, which is regarded as input of the fully learned FCM formed by step 5). A point $x(k+s)$ at $k+s$ time moment within the testing set, which is transformed into a tuples initial state vector $\mathbf{C}(k+s)$ according to (5), which indicate the activation level of all nodes in FCM model.

Step 7) Perform iteration computation on basis of (1). FCM model perform one step iteration from initial state vector $\mathbf{C}(k+s)$ and generates response vector $\hat{\mathbf{C}}(k+s+1)$. Results of iteration computation of FCM indicate the state of granular time series in the next time moment, in other words, the iteration process of FCM is actually the prediction process on the level of information granules.

Step 8) Reconstruction the numerical values according to the activation level of all nodes of FCM at a certain time moment. The numerical prediction is carried out according to (7) which reconstructs numerical values on basis of the activation value among all nodes of FCM at a certain time moment and prototypes formed by fuzzy c-means clustering algorithm prior.

$$\hat{x}(t)=\frac{\sum\limits_{j=1}^{c}\hat{c}_j(t)^m p_j}{\sum\limits_{j=1}^{c}\hat{c}_j(t)^m} \tag{7}$$

where c is the number of cluster, p_j $(j=1,2,\cdots,c)$ is prototypes, m is fuzzification coefficient, $\hat{c}_j(t)$ is an activation value of jth node of FCM model at t time moment, and $\hat{x}(t)$ is the predicted numerical value at t time moment.

5 Experimental study

In this section, two benchmark time series — the enrollment of university and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) time series, are used to carry out profound experiment for validating feasibility and effectiveness of proposed predicted method. The goal is to assess quality of the proposed method and compare with other method based on fuzzy sets theory.

5.1 Experimental setup

The proposed prediction method concern only one adjustable parameter,say the predefined number of clustering c. The all experiments for the two time series include the two tasks. One is that the quantification of the impact on the prediction accuracy of the proposed method being brought by the number of clusters c which implies the level of granularity. The other is that the predicted results are obtained by using our proposed method are compared with other prediction methods.

Considering a given time series and a predefined value of c, we first divide the original time series into the training set and the testing set. The training subset is used to develop HFCM model and the testing subset is used to perform prediction. In the sequence, the experiments are carried out by systematically sweeping through the values of the number of clusters c which start with 3 and increment it until the value of 8 is reached. To assure high confidence in the produced prediction results, for each of c, the experiments were repeated 50 times, the average root mean squared error (RMSE) are reported whose goal is to evaluate prediction accuracy and compare with other methods. Besides, the other one parameters of fuzzy c-means clustering algorithm — fuzzification coefficient m is set into 2 for each of experiments. The standard version of PSO algorithm is used to learn parameters of all candidate FCM models of the two time series. Parameters of PSO algorithm used in all experiments is shown in Table 2.

Table 2: Parameters of PSO algorithm used in all experiments

Description	Value
Population size	50
Acceleration constant ϕ_1	2
Acceleration constant ϕ_2	2
Inertial weight ζ	0.9
Initial positions	Random number
maximum number of iterations	1000
minimum objection function value	10^5

5.2 Experimental results and analysis

There are two benchmark time series are applied to validate and analyze the proposed prediction method. The first is the enrollment of university of Alabama during 1971–1992, which includes 2 time series which includes 22 observations. The time series is often used by many researchers [6]– [13]. The second time series involves the daily values of TAIEX from January 1, 2000 to December 30, 2000, which consists of 242 observations. The time series has been used in the related literature [9, 15].

In order to directly compare with other methods, for the first time series, its all observations are taken as not only the training data for constructing FCM model but also the testing data for validating prediction accuracy. Whereas for the second time series, data from 2000/1/4 to 2000/10/31 are used to construct FCM model and data from 2000/11/1 to 2000/12/30 are used to perform prediction. In this case, the comprehensive tests that involve the two data sets are carried out respectively according to the description in section 5.1, whose results are reported in Fig.4, Fig.5, Fig.6, Table 3 and Table 4.

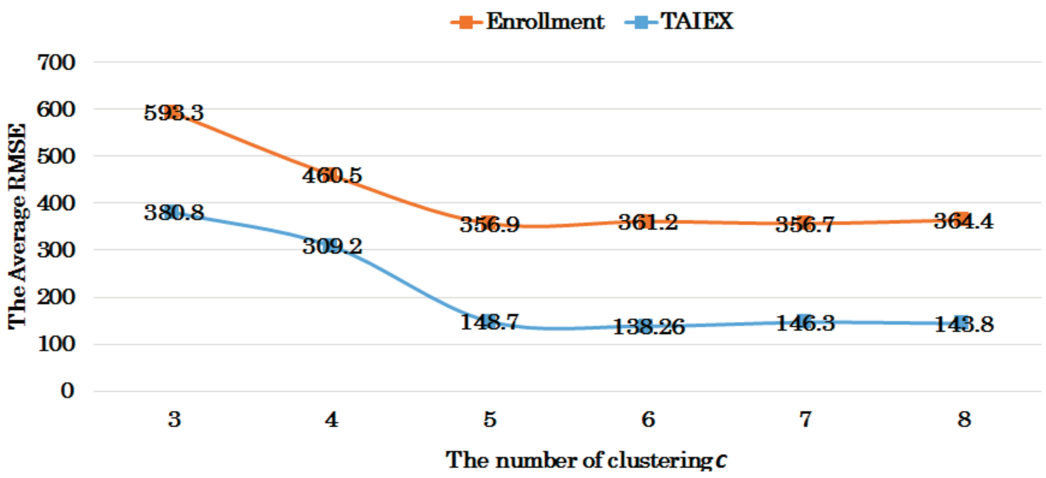


Figure 4: the plot of the average RMSE versus the number of clustering for the enrollment and TAIEX time series.

Table 3: The comparison result for enrollment time series.

Method	RMSE
Song's time-invariant method [6]	677
Chen's method [9]	663
Markov method [8]	638
Hwang's method [10]	567
Huarng's method [11]	489
Dan's method [13]	438
Proposed method (c=5)	346

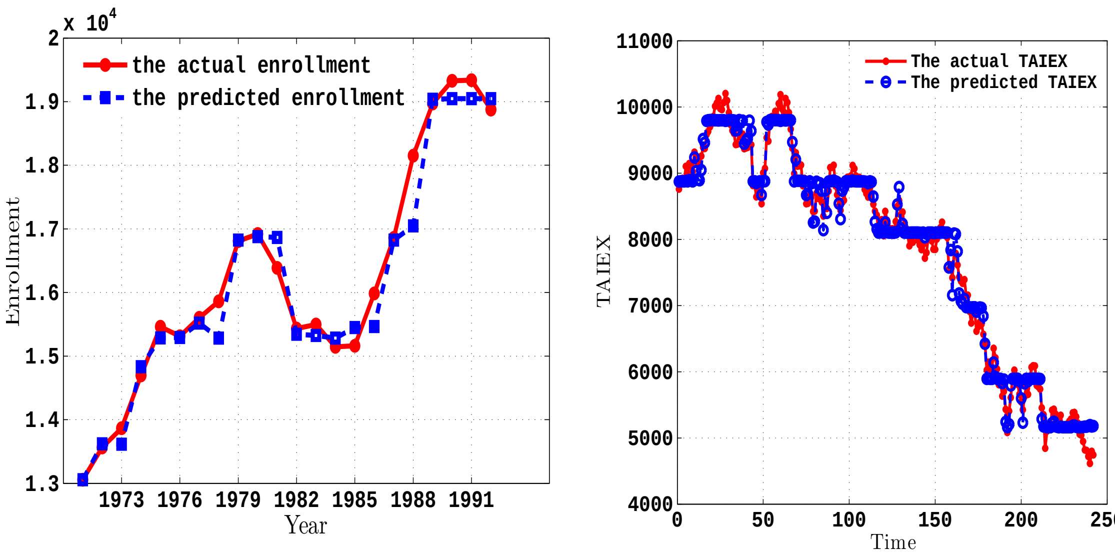
1) Experiments with respective to parameter c of proposed prediction method. Fig.4 represents a plot of the average RMSE versus the number of clustering c. Several interesting conclusions can be drawn when analyzing Fig.4. First, the prediction accuracy is highly sensitive to the choice of the number of clusters. The value of average RMSE get lower dramatically with the increasing number of clusters however the reduction of the average RMSE becomes less visible when going beyond a certain number of clusters i.e. 5 or 6. In other words, the prediction accuracy is not continuously increase with the increasing number of clusters. The significant increase in the prediction accuracy happens when moving from very low values of clusters to some higher values. The prediction accuracy changes unobviously with the change of clusters number when the value of cluster number c greater than 5 or 6. For example, as to enrollment time series, the average RMSE is 593.3 for c=3, whereas it is 356.9 for c=5 and 364.4 for c=8.

Table 4: The comparison result for TAIEX time series.

Method	RMSE
Chen's method [9]	176
Yu's method [23]	170
Huarng's method [24]	139
Proposed method (c=6)	132

There is smaller difference in RMSE at the higher number of cluster just 7.5, but there is larger difference in RMSE at the lower number of clusters reaching 236.4. For TAIEX time series, the situation is also similar and optimal values of c are located at the value of 6. The change of the average RMSE is slightly when the value of c over 6. The explanation for this is that when the number of clusters is low, very few granules are generated which is not sufficient to capture the character of time series and gives rise to the larger prediction error. But the number of clusters is larger, the granules may be more minute, which drastically reduces their interpretability.

2) Comparison with other prediction methods based on fuzzy sets theory. Table 3 and Table 4 reports comparison with other methods based on fuzzy sets theory respectively. The results in Table 3 show that the proposed prediction method with the best parameters (c=5) can obtain the better accuracy for enrollment time series. The RMSE of the proposed methods can achieve 346, while the best result Dans method [13] scored 438. At the same time, results for the TAIEX time series also show superiority of Huarngs method [24] . The raw time series and the selected best predicted results for enrollment and TAIEX time series are illustrated in Fig.5 and Fig.6 respectively.



(a) The plot of the actual and predicted enrollment. (b) The plot of the actual and predicted TAIEX.

Figure 5: The experimental results of using our proposal prediction method

6 Conclusions

In this paper, a time series prediction method based on fuzzy cognitive with information granules was proposed. It includes two important components — fuzzy clustering algorithm and fuzzy cognitive maps. The former extracts information granules from original time series data and then transform the original time series into the granular time series, while the latter is used to describe these granular time series and perform prediction. Comprehensive experiments have been carried out for two benchmark time series to validate feasibility and effectiveness of our proposed method. The results of experiment show predicted accuracy is related to the number of cluster. The proposed prediction method can obtain satisfying prediction accuracy in the case of the sound number of cluster. Additional, our proposed prediction method can automatically perform modeling and prediction of time series without more human intervention, its potential advantage is capability of handling modeling and prediction of large-scale time series.

Acknowledgements

The research is supported by the Natural Science Foundation of China under Grant 61175041 and The Major National Science and Technology Programs of China in the "Twelfth Five Year" Plan Period (No. 2011ZX05039-3-3).

Bibliography

[1] Kailath, T. (1980); *Linear Systems*, Prentice Hall.

[2] Papoulis, A. (1991); *Probability, Random Variables and Stochastic Processes*, Mcgraw-Hill College.

[3] Juditsky, A. et al (1994); *Wavelet in identification: wavelets, splines, neurons, fuzzies: how good for identification?*, INRIA reports No.2135.

[4] Kaplan, D.; Glass L. (1995); *Understanding Nonlinear Dynamics*, Springer Verlag.

[5] Song, Q.; Chissom, B.S. (1993); Fuzzy time series and its models, *Fuzzy Sets Systems*, 54(3): 269-277.

[6] Song, Q.; Chissom, B.S. (1993); Forecasting enrollments with fuzzy time series — Part I, *Fuzzy Sets Systems*, 54(1): 1-9.

[7] Song, Q.; Chissom, B.S. (1994); Forecasting enrollments with fuzzy time series — Part II, *Fuzzy Sets Systems*, 62(1): 1-8.

[8] Sullivan, J.; Woodall, W.H. (1994); A comparison of fuzzy forecasting and Markov modeling, *Fuzzy Sets Systems*, 64(3): 279-293.

[9] Chen, S.M. (1996); Forecasting enrollments based on fuzzy time series, *Fuzzy Sets Systems*, 81(3): 311-319.

[10] Hwang, J.R.; Chen, S.M.; Lee, C.H. (1998); Handling forecasting problems using fuzzy time series, *Fuzzy Sets Systems*, 100(1-3): 217-228.

[11] Huarng, K. (2001); Heuristic models of fuzzy time series for forecasting, *Fuzzy Sets Systems*, 123(3): 137-154.

[12] Chen, S.M. (2002); Forecasting enrollments based on high-order fuzzy time series, *Cybernetics and Systems: An International Journal*, 33(1): 1-16.

[13] Dan, J.; Dong F.; Hirota, K. (2011); Fuzzy Local Trend Transform based Fuzzy time series Forecasting Model, *International Journal of Computers, Communications & Control* , VI(4): 603-614.

[14] Kosko, B. (1986); Fuzzy Cognitive Maps, *International Journal of Man-Machine Studies*, 7: 65-75.

[15] Stach, W.; Kurgan, L.; Pedrycz, W. (2008); Numerical and Linguistic Prediction of Time Series With the Use of Fuzzy Cognitive Maps, *IEEE Transactions on Fuzzy system*, 16(1): 61-72.

[16] Pedrycz, W. (2010); The design of cognitive maps: A study in synergy of granular computing and evolutionary optimization, *Expert system with applications*, 37(10): 7288-7294.

[17] Pedrycz, W.; Vukovich, G. (2010); Abstraction and specialization of information granules, *IEEE Transactions on Systems Man and Cybernetics, Part B: Cybernetics*, 31(1): 106-111.

[18] Axelrod, R. (1976); *Structure of Decision: The Cognitive Maps of Political Elites*, Princeton University Press.

[19] Stach, W.; Kurgan, L.; Pedrycz, W.; Reformat M. (2005); Genetic learning of fuzzy cognitive maps, *Fuzzy Sets Systems*, 153(3): 371-401.

[20] Kennedy, J.; Eberhart, R. (1995); Particle Swarm Optimization, *Proceeding of IEEE International Conference on Neural network*, 4: 1942-1948.

[21] Shi, Y.H.; Eberhart, R. (1998); A modified particle swarm optimizer, *Proceeding of IEEE International Conference on Evolutionary Computation*, 69-73.

[22] Bezdek, J.C. (1981); *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press.

[23] Yu, H.K. (2005); A rened fuzzy time-series model for forecasting, *Physica A: Statistical Mechanics and its Applications*, 346(3-4): 657-681.

[24] Huarng, K.; Yu, H.K. (2005); A Type 2 fuzzy time series model for stock index forecasting, *Physica A: Statistical Mechanics and its Applications*, 353(1): 445-462.

[25] Lu, W.; Pedrycz, W; Liu, X.; Yang, J.; Li, P. (2014); The modeling of time series based on fuzzy information granules, *Expert Systems with Applications*, 41: 3799-3808.