

# 혼돈 시계열의 예측을 위한 Radial Basis 함수 회로망 설계

論 文

45~4~21

## Radial Basis Function Network Design for Chaotic Time Series Prediction

申 昌 溶\* · 金 宅 洙\*\* · 崔 允 浩\*\*\* · 朴 相 曠\*\*\*\*

(Chang-Yong Shin · Taek-Soo Kim · Yoon-Ho Choi · Sang-Hui Park)

**Abstract** - In this paper, radial basis function networks with two hidden layers, which employ the K-means clustering method and the hierarchical training, are proposed for improving the short-term predictability of chaotic time series. Furthermore the recursive training method of radial basis function network using the recursive modified Gram-Schmidt algorithm is proposed for the purpose. In addition, the radial basis function networks trained by the proposed training methods are compared with the X.D. He and A. Lapedes's model and the radial basis function network by nonrecursive training method. Through this comparison, an improved radial basis function network for predicting chaotic time series is presented.

**Key Words** : short-term prediction of chaotic time series, radial basis function network, recursive modified Gram-Schmidt algorithm, recursive training method, K-means clustering method

### 1. 서 론

예측(prediction)은 미지의 시스템의 과거 데이터로부터 미래의 상태를 추정하는 본래의 목적 이외에도 데이터를 발생시키는 시스템에 대한 중요한 정보를 제공한다[1]. 예측 문제에 대한 접근 방법으로서 데이터를 발생시키는 동력학 시스템을 결정하는 미분 방정식을 도출하고, 초기 상태 조건을 측정 한 후, 이 미분 방정식에 대한 해를 구하여 미래를 예측하는 방법이 있다. 그러나 실제로 접할 수 있는 것은 미지의 동력학 시스템의 일정한 관측점을 통해 얻은 시계열(time series) 데이터므로 위와 같은 방법을 사용하기 위한 기본 원리와 초기 상태 조건이 부족하거나 이용할 수 없는 경우가 많다.

한편, 관측 시계열의 랜덤성이 확률적인 시스템(stochastic system)이 아닌 결정론적 시스템(deterministic system)에 의해 발생한다고 보는 혼돈 이론에 근거하여 결정론적인 동력학 시스템을 모델링하여 예측하는 방법은 이러한 문제에 대한 하나의 해결 방법을 줄 수 있다. 혼돈적 특성을 갖는 결정론적 시스템에서의 동력학은 초기의 작은 차이가 증폭되어 결국에는 예측할 수 없는 상태가 되므로 혼돈 시계열의 장기 예측(long-term prediction)에는 근본적인 제한이 있지만 결정론적인 동력학 시스템의 모델링을 통하여 예측 한다면 단기 예측(short-term prediction)은 가능할 수 있다[2].

이러한 혼돈 시계열을 모델링하기 위해 여러가지 방법이 사용되었다[3~6]. 특히 그중에서도 최근 신경 회로망을 사용하기 시작하고 있으며 특히 A. Lapedes와 R. Farber[3]는 역전파 학

습 알고리즘(back propagation training algorithm)을 이용한 다층 퍼셉트론 회로망(multi-layer perceptron network)에 의한 혼돈 시계열의 예측을 시도하였으나, 이 방법은 학습시간이 길고 학습하는 동안 국부 최소치(local minimum value)에 빠질 수 있는 단점을 가진다. 한편, 이러한 단점을 극복하기 위한 신경 회로망 모델로서 radial basis 함수 회로망이 제시되어 이를 이용한 혼돈 시계열의 예측에 대한 많은 연구가 진행되고 있다 [4, 6, 7].

본 논문에서는 혼돈 시계열의 단기 예측 성능을 향상시키기 위해 X. D. He와 A. Lapedes가 제안한 radial basis 함수 회로망[6]을 기초로 K-means 클러스터링(clustering) 방법[8]과 계층적 학습[9]을 이용한 회로망과 회귀적인 변형 Gram-Schmidt 알고리즘[1, 10]을 이용한 radial basis 함수 회로망의 학습 방법을 제안한다. 그리고 혼돈 시계열의 단기 예측에 대해 제안된 학습 방법들로 학습한 회로망을 A. Lapedes 등의 모델 및 비회귀 학습 radial basis 함수 회로망과 비교하여 보다 개선된 radial basis 함수 회로망을 제시하고자 한다.

### 2. 혼돈 시계열

대부분의 시스템은 근사적으로 유한 차원의 결정론적 동력학 시스템으로 모델링할 수 있으며, 이 때 주어진 순간의 시스템 상태는  $k$ 차원 공간  $R^k$ 의 부분 집합인  $k$ 차원 어트랙터(tractor)  $M$ 상에 놓여 있는 한 점  $\vec{z}$ 에 의해 결정된다. 이 시스템의 시간 경과(time evolution)는 시간  $t_0$ 에서 시스템 상태가  $\vec{z}$ 이면 시간  $t_0 + t$ 에서의 상태는  $f^t(\vec{z})$ 인 사상  $f^t: M \rightarrow M$ 에 의해 주어진다. 그러나 실제의 경우에 있어서 상태  $\vec{z}$ 에 접근할 수 없고, 단지 이 상태의 한 함수  $\varphi(\vec{z})$ 를 관측할 수 있는데 이 함수  $\varphi: M \rightarrow R$ 은 시스템의 관측 가능한 성질의 측정을 나타내므로 측정 함수(measurement function)라고 부른다. 따라서 시간에 따른 이 측정치의 진행은  $\varphi(f^t(\vec{z}))$ 에 의

\* 準 會 員 : 延世大學校 工大 電氣工學科 碩士課程  
\*\* 正 會 員 : 延世大學校 工大 電氣工學科 博士課程修了  
\*\*\* 正 會 員 : 京畿大學校 工大 電子工學科 助教授, 工博  
\*\*\*\* 正 會 員 : 延世大學校 工大 電氣工學科 教授, 工博  
接受日字 : 1995年 9月 25日  
最終完了 : 1996年 3月 15日

해 주어지고 단지 이산적인 시간 구간에서 관측할 수 있으므로 이 구간들이 모두 균일하다고 가정하면 관측된 시계열은 식(1)과 같이 표현할 수 있다[1, 6].

$$x_n = \varphi(\vec{z}_n), \quad n = 1, 2, \dots \quad (1)$$

여기서  $\vec{z}_n = f^{\tau n}(\vec{z})$ 이고  $\tau$ 는 샘플링 시간이다.

한편  $\vec{z}_n$ 은 일반적으로 고차원 공간에 존재하는 반면 시계열  $\{x_n\}$ 이 일차원적 데이터이기 때문에  $x_n$ 이  $\vec{z}_n$ 의 상태에 관해 상대적으로 적은 정보를 포함하고 있고  $x_n$ 이 결정론적 과정(deterministic process)  $f^t$ 로부터 발생한다는 사실로부터 좌표 변환에 의해  $f^t$ 를 재구성하는 것이 가능하다는 것이 증명되었다[11]. 즉, 임베딩 차원(embedding dimension)  $d$ 에 대해 "지연 좌표 임베딩(delay coordinate embedding)"인 식(2)와 같은 사상  $\phi: M \rightarrow R^d$ 를 정의하면,

$$\begin{aligned} \vec{x}_n &= \phi(\vec{z}_n) \\ &= (\varphi(\vec{z}_n), \varphi(f^{-\tau}(\vec{z}_n)), \dots, \varphi(f^{-(d-1)\tau}(\vec{z}_n))) \\ &= (x_n, x_{n-1}, \dots, x_{n-(d-1)}) \end{aligned} \quad (2)$$

Takens의 정리에 의하면  $\phi$ 의 상(image)에 대해 변환  $\vec{F} = \phi \circ f \circ \phi^{-1}$ 를 정의할 수 있으므로  $\vec{F}$ 를  $\phi$ 에 의한 좌표 변환하에서  $f^t$ 와 동일한 동력학 시스템으로 간주할 수 있다. 특히  $\vec{F}$ 와  $f^t$ 의 모든 좌표 독립적 성질은 동일하다[1, 6].

위에서 언급한 바와 같이  $\vec{z}_n = f^{\tau n}(\vec{z})$ 인  $x_n = \varphi(\vec{z}_n)$ 에 의해 주어진 관측 시계열로부터 점  $\vec{y}_n \in R^d$ 를 지연 좌표(delay coordinate)로 정의하면 식(3)과 같이 쓸 수 있다.

$$\vec{y}_n = (x_n, x_{n-1}, \dots, x_{n-(d-1)}) \quad (3)$$

식(2)와 식(3)에 의해  $\vec{y}_n = \phi(\vec{z}_n)$  이고 따라서  $y_n$ 은  $\phi$ 의 상이다. 그러므로  $\vec{F}$ 를 식(4)와 같이  $\vec{y}_n$ 에 적용 할 수 있다.

$$\begin{aligned} \vec{F}(\vec{y}_n) &= \phi \circ f^t \circ \phi^{-1}(\vec{y}_n) \\ &= \phi \circ f^t(\vec{z}_n) = \phi(\vec{z}_{n+1}) = \vec{y}_{n+1} \end{aligned} \quad (4)$$

따라서  $\vec{F}$ 는  $R^d$ 에 있어서  $\vec{y}_n$ 의 동력학을 나타내고 이 동력학은 단지  $f^t$ 하에서  $\vec{z}_n$ 의 동력학의 가역적(invertible) 좌표 변환  $\phi$ 에 의한 상이다. 즉  $\vec{y}_n$ 의 동력학은 평활한(smooth) 가역적 좌표 변환에 의해 연관된  $\vec{z}_n$ 의 동력학과 동가가 된다 [1, 6]. 그러나  $\vec{z}_n$ 의 상태에는 접근할 수 없는 반면에  $\vec{y}_n$ 의 동력학은 시계열  $\{x_n\}$ 의 진행으로부터 관측할 수 있다. 따라서  $x_n$ 의 관점에서 식(4)를 식(5)와 같이 다시 쓸 수 있다.

$$\vec{F}(x_n, x_{n-1}, \dots, x_{n-d+1}) = (x_{n+1}, x_n, \dots, x_{n-d+2}) \quad (5)$$

그러므로 원래의  $f^t$ 와는 달리 변환  $\vec{F}$ 는 시계열  $\{x_n\}$ 의

샘플로부터 추정될 수 있다. 또한  $\vec{F}$ 를 첫번째 좌표 성분으로 제한하면  $x_n, x_{n-1}, \dots, x_{n-d+1}$ 값에 의해  $x_{n+1}$ 을 결정하는 식(6)과 같은 사상  $G: R^d \rightarrow R$ 을 얻을 수 있다[1].

$$G(x_n, x_{n-1}, \dots, x_{n-d+1}) = x_{n+1} \quad (6)$$

따라서 유한 차원의 동력학 시스템으로부터 발생한 시계열이 측정되면 위와 같은 지연 임베딩(delay embedding)을 사용하여 시계열의 진행에 대한 모델을 구성할 수 있다. 이러한 방법을 이용하면 동력학 시스템의 모델링 문제를 관측된 시계열과 관련된 모델링 문제로 간주할 수 있다[1, 6].

혼돈 시계열의 정량적 특성으로는 일반적으로 상관 차원(correlation dimension)과 Lyapunov 스펙트럼을 사용하는 데 이들 지수를 간단히 정의하면 다음과 같다.

상관 차원은 어트랙터의 복잡도(complexity)를 정량화하기 위해 주로 사용되는데 어트랙터의 상관 차원은 그 어트랙터에 대한 동력학을 나타내기 위해 요구된 상태 변수의 갯수에 대한 최저 한계를 주며 모델의 복잡도에 대한 지표가 된다[4].

이러한 상관 차원은 확률적 형태의 차원이며 식(7)과 같이 정의된다[12, 13].

$$D_c := \lim_{\epsilon \rightarrow 0} \frac{\log \sum_{i=1}^{N(\epsilon)} P_i^2}{\log \epsilon} \quad (7)$$

여기서  $N(\epsilon)$ 은 어트랙터를 채우기 위해 필요한 직경이  $\epsilon$  인 체적 요소(volume element)의 최소 갯수이고,  $P_i$ 는 궤적(trajjectory)이  $i$ 번째 체적 요소에 들어가는 상대 빈도이다.

식(7)에서 분자식의 의미를 살펴보기 위해 궤적상의  $N$ 개의 점이 관측되었다고 가정하면 상관 적분(correlation integral)은 식(8)과 같이 정의되고,

$$C(\epsilon) := \lim_{N \rightarrow \infty} \frac{1}{N^2} \{ \|x_i - x_j\| < \epsilon \text{ 인 점들의 쌍 } (x_i, x_j) \text{의 수} \} \quad (8)$$

상관 차원  $D_c$ 는 식(9)에서 구할 수 있다[12, 13].

$$D_c = \lim_{\epsilon \rightarrow 0} \frac{\log C(\epsilon)}{\log \epsilon} \quad (9)$$

한편, Lyapunov 지수는 위상 공간(phase space)에 있어서 바로 이웃하는 궤도(orbit)들의 발산도 혹은 수렴도의 평균 지수율이다. 바로 이웃하는 궤도는 거의 동일한 상태에 상응하므로 지수적인 궤도 발산은 미세한 초기치의 차이에도 시스템은 곧 상당히 다르게 행동한다는 것을 의미하고 따라서 예측 능력이 급격히 상실되어 간다는 것을 나타낸다. 하나 이상의 양의 Lyapunov 지수를 포함하고 있는 시스템은 혼돈적 특성을 갖게 되며, 이 지수의 크기는 시스템의 동력학이 예측 불가능하게 되는 시간 크기(time scale)와 관계한다[14].

Lyapunov 지수 스펙트럼은 다음과 같이 정의할 수 있다. 즉,  $n$ 차원 위상 공간에 있어서 연속적인 동력학 시스템이 주어졌을 때 초기 조건들로 구성된 극히 미소한  $n$ 차원 구의 장시간

동안의 진행을 관측하면 구는 시스템의 진행 흐름(flow)의 국부적인 변형 성질때문에  $n$ 차원 타원체가 된다. 그러면  $i$ 번째 1차원 Lyapunov 지수는 타원체의 주축  $p_i(t)$ 의 길이에 의해 식(10)과 같이 정의된다[14].

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log \frac{p_i(t)}{p_i(0)} \quad (10)$$

이러한 Lyapunov 스펙트럼은 초기 조건에 대한 동력학 시스템의 민감한 정도를 나타내며 이것들중 최대 양의 Lyapunov 지수는 예측 가능한 시간의 크기에 대한 지표가 된다[4].

### 3. 혼돈 시계열의 모델링을 위한 Radial Basis 함수 회로망 설계

#### 3.1 Radial Basis 함수 회로망

Radial basis 함수는 입력벡터  $\vec{x}$ 와 중심벡터  $\vec{c}$  사이의 거리  $r = \|\vec{x} - \vec{c}\|$ 에 의존하는 다차원 함수이다. 비선형 함수를 근사화하기 위한 가장 단순한 접근 방법은 고정된 비선형 basis 함수  $\{h_i(\vec{x})\}$ 의 선형 조합에 의해 구하고자 하는 비선형 함수를 식(11)과 같이 근사화하는 것이다.

$$F(\vec{x}) = \sum_{i=1}^N w_i h_i(\vec{x}) \quad (11)$$

이 회로망의 궁극적인 목적은 계수  $w_i$ 를 추정하는 것이다. 그러나 일반적으로 필요한 basis 함수의 수는 근사화할 함수의 차원에 따라 지수적으로 증가한다. 한편 radial basis 함수는 문제의 차원이 증가함에 따라 조절 가능한 매개 변수의 급격한 증가를 근본적으로 피할 수 있는 다차원 근사화 혹은 적합(fitting)을 위한 알맞은 수단이 될 수 있다.

출력이 1차원 radial basis 함수 회로망의 구조는 그림 1과 같으며 고차원 출력에 대한 회로망의 확장은 동일한 방식으로 이루어질 수 있다.

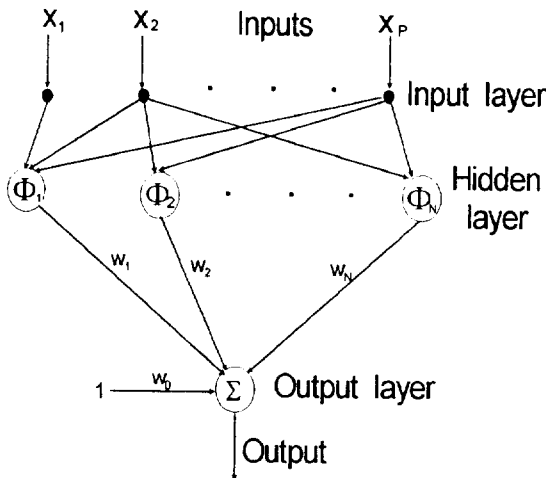


그림 1 Radial basis 함수 회로망  
Fig. 1 Radial basis function network

은닉층의 뉴런이 갖는 비선형성을 나타내는 radial basis 함수  $h(\vec{x}) = \phi(\|\vec{x} - \vec{c}\|)$ 로는 여러 종류의 함수들이 사용되나, 가장 많이 사용되는 것은 중심 벡터  $\vec{c}$ 에서 최대값을 가지고 중심으로부터 거리가 증가함에 따라 단조감소하는 가우시안 함수이다. Radial basis 함수로서 가우시안 함수를 사용할 때의 회로망은 다음과 같이 동작한다.

즉, 입력 벡터  $\vec{x}$ 에 대해 각 은닉층의 뉴런은 폭(width) 혹은 척도(scaling) 파라미터  $\sigma_i$ 와 중심 벡터  $\vec{c}_i$ 에 대해 식(12)와 같은 지수함수를 계산한다.

$$h_i = \exp[-D_i^2 / (2\sigma_i^2)], \quad D_i^2 = (\vec{x} - \vec{c}_i)^T (\vec{x} - \vec{c}_i) \quad (12)$$

이 때 은닉층 뉴런의 중심벡터는 입력 학습 벡터의 값으로 할당되고 출력 뉴런은 이러한 은닉층 뉴런들의 응답의 선형 조합으로 식(13)과 같은 추정치를 계산한다.

$$\hat{y} = w_o + \sum_{i=1}^N w_i \phi(\|\vec{x} - \vec{c}_i\|) \quad (13)$$

여기서  $\vec{x} = (x_1, x_2, \dots, x_p)$ 는 입력 벡터이고  $\vec{c}_i \in R^p (i = 1, 2, \dots, N)$ 는 각 은닉층 뉴런의 중심 벡터이며  $w_i (i = 0, 1, \dots, N)$ 는 출력층 연결 강도이다[15, 16].

#### 3.2 Radial Basis 함수 회로망을 이용한 혼돈 시계열 모델링

시계열  $x(t_i)$ 의  $N$ 개의 샘플에 대해 스칼라 시계열로부터 임베딩 공간에 있어서 지연 벡터(delay vector)를 구성할 수 있다. 이 때  $N$ 개의 샘플은 식(14)와 같이 다시 쓸 수 있다.

$$x_{-m+2}, \dots, x_0, x_1, \dots, x_{N+1} \quad (14)$$

이 샘플 데이터로부터 임베딩 차원이  $m$ 인  $N$ 개의 지연 벡터를 식(15)와 같이 얻을 수 있다.

$$\begin{aligned} \vec{x}_i &= \vec{x}(t_i) \\ &= (x(t_i), x(t_i - \tau), \dots, x(t_i - (m-1)\tau)) \\ &= (x_i, x_{i-1}, \dots, x_{i-m+1}), \quad i = 1, 2, \dots, N \end{aligned} \quad (15)$$

시계열의 모델은 입력 벡터  $\vec{x}_i$ 와 출력  $d_i = x_{i+1}$ 로 구성된 학습 데이터쌍  $(\vec{x}_i, d_i)$ 에 대한 함수  $\hat{f}_n$ 를 구함으로써 얻을 수 있다. 이러한  $\vec{x}_i$ 와  $d_i$ 는 함수  $\hat{f}_n$ 에 의해 식(16)과 같이 쓸 수 있다.

$$d_i = x_{i+1} = \hat{f}_n(\vec{x}_i) + \epsilon_i \quad (16)$$

이 때,  $\epsilon_i$ 는 적합 오차(fitting error)가 된다.

Radial basis 함수 회로망은 식(17)과 같이  $N$ 개의 radial basis 함수의 선형 조합으로서  $\hat{f}_n$ 을 나타낸다.

$$\hat{f}_n(\vec{x}) = \sum_{j=1}^N a_j \phi(\|\vec{x} - \vec{x}_j\|) \quad (17)$$

여기서  $\phi$ 는  $\phi: R^+ \rightarrow R$ 인 함수이고,  $\|\cdot\|$ 는  $m$ 차원 벡터 공간  $R^m$ 에서의 유클리드 놈(Euclidian norm)이며  $\phi(\|\vec{x} - \vec{x}_j\|)$ 는  $\vec{x}_j$ 에 중심을 갖는 radial basis 함수이다.

이 때, 계수  $a_j$ 는 조건  $\tilde{f}_n(\vec{x}_i) = d_i$ 를 이용함으로써 구할 수 있다. 이러한 강한 보간(strict interpolation) 방법은 주어진 모든 데이터를 중심으로 하는 radial basis 함수  $\phi(\|\vec{x} - \vec{x}_j\|)$ 와 이것으로 구성된  $N \times N$  행렬의 역행렬을 이용하여  $a_j$ 를 구하므로 실행하기 쉽고 적합 정밀도가 매우 높은 장점을 갖는 반면 데이터의 갯수가 많아지면 역행렬의 차수가 커지므로 계산시간이 많이 소요되고 높은 예측 정밀도를 보장할 수 없다는 단점이 있다.

한편, X. D. He와 A. Lapedes는 예측 능력과 학습 시간을 향상시키기 위해 다음과 같은 radial basis 함수 회로망을 제안했다[6]. 3.1절에서 언급한 회로망 구조가 하나의 은닉층을 갖는 반면 이 회로망은 그림 2와 같이 두개의 은닉층을 갖는다.

먼저 첫번째 은닉층이 구하고자 하는 함수의 대략적인 근사 함수들로 구성되고, 두번째 은닉층에서는 최종적인 모델 함수를 얻기 위해 첫번째 은닉층의 대략적인 근사 함수들의 최적 선형 조합이 되도록 학습된다.

X.D. He와 A. Lapedes의 회로망 구현은 다음과 같은 세 단계로 구성된다[6]. 즉,  $N$ 개의 데이터가 주어질 때, 식(18)과 같이  $N$ 개의 radial basis 함수를 이용한다.

$$\phi_j = \phi(\|\vec{x} - \vec{x}_j\|), \quad j = 1, 2, \dots, N \quad (18)$$

step 1 :  $N$ 개의 radial basis 함수  $\phi_j (j = 1, 2, \dots, N)$ 를 여러 개의 그룹으로 분할한다.

$$\frac{(\phi_1, \phi_2, \dots, \phi_L)}{\text{그룹1}}, \frac{(\phi_{L+1}, \phi_{L+2}, \dots, \phi_{2L})}{\text{그룹2}}, \dots, \frac{(\phi_{N-L+1}, \phi_{N-L+2}, \dots, \phi_N)}{\text{그룹}N'} \quad (19)$$

여기서  $L$ 은 각 그룹에 있어서의 radial basis 함수의 수이고  $N'$ 은 그룹의 수이다. 즉,  $LN' = N$ 이다.

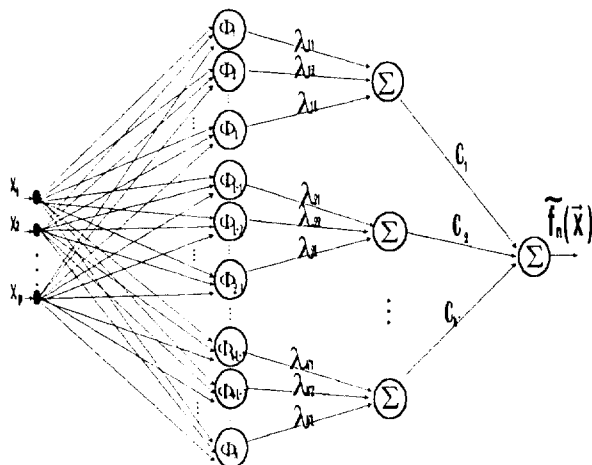


그림 2 X.D. He와 A. Lapedes에 의해 제안된 모델  
Fig. 2 The model proposed by X.D. He and A. Lapedes

step 2 : 최소 자승법을 사용하여 각 그룹에 대해  $L$ 개의 radial basis 함수의 최적 선형 조합으로 구성된 함수  $f_k(\vec{x})$ 를 구한다.

$$f_k(\vec{x}) = \sum_{j=1}^L \lambda_{kj} \phi(\|\vec{x} - \vec{x}_j\|), \quad k = 1, 2, \dots, N' \quad (20)$$

이 때  $k_j = (k-1)L + j$ 이다.

step 3 : 식(21)과 같이  $N'$ 개의 함수  $f_k$ 의 최적 선형 조합을 구하여 최종 모델  $\tilde{f}_n(\vec{x})$ 를 구한다.

$$\tilde{f}_n(\vec{x}) = \sum_{k=1}^{N'} c_k f_k(\vec{x}) \quad (21)$$

계수  $c_k$ 를 구하기 위해 조건  $\tilde{f}_n(\vec{x}_i) = d_i$ 를 이용하면 식(25)와 같이 쓸 수 있다.

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{pmatrix} = \begin{pmatrix} \tilde{d}_{11} & \tilde{d}_{12} & \dots & \tilde{d}_{1N'} \\ \tilde{d}_{21} & \tilde{d}_{22} & \dots & \tilde{d}_{2N'} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{d}_{N1} & \tilde{d}_{N2} & \dots & \tilde{d}_{NN'} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{N'} \end{pmatrix} \quad (22)$$

여기서  $\tilde{d}_{ik}$ 는 식(23)과 같다.

$$\tilde{d}_{ik} = f_k(\vec{x}_i) = \sum_{j=1}^L \lambda_{kj} \phi(\|\vec{x}_i - \vec{x}_{(k-1)L+j}\|) \quad (23)$$

계수  $c_k$ 는 식(22)의 최소 자승해를 구함으로써 얻을 수 있다. 결과적으로 X. D. He와 A. Lapedes의 회로망에 의해 구성된 모델은 식(24)와 같다.

$$\tilde{f}_n(\vec{x}) = \sum_{k=1}^{N'} c_k \sum_{j=1}^L \lambda_{kj} \phi(\|\vec{x} - \vec{x}_{(k-1)L+j}\|) \quad (24)$$

### 3.3 제안된 Radial Basis 함수 회로망

본 논문에서는 예측 능력을 향상시키기 위해 X. D. He와 A. Lapedes의 모델을 바탕으로 크러스터링 방법과 계층적 학습[9]을 이용하여 네 종류의 변형된 모델을 제안하였으며 각 모델은 다음과 같다.

A. Lapedes등의 모델이 각 그룹의 중심 벡터를 구성하기 위해 입력 벡터를 순서대로 그대로 이용하는 반면 이 모델은 그림 3과 같이 입력 벡터를 크러스터링 방법에 의해 여러 개의 크러스터(cluster)로 입력 벡터를 분할하여 각 크러스터당 하나의 그룹을 형성한다. 모델 I의 구현 절차는 다음과 같다.

step 1 :  $N$ 개의 radial basis 함수  $\phi_j (j = 1, 2, \dots, N)$ 를 크러스터링 방법에 의해 여러개의 그룹으로 분할한다.

$$\frac{(\phi_{11}, \phi_{12}, \dots, \phi_{1L_1})}{\text{그룹1}}, \frac{(\phi_{21}, \phi_{22}, \dots, \phi_{2L_2})}{\text{그룹2}}, \dots, \frac{(\phi_{N'1}, \phi_{N'2}, \dots, \phi_{N'L_{N'}})}{\text{그룹}N'} \quad (25)$$

여기서  $L_i$ 는 각 그룹에 있어서의 radial basis 함수의 수이고  $N'$ 은 그룹의 수이다. 즉  $\sum_{i=1}^{N'} L_i = N$ 이 된다.

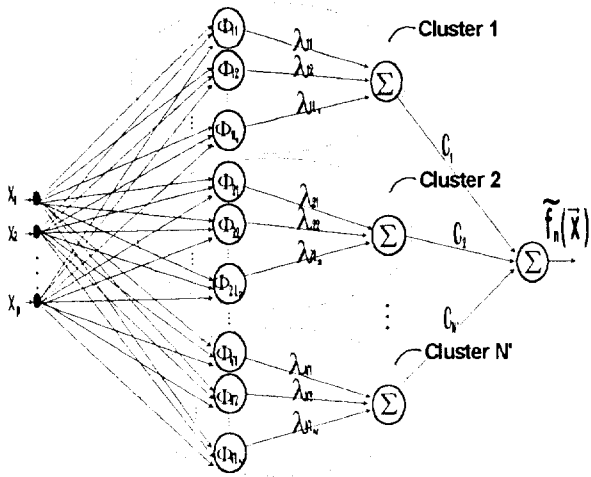


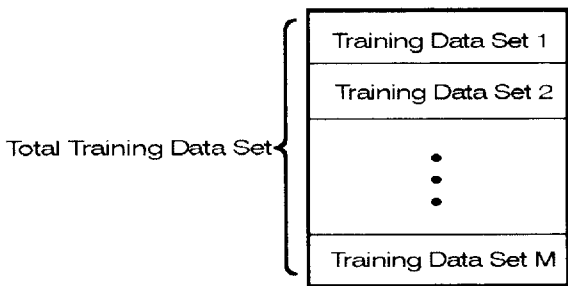
그림 3 모델 I  
Fig. 3 Model I

step 2 : Lapedes 모델의 2 단계와 동일하다.  
step 3 : Lapedes 모델의 3 단계와 동일하다.

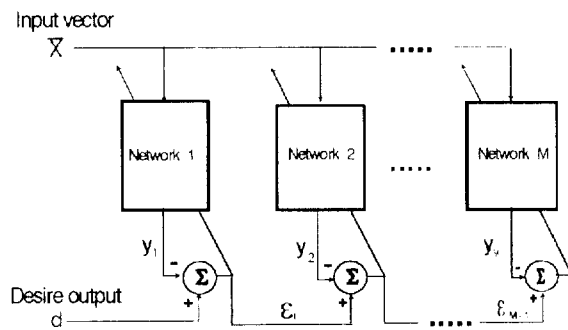
모델 I의 회로망에 의해 구성된 모델은 결과적으로 식(26)과 같이 된다.

$$\hat{f}_n(\vec{x}) = \sum_{k=1}^{N'} c_k \sum_{j=1}^{L_{sk}} \lambda_{skj} \phi(\|\vec{x} - \vec{x}_{skj}\|) \quad (26)$$

모델 II는 그림 4(a)와 같이 전체 입력 학습 데이터 집합을



(a) 학습 데이터의 분할



(b) 계층적 학습 구조

그림 4 모델 II  
Fig. 4 Model II

여러개의 부분 학습 데이터 집합으로 분할하여 각 부분 학습 데이터 집합에 대해 모델 I과 같은 회로망을 구성하여 그림 4(b)와 같이 회로망들을 계층적으로 학습시키므로써 구성된다. 모델 II의 구현 절차는 다음과 같다.

step 1 : 전체 입력 학습 데이터 집합을 순서대로 분할하여 동일한 갯수의 학습 데이터로 구성된 부분 학습 데이터 집합군을 형성한다.

step 2 : 각 부분 학습 데이터 집합에 대해 모델 I을 구성한다. 각 부분 회로망은 계층적으로 학습되며 i번째 부분 회로망의 학습에 있어서 요구된 출력은 식(27)과 같이 주어진다.

$$\begin{cases} d_i = d & : i = 1 \\ d_i = d_{i-1} - y_{i-1} & : i = 2, 3, \dots, M \end{cases} \quad (27)$$

여기서 d는 전체 회로망의 기대 출력이고 d\_i와 y\_i는 각각 i번째 부분 회로망의 기대 출력과 훈련된 i번째 부분 회로망의 출력을 나타낸다.

모델 II의 회로망에 의해 구성된 모델은 식(28)과 같이 정리할 수 있다.

$$\hat{f}_n(\vec{x}) = \sum_{s=1}^M \sum_{k=1}^{N'} c_{sk} \sum_{j=1}^{L_{sk}} \lambda_{skj} \phi(\|\vec{x} - \vec{x}_{skj}\|) \quad (28)$$

모델 III은 모델 II와 달리 그림 5와 같이 전체 학습 데이터 집합을 클러스터링 방법에 의해 여러개의 부분 학습 데이터 집합으로 분할하여 각 부분 학습 데이터 집합에 대해 모델 I과 같은 회로망을 구성하여 이 회로망들을 계층적으로 학습시키므로써 구성된다. 모델 III의 구현 절차는 다음과 같다.

step 1 : 전체 입력 학습 데이터 집합을 클러스터링 방법의 해 여러개의 부분 학습 데이터 집합으로 분할한다.

step 2 : 모델 II의 2 단계와 동일하다.

모델 III의 회로망에 의해 구성된 모델은 결과적으로 식(29)와 같이 된다.

$$\hat{f}_n(\vec{x}) = \sum_{s=1}^M \sum_{k=1}^{N'} c_{sk'} \sum_{j=1}^{L_{sk'}} \lambda_{sk'j} \phi(\|\vec{x} - \vec{x}_{sk'j}\|) \quad (29)$$

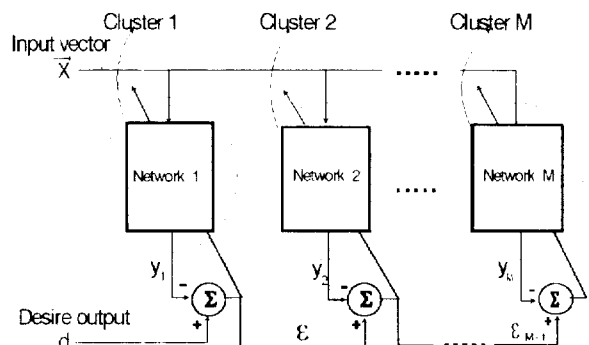


그림 5 모델 III  
Fig. 5 Model III

Radial basis 함수 회로망은 학습 방법에 따라 다양한 성능의 차이가 발생할 수 있다. 일반적으로 계산효율을 높이고 radial basis 함수의 국부적인 조절 특성을 이용하기 위해 앞 절에서 언급한 바와 같이 Moody 등이 제안한 두 단계로 구성된 혼합 학습 방법[7]이 주로 사용된다. 그리고 이러한 혼합 학습을 통해 radial basis 함수 회로망을 학습시키는데 있어서 일반적으로 비회귀적인 방법을 사용한다[15, 16]. 하지만 비회귀적 학습방법은 새로운 데이터가 관측되었을 때 이 데이터에 포함된 새로운 정보를 실시간 이내에 이용할 수 없다는 단점을 갖고 있다.

본 논문에서는 이러한 단점을 극복하기 위하여 J. Stark[1]가 학습 과정 없이 바로 radial basis 함수 회로망의 연결 강도를 회귀적으로 조절하기 위해 사용한 방법을 학습 과정에 도입하여 혼합 학습 과정으로 전환한 다음과 같이 두 단계로 구성된 회귀 혼합 학습방법을 제안한다.

step 1 : 은닉층의 고정된 중심의 수에 대해 중심 벡터  $\vec{c}_i$ 를 크러스터링 방법에 의해 구하고 폭 파라미터  $\sigma_i$ 는 식 (30)에 주어진 P-nearest neighbor heuristic 방법에 의해 구한다.

$$\sigma_i = \left[ \frac{1}{P} \sum_{j=1}^P \| \vec{c}_i - \vec{c}_j \|^2 \right]^{\frac{1}{2}} \quad (30)$$

여기서 중심 벡터  $\vec{c}_j$ 는 중심 벡터  $\vec{c}_i$ 와 가장 가까운 P개의 중심 벡터이다.

step 2 : 출력층의 연결 강도  $w_i$ 을 조절하기 위해 RLS 알고리즘보다 정확도가 높고 수치적으로 안정한 회귀적인 변형 Gram-Schmidt 알고리즘[1, 10]을 사용한다.

#### 4. 실험 및 결과 고찰

##### 4.1 예측 평가기준

본 논문에서는 radial basis 함수 회로망의 평가를 단 단계 예측(single step prediction)과 다 단계 예측(multi step prediction)의 두 가지 경우로 나누어 고려하였다.

단 단계 예측은 주어진 학습 데이터를 이용하여 학습된 모델이 주어졌을 때 입력 벡터  $\vec{x}_i = (x(t), x(t-\tau), \dots, x(t-(m-1)\tau))$ 가 입력단에 적용되어 출력으로 현재 입력 시계열의 바로 한 단계 미래의 예측된 값  $\hat{x}(t+\tau)$ 을 산출한다. 학습 데이터에 대한 적합 정확도를 측정하는 기준은 식(31)과 같고 검증 데이터에 대한 예측 정확도를 측정하는 기준은 식(32)와 같다[6, 9].

$$F_f = \log_{10} \left( \sqrt{\frac{E[(\hat{f}(\vec{x}) - f(\vec{x}))^2]}{E[(x(t) - E[x(t)])^2]}} \right) \quad (31)$$

$$F_p = \log_{10} \left( \sqrt{\frac{E[(\hat{f}(\vec{x}) - f(\vec{x}))^2]}{E[(x(t) - E[x(t)])^2]}} \right) \quad (32)$$

여기서  $f(\cdot)$ 는 혼돈 시계열의 비선형 사상을 나타내고  $\hat{f}(\cdot)$ 는 신경 회로망에 의해 구현된 모델을 나타낸다. 이 기준은 혼

돈 시계열의 예측 모델에 대한 평가 기준으로 일반적으로 가장 많이 사용되고 있다.

한편, 다 단계 예측은 주어진 학습 데이터를 이용하여 학습된 모델이 주어졌을 때 입력 벡터  $\vec{x}_i = (x(t), x(t-\tau), \dots, x(t-(m-1)\tau))$ 가 입력단에 적용되어 출력으로 현재 입력 시계열의 바로 한 단계 미래의 예측된 값  $\hat{x}(t+\tau)$ 를 산출하고 다음 단계 예측을 위해 입력 벡터  $\vec{x}_i = (\hat{x}(t+\tau), x(t), \dots, x(t-(m-2)\tau))$ 를 구성하여 입력단에 적용되어 출력으로 두 단계 미래의 예측된 값  $\hat{x}(t+2\tau)$ 를 산출한다. 동일한 방식으로 요구된 단계까지의 예측을 위해 예측된 값을 이용하여 입력 벡터를 구성하여 예측을 수행한다. 다 단계 예측의 성능 비교를 위해 사용한 기준은 J. Deppisch 등[9]에 의해 제안된 식(33)의 평가 기준을 사용하였다.

$$L_n = E \left[ \log \left| \frac{\hat{f}^n(\vec{x}_i) - f^n(\vec{x}_i)}{\sqrt{E[(x(t) - E[x(t)])^2]}} \right| \right] \quad (33)$$

##### 4.2 결과 고찰

실험에 사용할 혼돈 시계열로서 혼돈 비선형 시스템 및 시계열 모델링의 평가 기준으로 가장 많이 사용되는 로지스틱 사상(logistic map)과 Mackey-Glass 지연 미분 방정식(Mackey-Glass delay differential equation)으로부터 발생시킨 시계열을 사용했다.

로지스틱 사상은 식(34)와 같이 정의되는 이차 사상(quadratic map)이다.

$$x_{n+1} = rx_n(1 - x_n) \quad (34)$$

여기서  $r$ 은 외부 조절 변수이다. 로지스틱 사상은  $r$ 의 값에 따라 혼돈적 특성을 갖는 것으로 증명된 단순한 형태의 시스템이다. 본 실험에서는  $r = 4$ 를 사용하여 시계열을 발생시켰다. Mackey-Glass 지연 미분 방정식은 식(35)와 같이 정의된다.

$$\frac{dx}{dt} = \frac{ax(t - \Gamma)}{1 + x^{10}(t - \Gamma)} - bx(t) \quad (35)$$

본 실험에서는  $\Gamma = 17$ ,  $a = 0.2$ ,  $b = 0.1$ 일때 지연 미분 방정식으로부터 발생시킨 시계열을 사용했다.

한편, 상관 차원은 Grassberger-Procaccia 알고리즘[13]에 의해 추정하였고 최대 Lyapunov 지수는 Rosenstein 등의 방법[17]에 의해 추정하였으며, 각 시계열에 대한 추정치는 표 1과 같다.

표 1 추정된 상관 차원과 최대 Lyapunov 지수  
Table 1 Estimated correlation dimensions and the largest Lyapunov exponents

추정된 값	로지스틱 사상	Mackey-Glass 시계열
상관 차원	0.9 ~ 1.0	1.7 ~ 2.1
최대 Lyapunov 지수	0.6985	0.0073

4.2.1 단 단계 예측 결과

이 실험에서는 A. Lapedes 등의 실험[6]을 기준으로 하여 크러스터링 방법으로 K-means 크러스터링 방법(K-means clustering method)[8]을 사용하였다. 또한 로지스틱 사상의 경우 학습 데이터( $P_1$ )로 250개를 그리고 검증 데이터( $P_2$ )로 500개를 사용했으며 Mackey-Glass 시계열에 대해서는 학습 데이터( $P_1$ )로 500개를 그리고 검증 데이터( $P_2$ )로 500개를 사용하여 다음과 같은 회로망들의 성능을 비교하였다.

- ① X.D. He와 A. Lapedes의 모델 (RBFN 1).
- ② 제안 모델 I (RBFN 2).
- ③ 제안 모델 II (RBFN 3).
- ④ 제안 모델 III (RBFN 4).
- ⑤ 비회귀 학습 radial basis 함수 회로망 (RBFN 5).
- ⑥ 제안모델 IV (RBFN 6).

각 시계열에 대한 단 단계 예측 결과는 표 2 및 표 3과 같다. 여기서  $M$ 은 부분 회로망의 수,  $N'$ 은 첫번째 은닉층의 크러스터 수 그리고  $L$ 은 각 크러스터 당 은닉층 노드의 수이다.

실험 결과, 학습 입력 벡터를 모두 radial basis 함수의 중심으로 사용하는 경우는 Lapedes 등의 모델과 같이 학습 입력 벡터를 순서대로 사용하여 이중의 은닉층을 구성하는 것 보다는

표 2 로지스틱 사상에 대한 단 단계 예측 결과  
Table 2 Results of single step prediction for the Logistic map

적합 및 예측 성능	$F_f$	$F_p$	$\sigma^2$	기타
RBFN 1	-8.508	-8.421	0.02	$N' = 50, L = 5$
RBFN 2	-9.034	-8.941	0.20~0.60	$N' = 50$
RBFN 3	-11.259	-11.186	0.50~2.10	$M = 2, N' = 70$
RBFN 4	-11.495	-11.395	0.25~2.25	$M = 3, N' = 50$
RBFN 5	-12.387	-12.374	0.05~0.20	은닉층 노드수 = 50
RBFN 6	-11.216	-11.877	0.05~0.20	은닉층 노드수 = 50

표 3 Mackey-Glass 시계열에 대한 단 단계 예측 결과  
Table 3 Results of single step prediction for the Mackey-Glass time series

적합 및 예측 성능	$F_f$	$F_p$	$\sigma^2$	기타
RBFN 1	-2.056	-2.041	1.00	$N' = 100, L = 5$
RBFN 2	-2.060	-2.040	0.35~1.00	$N' = 100$
RBFN 3	-2.118	-2.085	0.10~0.28	$M = 4, N' = 100$
RBFN 4	-2.169	-2.082	0.07~0.30	$M = 2, N' = 100$
RBFN 5	-2.066	-2.055	0.08~0.20	은닉층 노드수=100
RBFN 6	-2.062	-2.282	0.08~0.20	은닉층 노드수=100

이 벡터들을 크러스터링하여 각각의 그룹으로 만든 후 그룹당 하나의 radial basis 함수 회로망을 구성하여 계층적으로 학습 시킴으로써 성능을 더 향상시킬 수 있다는 것을 알 수 있었다.

그러나 위와 같이 전체 입력 학습 벡터를 중심으로 사용한 경우보다 입력 학습 데이터를 크러스터링 하여 각 크러스터의 중심을 은닉층 노드의 중심으로 사용한 경우가 계산 효율도 높고 성능 또한 향상됨을 알 수 있었다. 왜냐하면 이러한 임베딩 공간이 모델의 입력 벡터의 공간과 동일하게 되어 각 은닉층 노드의 radial basis 함수의 중심으로 전체 학습 입력 벡터를 사용하는 것보다 임베딩 공간에 재구성된 어트리뷰터상에 적절하게 분포된 중심을 사용하여 은닉층을 구성하는 것이 효율적이기 때문이다.

따라서 혼돈 시계열의 예측을 위한 회로망의 구조적인 측면에 있어서는 전체 입력 학습 벡터를 은닉층의 중심으로 사용하는 것보다 입력 학습 데이터를 크러스터링 하여 각 크러스터의 중심을 은닉층 노드의 중심으로 사용하는 것이 효율적이다.

한편 위와 같은 크러스터링한 은닉층 노드로 구성된 회로망 구조의 출력층 연결 강도를 학습 시키는 방법으로 회귀적인 변형 Gram-Schmidt 알고리즘을 사용하여 학습시킨 경우, 로지스틱 사상에 대해서는 비회귀 학습보다 다소 성능이 저하되지만 학습 데이터 갯수를 500개로 증가시켜 실험한 결과, 비회귀 학습을 사용한 경우는 적합 성능  $F_f = -12.387$ 이고 예측 성능  $F_p = -12.369$ 인 반면 제안된 회귀 학습을 사용한 경우에는 적합 성능  $F_f = -12.271$ 이고 예측 성능  $F_p = -12.303$ 으로 비슷한 성능을 나타냄을 알 수 있었다. 또한 Mackey-Glass 시계열에 대해서는 전반적으로 비회귀 학습보다 제안된 회귀 학습이 우수함을 확인 할 수 있었다. 왜냐하면 크러스터링 방법으로 구성된 은닉층을 갖는 radial basis 함수 회로망에 대해 비회귀 학습 방법에 의해 출력층의 연결 강도를 조절하는 경우 새로운 관측 데이터에 대해 연결 강도에 새로이 관측된 정보가 포함되도록 실시간내에 갱신할 수 없는 반면에 제안된 회귀적인 변형 Gram-Schmidt 알고리즘에 의한 회귀 학습방법은 새로운 관측 데이터에 대해 수치적으로 정확하고 안정되게 실시간내에 연결 강도를 갱신할 수 있기 때문이다.

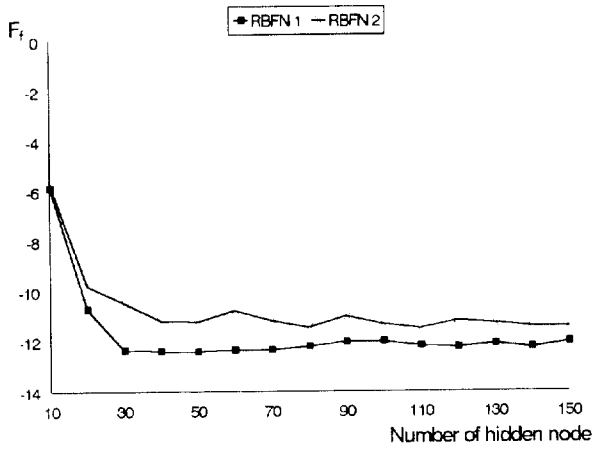
결과적으로 입력 학습 데이터를 크러스터링하여 각 크러스터의 중심을 은닉층 노드의 중심으로 사용하고 출력층의 연결 강도를 제안된 회귀 학습 방법을 사용하여 학습시키는 radial basis 함수 회로망이 혼돈 시계열의 예측을 위해 적합하다고 결론 지을 수 있다.

한편 은닉층 노드의 수에 따른 회로망의 성능을 알아보기 위해 각 시계열에 대해 실험한 결과는 그림 6 및 7과 같다.

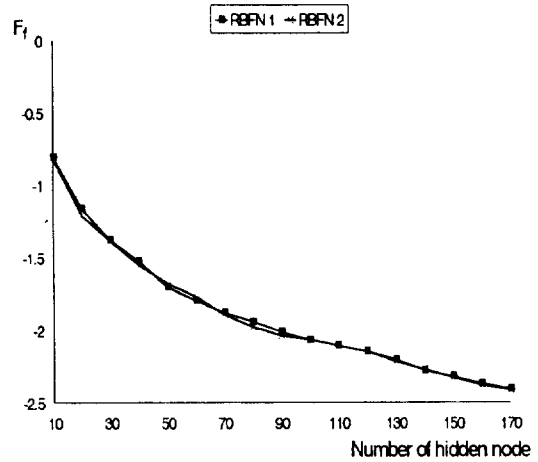
그림에서 RBFN 1은 비회귀 학습 radial basis 함수 회로망을 나타내며 RBFN 2는 제안된 회귀 학습 방법을 사용한 radial basis 함수 회로망을 나타낸다.

결과에서 보는 바와 같이 은닉층 노드의 수가 증가할수록 예측 성능이 좋아지는 것이 아니라 적절한 은닉층 노드의 수가 존재함을 알 수 있다.

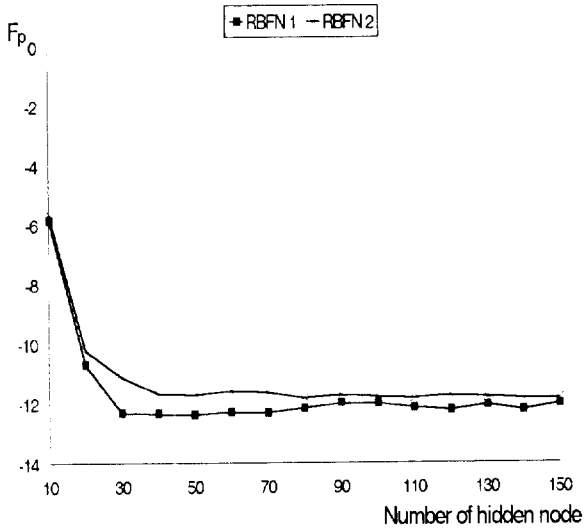
즉, 은닉층 노드의 수가 작으면 학습 데이터의 비선형 사상 관계를 학습할 수 없고 은닉층 노드의 수가 너무 많으면 강한 보간의 성격을 갖게 되므로 적합 성능은 상당히 좋아지나 일반화 특성이 약화되어 예측 성능은 저하되게 된다. 따라서 로지스틱 사상의 경우 30~70개의 은닉층 노드가 적당하고



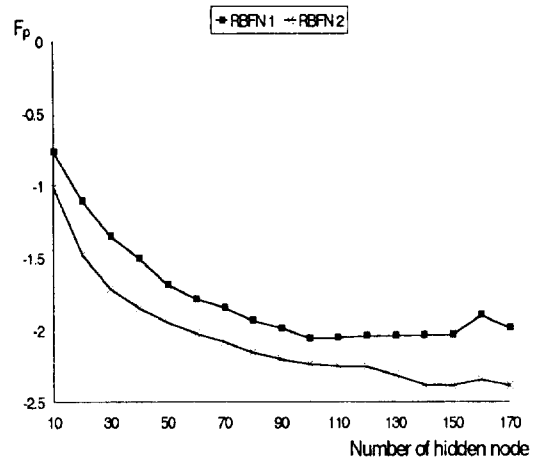
(a) 적합 성능



(a) 적합 성능



(b) 예측 성능



(b) 예측 성능

그림 6 로지스틱 사상에 대한 은닉층 노드수에 따른 성능  
Fig. 6 Performance for the number of hidden nodes in the Logistic map

Mackey-Glass의 경우는 100~150개의 은닉층 노드가 적정함을 알 수 있었다. 특히 제안된 회귀 학습을 이용하는 회로망이 로지스틱 사상의 경우 비회귀 학습 회로망에 비해 예측 성능이 다소 저하되거나 비슷하였지만 앞에서 언급한 바와 같이 학습 데이터의 갯수를 증가시킴으로써 성능을 개선할 수 있고 Mackey-Glass 시계열에 대해서는 우수함을 알 수 있었으며 은닉층 노드의 수가 증가하더라도 비회귀 학습 회로망과 같은 과도 학습에 의한 예측 성능의 저하가 발생하지 않고 일정한 값에서 포화되었다. 이것은 제안된 회귀 학습에 의한 회로망의 연결 강도 갱신 특성에서 기인한다고 볼 수 있다.

4.2.2 다 단계 예측 결과

이 실험에서는 단 단계 예측 결과를 기초로 아래에 주어진 4 가지 회로망을 선정하여 로지스틱 사상과 Mackey-Glass 시계열 모두에 대해 학습 데이터로 500개를 사용하고 은닉층 노드

그림 7 Mackey-Glass 시계열에 대한 은닉층 노드수에 따른 성능

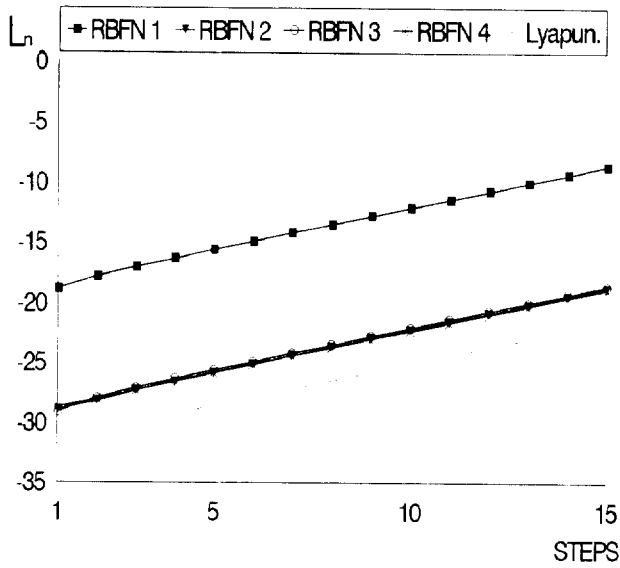
Fig. 7 Performance for the number of hidden nodes in the Mackey-Glass time series

의 수로 각각 50개와 100개를 사용하여 로지스틱 사상의 경우는 15 단계 예측을 시도했고 Mackey-Glass 시계열의 경우에는 30 단계 예측을 시도했다.

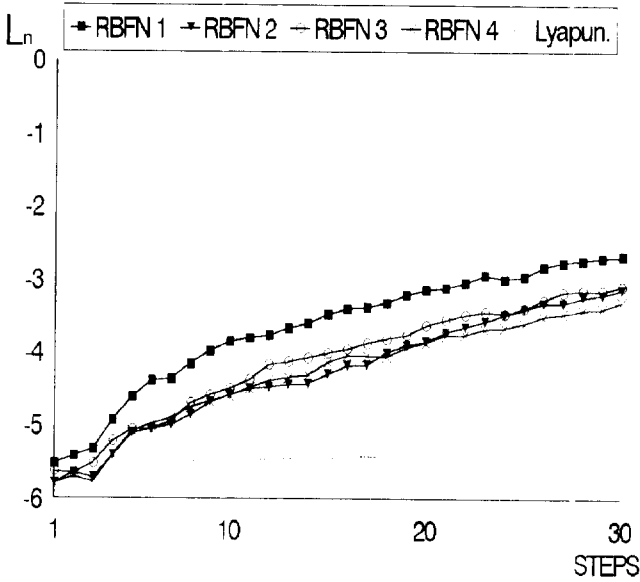
- ① X.D. He와 A. Lapedes의 모델 (RBFN 1).
- ② 비회귀 학습 radial basis 함수 회로망 (RBFN 2).
- ③ 회귀적인 변형 Gram-Schmidt 알고리즘을 이용한 회귀 학습 radial basis 함수 회로망 (RBFN 3).
- ④ 학습과정 동안만 회귀적인 변형 Gram-Schmidt 알고리즘에 의해 회귀 학습한 radial basis 함수 회로망 (RBFN 4).

실험 결과는 그림 8과 같으며 각 시계열에 대해 다 단계 예측의 경우에도 A. Lapedes의 모델보다 크러스터링하여 은닉층 노드를 구성한 회로망이 예측 성능이 우수함을 알 수 있다. 또한 다 단계 예측하에서 각 예측 단계마다 발생한 오차가 혼돈 시계열의 최대 Lyapunov 지수와 동일한 비율로 증가하므로 식(36)이 유효하리라 예상된다[4].





(a) 로지스틱 사상



(b) Mackey-Glass 시계열

그림 8 다 단계 예측 결과

Fig. 8 Results of multi-step predictions

$$\epsilon((\hat{f}_N)^i) = e^{i\lambda_{\max}} \epsilon(\hat{f}_N) \quad (36)$$

여기서  $\hat{f}_N$ 는 단 단계 예측 모델,  $i$ 는 예측 단계,  $(\hat{f}_N)^i$ 는 단 단계 예측 모델의  $i$ 번째 예측 단계로의 반복,  $\epsilon(\cdot)$ 는 예측 오차, 그리고  $\lambda_{\max}$ 는 모델링하고자 하는 혼돈 시계열의 최대 Lyapunov 지수이다.

로지스틱 사상의 경우, 그림 8(a)에서 점선으로 나타나 있는 최대 Lyapunov 지수와 동일한 기울기로 예측 오차가 증가함을 알 수 있으나 그림 8(b)의 Mackey-Glass 시계열의 경우는 최대 Lyapunov 지수보다 큰 기울기로 예측 오차가 증가하고 있다. 따라서 로지스틱 사상에 대해서는 최적의 모델에 접근했다

고 볼 수 있으나 Mackey-Glass 시계열의 경우는 더 향상된 모델의 존재 가능성을 암시한다.

또한 다 단계 예측에 있어서는 비회귀 학습 회로망과 회귀 학습 회로망이 유사한 성능을 나타내었으며 특히 연결 강도를 갱신하기 위해 예측된 값을 사용하여 회귀 학습한 RBFN 3 모델의 경우에는 각 단계에 있어서의 예측 오차로 인한 잘못된 정보가 출력층 연결 강도 갱신에 사용되므로 학습 과정에서만 회귀 학습을 사용한 RBFN 4 모델보다 예측 오차의 기울기가 미소하지만 증가함을 볼 수 있었다. 이것 또한 혼돈 시계열에 있어서 예측 오차가 최대 Lyapunov 지수와 동일한 비율로 발산하는 특성에서 기인한다고 생각된다.

### 5. 결 론

본 논문에서는 혼돈 시계열의 단기 예측 성능을 향상시키기 위해 두개의 은닉층 구조를 갖는 radial basis 함수 회로망을 기초로 K-means 클러스터링 방법과 계층적 학습을 이용한 회로망과 회귀적인 변형 Gram-Schmidt 알고리즘을 이용한 radial basis 함수 회로망의 회귀 학습 방법을 제안하였다. 그리고 혼돈 시계열의 단기 예측에 대해 제안된 학습 방법들로 학습한 회로망을 X.D. He와 A. Lapedes의 모델 및 비회귀 학습 radial basis 함수 회로망과의 비교를 통하여 개선된 radial basis 함수 회로망을 제시하였다.

얻어진 결과는 다음과 같다.

(1) 혼돈 시계열의 임베딩 공간은 radial basis 함수 회로망 모델의 입력 공간과 동일하므로 임베딩 공간에 재구성된 어트랙터상에 적절하게 분포된 중심을 이용하여 radial basis 함수 회로망의 은닉층 노드를 구성함으로써 계산 효율을 높일 수 있고 단기 예측 성능도 향상되었다.

(2) 단 단계 예측의 경우 radial basis 함수 회로망의 출력층 연결강도가 새로운 관측 데이터의 정보를 이용하기 위해서는 비회귀 학습방법보다는 제안된 회귀 학습방법을 이용하여 학습하는 것이 예측 성능을 향상시킬 수 있으며 또한 제안된 모델이 Lapedes등의 모델 보다 예측 성능이 우수하였다.

(3) 다 단계 예측의 경우 제안된 모델이 A. Lapedes등의 모델 보다 예측 성능이 우수하고 단 단계 예측 오차가 매우 작으면서 최대 Lyapunov 지수와 동일한 오차 발산율을 갖도록 구성된 모델이 최적의 모델과 근사하였다. 또한 혼돈 시계열의 오차 발산율에 의해 제안된 회귀 학습방법을 이용한 회로망이 비회귀 학습방법에 의한 회로망과 비슷한 예측 성능을 나타내거나 다소 저하될 수 있음을 볼 수 있었다.

### 참 고 문 헌

[1] J. Stark, "Recursive Prediction of Chaotic Time Series", J. of Nonlinear Science, Vol. 3, pp. 197-223, 1993.  
 [2] J. D. Farmer and J. J. Sidorowich, "Predicting Chaotic Time Series", Phys. Rev. Lett. Vol. 59, pp. 845-848, 1987.  
 [3] A. S. Lapedes and R. Farber, "Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling", Technical Report, LA-UR-87-2662, Los Alamos National Laboratory, 1987.

- [4] M. Casdagli, "Nonlinear Prediction of Chaotic Time Series", *Physica D*, Vol. 35, pp. 335-356, 1989.
- [5] J. H. Kim and J. Stringer, *Applied Chaos*, John Wiley & Sons, inc., 1992.
- [6] 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.
- [7] J. Moody and C. J. Darken, "Fast Learning in Networks of Locally-Tuned Processing Units", *Neural Computation*, Vol. 1, pp. 281-294, 1989.
- [8] D. H. Hush and B. Horne, "Progress in Supervised Neural Networks", *IEEE Signal Processing Magazine*, pp 8-39, Jan, 1993.
- [9] J. Deppisch, H. U. Bauer and T. Geisel, "Hierarchical Training of Neural Networks and Prediction of Chaotic Time Series", *Physics Letters A*, Vol. 158, pp. 57 - 62, 1991.
- [10] F. Ling, D. Manolakis and J.G. Proakis, "A Recursive Modified Gram-Schmidt Algorithm for Least-Squares Estimation", *IEEE Trans. on ASSP*, Vol. ASSP-34, No. 4, pp. 829-836, Aug. 1986.
- [11] F. Takens, "Detecting Strange Attractors in Turbulence", in *Dynamical Systems and Turbulence*, Warwick, ed. D. A. Rand, and L. S. Young, *Lecture Notes in Mathematics*, Vol. 898, New York : Springer-Verlag, pp. 366-381, 1980.
- [12] T. S. Parker and L. O. Chua, "Chaos : A Tutorial for Engineers", *Proceedings of the IEEE*, pp. 982-1008, Aug. 1987.
- [13] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer-Verlag New York Inc., 1989.
- [14] A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastano, "Determining Lyapunov Exponents from a Time Series", *Physica D*, Vol. 16, pp. 285-317, 1985.
- [15] P. D. Wasserman, *Advanced Methods in Neural Computing*, Van Nostrand Reinhold, 1993.
- [16] A. Cichochi and R. Unbehauen, *Neural Networks for Optimization and Signal Processing*, John Wiley & Sons, Inc., 1993.
- [17] M. T. Rosenstein, J. J. Collins and C. J. De Luca, "A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets", *Physica D*, Vol. 65, pp. 117 - 134, 1993.

## 저 자 소 개



**신창용 (申昌溶)**

1969년 12월 30일생. 1993년 연세대 공대 전기공학과 졸업. 1995년 연세대 대학원 전기공학과 졸업(석사)



**김택수 (金宅洙)**

1966년 4월 5일생. 1990년 연세대 공대 전기공학과 졸업. 1992년 연세대 대학원 전기공학과 졸업(석사). 1996년 연세대 대학원 전기공학과 박사과정



**최윤호 (崔允浩)**

1957년 1월 14일생. 1980년 연세대 공대 전기공학과 졸업. 1982년 연세대 대학원 전기공학과 졸업(석사). 1991년 연세대 대학원 전기공학과 졸업(공박). 1996년 현재 경기대 공대 전자공학과 조교수



**박상희 (朴相暻)**

1939년 8월 25일생. 1962년 연세대 공대 전기공학과 졸업. 1964년 동 대학원 졸업(석사). 1971년 동 대학원 졸업(공박) 및 동·전기공학과 조교수. 1982년 미국 워싱턴대 교환교수. 현재 연세대 공대 전기공학과 교수. 1994년~1995년 당학회 부회장