

# 신경 회로망을 이용한 비선형 동적 시스템의 적응 제어

\*장성환, \*조현석<sup>0</sup>, \*김기철, \*최봉식, \*\*유인호,  
\*원광대학교 전기공학과, \*\*이리농공전문대학 전기과

## Adaptive Control of Non-linear Dynamic System using Neural Network.

Seong-Whan Jang, \*Hyeon-Seob Cho, \*Ki-Cheol Kim, \*Bong-Shik Choi, \*\*In-Ho Yu  
\*Dept. of Electrical Eng., Won-Kwang Univ.  
\*\*Dept. of Electrical Eng. Iri National Agricultural & Technical College

### ABSTRACTS

Studied on identification of nonlinear system with unknown variables and adaptive control were successful. We need a mathematical model when control a dynamic system using adaptive control technique, but it is very difficult due to its nonlinearity. In this paper, we described about performance improvement of error back-propagation algorithm and learning algorithm of non-linear dynamic system. We examined the proposed back-propagation learn algorithm through an experiment.

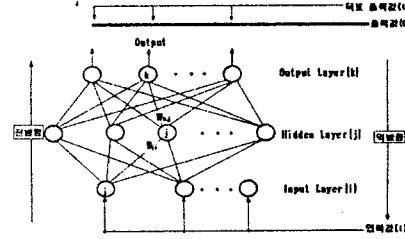
### I. 서 론

미지의 변수를 가진 선형 시스템의 식별과 적응 제어에 관한 연구는 지난 30년간 꾸준히 연구되어 왔으며 갈등할 만하게 진보되었다. 하지만 기존의 적응제어 기법은 동적 시스템의 수학적 모형을 필요로 하지만 동적 시스템은 비선형성, 불확정성, 시간 지연, 시변 시스템의 변수, 구조등에 의해서 정확하게 기술한다는 것은 불가능하다. 컴퓨터의 발달과 더불어 인간의 두뇌와 유사한 기능을 가진 신경망(Neural Network)에 관한 연구와 관심이 고조되고 있다 [1]. 역치를 포함한 간단한 형태의 인공 신경회로망이 일반 함수를 원하는 만큼 정확하게 근사화 할 수 있는 장점을 지니고 있기 때문에, 최근 비선형 시스템에 폭넓게 적용하기 위하여 다중의 신경 회로망을 이용한 비선형의 동적 특성을 표현하는 연구가 진행되고 있다 [2]. 또한 신경세포의 동적 특성은 인공 신경모형보다 더 복잡하여 뇌의 동적 특성을 더욱 더 정량적으로 기술하는 동적 신경모형이 연구중이다 [3]. 일반적으로 신경 회로망의 기본 모델인 인간의 정보처리 방법은 다양한 입력 기관을 통해서 들어온 정보를 '학습'이라는 고도의 분산, 병렬처리 메카니즘을 통하여 뇌 속에 이미 구조적으로 결합된 뉴런과 그들 사이의 연결 강도들을 어려운 시행착오를 거치면서 학습의 목적이 가장 잘 부합되도록 조정해 나간다 [4]. 또한 한번 학습되어진 문제에 대해서는 목적에 부합되어 조정되어진 뉴런들 사이의 연결 강도에 의해 입력 기관을 통해 재시되어진 입력 정보는 자동적으로 optimal한 출력 정보를 구해내게 되는데 가장 중요한 사항은 하나의 작업에 대해서만 적용성을 가지지 않고 무한히 많은 작업에 대해서 위의 학습 과정을 수행해 나갈 수 있다 [5]. 본 논문에서는 시뮬레이션(Simulation)을 통하여 제안된 미지의 비선형 동적 시스템의 역전파 학습을 고찰하였다. 역전파 학습 알고리즘의 성능 개선을 위해 학습률과 모멘텀 항의 수정된 역전파 학습 알고리즘을 사용함으로서 최적 규칙으로의 수렴 특성의 향상을 고찰하고 부하 외란이 인가된 미지의 비선형 동적 시스템에 실험적으로 적용함으로써 다양한 응용 문제가 해결될 수 있음을 보이고 그에 따른 결과들을 분석 평가하였다.

### II. 역전파 학습 알고리즘

역전파 학습은 이들 유니트가 그림 2-1과 같이 연결된 구조으로 한다. 유니트  $i$ 의 출력  $I_i$ 가, 결합 계수  $W_{ij}$ 로의 합중합에 의해 유니트  $j$ 의 입력으로 되고 임계치  $\theta_j$ 와 출력을 규격화 하는 함수  $f$ 에 의해 유니트  $j$ 의 출력  $O_j$ 가 결정되는 모델은 식 2-1과 같다.

$$O_j = f \sum_i W_{ij} \cdot I_i + \theta_j \quad (2-1)$$



$W_{ij}$  : 출력층 뉴런  $k$ 와 중간층 뉴런  $j$ 사이의 가중치

$\theta_j$  : 중간층 뉴런  $j$ 와 입력층 뉴런  $i$ 사이의 가중치

그림 2-1. 3층 역전파 학습의 구조

그림 2-1에서는 중간층 유니트  $j$ 의 출력을  $H_j$ , 출력층 유니트  $k$ 의 출력을  $O_k$ , 출력층  $k$ 의 음셋을  $\gamma_k$ , 중간층 유니트  $j$ 와의 결합계수를  $W_{kj}$ 로 하고 있다. 함수  $f(x)$ 는 뉴런의 활성화 함수(Activation function)라 하고 비선형 포화형 응답특성을 가진 시그모이드(Sigmoid)모델 함수를 이용한다. 시그모이드 함수의  $a$ 는 시그모이드 함수의 기울기를 결정하는  $x$ 의 계수이다. 즉, 이 계수는 그 미분이 본래의 시그모이드 함수로 표현할 수 있는 특징을 갖고 있다. 시그모이드 모델은 그림 2-2와 같다.

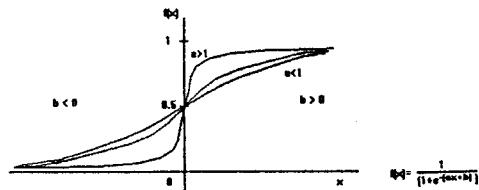


그림 2-2. 시그모이드 함수와 함수 형태

역전파 학습의 구체적인 학습 알고리즘은 다음과 같다.

[1] 네트워크의 상태를 결정하는 연결강도( $W_{ij}$ ), ( $W_{kj}$ )와 음셋(Offset)( $\theta_j$ ), ( $\gamma_k$ )를 각각 작은 값의 난수치로 초기화.

[2] 학습패턴을 설정.

[3] 학습 패턴의 값을 입력층 유니트의 출력( $I_i$ )에 넣고, 입력층에서 중간층에의 결합 하중계수( $W_{ij}$ )와 중간층 유니트  $j$ 의 음셋  $\theta_j$ 를 이용해서 중간층 유니트  $j$ 의 입력  $U_j$ 를 구하고 입력  $U_j$ 와 그림 2-2에 보여주는 것과 같은 시그모이드 함수  $f$ 를 이용하여 중간층 유니트  $j$ 의 출력  $H_j$ 를 구한다.

$$U_i = \sum_j W_{ij} \cdot I_j + \delta_i \quad (2-2)$$

$$H_i = f(U_i) \quad (2-3)$$

[4] 중간층 유니트의 출력 ( $H_i$ )과 중간층 출력층에의 결합형 계수 ( $W_{kj}$ )와 출력층 유니트  $k$ 의 음셋  $\gamma_k$ 를 이용해서 출력층 유니트  $k$ 에의 입력  $S_k$ 를 구하고 입력  $S_k$ 와 시그모이드 함수  $f$ 에 의해 출력층 유니트  $k$ 의  $O_k$ 를 구한다.

$$S_k = \sum_i W_{ki} \cdot H_i + \gamma_k \quad (2-4)$$

$$O_k = f(S_k) \quad (2-5)$$

[5] 학습 패턴의 교사 신호  $T_k$ 와 출력층의 출력  $O_k$ 와의 차로부터의 출력층의 유니트  $k$ 에 연결되는 결합 계수와 중간층 유니트  $k$ 의 음셋에 대한 오차  $\delta_k$ 를 구한다.

$$\delta_k = (O_k - T_k) \cdot O_k \cdot (1 - O_k) \quad (2-6)$$

[6] 오차  $\delta_k$ 와 중간층에서 출력층에의 결합계수 ( $W_{kj}$ )와 중간층의 출력  $H_i$ 으로부터 중간층 유니트  $j$ 에 연결되는 결합계수와 중간층 유니트  $j$ 의 음셋에 대한 오차  $\sigma_j$ 를 구한다.

$$\sigma_j = \delta_k \cdot W_{kj} \cdot H_j + (1 - H_j) \quad (2-7)$$

[7] [5]에서 구한 출력층 유니트  $k$ 에서의 오차  $\delta_k$ 와 중간층 유니트  $j$ 의 출력  $H_j$ 와 상수  $\alpha$ 와의 곱을 가산하는 것으로 입력층 유니트  $i$ 에서 중간층 유니트  $j$ 에 연결되는 결합계수  $W_{ij}$ 를 수정한다. 또 오차  $\sigma_j$ 와 상수  $\beta$ 와의 곱을 가산하는 것으로 중간층 유니트  $j$ 의 음셋  $\theta_j$ 를 수정한다.

$$W_{ij} = W_{ij} + \alpha \cdot \delta_k \cdot H_j \quad (2-8)$$

$$\theta_j = \theta_j + \beta \cdot \sigma_j \quad (2-9)$$

[8] 중간층 유니트  $j$ 에서의 오차  $\sigma_j$ 와, 입력층 유니트  $i$ 의 출력  $I_i$ 와 상수  $\alpha$ 와의 곱을 가산하는 것으로 입력층 유니트  $i$ 에서 중간층 유니트  $j$ 에 연결되는 결합계수  $W_{ij}$ 를 수정한다. 또 오차  $\sigma_j$ 와 상수  $\beta$ 와의 곱을 가산하는 것으로 중간층 유니트  $j$ 의 음셋  $\theta_j$ 를 수정한다.

$$W_{ij} = W_{ij} + \alpha \cdot \sigma_j \cdot I_i \quad (2-10)$$

$$\theta_j = \theta_j + \beta \cdot \sigma_j \quad (2-11)$$

[9] 다음의 패턴을 학습 패턴으로 한다.

[10] 학습 패턴이 종료될 때까지 [3]에 되돌아간다.

[11] 학습이 반복 회수를 간다.

[12] 학습이 반복 회수를 제한 횟수 이하라면 [2]로 간다. 그리고 음셋의  $\theta$ ,  $\gamma$ 는 항상 1을 출력하는 유니트와 결합계수로서 결합 계수  $W_{ij}$ ,  $W_{kj}$ 와 똑같이 학습할 수 있다. 따라서 식(2-8)과 식(2-10)을 이용하여 각 층에서의 새로운 가중치들은 최종적으로 다음 식(2-12)에 의해 조정된다.

$$\begin{aligned} W_{ij}(t+1) &= W_{ij}(t) + \alpha W_{ij} + \zeta [W_{ij}(t) - W_{ij}(t-1)] \\ W_{kj}(t+1) &= W_{kj}(t) + \alpha W_{kj} + \zeta [W_{kj}(t) - W_{kj}(t-1)] \end{aligned} \quad (2-12)$$

여기에서,  $\zeta$ 는 상수이며 오차 전동을 계산하여 수렴 속도를 빨리하기 위하여 첨가한 모멘텀 항(momentum term)이다. 식(2-12)을 출력 단에서 입력 단까지 계속적으로 반복하면 각 뉴런의 출력 오차가 감소하도록 가중치가 변하게 되며, 마지막에는 오차가 없게 되는 역전파 학습이 이루어진다.

### III 역전파 학습 알고리즘의 문제점 및 개선방법

#### III-1 역전파 학습 알고리즘의 문제점

다중 신경회로망은 중간층에서의 비선형 변환에 의한 특성으로 임의의 입출력에 대한 비선형 Mapping이 가능한 장점을 가지고 있으나 학습속도가 느리다는 문제점을 갖는다. 식(3-1)의 오차  $E_p$ 를 감소시키기 위한 출력단 가중치들의 변화량은 미분의 chain rule을 여러번 반복적으로 적용함으로써 다음과 같이 구할 수 있다.  $E_p = \frac{1}{2} \sum_i (T_i - O_i)^2$

식(3-2)와 식(3-3)에서  $\delta_k$ 와  $\delta_i$ 값을 결정할 때 오차의 크기에 상관없이 가중치들의 합이 부적절하게 위치하면 학습이 이루어지지 않는 상태에서도 활성화 함수의 미분 출력값이 0에 가깝게 되며 이때는 식(3-2)과 식(3-3)에서 오차값에 관계없이 가중치 변화량이 작아지게 되는 이른바 조기포화 상태(Premature saturation)에 빠지게된다 [6].

$$\Delta W_k = \eta \cdot \delta_k \cdot H_i \quad (2-2)$$

$$\begin{aligned} \delta_k &= -\delta E / \delta S_k \\ &= -\delta E / \delta O_k \cdot \delta O_k / \delta S_k \\ &= \delta_k \cdot \hat{O}_k \cdot (1 - \hat{O}_k) \end{aligned} \quad (3-2)$$

$$\begin{aligned} \Delta W_i &= -\eta \cdot \delta_i \cdot \hat{I}_i \\ &= \eta \cdot \delta_i \cdot \hat{I}_i \\ \text{단, } \hat{I}_i &\text{ 학습율} \end{aligned} \quad (3-3)$$

조기포화 상태에서 가중치들의 변화량은 매우 작아지게 되며 전체적인 학습 속도에 큰 영향을 미치게 된다. 따라서 이러한 조기 포화 상태에서 활성화 함수의 기울기를 조절해 주면 용이하게 문제의 조기포화 상태를 벗어나게 할 수 있다. 그림 3-1은 활성화 함수  $f(a, x)$ 에서 기울기 계수  $a$ 의 변화에 따른 활성화 함수들과 그 일차 미분값, 즉 기울기 변화를 보여준다.

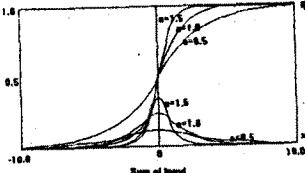


그림 3-1. 뉴런의 활성화 함수와 기울기

#### III-2. 역전파 학습 알고리즘의 개선방법

본 논문에서는 뉴런들의 조기포화 상태를 벗어나게 하기 위한 2가지 방법을 제안한다. 첫째 학습률  $\eta$ 의 값은 학습 속도 연결 강도의 변화량이 큰 값을 가지게 된다. 따라서 학습 속도는 증가하지만 신경망이 수렴하지 못하고 진동할 우려가 있다. 반면에  $\eta$ 의 값이 너무 작으면 진동은 방지되지만 학습 속도가 느리다. 식(3-2), 식(3-3)을 학습을 표현식으로 바꾸면 식(3-4)과 같다. 모멘텀 항  $\zeta$ 는 학습률  $\eta$ 과 유사한 역할을 하는 것으로 모멘텀의 크기를 제어하기 위한 상수이다 [7].

$$\Delta W_k(t+1) = \eta \cdot \delta_k \cdot H_i + \alpha \Delta W_k(t) \quad (3-4)$$

$$\Delta W_i(t+1) = \eta \cdot \delta_i \cdot I_i + \alpha \Delta W_i(t) \quad (3-4)$$

둘째 활성화 함수의 기울기를 변화 시켜주는 방법으로 그림 2-2에서 상수  $b$ 는 시그모이드 함수의 좌우 이동을 결정하는 상수로써,  $b$ 의 값이 0보다 크면 시그모이드 함수는 오른쪽으로 이동하게 되고,  $b$ 의 값이 0보다 작으면 시그모이드 함수는 왼쪽으로 이동하게 된다. 결과,  $b$ 는 역치 함수에서 역치와 같은 역할을 한다. 따라서 식(3-4)의 모멘텀 항에 바이어스(bias)를 추가시키기 위한 식(3-5)와 같다.

$$bias(t+1)_k = bias(t)_k + \alpha \cdot \delta_k \cdot 1 + \beta \Delta bias_k(t) \quad (3-5)$$

$$bias(t+1)_i = bias(t)_i + \alpha \cdot \delta_i \cdot 1 + \beta \Delta bias_i(t) \quad (3-5)$$

바이어스는 입력층을 제외한 모든 신경세포가 가지는 특성으로서 그림 2-2에 보인 시그모이드 함수  $b$ 에 해당하는 역할을 한다. 또한 이것은 신경세포의 활성화에 참여할 뿐만 아니라 다른 연결 가중치들과 마찬가지로 학습에 의해 조절된다.

#### IV 컴퓨터 시뮬레이션 및 결과 고찰

본 논문에서 제안한 방법에 대한 타당성 조사는 역전파 학습 알고리즘을 이용하여 학습한 비선형 동적 시스템이 학습률과 바이어스 변수의 변화에 대한 학습 속도와 조기포화 상태를 벗어나게 할 수 있는지에 중점을 두었다. 시뮬레이션에 사용되는 역전파 학습은 그림 2-1과 같은 3층으로 구성되고, 미지의 비선형 플랜트에 요구되는 응답에 대해 입력 신호  $x(k)$ 를 검토하며, 차분 방정식으로 기술되는 비선형 동적 플랜트를 고찰한다. 단일 입력값을 갖는 다음과 같은 형태의 비선형 동적 시스템이 사용된다.

$$\begin{aligned} x(k+1) &= f(x(k), u(k-1), \dots, x(k-n+1), u(k), u(k-1), \dots) \\ &= f(x(k-m+1)) \end{aligned} \quad (4-1)$$

단,  $[x(k), u(k)]$ 는 이산시간  $k$ 에서 단일 입력 출력 플랜트의 입력 출력 쌍을 나타낸다. 이러한 제어 시스템의 구성이 그림 4-1에 나타나 있다.

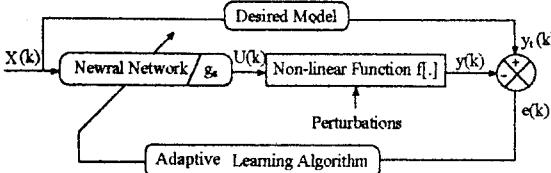


그림 4-1 역전파 학습알고리즘을 이용한 비선형 동적 시스템의 제어기법  
 여기에서,  $g_s$ 는  $\eta$ (학습율)과  $b$ (바이어스)를 나타내며 역전파 학습의 활성화 함수에 새로운 기울기값으로 오차 진동을 적게하여 수렴 속도를 빨리하기 위하여 첨가한 모멘텀(momentum term)이다. 만약 요구되는 출력이  $y_t(k)$ 이고 제어중인 플랜트의 실제 출력이  $y(k)$  일 때 오차가  $e(k) = y_d(k) - y(k)$  이면 제어의 목표는 다음과 같은 제한된 제어입력  $u(k)$ 를 결정하는 것이다.  $\lim [y_d(k) - y(k)] = e(k) = 0$  (4-2)

제어되는 플랜트는 다음 차분 방정식으로 기술 된다.

$$y(k) = f[y(k-1), y(k-2), u(k), u(k-1), u(k-2)] \quad (4-3)$$

이때 미지의 비선형 함수는

$$f[\cdot] = [\sin(\pi(y^2(k-2)+0.5))] + 0.3\sin(2\pi u(k)) / (1+u^2(k-1)+u^2(k-2)) \quad k < 250 \quad (4-4)$$

이고 제어공정이 진행되는 동안  $f[\cdot]$ 는 다음으로 변화되었다.

$$f[\cdot] = [\sin(\pi(y^2(k-2))) + \sqrt{1 + (u^2(k) + u^2(k-1) + u^2(k-2))}] \quad k \geq 250 \quad (4-5)$$

시스템 입력은  $x(k) = \sin(2\pi k/250)$  이었다.

이때 얻어진 시뮬레이션 결과가 그림 4-2, 4-3, 4-4, 4-5에 나타나 있다. 그림 4-2는 시스템 입력  $x(k)$ 와 미지의 비선형 함수  $f[\cdot]$  특성을 보여 준다. 초기 가중치는 0.8694로 고정하고 학습율과 바이어스값을 변화시켜 보았다. 이때 사용한 학습율과 바이어스의 변수는 표 4-1와 같다. 본 연구에서의 최적 학습율과 바이어스는 각각 0.0003과 0.001으로 나타났다. 따라서 그림 4-5를 기준으로 할 때 그림 4-3는 바이어스를 0.01으로 0.009의 변화에 대한 오차 응답에 미치는 영향은 매우 민감함을 볼 수 있고, 또한 그림 4-5를 기준으로 할 때 그림 4-4는 학습율을 0.0006으로 0.0003의 변화에 대한 오차응답을 보여준다.

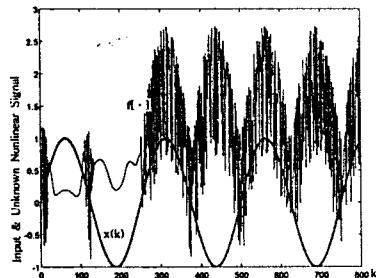


그림 4-2 시스템 입력  $x(k)$ 와 미지의 비선형 함수  $f[\cdot]$  특성

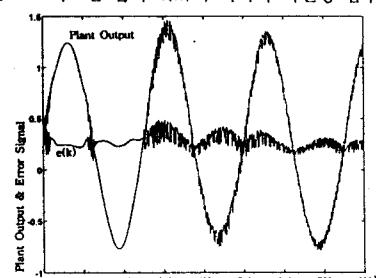


그림 4-3 초기가중치  $w=0.8694$ , 바이어스  $b=0.01$ , 학습율  $\eta=0.0003$  일 때 플랜트 출력과 오차응답  $e(k)$

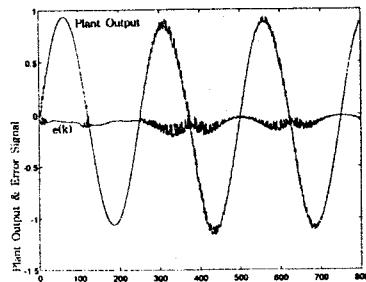


그림 4-4 초기가중치  $w=0.8694$ , 바이어스  $b=0.001$ , 학습율  $\eta=0.0006$  일 때 플랜트 출력과 오차응답  $e(k)$

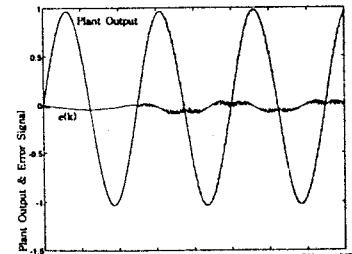


그림 4-5 초기가중치  $w=0.8694$ , 바이어스  $b=0.01$ , 학습율  $\eta=0.0003$  일 때 플랜트 출력과 오차응답  $e(k)$

표 4-1 역전파 학습에 사용된 학습율과 바이어스 변수값

	학습율 ( $\eta$ )	바이어스 ( $b$ )	초기 가중치 ( $w$ )
1(그림 4-3)	0.0003	0.01	0.8694
2(그림 4-4)	0.0006	0.001	0.8694
3(그림 4-5)	0.0003	0.001	0.8694

## V. 결론

본 연구에서는 역전파 학습알고리즘을 이용하여 비선형 동적 시스템 제어에 관해서 고찰하였다.

본 연구에서는 다음과 같은 결론을 얻을 수 있었다.

1. 역전파 학습알고리즘을 이용하여 비선형 동적 시스템에 적용한 결과 신경망 학습의 조기 포화 상태의 문제점을 해결함으로서 시스템의 출력이 요구되는 신호에 잘 추종함을 알 수 있었다.
2. 초기 오차가 커서 최초의 주기 동안에는 신경망의 훈련이 느렸음이 컴퓨터 시뮬레이션을 통해 관측된다. 그러나 시스템 응답에 나타난 외란의 영향은 초기 오차를 제외하고 훈련이 진행됨에 따라 비선형 동적 플랜트는 오차가 작아진 것의 결과로 플랜트 응답이 요구되는 신호에 균형함이 관측된다. 따라서 역전파 학습 알고리즘을 이용한 비선형 동적 시스템 설계시 학습율과 바이어스의 조절 가능한 매개변수( $\eta$ , bias)을 초기에 적절히 조절함으로써 오차와 진동을 줄이고 향상된 시스템 응답을 얻을 것으로 사료된다.

## 참고문헌

- [1] 장성환, 자동제어, 1992
- [2] 김대수 신경망이론과 응용, 1994
- [3] S. Y. KUNG, Digital Neural Network, 1993
- [4] M. M. Guta, D. H. Rao, and Gao, Learning and adaption in neural control of highorder linear systems, ACC pp. 3044-3048, 1992.
- [5] M. M. Guta, D. H. Rao, and H. G. Wood, "Learning and adaptive neural controller," The Int. Joint Conference on Neural Networks, Singapore, Nov. 1991.
- [6] S. A. Billings, H. B. Jamaliudding, and S. Chen, "Properties of Neural Networks with Applications to Modeling Nonlinear Dynamical Systems," Int. J. Control, 1992
- [7] K. s Narendra and K. Parthasarathy, "Identification and Control of Dynamical System Using Neural Network," IEEE Trans. Neural Network, 1, 4(1990)