연속된 데이터의 퍼지학습에 의한 비정상 시계열 예측

論 文
50D-5-4

Predicting Nonstationary Time Series with Fuzzy Learning Based on Consecutive Data

金 仁 澤
(Intaek Kim)

Abstract - This paper presents a time series prediction method using a fuzzy rule-based system. Extracting fuzzy rules by performing a simple one-pass operation on the training data is quite attractive because it is easy to understand, verify, and extend. The simplest method is probably to relate an estimate, $x(n+k)$, with past data such as $x(n)$, $x(n-1)$, ..$x(n-m)$, where k and m are prefixed positive integers. The relation is represented by fuzzy if-then rules, where the past data stand for premise part and the predicted value for consequence part. However, a serious problem of the method is that it cannot handle nonstationary data whose long-term mean is varying. To cope with this, a new training method is proposed, which utilizes the difference of consecutive data in a time series. In this paper, typical previous works relating time series prediction are briefly surveyed and a new method is proposed to overcome the difficulty of predicting nonstationary data. Finally, computer simulations are illustrated to show the improved results for various time series.

Key Words : Time series prediction, Nonstationary data, Fuzzy rule-based system.

1. Introduction

A time series is defined as a sequential set of data measured over time. The sequence is important in time series analysis, because the information on the source is embedded in it. The most crucial assumption that we use in time series analysis is that a source of a time series is governed by a deterministic dynamic system.

Time series prediction involves forecasting the future by understanding the past. It has been widely studied by many researchers. Most of works on prediction have been conducted from the viewpoint of stochastic models such as MA (moving average), IMA (integrated moving average) and ARIMA (autoregressive integrated moving average) [1-2]. One of the breakthroughs in this area began with the development of computational intelligence paradigms, including fuzzy logic systems and neural networks in the 1980s [3]. Both methodologies are well known not only as universal approximator [4-5] but also for their capability of learning. The most important aspect of universal approximator is that it can be applied to any nonlinear modeling problem. However it should be noticed that there are many other types of universal approximator and the advantage of fuzzy logic system (as well as neural networks) as a universal approximator lies in the fact that it can be easily and efficiently implemented.

In this paper, we address the time series prediction using fuzzy rule-based system. Extracting fuzzy rules from a time series means expressing the dynamic system in terms of fuzzy if-then rules. The first procedure for building fuzzy rule-based system is training from data. There are many techniques for training of fuzzy logic systems. For example, Wang [6] shows some of them using back-propagation, orthogonal least squares, nearest neighborhood clustering, and one-pass operation. In this paper, we adopt the idea of one-pass operation. The technique based on the one-pass operation produces if-then fuzzy rules from every input and output pair. The simple training can avoid computationally expensive procedure, but the range of input values should be specified a priori and no optimization can be made after training.

A system with unknown dynamics is approximated or represented by a number of fuzzy rules. Extracting fuzzy rules from data has been investigated [7-8]. This is true as long as the dynamics can be described in terms of fuzzy rules. Due to the intelligence of describing nonlinear systems by fuzzy logic, the prediction of unknown systems using fuzzy rule-based systems has

* 正 會 員 : 明知大 電氣情報制御工學部 副敎授 · 工博
接受日字 : 2001年 2月 21日
最終完了 : 2001年 4月 20日

been successful in nonlinear areas [8-10]. Additional refinements coupled with genetic algorithms are reported in the literature [11-12]. Improvement using genetic algorithms is resulted from fine tuning of fuzzy membership function. Early work of chaotic time series prediction using fuzzy if-then rules is found in [13]. Fusing neural network architecture to fuzzy system, there have been many attempts to predict time series. Kozma et al. [14] show application of fuzzy neural network (FuNN) to utilize and generate knowledge during an iterative learning and adaptation procedure. Maguire et al. [15] propose an architecture employing fuzzy reasoning system with efforts to reduce the dimensions of the network, where the author claim the advantages of the architecture for hardware implementation. As a practical application, Chiu [16] presents a prediction system for the Taipeis unemployment rate series. The author uses the idea of fuzzy system to modeling data relating unemployment and prediction is made using neural network architecture.

Many applications, however, show that a time series is not necessarily stationary. Examples of this include many indices related to real life taken from economy, consumption, production, weather, biomedical engineering, and many other fields. A nonstationary system, in strict sense, cannot be represented by fuzzy rules, because a fixed number of rules can describe a time invariant system, which obviously rules out the nonstationarity. Recently intensive efforts have been made to analyze nonstationary time series: Yee and Haykin [17] report a comprehensive work on a dynamic regularized radial basis function network for nonlinear and nonstationary time series prediction. Fitting time series model to nonstationary processes is presented by Dahlhaus [18]. For specific areas, Hori et al. [19] address nonstationary property of biomedical signal, while Lesch and Lowe [20] present a framework combining stochastic and deterministic description for nonstationary financial time series.

In this paper, we propose a method of building fuzzy rules for prediction, which utilizes the difference of consecutive data in a time series. Computer simulations show that a nonstationary signal can be represented by new fuzzy rules and the performance of prediction is improved by applying the proposed method to time series data such as Mackey Glass time series and Lorenz data.

2. Nonstationarity

Stationarity has always played a major role in time series analysis. The power spectrum, for example, is defined for stationary processes and time series analysis in the frequency domain depends on the assumption of stationarity. The important ARMA model is also a stationary time series model.

Unfortunately, many empirical data show that they have no fixed average, which means they are nonstationary. Nonstationary data cannot be expressed in terms of a fixed number of fuzzy rules since no rule exists for the data that are not observed in the training process. However, if nonstationary data shows some local homogeneity, then stationary model can be derived by supposing some proper difference of the nonstationary process. These models are called autoregressive integrated moving average (ARIMA) processes [1]. The ARMA model and ARIMA model can be expressed respectively as

$$x(n) = \sum_{p=1}^{P} a(p)x(n-p) + \sum_{q=1}^{Q} b(n)e(n-q) \tag{1}$$

$$\nabla x(n) = \sum_{p=1}^{P} a(p)\nabla x(n-p) + \sum_{q=1}^{Q} b(n)e(n-q) \tag{2}$$

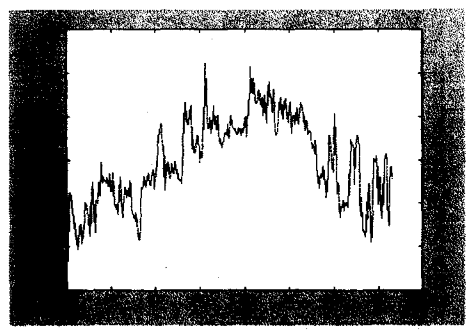


Fig. 1. Time series of daily temperature from [2].

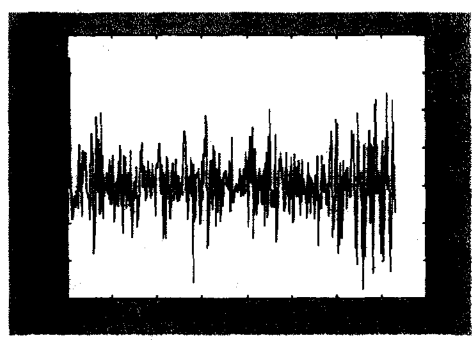


Fig. 2. Difference of the temperature data shown in Fig. 1.

where $x(n)$ is the time series, $e(n)$ is the external input, and $x(n)$, is defined as $x(n)-x(n-1)$. The first part of the right-hand side of (1) is autoregressive part, and the second part is moving average part. In Fig. 1, it is illustrated that $x(n)$'s form a nonstationary time series data, while its difference in Fig. 2, $x(n)$'s become a stationary time series. We exploit this observation. Equations (1) and (2) do not reflect the exact structure we want to use in this paper, since the time series data $x(n)$ is designated as a fuzzy linguistic value and $\alpha(p)$ is a degree of fulfillment in the training process. More importantly, only autoregressive part of (2) is used and fuzzified $x(n)$ can be expressed as a nonlinear combination of fuzzified $x(n-1)$, $x(n-2),..,x(n-P)$. Therefore, the fuzzy rules derived in the training process contain information regarding the difference of data, more specifically, the difference of consecutive data in a time series.

3. Fuzzy Predictor

A k-step ahead fuzzy prediction can be mathematically described as

$$\hat{x}(n+k) = F\{x(n), x(n-1),..,x(n-m)\} \tag{3}$$

where F represents a nonlinear mapping from a set of time series, $\{x(n),\ x(n-1),\ ..\ x(n-m)\}$ onto k-step ahead estimate, $\hat{x}(n+k)$

It should be noted in (3) that input sampling time is normalized for simplicity and not all of the m inputs are necessarily used as an input variable to F. When k is small, it is called short-term prediction. Otherwise, it is long-term prediction and it is, in general, very difficult to produce good results. The mapping of F can be represented by fuzzy rules such as

If $x(n)$ is $T_x(n)$ and $x(n-1)$ is $T_x(n-1)$ .. and $x(n-m)$ is $T_x(n-m)$, then $x(n+k)$ is $T_x(n+k)$ (4)

where x's are input variables and T_x's are fuzzy linguistic values partitioned according to the range of the input variables. Conventional fuzzy prediction methods [6-7] based on the above rule, are applicable only to the time series whose statistics is stationary or wide sense stationary. Otherwise, they fail because fuzzy rules are not able to describe the dynamics of nonstationary time series. In nonstationary time series, the average tends to change, and the fuzzy predictor may not accommodate the changes.

To resolve this difficulty, we restate the k-step ahead prediction as

$$\nabla\hat{x}(n+k) = F_k\{\nabla x(n), \nabla x(n-1),.., \nabla x(n-m)\} \tag{5}$$

where $x(n)$ = $x(n)-x(n-1)$ and the mapping of Fk should be determined from the difference of consecutive data in a time series. Accordingly, the rule for a new mapping of Fk should be modified as

If $x(n)$ is $T_d(n)$ and $x(n-1)$ is $T_d(n-1)$ .. and $x(n-m)$ is $T_d(n-m)$, then $x(n+k)$ is $T_d(n+k)$ (6)

As shown in the rule (6), the proposed method learns the difference of consecutive data in a time series rather than values of time series. Therefore, any time series showing some tendency of drifting becomes predictable by the proposed method. It indicates that the differences of consecutive time series are valuable information in describing dynamics of systems, particularly nonstationary ones. An example is detailed in the next section. An additional improvement obtained is the reduction of prediction error, which will be described in the following section.

The procedure for building fuzzy rules for the fuzzy predictor can be described as follows:

Step 1: Divide the range of difference of consecutive data in a time series.

Find the distribution of x's in a time series. Assume they are in the interval [dmin, dmax], then divide it into 2n+1 regions and assign fuzzy membership function to each region.

Step 2: Obtain fuzzy rules from data pairs.

In the training stage, the input, the difference of consecutive data, is assigned to a suitable fuzzy membership function whose fulfillment of degree is the highest. Since the fuzzy membership functions are overlapped in most cases, a particular input value may have multiple membership function. For example, if x_i is 0.6 in T_i and 0.4 in T_j , then x_i belongs to the fuzzy membership function T_i . Fuzzy rules are obtained by assigning each data to a fuzzy membership function and can be expressed as (6).

Step 3: Resolve conflicting rules.

In the above step, we can obtain fuzzy rules that have the same premise with different consequence. In order to resolve such conflict, we choose the rule whose minimum degree is the largest. Minimum degree of a rule is defined as the smallest degree of fulfillment to the premise and the consequence part of a rule. For example, assume we have two conflicting rules as follows:

Rule1: If x_1 is FA and x_2 is FB, then y is FY1 with degree of fulfillment (FA(x_1)=0.6, FB(x_2)=0.7, FY1(y)=0.9),

Rule2: If x_1 is FA and x_2 is FB, then y is FY2 with degree of fulfillment (FA(x_1)=0.7, FB(x_2)=0.7, FY2(y)=0.8).

Since the minimum degree of the first rule is 0.6 and the second is 0.7 and the latter, Rule 2, is preferred as a rule. Therefore, in the training process, the conflicting rules are resolved by choosing the largest minimum degree of a rule. Wang [8] use another method to resolve the problem by assigning a degree of product of membership functions and selecting one with the largest product.

4. Simulations

In this section, we present two experiments that show the advantage of the proposed method. First, the representation of a nonstationary signal is attempted using fuzzy rules. Second, the prediction of the well-known time series including Mackey Glass time series, Lorenz data, and temperature data [2] is illustrated. In both experiments except temperature data, we used the first 700 data for training and applied the algorithm to following 300 data for the purpose of prediction. Mandanis fuzzy reasoning is employed for the inference procedure and the center of area (COA) method is utilized for a defuzzifier.

4.1. Representation of nonstationary signals

Figure 2 shows the data used in the simulation. The data can be expressed as

$$y = \sin(ax) + bx \tag{7}$$

where a and b are positive numbers. It is therefore an increasing function with periodicity. The first 700 data are used to build the fuzzy rules described in (4) and (6), and their results are depicted in Figure 3 and 4, respectively. With the conventional fuzzy rules on (4), rules are missing in the points where the outputs are zero. This is to say, that the corresponding inputs at these points have no match in the premise part of the fuzzy rules based on (4). In contrast, representing (7) with the fuzzy rules based on (6) works fine as shown in Figure 4. Equation (5) explains why the nonstationary signal of (7) can be trained into fuzzy inference system.

4.2 Prediction of time series

Although the proposed fuzzy prediction method demonstrates better performance in representing the nonstationary signal in the previous section, it is not evident if it holds true for time series prediction. In order to investigate the issue, we carry out computer simulation with two sets of data.

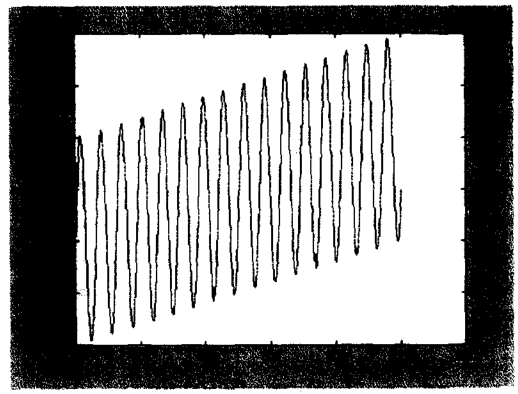


Fig. 3. Data generated by (7) for experiment.

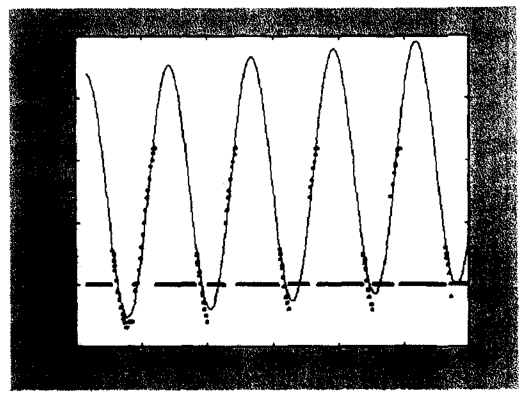


Fig. 4. Representation of data in [701, 1000] in Fig. 3 using rule (4). [Dotted line is prediction. The points producing series of zeros indicate that no rule is applicable for given input.]

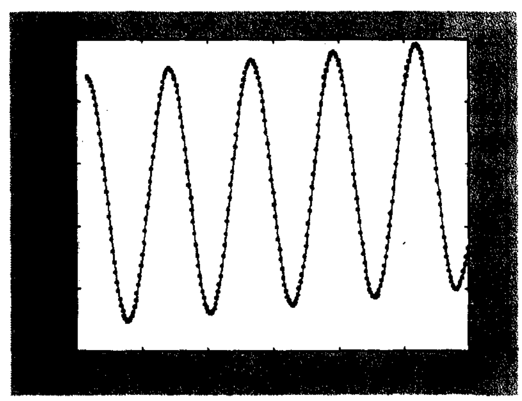


Fig. 5. Representation of data in [701, 1000] in Fig. 3 using rule (6). [Dotted line is prediction.]

[Mackey-Glass time series]

Mackey-Glass time series are obtained from the following equation:

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \qquad (8)$$

The first 700 data are used to build fuzzy rules and 291 points (from n=710 to n=1000) are predicted as shown in Figs. 6 and 7, where m=8 and the number of partition for the inputs is 14. Figs. 6 and 7 are obtained from the rules based on (4) and (6), respectively.

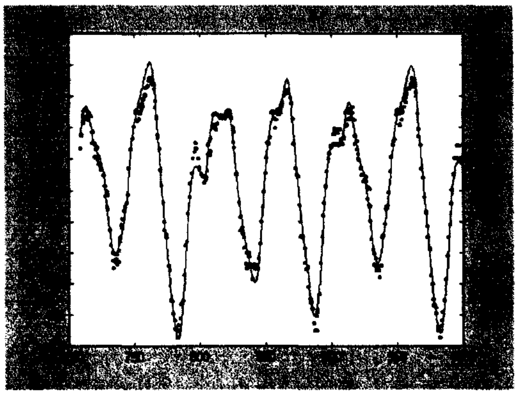


Fig. 6. Prediction of Mackey-Glass time series using rule (4). [Dotted line is prediction, solid line is real value.]

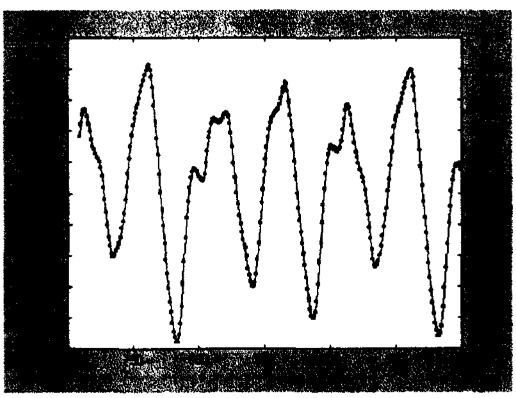


Fig. 7. Prediction of Mackey-Glass time series using rule (6). [Dotted line is prediction, solid line is real value.]

[Lorenz data]

Lorenz data is generated from the well-known Lorenz chaotic equation and we observe only x variable in the simulation. In the same way, we compare the performance of the prediction and the results are shown in Figs. 7 and 8, where m=8 and the number of partition for the inputs is 9.

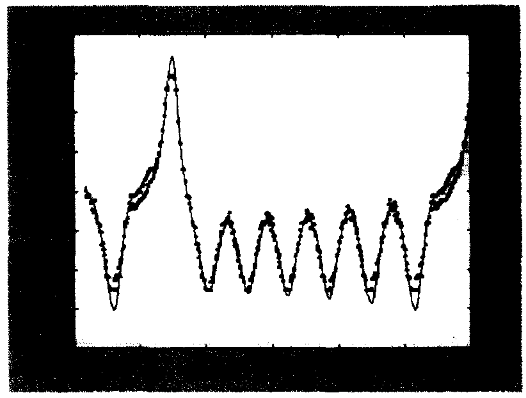


Fig. 8. Prediction of Lorenz data using rule (4). [Dotted line is prediction, solid line is real value.]

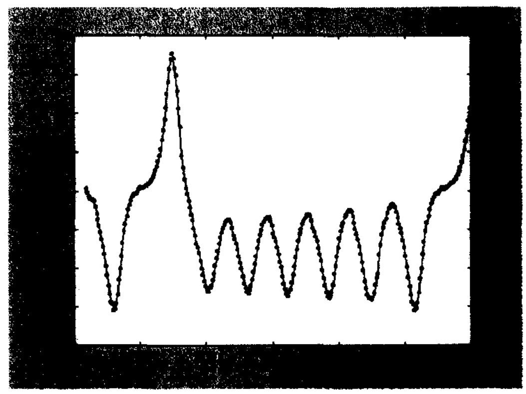


Fig. 9. Prediction of Lorenz data using rule (6). [Dotted line is prediction, solid line is real value.]

[Daily Temperature Data]

In the previous examples, we obtain data from the equations and attempt to build the fuzzy rules to describe the dynamics. The results show difference of data is more valuable in predicting a time series. At this point, however, we may have questions: Does it work well on nonstationary data? Isnt it due to the deterministic equation that prediction fits well to real data? In fact, both data Mackey-Glass time series and Lorenz data are not quite evident if they are nonstationary, and the difference of data is always used when we solve differential equations with computer.

Temperature data shown in Fig. 1. is a proper example in those aspects. It is nonstationary data with seasonal effects and not generated by deterministic equations. Figs. 10 and 11 are the results of prediction based on rule (4) and (6), respectively. The first 200 data in Fig. 1 and 2 are used to build fuzzy rules, and the remaining 165 points are predicted using m=2 and k=1 in (4) and (6).

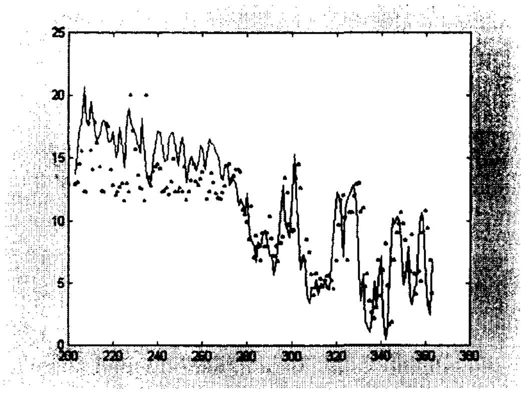


Fig. 10. Prediction of daily temperature data using rule (4).

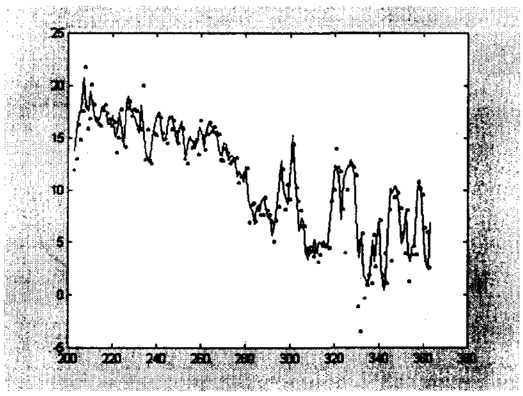


Fig. 11. Prediction of daily temperature data using rule (6).

Tables 1-3 summarize the prediction error in terms of RMSE (Root Mean Square Error). The first column of the table shows the number of membership functions. The second column has three sub-columns showing the prediction error calculated by the conventional method and the proposed method using two different ranges of membership functions denoted by [xmin, xmax], respectively. The first range used in the conventional method is derived from the minimum and the maximum values of time series (CASE 1). The second range is derived is from the minimum and the maximum of the difference of any two values (CASE 2). The last range is obtained from the minimum and the maximum of two consecutive data of the time series (CASE 3). In general the proposed method reduces the prediction error. Although it is not explicitly described in this paper, an important finding was that the new method generated fewer rules in the prediction of CASE 2 than in the prediction of CASE 1. CASE 3 generated the same or more rules than CASE 2.

Table 1. Prediction error for Mackey-Glass time series in RMSE

Number of input membership functions	Range of inputs		
	Error by the conventional method (CASE 1)	Error by the proposed method (CASE 2)	(CASE 3)
	[0:1.5]	[-1.5:1.5]	[-0.08:0.06]
6	0.0516	0.1054	0.0080
14	0.0274	0.0124	0.0041
29	0.0156	0.0092	0.0156

Table 2. Prediction error for Lorenz data in RMSE.

Number of input membership functions	Range of inputs		
	Error by the conventional method (CASE 1)	Error by the proposed method (CASE 2)	(CASE 3)
	[-16:18]	[-34:34]	[-1.8:2.4]
9	0.8694	0.8596	0.2837
16	0.4845	0.3028	0.3663
33	1.1927	0.4960	0.5151

Table 3. Prediction error for daily temperature data in RMSE

Number of input membership functions	Range of inputs		
	Error by the conventional method (CASE 1)	Error by the proposed method (CASE 2)	(CASE 3)
	[-2:24]	[-26:26]	[-5.5:6.8]
9	3.0975	2.0153	2.0970
15	2.8929	2.1842	2.1307
32	3.3411	2.0628	2.0659

5. Conclusions

Time series analysis has been studied by many researchers for a long time. Modeling and prediction of time series are the main streams of the research. However many time series in real life have both nonlinear and nonstationary properties and it has been a big challenge to take them into proper consideration in the analysis. Recently developed computational intelligence paradigms including neural networks and fuzzy logic systems have addressed some of the issues relating those properties.

In this paper, a new fuzzy learning method for time series prediction utilizing the difference of consecutive data has been proposed. Training of time series is accomplished in the one-pass operation and two simulations were carried out: First, the representation of a nonstationary signal is attempted using fuzzy rules. Second, the prediction of the well-known time series including Mackey Glass time series, Lorenz data, and daily temperature data is illustrated.

Result showed that a nonstationary signal given by (5), an increasing sinusoidal function, can be represented by the fuzzy rules based on the difference of consecutive data. Reduction of prediction error was also achieved by applying the proposed method in the second simulation.

Future work will be aimed to apply the idea of difference of consecutive data to the various prediction schemes based on fuzzy logic paradigm. We expect enhanced performance in modeling and prediction, because the differences of consecutive data in time series are more adequate in describing nonstationarity than the values themselves.

감사의 글

이 논문은 2000년 명지대학교 교내연구비에 의하여 지원되었음.

참 고 문 헌

[1] George E.P. Box and Gwilym M. Jenkins, Time Series Analysis: forecasting and control, Revised Ed. Prentice Hall, 1976.

[2] Gareth Janacek and Louise Swift, Time Series forecasting, simulation, applications, Ellis Horwood, 1993.

[3] Andreas S. Weigend and Neil A. Gershenfeld, Ed. TIME SERIES PREDICTION: Forecasting the Future and Understanding the Past, Addison-Wesley, 1994.

[4] L. X. Wang, Fuzzy systems are universal approximators, Proc. IEEE International Conf. On Fuzzy Systems, pp. 1163-1170, 1992.

[5] F. Scarselli and A. C. Tsoi, Universal Approximation Using Feedforward Neural Networks A Survey of Some Existing Methods, and Some New Results, Neural Networks, vol. 11, No. 1, pp. 15-37, 1998.

[6] L. X. Wang, Adaptive Fuzzy Systems and Control: Design and Stability Analysis, Prentice-Hall, 1994.

[7] T. Takagi and M. Sugeno, Fuzzy Identification of Systems and its Application to Modeling and Control. IEEE Trans. Syst., Man, Cybern., vol.15, pp.116-132, 1985

[8] L. X. Wang and J. M. Mendel, Generating fuzzy rules by learning from examples, IEEE Trans. Syst., Man, Cybern., vol. 22, pp. 1414-1427, Nov. 1992.

[9] S. H. Lee and I. Kim, Time sereis analysis using fuzzy learning, in Proc. Int. Conf. Neural Inform. Processing, Seoul, Korea, Oct. 1994, vol. 6 pp. 1577-1582.

[10] C.C. Lee, Fuzzy logic in control systems: Fuzzy logic controller Parts I and II, IEEE Trans. Systems Man Cybernetics. Vol. 20. No.2, pp.404-435, 1990.

[11] J. L. Castro, M. Delgade, and F. Herrera, A learning method of fuzzy reasoning by genetic algorithms, in Proc. 1 st Eur. Congress Fuzzy Intell. Technol., Aachen, Germany, Sept. 1993, pp.804-809.

[12] Daijin Kim and Chulhyun Kim, Forecasting Time Series with Genetic Fuzzy Predictor Ensemble, in IEEE Trans. Fuzzy Systems, vol. 5 no.4, Nov. 1997, pp.523-535

[13] R.J.S. Jang, Predicting chaotic time series with fuzzy IF-THEN rules, Proc. IEEE Second Int. Conf. On Fuzzy Systems, vol. 2, pp.1079-1084, 1993.

[14] R. Kozma, N.N.Kasabov, and J.S.Kim, Integration of Connectionist Methods and Chaotic Time-Series Analysis for the Prediction of Process Data, International Journal of Intelligent Systems, Vol. 13, pp. 519-538, 1998.

[15] L.P. Maguire, B. Roche, T.M. McGinnity, and L.J. McDaid, Prediction a chaotic time series using a fuzzy neural network, Information Sciences, 112, pp.125-136, 1998.

[16] C. Chiu, Prediction unemployment rates using fuzzy time series model and neural network, Proc. the IASTED International ConferenceArtificial Intelligence and Soft Computing, pp.496-499, May 27-30, 1998.

[17] P. Yee and S. Haykin, A Dynamic Regularized Radial Basis Function Network for Nonlinear, Nonsta- tionary Time Series Prediction, IEEE Trans. on Signal Processing, vol. 47, no. 9 pp. 2503-2521, Sep., 1999.

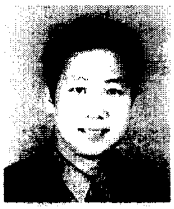
[18] R. Dahlahaus, Fitting Time Series Models to

Nonstationary Processes, The Annals of Statistics, vol. 25, no. 1, pp. 1-37, 1997.

[19] J. Hori, Y. Saitoh, and T. Kiryu, Real-Time Restoration of Nonstationary Biomedical Signal under Additive Noise, IEICE Trans. INF. & SYST., vol.E82-D, no.10, pp1409-1416, Oct., 1999.

[20] R. Lesch and D. Lowe, Towards a Framework for Combining Stochastic and Deterministic Descriptions of Nonstationary Financial Time Series, Proc. 1998 IEEE Workshop Signal Processing, pp. 587-596, 1998.

저 자 소 개



김 인 택 (金 仁 澤)

1960년 12월 25일 생. 1984년 서울대 공과대학 전자공학과 졸업. 1986년 서울대 대학원 전자공학과 졸업 (석사). 1992년 Georgia Institute of Technology, 전기공학과 졸업 (공박). 1993년~1995년 LG 전자기술원 선임연구원. 1995년~현재 명지대 전기정보제어공학부 부교수.

Tel : 031-330-6475, Fax : 03033-140-1411

E-mail : kit@mju.ac.kr