

논문 2011-4-13

유전 알고리즘을 이용한 퍼지신경망의 시계열 예측에 관한 연구

A Study on the Prediction of the Nonlinear Chaotic Time Series Using Genetic Algorithm based Fuzzy Neural Network

박인규*

In-Kyu Park

요 약 본 논문에서는 Mackey-Glass 시계열의 예측에서 유전자알고리즘을 이용하는 구조적인 동정과 뉴로퍼지에 의한 파라미터 동정의 학습방법과 하이브리드 시스템을 제안하였다. 본 방법은 두 가지로 구성되었다. 하나는 입력공간에 대한 분할을 통하여 유전 알고리즘을 이용하여 퍼지 규칙베이스를 구축하고 다른 하나는 이 규칙베이스를 토대로 기울기 최하강법을 이용하여 제어규칙의 변수에 대한 파라미터 동정이다. 제안된 방법을 성능을 검증하기 위하여 입력의 패턴을 시간간격에 따라서 $x(t-3)$, $x(t-6)$ 과 $x(t-9)$ 의 세 가지로 구성하였다. 많은 시뮬레이션을 통하여 유전알고리즘에 의한 구조적인 동정으로 인하여 학습초기에 오차가 작은 것을 알 수 있었다. 표2에서와 같이 성능을 확인 할 수 있었다.

Abstract In this paper we present an approach to the structure identification based on genetic algorithm and to the parameter identification by hybrid learning method in neuro-fuzzy-genetic hybrid system in order to predicate the Mackey-Glass Chaotic time series. In this scheme the basic idea consists of two steps. One is the construction of a fuzzy rule base for the partitioned input space via genetic algorithm, the other is the corresponding parameters of the fuzzy control rules adapted by the backpropagation algorithm.

In an attempt to test the performance the proposed system, three patterns, $x(t-3)$, $x(t-6)$ and $x(t-9)$, was prepared according to time interval. It was through lots of simulation proved that the initial small error of learning owed to the good structural identification via genetic algorithm. The performance was showed in Table 2

Key Words : Genetic Algorithm, Fuzzy Input Partition, Backpropagation Neural Networks.

1. 서 론

신경회로망은 다소의 복잡한 형태의 함수를 학습할 수 있는 능력 때문에 비선형 시스템의 제어에 자주 이용되고 있다. 대표적인 역전파 알고리즘을 이용한 다층퍼셉트론구조의 신경회로망은 많은 적용의 예를 가지고 있다. 그러나 이 회로망의 단점인 국부해의 수렴성 문제를

가지고 있다. 또한 퍼지 모델링은 전문가의 경험적 지식과 지식의 정성적인 언어형태의 규칙을 이용하여 복잡하고 불확실하며 비선형성을 포함한 시스템에 대해서 일반적으로 수행할 수 있다. 따라서 이들의 장점을 융합한 뉴로 퍼지 시스템은 퍼지이론의 if-then 규칙의 지식표현과 신경회로망의 학습능력, 적응성, 결함극복 능력, 병렬처리 능력과 같은 특성을 융합한 시스템 식별의 형태를 포함하고 시스템의 동정에는 구조 동정(structure identification)과 파라미터 동정(parameter identification)으로 구성된다. 이를 통하여 최적의 규칙과 최적의 멤버십 함수를 조

*정회원, 중부대학교 컴퓨터학과
접수일자 2011.6.16, 수정일자 2011.7.22
게재확정일자 2011.8.12

정하게 되는데, 이러한 최적의 규칙과 멤버쉽함수의 결정하는 문제는 아직도 응용에서 많은 문제점을 야기하고 있다. 결국 시스템의 추론의 결과는 규칙에 좌우되기 때문에 신경망의 학습능력을 바탕으로 제어규칙, 멤버쉽함수등의 파라미터를 적응시키는 방향으로의 연구가 활성화 되고 있다.

퍼지 제어규칙의 조건부와 결론부의 매개변수 동정에는 여러 가지의 파라미터 동정기법들이 있다. 이에 대한 연구로는 학습기능을 이용하는 Wang과 Mendal의 방법^[1], Ichitashi에 의한 방법^[2], Araki에 의한 방법^[3]등이 있다. 또한 제어규칙의 결론부가 입출력의 선형조합으로 나타나는 Sugeno에 의한 방법^[4]과 기울기 하강법의 학습에 의한 Nomura에 의한 방법^[5], 퍼지 클러스터링을 이용한 여러 방법등이 있다.

따라서 본 논문에서는 입력공간 분할 방법을 이용하여 이미 알고 있는 지식없이 입출력 데이터를 근거로 유전알고리즘을 이용하여 최적의 규칙과 멤버쉽 함수를 동정한다. 동정된 규칙과 멤버쉽 함수를 가지고 역전과 알고리즘과 최소자승법을 결합한 하이브리드 학습에 의해 최적에 가깝도록 파라미터의 동정을 수행할 수 있는 적용형 뉴로-퍼지-유전자 추론 시스템을 구축한다. 시뮬레이션으로 비선형의 시스템을 검증하는 전형적인 예인 Mackey-Glass Chaotic 시계열 Data Set을 통해 제안된 방법의 성능을 보이고자 한다.

II. 입력공간의 분할

1. 입력공간과 제어 규칙

뉴로퍼지 시스템을 구축하기 위하여 입력변수의 선택은 공학적인 상식으로 알려져 있다고 가정하고 이미 알고 있는 지식이 없이 입출력 데이터를 근거로 한 입력공간 분할에 따른 퍼지규칙과 소속함수의 생성에 대하여 다룬다. m개의 패턴 $X_p = (x_{p1}, \dots, x_{pn})$, $p=1,2,\dots,m$ 에 대하여 M 개의 그룹으로 분할문제의 경우에 식(1)과 같은 퍼지규칙을 정의할 수 있다.

$$R_q : \text{If } x_1 \text{ is } A_{q1} \text{ and } \dots \text{ and } x_n \text{ is } A_{qn} \text{ (1)} \\ \text{then } C_q \text{ with } CF_q$$

R_q 는 q번째 퍼지 규칙이고, $X = (x_1, \dots, x_n)$ 는 n차원의

패턴 벡터이다. A_q 는 조건부 퍼지집합이고, C_q 는 동작부 클래스이다. CF_q 는 규칙의 가중치를 나타낸다.

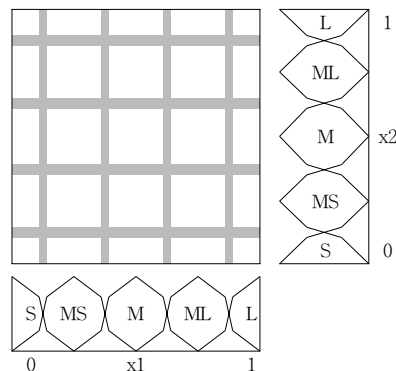


그림 1. 5*5 입력공간
Fig. 1 5*5 input space

S가 그림 1과 같이 15³개의 퍼지규칙의 집합이라고 할 경우에 최적의 규칙을 얻기 위하여 식(2)와 같은 최적화 문제를 고려할 수 있다.

$$\text{Maximize } f_1(S) = \text{minimize } f_2(S), f_3(S) \text{ (2)}$$

여기서 $f_1(S)$ 는 S에 의한 잘 분류된 학습패턴의 수, $f_2(S)$ 는 S에 있는 규칙의 수이고, $f_3(S)$ 는 S에 있는 규칙의 총수이다.

2. 유전알고리즘

유전알고리즘은 목적함수의 최적화를 위하여 자주 이용되어져 왔다^[6]. 본 논문에서는 세 개의 목적함수를 이용하였다. 각각의 규칙에 존재하는 각 스트링을 계산하기 위한 가변의 랜덤 가중치를 가지는 스칼라 평가함수를 통하여 교배와 가중치의 갱신을 수행한다. S에 존재하는 각각의 규칙에 대한 적합도는 식(3)으로 정의 될 수 있다.

$$\text{Fitness}(S) = w_1 f_1(S) - w_2 f_2(S) - w_3 f_3(S) \text{ (3)}$$

여기서 w_1, w_2 , 와 w_3 은 식(4)와 (5)의 조건을 만족하는 가중치이다.

$$w_1, w_2, w_3 \geq 0 \text{ (4)}$$

$$w_1 + w_2 + w_3 \geq 1 \quad (5)$$

모집단 스트링에서 선택이 된 뒤에 이러한 가중치들은 갱신된다. 선택은 식 (3)에 의한 스칼라 적합도 함수 (scalar fitness function)를 이용하여 모집단의 스트링을 선택하는데 이진 토너먼트 선택을 이용하였다. 교배는 새로운 스트링을 만들기 위하여 단일교배(uniform crossover)를 이용하였다. 또한 제어규칙의 수를 줄이기 위하여 편이된 돌연변이(biased mutation) 연산자를 이용하였다. 이 연산에서는 확률을 높이기 위하여 0에서 1로의 돌연변이 보다는 1에서 0으로의 돌연변이를 수행하였다.

최종적으로 얻어진 스트링으로 퍼지 제어 규칙의 조건부와 동작부의 퍼지집합을 구축하여 그림 2의 역전과 신경회로망(back propagation neural network)의 (b)계층의 개수와 계수를 결정하고 학습을 위한 단계로 진입하게 된다.

III. 퍼지 신경망(Fuzzy Neural Network)

1. 퍼지 신경망의 구조

퍼지논리 제어기의 설계는 인간의 경험과 지식에 기초한 추론과정을 모델링하는 것으로 간주할 수 있다. 일반적으로 퍼지규칙의 결론부는 하나의 퍼지집합이다. 여기서 k번째 규칙은 식(6)과 같이 표현된다^[7].

$$\begin{aligned} &IF x_1 \text{ is } A_1^k \text{ and } x_2 \text{ is } A_2^k \text{ and } \dots \text{ and } x_n \text{ is } A_n^k \\ &THEN y \text{ is } \beta_k \end{aligned} \quad (6)$$

여기서 A_{ik} 와 β_k 는 각각 R에 속하는 U와 V의 퍼지 집합이다. $x = (x_1, \dots, x_n) \in U_1 \times \dots \times U_n$ 와 $y \in V$ 는 퍼지추론 시스템의 각각 입력과 출력이다. 위의 식과 같은 퍼지 규칙에 대해 다음에 설명될 추론과정을 신경망으로 구현한 것이 그림 2에 나타나 있다.

퍼지 신경망의 입력은 (a)층과 같이 네 개의 입력 x_1, x_2, x_3 과 x_4 로 이루어져 있다. 그리고 (b)층은 규칙의 조건부에 해당하는 언어변수로서 각 변수에 따른 소속 함수는 퍼지 엔트로피에 의해 분할된 각각의 퍼지집합들로

이루어져 있다. 그리고 소속함수의 전체집합은 정규화되어 있다. $G_i(x_i)$ 은 입력 x_i 이 퍼지변수 G_i 의 언어적 속성에 일치하는 소속정도를 의미한다.

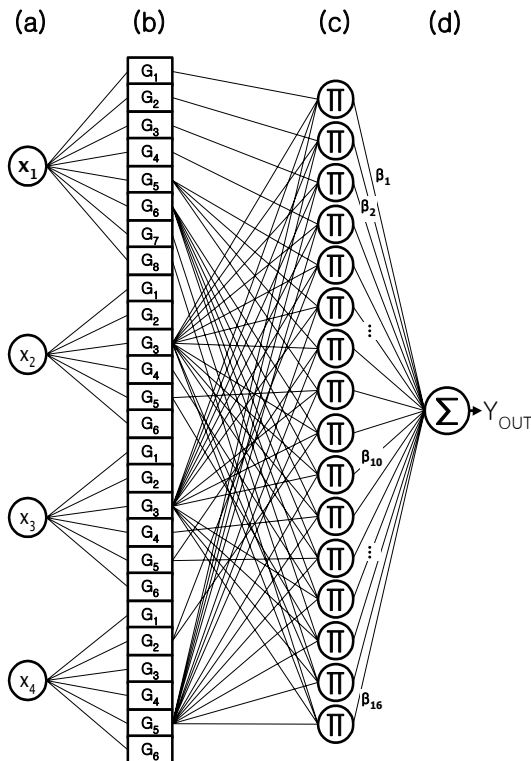


그림 2. 뉴로-퍼지 신경망의 구조
Fig. 2. The architecture of the neuro-fuzzy network

이 소속 함수는 사각형의 적응노드로서 신경회로망에 의해 적용될 중심과 폭의 파라미터를 가진다. (c)층은 규칙층으로 조건부의 해당 퍼지 부공간의 중요도를 출력한다. 결국 퍼지규칙의 진리값이 결정된다. 제어규칙의 결론부는 (c)층과 (d)층을 연결하는 β 로서 결론부의 퍼지 집합을 퍼지 단일 값으로 나타낸다. 결국 각 퍼지규칙의 단일 값에 대해 (c)층의 출력과의 무게 중심법에 의한 비퍼지화의 과정을 통하여 식(7)과 같이 최종 추론 값을 구할 수 있다.

$$y_{out} = \frac{\sum_{k=1}^m \beta^k \left[\prod_{i=1}^n \mu_{A_i^k}(x_i) \right]}{\sum_{k=1}^m \prod_{i=1}^n \mu_{A_i^k}(x_i)} \quad (7)$$

전방향의 신호의 전파와 각 계층의 각 노드들의 기능은 다음과 같다. 여기서 net_{ij} 와 f_{ij} 는 각각 j번째 노드의 입력 가중합과 활성화함수를 나타낸다. 또한 x_{ij} 와 y_{ij} 는 i번째 계층의 j번째 노드의 입력과 출력벡터를 나타낸다. 각 층에 관한 전방향의 전파는 식(8)~(11)과 같다.

$$(a) \text{층} : net_j^1 = x_j = x_j, y_j^1 = f_j^1(net_j^1) = net_j^1 \quad (8)$$

$$(b) \text{층} : (net_{ij}^2) = -\frac{(x_i^2 - c_{ij})^2}{(w_{ij}^2)^2},$$

$$y_{ij}^2 = f_{ij}^2(net_{ij}^2) = \exp(net_{ij}^2) \quad (9)$$

$$(c) \text{층} : net_j^3 = \prod_{i=0}^n x_i^3, y_j^3 = f_j^3(net_j^3) = net_j^3 \quad (10)$$

$$(d) \text{층} : net^4 = \sum_{i=1}^m \beta_i^4 x_i^4,$$

$$y^4 = \frac{f^4 net^4}{\sum_{i=1}^m x_i^4} = \frac{net^4}{\sum_{i=1}^m x_i^4} \quad (11)$$

제안된 시스템의 적응 파라미터는 퍼지 제어규칙의 조건부의 멤버쉽 함수를 결정하는 파라미터 즉, 가우시안 멤버쉽 함수의 평균에 해당하는 중심과 분산에 해당하는 폭이다. 또 하나는 결론부의 퍼지집합의 단일 값을 나타내는 연결강도이다. 시스템이 초기화 된 후 역전과 알고리즘을 이용하여 일단의 학습 패턴에 대해 시스템의 각 파라미터를 적응시킨다. 일반적으로 역전과 신경망의 목적함수는 식(12)와 같이 정의된다^[8,9].

$$E = \frac{1}{2} \sum_{j=1}^p (T_j - O_j)^2 \quad (12)$$

여기서 출력값 O_j 는 j번째 출력 노드 y_j 의 출력이고 T_j 는 목표값이다. 이것을 감안하여 식(10)을 y_j 에 대해 편미분하면 식(13)과 같이 출력노드의 활성화 값에 대한 오차의 편미분 값을 얻게 된다.

$$\frac{\partial E}{\partial y_j} = d_j - y_j \quad (13)$$

이 값이 역전과 학습알고리즘의 역전과 출발점이 된다. 그리고 j번째 출력노드에 의해서 전달되어지는 오차항은 식(12)를 net_{ij} 에 대해 편미분하면 식(14)의 오차항을 얻을 수 있다.

$$\delta_{ij} = \frac{-\partial E}{\partial net_{ij}} = \frac{-\partial E}{\partial y_j} \frac{\partial y_j}{\partial net_{ij}} = \frac{t_j - y_j}{\sum_i x_i} \quad (14)$$

또한 하위 계층에 대한 역전과 과정은 기존의 알고리즘과 동일하다. 따라서 최종적으로 b층의 중심과 폭에 해당하는 적응 파라미터에 대한 c_{ij} 와 w_{ij} 에 대한 적응 규칙은 식(15)와 (16)과 같다. 학습 중에 에러를 줄이고 진동을 줄이기 위하여 모멘텀 항을 사용하였다.

$$c_{ij}(t+1) = c_{ij}(t) + \eta \delta_{ij} \left[\frac{2(x_i - c_{ij})}{w_{ij}^2} \right] + \alpha \Delta c_{ij}(t) \quad (15)$$

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{ij} \left[\frac{2(x_i - c_{ij})^2}{w_{ij}^3} \right] + \alpha \Delta w_{ij}(t) \quad (16)$$

IV. 실험 및 결과

그림 3은 비선형 시계열의 예측 문제에서 많이 쓰이는 Mackey-Glass Chaotic 시계열 Data Set를 식 (17)에 의해 4차 Runge-Kuta법을 이용하여 1200개의 발생된 데이터를 나타낸 것이다. $t < 0$ 인 구간에서는 $x(t) = 0$ 이며 초기값은 1.2로 하였다. 그림 5에서 발생된 1200개의 데이터에서 500개는 학습에 사용하였고, 나머지 600개는 예측에 사용하였다.

$$x(t) = \frac{0.2x(t-\tau)}{1 + x^{10}(t-\tau)} - 0.1x(t) \quad (17)$$

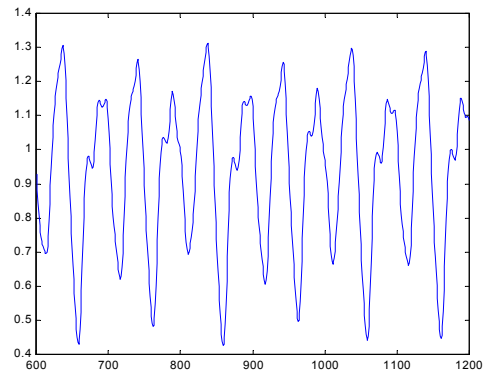


그림 3. Mackey-Glass 시계열 데이터 열
Fig. 3. Mackey-Glass time series data set

입출력은 표 1과 같이 시계열에서 t=0에서 입력 x_1, x_2 ,

x_3, x_4 를 구성하고 출력은 y_1 으로 구성하였다. 각 패턴에 따라 간격은 3, 6과 9간격을 유지 하도록 구성 하였다. 제안된 시스템의 성능평가를 위하여 시간축 상의 간격을 세 단계로 나누어 시스템의 성능을 측정하였다.

표 1 입출력 패턴의 구성
Table. 1. The structure of input/output pattern

	입 력				출 력
	x_1	x_2	x_3	x_4	
패턴 1	$x(t-9)$	$x(t-6)$	$x(t-3)$	$x(t)$	$x(t+3)$
패턴 2	$x(t-18)$	$x(t-12)$	$x(t-6)$	$x(t)$	$x(t+6)$
패턴 3	$x(t-27)$	$x(t-18)$	$x(t-9)$	$x(t)$	$x(t+9)$

그림 4는 유전알고리즘에 의하여 동정된 규칙과 멤버쉽 함수를 가지고 학습을 완료한 후에 적용된 멤버쉽 함수를 나타내고 있다. 퍼지 규칙에서 사용된 조건부의 적용 파라미터는 30개이고, 결론부는 16개로 총 46개의 파라미터이다.

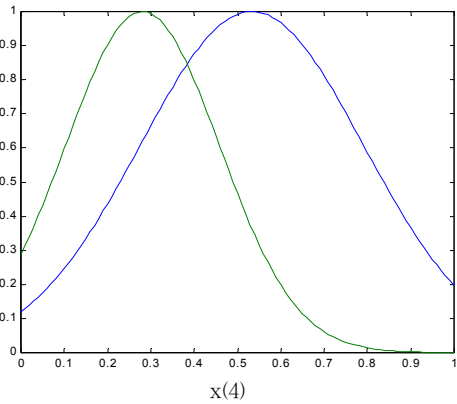
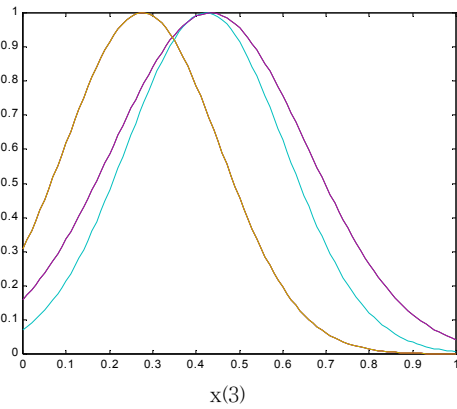
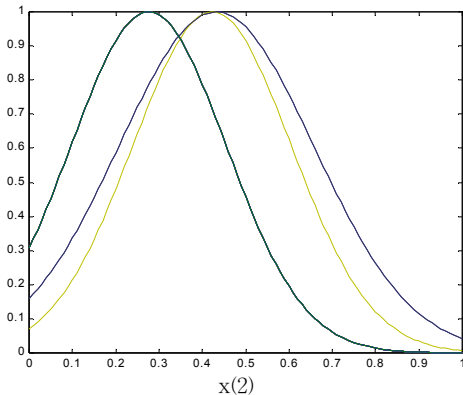
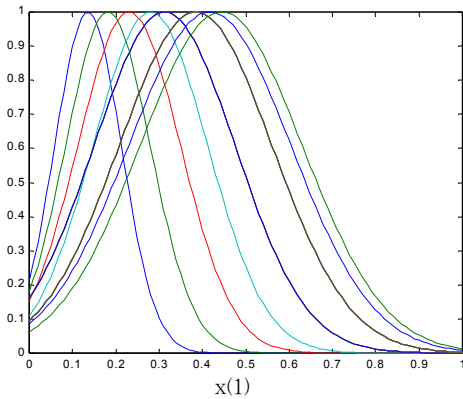


그림 4. 학습이전의 멤버쉽 함수
Fig. 4. Membership function before learning

그림 5, 6과 7에는 시계열의 예측결과를 나타내었다. 실선이 원래의 값이고 점선이 예측된 값이다.

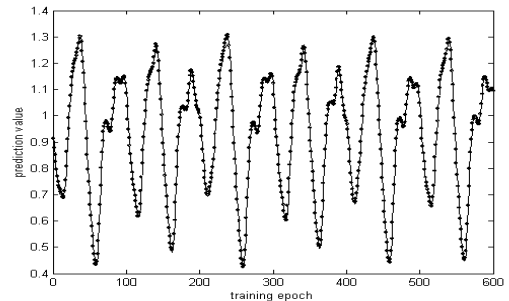


그림 5. 간격이 3인 경우의 예측결과
Fig. 5. The prediction of time interval 3

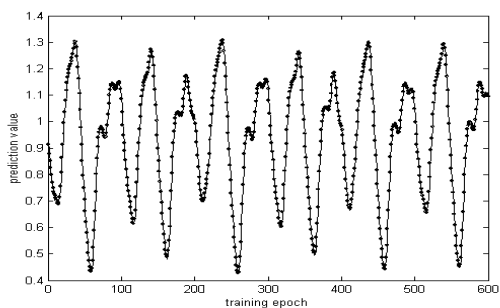


그림 6. 간격이 6인 경우의 예측결과
Fig. 6. The prediction of time interval 6

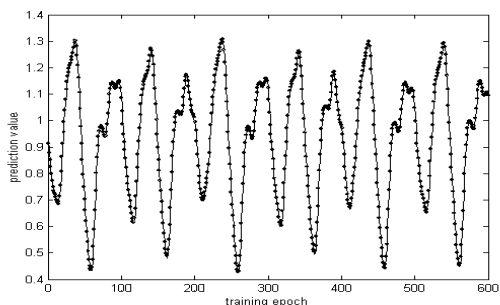


그림 7. 간격이 9인 경우의 예측결과
Fig. 7. The prediction of time interval 9

그림을 통하여 학습 초기에 상당히 오차가 적은 상태에서 시작하는 것을 알 수 있는데 이는 유전 알고리즘을 통하여 시스템을 기술하는 제어규칙을 구조적으로 비교적 잘 동정한 것임을 보여주고 있다.

또한 제안된 시스템의 정확도를 평가하기 위하여 실제 데이터 출력과 동정모델의 출력간의 차의 평균을 제공하는 RMSE(root mean square error)를 이용하였다.

$$R.M.S.E = \sqrt{1/N \sum_{k=1}^N (T(k) - O(k))^2} \quad (16)$$

여기에서 N은 사용된 데이터 수, T(k)는 실제 데이터의 출력 값이고 O(k)는 동정된 모델로부터 추론된 값이다. 표 2에는 동정용 데이터를 사용하여 식(16)에 의해 계산된 RMSE 학습 오차를 나타내었다. 동정된 오차에서도 기존의 방법보다 제안된 방법의 비교우위를 확인할 수 있었다.

표 2. 학습시의 RMSE

Table 2. RMSE of training

	RMSE(root mean square error)		
	ANFIS	WANG	제안된방법
간격 x(t-3)	0.000426	0.000420	0.000415
간격 x(t-6)	0.000228	0.000228	0.000112
간격 x(t-9)	0.00021	0.00020	0.00013

V. 결론

본 논문에서는 비선형 시계열 시스템으로부터 발생한 시계열 데이터를 뉴로퍼지 모델을 사용하여 예측하는 기법을 제안하였다. 모의 실험에서 시계열 데이터로 Mackey-Glass지연 미분 방정식을 사용하였으며, 실험을 통해 제안된 모델은 입력공간의 효율적인 분할과 마라미터의 학습을 통해 시계열 데이터를 정확하게 예측할 수 있었다. 시간간격이 커질수록 예측의 척도는 작아지는 것을 알 수 있었다. 앞으로도 시계열 데이터의 종류의 특성에 따라 다양한 예측 기법들이 연구되어야 할 것이다.

참고 문헌

- [1] Li-Xin Wang and Jerry M. Mendal, "Generating fuzzy rules by learning from examples", IEEE Trans. SMC, vol. 22, no. 6, pp. 1414- 1427, 1992.
- [2] H. Ichihashi, "Iterative fuzzy modeling and a hierarchical network", Proc. of the 4th IFSA Congress, vol. Eng., Brussels, 49-52, 1991.
- [3] D. Araki, H. Nomura, I. Hayashi and N. Wakami, "A self generating method of fuzzy inference rules", Fuzzy engineering toward human friendly systems, pp. 1047-1058, eds. T. Terano et al., 1991.
- [4] M. Sugeno and G. T. Kng, "Structure identification of fuzzy model", Fuzzy sets and systems, vol. 28, pp. 15-23, 1988.
- [5] H. Nomura, I. Hayashi and N. Wakami, "A learning method of fuzzy inference rules by descent method", IEEE Int. Conf. on Fuzzy Systems(San

Diego, 1992) 203, 210.

- [6] Hisao Ishibuchi and T. Yamamoto, "Fuzzy Rule Selection by Data Mining Criteria and Genetic Algorithm", H. Ishibuchi et.al. Genetic and Evolutionary Computation ..., 2002 - Citeseer
- [7] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control", IEEE Trans. Syst. Man, Cybern., vol. 15, pp. 116-132, 1985.
- [8] W. Pedrycz, "Fuzzy Control and Fuzzy Systems", New York: Wiley, 1989.
- [9] Chin-Teng Lin and C. S. George Lee, "Neural Fuzzy Systems", Prentice-Hall, 19

저자 소개

박 인 규(정회원)



- 제10권 5호 참조
- 현 중부대학교 컴퓨터학과 교수
- <주관심분야 : 영상처리, 인공지능>