

자기 회귀 웨이블릿 신경 회로망을 이용한 비선형 혼돈 시계열의 예측에 관한 연구

\*이혜진, \*박진배, \*\* 최윤호  
 \*연세대학교 전기전자공학부, \*\*경기대학교 전자공학부

A Study on the Prediction of the Nonlinear Chaotic Time Series Using a Self-Recurrent Wavelet Neural Network

\*Hye Jin Lee, \*Jin Bae Park, \*\*Yoon Ho Choi  
 \*Dept. of Electrical and Electronic Eng. Yonsei Univ., \*\*School of Electronic Eng. Kyonggi Univ.

**Abstract** - Unlike the wavelet neural network, since a mother wavelet layer of the self-recurrent wavelet neural network(SRWNN) is composed of self-feedback neurons, it has the ability to store past information of the wavelet. Therefore we propose the prediction method for the nonlinear chaotic time series model using a SRWNN. The SRWNN model is learned for the modeling of a function such that the inputs are known values of the time series and the output is the value in the future. The parameters of the network are tuned to minimize the difference between the nonlinear mapping of the chaotic time series and the output of SRWNN using the gradient-descent method for the adaptive backpropagation algorithm. Through the computer simulations, we demonstrate the feasibility and the effectiveness of our method for the prediction of the logistic map and the Mackey-Glass delay-differential equation as a nonlinear chaotic time series.

RNN의 동특성과 WNN의 빠른 수렴성을 갖는 자기 회귀 웨이블릿 신경 회로망(SRWNN)을 이용한 혼돈 시계열 예측 방법을 제안한다. SRWNN은 WNN의 수정된 모델로, 은닉층이 자기 회귀 뉴런으로 된 모체 웨이블릿으로 구성되어 있으며 자기 회귀 뉴런은 신경망의 과거 정보를 저장할 수 있기 때문에 동역학 시스템에서 발생한 혼돈 시계열의 특성을 정확히 해석할 수 있다.

2. 혼돈 시계열 모델

비선형 시스템의 상태 변수들간의 비선형적인 관계로 인하여 불규칙하게 보이는 현상을 혼돈 현상이라고 하는데, 이것은 단순히 불규칙한 해를 갖는 결정론적 방정식의 특성을 의미하므로 불규칙성이 비교적 단순한 미분 방정식에 의해 설명될 수 있다. 그러므로 비주기적인 변동을 나타내는 동력학적 시스템을 해석하는데 유용하고, 단순한 결정론적 방정식을 이용하여 복잡한 변동을 모델링하는 것이 가능하다면 최소한 단기간의 미래 변동을 예측하고 제어하는 것이 가능하다.

이러한 혼돈 현상을 갖는 시계열 데이터의 정량적 특성을 나타내는 대표적인 지표로는 상관 차원(correlation dimension)과 Lyapunov 지수가 있다[10]. 모의 실험에서는 혼돈적인 비선형 시스템과 시계열 모델링의 평가 기준으로 가장 많이 사용되는 Logistic map과 Mackey-Glass 지연 미분 방정식으로부터 발생시킨 시계열을 사용하였으며 각각 다음과 같이 정의된다.

$$x_{n+1} = rx_n(1 - x_n) \tag{1}$$

$$\frac{dx}{dt} = \frac{ax(t-\Gamma)}{1 + x^{10}(t-\Gamma)} - bx(t) \tag{2}$$

두 혼돈 시계열 모델의 상관 차원과 최대 Lyapunov 지수는 표 1과 같이 비교할 수 있다.

표 1. 추정된 상관 차원과 최대 Lyapunov 지수

추정된 값	Logistic map	Mackey Glass 시계열
상관 차원	0.9 ~ 1.0	1.7 ~ 2.1
최대 Lyapunov 지수	0.6985	0.0073

표 1을 보면 상관 차원은 시스템의 상태 변수와 연관되므로 Logistic map의 경우는 단순한 시스템인 반면 Mackey-Glass 시계열은 비교적 복잡하다는 것을 알 수 있다. 또한 최대 Lyapunov 지수는 예측 오차의 발산율과 연관되므로 다 단계 예측에 있어서 Logistic map은 처음 단계의 예측 오차가 빨리 발산하여 신뢰할 수 있는 예측 시간 구간이 비교적 짧은 반면 Mackey-Glass 시계열은 신뢰할 수 있는 예측 시간 구간이 다소 길다는 것을 확인할 수 있다.

3. SRWNN을 이용한 시계열 예측

1. 서 론

시계열 예측이란 시계열 데이터가 과거 관찰값들의 변화 패턴과 비슷하게 현재 시점 이후에도 변화될 것이라는 전제하에서 미래 시점의 값을 추정하는 기법이다. 기존의 시계열 예측 방법으로는 ARMA, ARIMA 모델과 같은 통계적 기법과 퍼지 이론, 유전 알고리즘, 신경 회로망(NN) 등의 기능 기법들이 활발히 연구되어 왔다 [1]-[3]. 특히 판별력과 예측력이 우수한 것으로 밝혀진 신경 회로망을 활용한 예측이 많이 시도되고 있다. 그러나 NN은 학습시간이 길고 학습하는 동안 국부 최소치(local minimum)에 빠질 수 있으며, 혼돈 시계열의 혼돈 특성을 정확히 반영하지 못하는 단점이 있다. 따라서 이러한 단점을 극복하기 위한 방법으로 방사 기저 함수 회로망(RBFN)과 웨이블릿 신경 회로망(WNN) 등이 제시되어 혼돈 시계열의 예측에 대한 많은 연구가 진행되어 왔다[4]-[6].

RBFN은 입력층과 은닉층이 선형으로 연결되어 있기 때문에 구조가 간결하여 수학적 표현이 명료하고 역행력 방법을 이용할 수 있기 때문에 연산 속도가 매우 빠른 장점이 있다[7]. WNN은 데이터를 분석하는 해상도를 진폭축과 시간축에 따라 다양하게 구성하기 때문에 시계열 데이터들의 장기적 패턴 뿐만 아니라 미세한 불규칙적인 패턴도 동시에 해석해낼 수 있다. 또한 한 개의 은닉층으로 구성되어 구조가 단순하며 빠르게 수렴하는 장점이 있다[8].

그러나 NN, RBFN, WNN 모형 모두 정적 사상(static mapping)이므로 동역학 시스템에서 발생한 혼돈 시계열을 정확히 나타내지 못한다. 따라서 은닉층 각각의 뉴런에서 현재의 결과값을 저장하고 다음 학습 단계에서 사용함으로써 현재의 동특성(dynamics)을 보존하는 자기 회귀 신경 회로망(RNN)이 제안되었다[9]. 본 논문에서는

혼돈 시계열 예측을 위해 본 논문에서 제안한 SRWNN의 구조는 그림 1과 같다.  $N_i$ 개의 입력과 1개의 출력, 그리고  $N_w$ 개의 웨이블릿 노드를 가지고 있으며 입력층(A층), 모체 웨이블릿 층(B층), 다차원 웨이블릿 층(C층), 출력층(D층)으로 구성된다. 모체 웨이블릿 층의 각 노드는 모체 웨이블릿과 자기 회귀 루프로 이루어지며, 모체 웨이블릿은 식 (3)과 같이 가우시안(Gaussian) 함수의 1차 미분 방정식으로 구성된다.

$$\phi(x) = -x \exp\left(-\frac{1}{2}x^2\right) \quad (3)$$

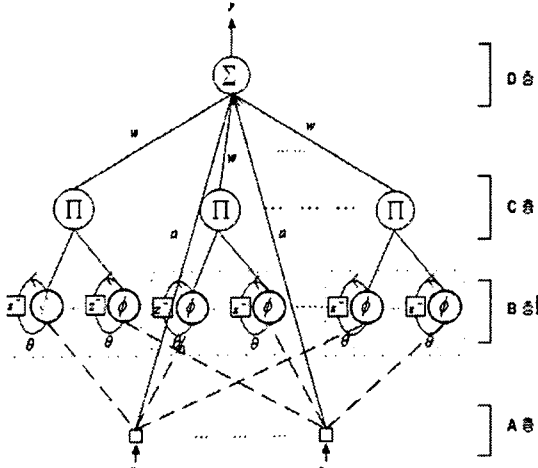


그림 1. SRWNN의 구조

A층의  $N_i$ 개의 입력에 대한 B층의 웨이블릿  $\phi_{jk}$ 는 식 (4)과 같이 정의된다.

$$\phi_{jk} = \phi\left(\frac{u_{jk} - m_{jk}}{d_{jk}}\right), \quad \text{with } z_{jk} = \frac{u_{jk} - m_{jk}}{d_{jk}} \quad (4)$$

여기서  $m_{jk}$ 와  $d_{jk}$ 는 각각 웨이블릿의 변형 인자와 확장 인자이며  $jk$ 는  $j$ 번째 웨이블릿의  $k$ 번째 입력을 나타낸다.

B층의 각 웨이블릿에 대한 입력은 식 (5)과 같이 정의된다.

$$u_{jk}(n) = x_k(n) + \phi_{jk}(n-1) \cdot \theta_{jk} \quad (5)$$

여기서  $\theta_{jk}$ 는 자기 회귀 루프의 가중치를 나타내며,  $\theta_{jk}(n-1)$ 은 신경망의 과거 정보를 저장하는 메모리 부분으로 WNN과 SRWNN 사이의 차이를 보여준다.

C층은 모체 웨이블릿의 곱으로써 식 (6)과 같이 나타낼 수 있다.

$$\begin{aligned} \Phi_j(X) &= \prod_{k=1}^{N_i} \phi(z_{jk}) \\ &= \prod_{k=1}^{N_i} \left[ -\left(\frac{u_{jk} - m_{jk}}{d_{jk}}\right) \exp\left(-\frac{1}{2}\left(\frac{u_{jk} - m_{jk}}{d_{jk}}\right)^2\right) \right] \end{aligned} \quad (6)$$

D층은 출력층으로 최종 출력은 식 (7)과 같이 나타낼 수 있다.

$$y(n) = \sum_{j=1}^{N_w} w_j \Phi_j(X) + \sum_{k=1}^{N_i} a_k x_k \quad (7)$$

여기서  $w_j$ 는 C층과 D층사이의 연결 가중치이고  $a_j$ 는 A층과 D층 즉, 입력층과 출력층 사이의 가중치이다.

SRWNN에서 학습을 통하여 갱신되는 파라미터들의 집합  $W$ 는 다음과 같다.

$$W = [a_k, m_{jk}, d_{jk}, \theta_{jk}, w_j]^T \quad (8)$$

본 논문에서는 학습 방법으로 식 (9)와 같은 비용함수를 최소화하는 과정을 통해 파라미터들을 갱신한다.

$$J(n) = \frac{1}{2} [y_c(n) - y(n)]^2 = \frac{1}{2} e^2(n) \quad (9)$$

여기서  $y_c(n)$ 은 혼돈 비선형 시스템에서 발생한  $n$ 번째 시계열이고,  $y(n)$ 은 SRWNN의  $n$ 번째 출력이다.

최소화 과정은 반복적인 경사하강법(gradient-descent method)을 사용하며 식 (10)과 같은 방법으로 파라미터들을 갱신한다.

$$\begin{aligned} W(n+1) &= W(n) + \Delta W(n) \\ &= W(n) + \eta \left( -\frac{\partial J(n)}{\partial W(n)} \right) \end{aligned} \quad (10)$$

여기서  $\eta$ 는 SRWNN 파라미터의 학습률을 의미하고,  $\frac{\partial J(n)}{\partial W(n)}$ 는  $W(n)$ 에 대한 비용함수의 편미분으로 다음과 같이 계산한다.

$$\frac{\partial J(n)}{\partial W(n)} = e(n) \frac{\partial e(n)}{\partial W(n)} = -e(n) \frac{\partial y(n)}{\partial W(n)} \quad (11)$$

식 (10)과 식 (11)에 의해 식 (8)의 각 파라미터들은 다음과 같이 갱신된다.

$$\frac{\partial y(n)}{\partial a_k(n)} = x_k \quad (12)$$

$$\frac{\partial y(n)}{\partial m_{jk}(n)} = -\frac{w_j}{d_{jk}} \frac{\partial \Phi_j(X)}{\partial z_{jk}} \quad (13)$$

$$\frac{\partial y(n)}{\partial d_{jk}(n)} = -\frac{w_j}{d_{jk}^2} z_{jk} \frac{\partial \Phi_j(X)}{\partial z_{jk}} \quad (14)$$

$$\frac{\partial y(n)}{\partial \theta_{jk}(n)} = \frac{w_j}{d_{jk}} \phi_{jk}(n-1) \frac{\partial \Phi_j(X)}{\partial z_{jk}} \quad (15)$$

$$\frac{\partial y(n)}{\partial w_j(n)} = \Phi_j(X) \quad (16)$$

여기서  $\frac{\partial \Phi_j(X)}{\partial z_{jk}}$ 와  $\phi(z_{jk})$ 는 다음과 같다.

$$\frac{\partial \Phi_j(X)}{\partial z_{jk}} = \phi(z_{j1}) \phi(z_{j2}) \cdots \phi(z_{jk}) \cdots \phi(z_{jN_i}) \quad (17)$$

$$\phi(z_{jk}) = \frac{\partial \phi_j}{\partial z_{jk}} = (z_{jk}^2 - 1) \exp\left(-\frac{1}{2}z_{jk}^2\right) \quad (18)$$

#### 4. 모의 실험

본 논문의 모의 실험에서는 SRWNN을 이용한 비선형 혼돈 시계열 예측의 성능을 평가하기 위해 2장에서 소개한 Logistic map과 Mackey-Glass 지연 미분 방정식에 대하여 기존의 예측 방법인 RBFN 및 WNN의 성능과 비교 실험하였다. 예측의 정확성을 평가하기 위해 평균 제곱오차(MSE)와 결정 계수(coefficient of determination)를 다음과 같이 계산하였다.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (19)$$

$$\text{결정 계수} : r^2 = 1 - \frac{\sum_{i=1}^n (x_i - \bar{x}_i)^2}{\sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (20)$$

여기서  $x_i$  및  $\hat{x}_i$ 는 각각 실제 시계열 및 예측된 값을 나타내며, 예측 오차는  $(x_i - \hat{x}_i)$ 이다. 또한  $\bar{x}_i$ 는 실제 시계열 값들의 평균이다. 결정 계수는 실제 데이터의 분산에 의해 정규화된 MSE의 함수이며 예측의 정확성을 평가할 때 중요한 지표로서 1에 가까울수록 예측 정확성이 높다.

2장에서 서술한 것처럼 Logistic map은 신뢰할 수 있는 구간이 비교적 짧기 때문에 시간 지연을 1로 하여 식 (21)과 같은 방법으로 예측하였다.

$$x(t+1) = F(x(t), x(t-1), x(t-2), x(t-3)) \quad (21)$$

모의 실험 조건에서 SRWNN 구조에 대하여  $\eta=0.35$ ,  $N_f=4$ ,  $N_w=1$ 로 하였고, 샘플링 시간은 1초로 하였다.

Logistic map의 예측에 대하여 RBFN, WNN과 SRWNN을 비교한 MSE와 결정 계수의 결과는 표 2와 같으며, SRWNN을 이용한 예측 결과는 그림 2와 같다.

표 2. MSE와 결정 계수 비교 (Logistic map)

	MSE	결정 계수
RBFN	0.1371	0.7406
WNN	0.0010	0.9980
SRWNN	0.00035	0.9993

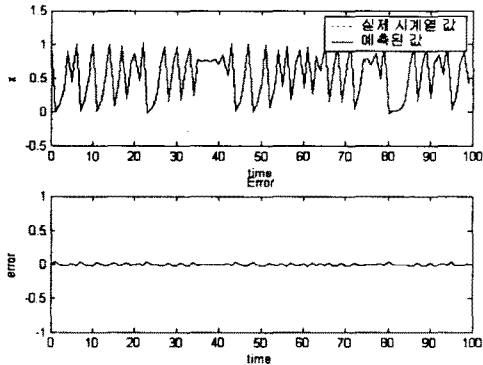


그림 2. Logistic map 예측 결과  
( $x(1) = 0.3$ ,  $x(t) = 0$  for  $t < 0$ , and  $r = 4$ )

Mackey-Glass 지연 미분 방정식은 신뢰할 수 있는 구간이 비교적 길기 때문에 시간 지연을 6으로 하여 식 (22)와 같은 방법으로 예측하였다.

$$x(t+6) = F(x(t), x(t-6), x(t-12), x(t-18)) \quad (22)$$

Mackey-Glass 지연 미분 방정식의 예측 결과는 표 3과 그림 3과 같다.

표 3. MSE와 결정 계수 비교 (Mackey-Glass 시계열)

	MSE	결정 계수
RBFN	0.0015	0.9984
WNN	0.0012	0.9987
SRWNN	0.000414	0.9995

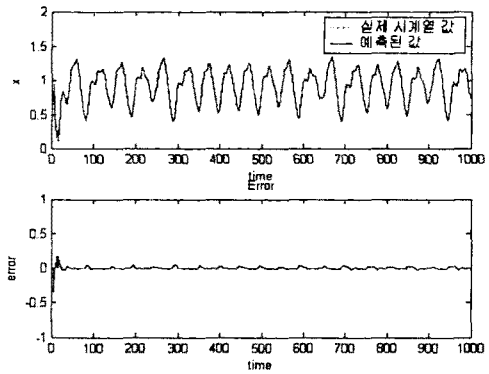


그림 3. Mackey-Glass 시계열 예측 결과  
( $x(1) = 0.3$ ,  $x(t) = 0$  for  $t < 0$ ,

## 5. 결 론

본 논문에서는 비선형 혼돈 시스템으로부터 발생한 시계열 데이터를 SRWNN 모델을 사용하여 예측하는 기법을 제안하였다. 모의 실험에서 시계열 데이터로 Logistic map과 Mackey-Glass 지연 미분 방정식을 사용하였으며, 실험을 통해 SRWNN은 자기 회귀 뉴런으로 인한 동특성과 WNN의 빠른 수렴성으로 다른 신경망 모델보다 시계열 데이터를 더 정확하게 예측할 수 있음을 확인하였다. 그러나 여전히 모든 종류의 시계열 데이터에 우월한 단 하나의 예측 기법은 존재하지 않으며, 복잡하거나 예측 기법간의 결합이라고 해서 단순한 기법보다 항상 정확한 것은 아니다. 따라서 앞으로도 시계열 데이터의 종류의 특성에 따라 다양한 예측 기법들을 연구되어야 할 것이다.

감사의글: 본 논문은 한국과학재단의 목적기초연구사업(R01-2001-000-00316)에 의해 지원되었습니다.

## [참 고 문 헌]

- [1] G. Box, G. Jenkins, and G. Reinsel, *Time Series Analysis Forecasting and Control, 3rd ed.*, Englewood Cliffs, N.J., Prentice Hall, 1994.
- [2] D.A. White, and D.A. Sojga, *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, Van Nostrand Reinhold, New York, 1992.
- [3] J. Nie, "A Fuzzy Neural Approach to Time Series Prediction", *IEEE International Conference on Neural Networks*, Vol. 5, pp. 3164-3169, 1994.
- [4] J. Moody and C. J. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units", *Neural Computation*, Vol. 1, pp. 281-294, 1989.
- [5] X. D. He and A. Lapedes, "Nonlinear Modeling and Prediction by Successive Approximation using Radial Basis Functions", *Physica D*, Vol. 70, pp. 289-301, 1994.
- [6] L. Cao, and Y. Hong, "Prediction Chaotic Time Series with Wavelet Networks", *Physica D*, Vol. 85, pp.225-237, 1995.
- [7] M. J. L. Orr, *Introduction to Radial Basis Function Networks*, Centre for Cognitive Science, 1996.
- [8] J. Zhang, G. G. Walter, Y. Miao, and W. N. W. Lee, "Wavelet Neural Networks for Function Learning," *IEEE Trans. on Signal Processing*, Vol. 43, No. 6, pp. 1485-1497, 1995.
- [9] C. C. Ku and K. Y. Lee, "Diagonal Neural Networks for Dynamics Systems Control", *IEEE Trans. Neural Networks*, Vol. 6, pp. 144-156, 1995.
- [10] M. Casdagli, "Nonlinear Prediction of Chaotic Time Series", *Physica D*, Vol. 35, pp. 335-356, 1989.