

Optimization of Polynomial Neural Networks: An Evolutionary Approach

金東源* · 朴貴泰**
(Dong-Won Kim · Gwi-Tae Park)

Abstract - Evolutionary design related to the optimal design of Polynomial Neural Networks (PNNs) structure for model identification of complex and nonlinear system is studied in this paper. The PNN structure is consisted of layers and nodes like conventional neural networks but is not fixed and can be changable according to the system environments. three types of polynomials such as linear, quadratic, and modified quadratic is used in each node that is connected with various kinds of multi-variable inputs. Inputs and order of polynomials in each node are very important element for the performance of model. In most cases these factors are decided by the background information and trial and error of designer. For the high reliability and good performance of the PNN, the factors must be decided according to a logical and systematic way. In the paper evolutionary algorithm is applied to choose the optimal input variables and order. Evolutionary (genetic) algorithm is a random search optimization technique. The evolved PNN with optimally chosen input variables and order is not fixed in advance but becomes fully optimized automatically during the identification process. Gas furnace and pH neutralization processes are used in conventional PNN version are modeled. It shows that the designed PNN architecture with evolutionary structure optimization can produce the model with higher accuracy than previous PNN and other works.

Key Words: Polynomial Neural Network, Optimization, Evolutionary approach, Gas furnace and pH neutralization process

1. 서 론

시스템 분석 및 제어와 자동화를 위해 시스템 모델링은 매우 중요하다. 선형시스템의 모델링 방법은 그동안 많이 발표되었으나, 상대적으로 비선형 시스템에 대한 것은 여전히 초기단계에 있다 [1]. 우리가 접하는 대부분의 시스템은 비선형이다. 따라서 최근에는 이러한 비선형 시스템의 모델링 방법들이 다양하게 연구되고 있으며, 그 중에서 특히 퍼지시스템과 신경망을 이용한 모델링방법에 많은 관심이 집중되고 있다 [18-24]. 퍼지 시스템은 규칙기반 시스템으로 추론기능이 우수하며, 신경망은 강력한 학습기능으로 미지의 시스템을 모델링 하는데 효율적이다. 또 다른 방법으로, Group Method of Data Handling (GMDH) [2] 타입 알고리즘이 있다. GMDH는 1970년 초에 Ivakhnenko 의해 소개되었으며, 1970년대 중반부터 복잡한 비선형 공정의 모델링과 예측을 위해 활발히 이용되었다 [3-7]. GMDH 타입 알고리즘은 자기구성의 형태를 보이며, 입출력 변수 사이에서 선택적인 경험이나 지식을 사용하지 않고 중요한 입력변수를 선

택할 수 있는 것이 특징이다 [7-8]. 이것은 복잡한 비선형 시스템에서 파악하기 힘든 변수들에 대하여 설계자의 경험과 사전지식에 의존하여 모델설계를 행하던 기존의 방법에 비하여 커다란 장점이 되었다. 다항식 뉴럴네트워크(polynomial neural networks, PNNs) [9-10]는 이러한 GMDH 타입 알고리즘의 한 형태이다. PNN 알고리즘은 입출력 데이터를 분할하고, 입력변수를 선택하며, 선택된 입력변수로 부분표현식을 이용하여 비선형 시스템을 모델링하는 방법이다. PNN 구조는 기존의 뉴럴 네트워크처럼 층과 노드로 구성되며 각 노드의 출력은 다수의 입력변수들의 1차, 2차, 및 변형된 2차 다항식으로 계산된다. PNN 알고리즘은 전체 입력 변수에서 임의의 입력변수 수를 선택하고, 이로 생성되는 조합의 총수를 노드의 개수로 정의한다. 또한 위의 3 가지 형태의 차수 중 설계자에 의하여 미리 정의된 차수와 선택된 입력변수로 부분표현식을 구성한다. 그러므로 노드의 입력 변수와 차수는 부분표현식을 구성하는 핵심요소로 PNN 모델의 구조와 성능에 직접적으로 연관된다. 대부분의 경우, 이러한 요소들은 과중한 연산 시간과 저효율의 부담을 안고 설계자의 시행착오에 의하여 결정된다. 따라서 설계자의 경험적인 방법으로 생성된 최종적인 PNN 모델은 모델링과 예측성능을 위한 좋은 결과를 가진 네트워크는 될 수 있어도 비선형 시스템에 대한 최상의 모델을 보장할 수는 없다. 최적의 성능을 가진 합리적인 모델설계를 위해 각 노드의 부분표현식을 구성하는 입력변수와 차수는 객관적이고 신뢰성 높은 자동 선택방법으로 결정해야 한다.

* 正 會 員 : 高麗大學校 工科學科 電氣工學科 博士課程
 **正 會 員 : 高麗大學校 工科學科 電氣工學科 教授 · 工博
 接受日字 : 2002年 10月 16日
 最終完了 : 2003年 5月 3日

본 논문에서는 PNN 알고리즘의 이러한 단점을 극복하고 부분표현식을 구성하는 중요 요소인 입력변수와 차수의 최적선택을 위해 유전 알고리즘을 이용한다. 유전 알고리즘 (genetic algorithm, GA) [11]은 자연계에 존재하는 적자생존의 원리와 유전학에 기초하여 집단내 개체들을 모의 진화시킴으로써 전역해(global minima)를 찾아가는 최적화 기법이다. 유전 알고리즘이 기존의 최적화 이론들과 다른 점은 도함수를 이용하지 않으며 문자열로 나타낸 집단을 사용한다는 것이다. 또한 전역탐색을 수행함으로 부분극소점(local minima)에 빠질 가능성이 적다. 이러한 장점 때문에 최근 주목받는 최적화 기법의 하나가 되었으며 신경망, 퍼지이론과 결합하여 적용된 다양한 사례들이 발표되고 있다. Lin [12]은 유전 알고리즘을 신경망의 연결강도를 최적화하는데 이용하였고, Homaifar [13]은 유전 알고리즘을 이용하여 퍼지 제어기의 소속함수와 제어규칙을 동시에 최적화하였다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 PNN 구조 [9]와 알고리즘을 소개하고 그 문제점을 제시한다. 3장에서는 진화론적 방법을 이용하여 새로운 PNN의 최적구조 설계방법을 제안한다. 4장에서는 제안된 모델의 성능평가를 위해, 앞서 연구된 다양한 퍼지 및 지능형 모델링 방법에 적용된 가스로 공정의 시계열데이터 [14]와 화학공정이나 오·폐수처리, 생물공학 등 많은 분야에서 널리 사용되는 pH중화 공정 [15]을 모델링 하여 그 결과를 고찰한다. 또한 기존 PNN 모델과 성능 비교를 통해 새로 제안한 모델의 성능이 보다 우수함을 보인다. 마지막으로 5장에서 결론을 맺는다.

2. 다항식 뉴럴네트워크 알고리즘

PNN 알고리즘 [9]은 다변수 입력변수에서의 특정 변수와 차수의 선택, 부분표현식의 구성, 매개변수 추정, 알고리즘의 종료판정 등의 일련의 과정에 의해 모델을 선정하게 되고, 계층적 모델추정에 의하여 최종적으로 모델 추정식을 얻는다. 먼저 다음과 같은 입력력 관계식을 만족하는 데이터 군이 주어졌다고 한다.

$$y = f(x_1, x_2, \dots, x_N) \tag{1}$$

$$(x_i, y_i) = (x_{1i}, x_{2i}, \dots, x_{Ni}, y_i), i=1, 2, \dots, n \tag{2}$$

여기서, N 과 n은 각각 전체 입력변수의 수와 각 입·출력 데이터 쌍의 개수를 의미한다. PNN 알고리즘은 관계식 f 의 추정 모델 \hat{f} 를 다음과 같은 다항식으로 표현하는 방법이다.

$$\hat{y} = \hat{f}(x_1, x_2, \dots, x_N) = c_0 + \sum_{k_1} c_{k_1} x_{k_1} + \sum_{k_1, k_2} c_{k_1, k_2} x_{k_1} x_{k_2} + \sum_{k_1, k_2, k_3} c_{k_1, k_2, k_3} x_{k_1} x_{k_2} x_{k_3} + \dots \tag{3}$$

여기서 \hat{y} 와 $c_{k_1, k_2, \dots}$ 는 각각 출력 y의 추정치와 계수이며, $\hat{f}(\cdot)$ 는 추정함수이다.

[9]에 있는 기존 PNN 알고리즘의 설계순서를 간략히 기술하면 다음과 같다.

[단계 1] 입력변수 설정

출력변수 y에 관련된 임의의 N개의 입력변수를 설정한

다. 필요하면 데이터를 정규화 한다. 이 입력변수를 x_1, x_2, \dots, x_N 이라고 하자.

[단계 2] 데이터 분할

n 개의 입출력 데이터 $(x_i, y_i) = (x_{1i}, x_{2i}, \dots, x_{Ni}, y_i), i=1, 2, \dots, n$ 을 학습용 데이터와 평가용 데이터로 분할한다. 학습용 데이터는 모델동정을 위해 사용하는 데이터이고, 평가용 데이터는 동정된 모델을 평가하기 위해 사용된다. 학습용 데이터와 평가용 데이터의 개수를 각각 n_L 개와 n_C 개로 한다. 즉, $n = n_L + n_C$ 이다.

[단계 3] PNN 구조 선택

PNN 구조는 첫 번째 층과 그 이후 층의 부분표현식을 구성하는 방법에 따라 그 구조가 정해진다. 첫 번째 층과 그 이후 층의 각 노드의 입력변수 수가 동일한 경우와 그렇지 않은 경우의 두 구조로 분류되며, 또 각 구조는 노드의 부분표현식 (partial description, PD)의 차수에 따라 각각 Case 1과 Case 2로 다시 세분화된다. 표 1은 PNN구조를 정리한 것이다.

- (a) 기본 PNN 구조 - Case 1, Case 2
- (b) 변형된 PNN 구조 - Case 1, Case 2

표 1 전형적인 PNN 구조

Table 1 Structure of conventional PNN

Layer	PD type	No. of input to each node	order	PNN structure
1st layer		i	Type I	▶ i=j: Basic PNN · I=J: Case 1 · I≠J: Case 2
2nd layer or higher		j	Type J	▶ i≠j: Modified PNN · I=J: Case 1 · I≠J: Case 2

$$(i, j=2, 3, \dots; I, J=1, 2, 3)$$

[단계 4] 노드의 차수 및 입력변수 수 선택

N개의 입력변수 중에서 각 노드에 적용할 입력변수 수를 선택한다. N에서 r개의 변수를 각 노드의 입력으로 선택하면 생성되는 노드의 총 개수는 $\frac{M}{(N-r)!r!}$ 가 되며, N은 전체 입력변수의 수이고, r은 설계자가 택한 입력변수의 수이다. 노드의 부분표현식 차수 또한 설계자가 정한다. 노드의 차수 및 입력변수 수에 따른 PD는 표 2와 같다.

[단계 5] 부분표현식의 매개변수 추정

[단계 4]에서 노드의 PD형태가 정해지면 학습용 데이터를 사용하여 매개변수 (c_1, c_2, \dots) 값을 추정한다. 최소자승법에 의해 (4)를 최소화하는 매개변수 값을 구한다.

$$E_k = \frac{1}{2} \sum_{i=0}^{n_i} (y_i - z_{ki})^2, \quad k=1, 2, \dots, \frac{M!}{(N-r)!r!} \tag{4}$$

여기서 z_{ki} 는 k 번째 노드의 PD에서 i 번째 입력 데이터를 사용했을 때의 값을 나타낸다.

[단계 6] 평가 후 노드의 선택

[단계 5]에서 구한 매개변수로 이루어진 각 노드의 PD에

평가용 데이터 (n_c)를 사용하여 (4)의 E_k 를 계산한다. 여기서 얻은 $E_1, E_2, \dots, E_{N/(N-r)!}$ 에 대하여, 값이 작은 순으로 정렬하고 이들 중 값이 작은 순으로 w 개의 노드를 선택하고 나머지 노드는 버린다. 여기서, w 는 설계자가 정한 값이다.

$$\frac{M}{(N-r)!} < w :$$

다음세대 노드의 개수는 $\frac{M}{(N-r)!}$ 가 됨

$$\frac{M}{(N-r)!} \geq w :$$

다음세대 노드의 개수는 w 가 됨

표 2 입력변수 수에 따른 부분표현식

Table 2 PD according to a number of input variables

No. of inputs	Polynomial Order	Partial Description (PD)
2	Type 1	$c_0+c_1x_1+c_2x_2$
	Type 2	$c_0+c_1x_1+c_2x_2+c_3x_1^2+c_4x_2^2+c_5x_1x_2$
	Type 3	$c_0+c_1x_1+c_2x_2+c_3x_1x_2$
3	Type 1	$c_0+c_1x_1+c_2x_2+c_3x_3$
	Type 2	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_1^2+c_5x_2^2+c_6x_3^2+c_7x_1x_2+c_8x_1x_3+c_9x_2x_3$
	Type 3	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_1x_2+c_5x_1x_3+c_6x_2x_3$
4	Type 1	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_4$
	Type 2	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_4+c_5x_1^2+c_6x_2^2+c_7x_3^2+c_8x_4^2+c_9x_1x_2+c_{10}x_1x_3+c_{11}x_1x_4+c_{12}x_2x_3+c_{13}x_2x_4+c_{14}x_3x_4$
	Type 3	$c_0+c_1x_1+c_2x_2+c_3x_3+c_4x_4+c_5x_1x_2+c_6x_1x_3+c_7x_1x_4+c_8x_2x_3+c_9x_2x_4+c_{10}x_3x_4$
multi-input

[단계 7] 종료 판정

PNN 구조는 설계자가 미리 정한 층까지 노드의 PD가 결정되었을 때 종료된다. 최종 층의 노드 중에서 (4)의 E_k 값이 가장 적은 노드를 택하고 나머지는 버린다. 그 이전 층에서는 역순으로 최종 층의 노드출력에 영향을 미치는 노드만 남기고 나머지는 모두 버리면 PNN 구조가 완성된다.

[단계 8] 새로운 입력변수 선택

종료조건을 만족하지 않을 경우, 전 층의 출력값이 다음 층의 입력값이 되어 새로운 입출력 데이터를 구성하고, 단계 4로 돌아간다. 이후, 단계 4부터 단계 8까지를 반복한다.

PNN 구조의 전체 개요도는 그림 1과 같다. 그림 1에서 보듯이 PNN 구조는 층과 노드로 구성되었으나 고정되지 않는다. 또한 각 노드는 다변수 입력과 다양한 차수를 사용할 수 있다. 그러므로 기존의 GMDH보다 성능이 향상되었다.

PNN 알고리즘은 여러 가지 복잡한 공정에 적용한 결과 우수한 모델링 특성을 가짐을 확인할 수 있다 [9-10]. 그러나 위와 같은 이점에도 불구하고 개선해야 할 중요한 문제점들이 있다. 첫째, 각 층에서 PD (노드)의 총 개수는 설계자가 정한 노드의 입력 변수 수로 생성될 수 있는 모든 경우의 수를 조합하여 얻는다. 따라서, 만약 시스템의 입력변수와 데이터가 많다면 그에 따른 네트워크는 매우 복잡해진다. 이것은 해공간이 입력 변수의 규모에 따라 기하급수적으로 늘어나며 따라서 매우 많은 연산 시간을 소요한다. 이

와 같은 이유로 입력변수와 데이터가 많은 경우 최종적인 추정모델을 얻기 위해 상당히 복잡한 네트워크를 생성할 수밖에 없다. 반대로 입력변수의 수가 적다면 원활하게 네트워크 생성을 하지 못하여 좋은 성능을 유지하지 못한다 [17,25]. 둘째로 PNN은 PD의 형태에 따라 모델 성능이 각각 달라진다. 따라서 모델링 과정이 진행되기 전에 PD를 구성하는 특정한 입력변수의 수와 차수의 형태를 설계자가 미리 정의해야만 한다 [9,25]. 셋째, PNN은 설계자가 정한 모델링 조건 즉 PD를 구성하는 입력변수 수와 차수에 따라 그 구조가 결정되기 때문에 결정된 노드의 입력변수의 수와 차수가 일반적으로 최적이라고 보장할 수 없다 [25]. 따라서 본 논문에서는 이러한 단점을 해결하기 위해 유전자 알고리즘을 이용한 PNN의 진화론적 설계방법을 제안한다.

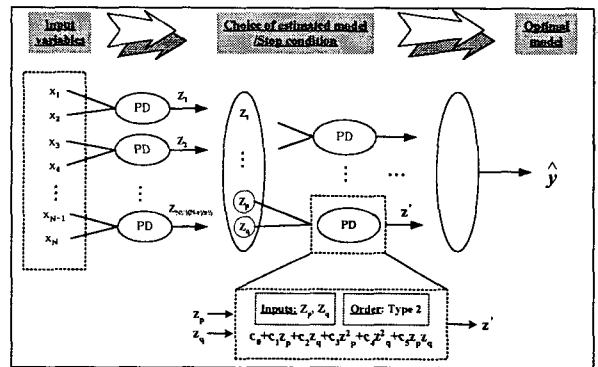


그림 1 다항식 뉴럴네트워크 구조의 전체 개요도
Fig. 1 Overall architecture of the Polynomial Neural Network

3. 다항식 뉴럴네트워크의 진화론적 설계

3.1 유전자 알고리즘에 의한 부분표현식(PD) 설계

최근 들어 유전자 알고리즘은 자연 선택과 유전학에 기반한 확률적인 탐색방법으로써 생산, 교차, 돌연변이의 과정을 수행하는 탐색법이며, 비선형 최적화 이론으로 탁월한 성능을 발휘하고 있다. 본 논문에서는 이러한 유전 알고리즘을 이용하여 2장에서 언급한 문제들을 해결하고자 한다. 다시 말하면 전체 입력변수 중에서 각 노드에 최적의 입력변수의 개수와 이에 해당되는 입력변수는 어떤 것이며, Type 1, 2, 3 중에서 최적의 차수는 어떤 것인가를 유전 알고리즘을 통해 자동 선택하는 것이다.

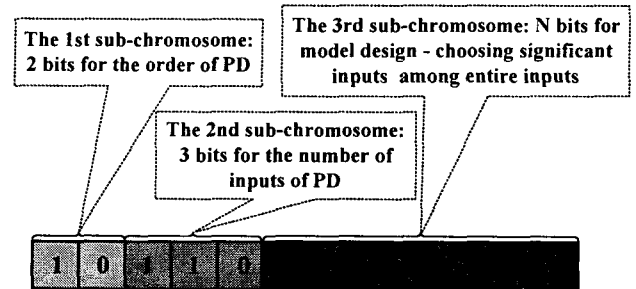


그림 2 부분표현식을 이진스트링으로 표현한 염색체의 구조
Fig. 2 Structure of binary chromosome for PD

각 노드의 부분표현식에 대한 정보는 이진 스트링의 염색체로 표현하며, 이 염색체는 3개의 부염색체로 구성한다. 이진 스트링으로 표현된 PD의 염색체구조는 그림 2와 같다. 첫 번째 부염색체는 2 비트로 구성되며 표 3과 같이 PD의 차수를 결정하는 정보를 제공한다. 3 비트로 구성된 두 번째 부염색체는 PD의 입력 변수의 수를 정하며 표 4와 같다. 마지막 N 비트로 이루어진 세 번째 부염색체는 출력변수 y에 관련된 N개의 입력변수를 나타낸 것으로 두 번째 부염색체에서 결정된 입력변수의 수만큼 실제로 선택된다. 여기서 염색체의 특정위치의 비트가 1이 됨은 해당되는 입력변수가 PD를 구성하는 입력변수로 사용됨을 의미하며, 0은 반대의 의미를 가진다.

표 3 첫 번째 부염색체에 따른 부분표현식의 차수
Table 3 Order of PD represented by the 1st sub-chromosome

Bits in the 1st sub-chromosome	Order of PD
00	Type 1
01	Type 2
10	
11	Type 3

표 4 두 번째 부염색체에 따른 부분표현식의 입력변수 수
Table 4 Number of inputs of PD represented by the 2nd sub-chromosome

Bits in the 2nd sub-chromosome	No. of inputs of PD
000	1
001	2
010	
011	3
100	
101	4
110	
111	5

진화 알고리즘에 의한 PD의 구성절차는 그림 3과 같다. 하나의 PD에 선택될 수 있는 최대 입력변수 수는 m으로 정의하며 시스템의 특성과 설계스펙을 고려하여 설계자가 결정한다. 만약 m보다 많은 수의 입력변수가 PD에 인가된다면 표 4와 같이 입력변수 수는 재 선택된다. PD의 염색체와 이에 해당되는 노드를 예로 들면 그림 4와 같다. PD노드에서 좌상의 2는 다항식의 차수로 Type 2를 의미하고, 좌하의 2는 PD로 인가되는 입력변수의 수를 나타낸다. 이것은 해당 염색체의 정보에 의하여 자동적으로 결정된다. 또 선택된 입력변수로는 1의 비트에 해당되는 입력으로 x_1 과 x_N 이 선택된다. 그 결과 노드의 출력, 즉 y의 추정치 \hat{y} 는 다음과 같다.

$$\hat{y} = \hat{f}(x_1, x_N) = c_0 + c_1x_1 + c_2x_N + c_3x_1^2 + c_4x_N^2 + c_5x_1x_N \quad (5)$$

여기서, 계수 c_0, \dots, c_5 는 학습데이터를 이용하여 2장의 [단계 5]에서처럼 구한다.

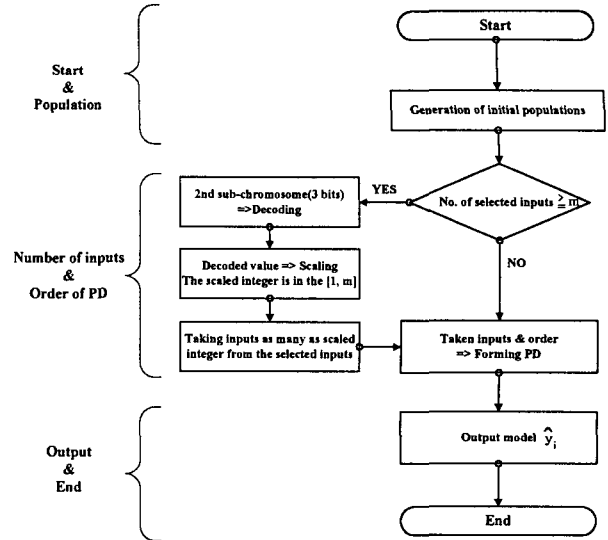
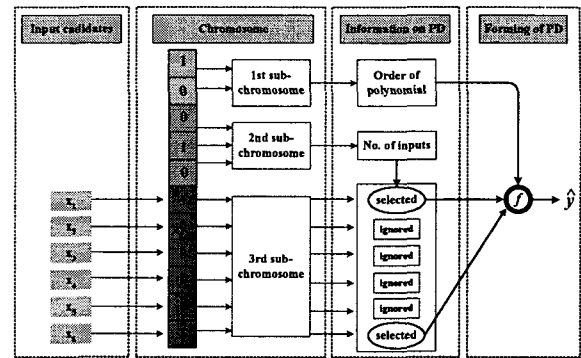
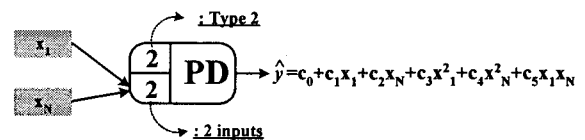


그림 3 진화알고리즘에 의한 PD의 구성절차
Fig. 3 Procedure of forming PD by evolutionary algorithm



(a) 부분표현식의 염색체



(b) 염색체에 해당되는 PD를 표현한 노드

그림 4 부분표현식의 염색체와 이에 해당하는 노드 표현식
Fig. 4 Chromosome of PD and its corresponding node description

유전 알고리즘에 기반한 모델의 최적화 과정은 그림 5와 같다. 초기화 단계는 무작위로 개체들의 초기 집단을 형성한다. 여기서 유전 알고리즘으로 최적화하려는 파라미터들은 염색체로 인코딩 된다. 인코딩된 염색체들은 하중값을 가진 목적함수로 평가되며 각각의 적합도 값을 가진다. 다음 단계는 종료조건을 판정하는 것으로 유전 알고리즘은 각 층에서 정해진 세대수까지 수행한다. 종료조건을 만족하지 않는다면 유전 연산자에 의해 새로운 염색체가 다시 형성되고 목적함수 또한 재평가된다. 엘리트 전략이 가미된 룰렛 휠 방법으로 재생산이 이루어지며, 교배연산은 일점교배가 사용되었고, 돌연변이 연산은 0을 1로 혹은 1을 0으로 바꾸

는 인버트 형식을 사용했다. 만약 각 층에서 정한 종료조건을 만족한다면 다음 층의 새로운 입력변수를 위해 좋은 적합도 값을 가진 염색체들이 선택된다. 각 층에 대한 파라미터는 표 5와 같다. 이와 같은 일련의 연산과정이 끝나고 최종적으로 3층에 도달한다면 가장 좋은 성능지수를 가진 하나의 염색체가 최종 출력값으로 선택된다. 이처럼 2장에서 단계 3인 PNN 구조선택과 단계 4인 노드의 차수 및 입력변수 수 선택은 유전 알고리즘을 기반으로 하여 최적으로 선택되며 단계 5와 단계 6 또한 자동 실행된다.

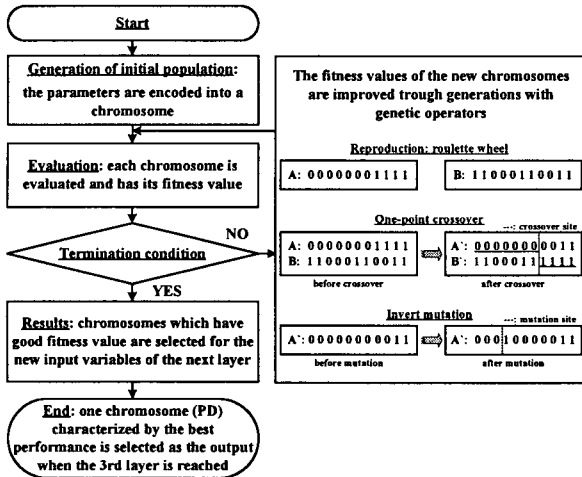


그림 5 유전알고리즘에 기반한 최적화 과정
Fig. 5 Optimization process using evolutionary algorithm

3.2 하중값을 가진 목적함수

입출력 데이터로 시스템을 모델링할 때 가장 문제가 되는 것은 특정한 데이터만을 이용해서 모델을 설계하는 것이다. 이는 주어진 데이터에 의해서만 성능향상을 꾀하였기 때문에 실 공정에 적용하면 많은 오차를 발생시키고 정확한 예측능력을 가지지 못한다. 이러한 현상은 비선형 공정의 경우 더욱 큰 문제가 된다. 그것을 해결하기 위해 본 논문에서는 주어진 학습 데이터를 이용하여 모델을 구축하고 테스트 데이터를 사용하여 설계된 모델을 평가한다. 또한 모델 평가의 기준인 근사화 및 일반화 능력의 상호 균형을 위해 목적 함수를 이용한다. 이를 위해 학습 및 테스트 데이터의 성능 결과에 대한 하중 계수값 [26]을 가진 목적함수 [10]를 다음과 같이 정의한다.

$$f(PI, EPI) = \theta \times PI + (1 - \theta) \times EPI$$

$$fitness = \frac{1}{1 + f(PI, EPI)} \quad (6)$$

PI는 학습데이터에 대한 성능지수, EPI는 테스트데이터에 대한 성능지수, 그리고 θ 는 PI와 EPI에 대한 하중계수를 나타낸다. (6)은 하중계수값 설정에 따라 다음과 같은 특징을 가진다.

1. $\theta = 1$, $f(PI, EPI) = PI$ 이므로 학습 데이터를 기반으로 최적화된다.
2. $\theta = 0.5$, $f(PI, EPI) = 0.5PI + 0.5EPI$ 이므로 학습과

테스트 모두 같은 비중으로 최적화된다.

3. $\theta = 0$, $f(PI, EPI) = EPI$ 이므로 테스트 데이터를 기반으로 최적화된다.
4. $\theta \in [0, 1]$ 이면 모델의 근사화 능력과 일반화 능력을 상호 고려한 목적함수의 형태로 적절한 하중계수 설정에 따라 최적의 모델을 구축할 수 있다.

4. 시뮬레이션 및 결과 고찰

본 논문에서 제안한 모델을 가스로 공정 [9-10,14]과 pH 중화 공정 [9,15,18]의 모델 동정에 적용하고 그 결과를 여러 가지 기존모델링 기법의 결과와 비교함으로써 제안한 방법의 우수성을 알아본다.

4.1 가스로 공정

본 논문에서 제안한 모델의 성능평가를 위해 Box와 Jenkin이 사용한 가스로 시계열 데이터를 이용한다. 시스템의 입·출력변수는 가스 흐름율 $u(t)$ 과 연소된 이산화탄소의 농도 $y(t)$ 이며 다음과 같은 벡터형태를 사용하였다.

[입력: 출력 변수] $\rightarrow [u(t-3), u(t-2), u(t-1), y(t-3), y(t-2), y(t-1): y(t)]$

총 296개의 데이터 중 처음 148개의 데이터는 학습에 이용하고 나머지 148개의 데이터는 테스트에 이용한다.

진화론적 설계에는 다음과 같이 이루어진다. 1, 2, 3층에서 각각 20, 60, 80개의 염색체를 생성한다. 생성된 염색체는 각 층에서 40, 60, 80세대 동안 진화과정을 수행한다. 1층에서 20개의 염색체는 유전자알고리즘에 의해서 정의된 입력변수와 차수를 가지고 각각의 PD를 구성한다. 즉 전체 6 시스템입력 중 노드의 입력변수와 차수의 형태를 염색체로 자동 선택하며 이에 따라 여러 가지 형태의 PD를 구성한다. 각 PD는 정해진 세대수까지 진화하면서 최적의 입력변수와 차수를 탐색한다.

표 5 진화론적 설계를 위한 유전알고리즘의 매개변수

Table 5 Parameters for evolutionary design

parameters	1st layer	2nd layer	3rd layer
Maximum generation	40	60	80
Population size	20: (15)*	60: (50)*	80
Chromosome length	11	20	55
Crossover rate	0.85		
Mutation rate	0.05		
weighting factor: θ	0.1~0.9		
Type (차수)	1~3		

(*) : 다음 층의 입력변수로 선택될 노드 수

염색체의 진화를 통해 탐색된 최적의 정보로 구성된 PD는 학습데이터와 테스트데이터를 이용하여 성능지수 (동정 오차)가 구해진다. 동정오차는 평균자승오차 (MSE)로 모델의 정밀도를 평가할 수 있으며 (7)로 정의된다.

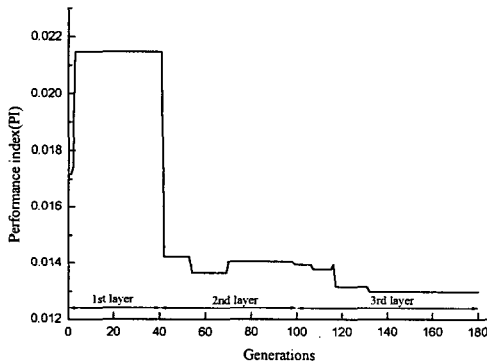
$$PI(EPI) = \frac{1}{148} \sum_{i=1}^{148} (y_i - \hat{y}_i)^2 \quad (7)$$

각 PD는 성능지수뿐만 아니라 목적함수에 의해 평가되며, 각각의 적합도 값에 따라 순위가 정해진다. 여기서 우수한 적합도 값을 가진 PD는 선택되어 다음 층까지 유지되며 그렇지 못한 것은 제거된다. 각 층에서 선택되는 PD의 수는 1, 2, 3층에 따라 각각 15, 50, 80개가 된다. 선택된 PD의 출력값은 다음 층의 입력값이 되어 새로운 입출력 데이터를 구성한다. 이제 새로운 층의 PD정보가 다시 탐색되며 위의 과정이 반복된다. 이 과정은 종료판정 기준인 3층까지 반복 진행되며 3층에서의 가장 좋은 성능지수가 최종 출력값이 된다. 최적의 성능을 얻기 위한 유전 알고리즘의 매개변수와 설정치는 표 5와 같다.

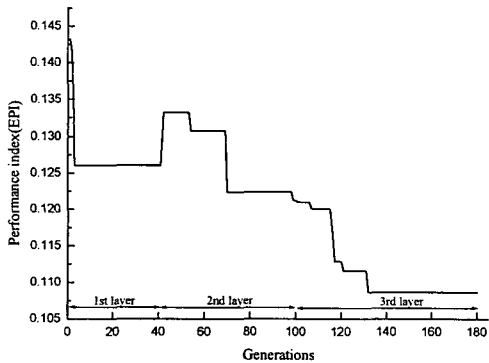
표 6 하중계수에 따른 제안된 방법의 성능지수

Table 6 Performance index of proposed method according to the various weighting factors

θ	1st layer		2nd layer		3rd layer	
	PI	EPI	PI	EPI	PI	EPI
0.1	0.0214	0.1260	0.0200	0.1231	0.0199	0.1228
0.25	0.0214	0.1260	0.0149	0.1228	0.0145	0.1191
0.5	0.0214	0.1260	0.0139	0.1212	0.0129	0.1086
0.75	0.0214	0.1260	0.0139	0.1293	0.0138	0.1235
0.9	0.0173	0.1411	0.0137	0.1315	0.0129	0.1278



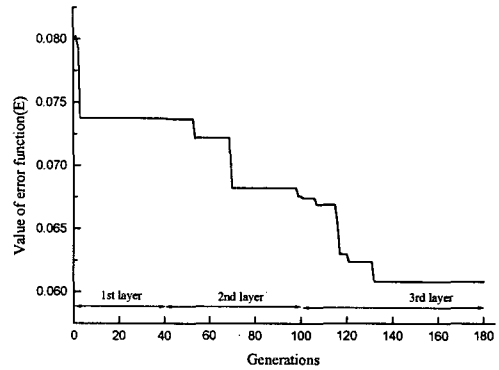
(a) PI



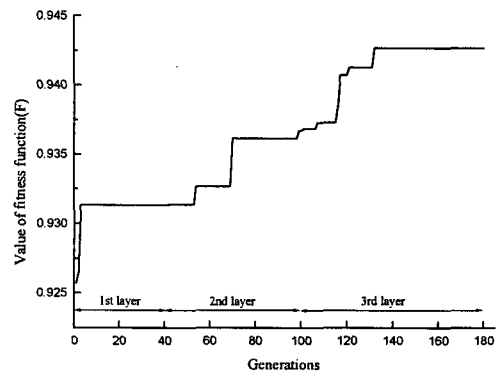
(b) EPI

그림 6 유전자 알고리즘에 의한 성능지수의 수렴 ($\theta=0.5$)
Fig. 6 Convergence process of performance index by GA ($\theta=0.5$)

다양한 하중계수에 따른 결과는 표 6과 같다. 표에서 보는 바와 같이 가장 우수한 성능의 결과 ($PI=0.0129$, $EPI=0.1086$)는 $\theta=0.5$ 일 때이다.



(a) 목적함수



(b) 적합도 함수

그림 7 유전자 알고리즘에 의한 적합도의 수렴과정 ($\theta=0.5$)
Fig. 7 Convergence process of fitness value by GA ($\theta=0.5$)

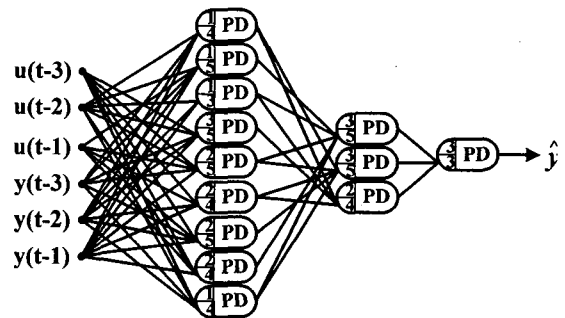
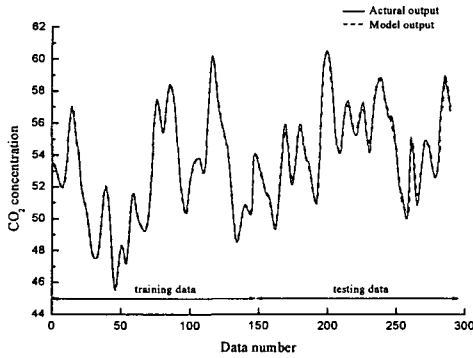


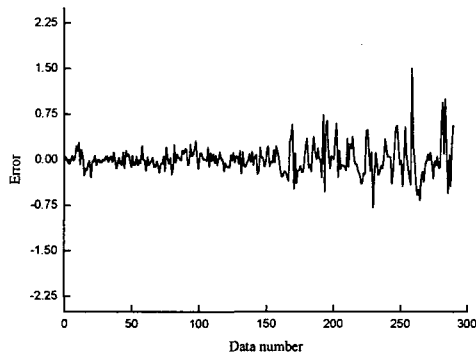
그림 8 유전자 알고리즘을 통해 생성된 최적구조 ($\theta=0.5$)
Fig. 8 Optimal structure generated via GA ($\theta=0.5$)

$\theta=0.5$ 일 때, 각 층에 따른 GA의 연속된 세대수의 성능지수가 수렴하는 과정을 그림 6에서 나타내었다. 성능지수는 각각 PI와 EPI의 형태로 구성되었다. 마찬가지로 이에 대한 목적함수와 적합도 함수값은 그림 7과 같다. 또한 GA를 통해 생성된 최종적인 최적 네트워크는 그림 8과 같다. 최종적인 네트워크의 출력값과 실공정값 사이의 오차를 그

림 9에서 나타내었다. 표 7에서 기존의 PNN 모델링 방법을 포함하여 이전 방법들과 제안된 방법의 성능지수를 나타내었다. 표에서 보듯이 제안된 모델은 시스템의 비선형성, 입력변수에 따른 복잡성등 데이터의 특징을 모델에 반영하고 유전자 알고리즘에 의한 최적의 모델을 자동선정 하여 정밀도가 높고 구조 또한 간단함을 알 수 있다.



(a) actual output versus model output



(b) error

그림 9 모델의 출력값과 실공정값 사이의 비교(θ=0.5)

Fig. 9 Comparison of the model output with actual output (θ=0.5)

표 7 기존 모델방법과의 동정 오차비교

Table 7 Comparison of identification error with previous modeling methods

Model	MSE	
	PI	EPI
Sugeno's model [19]	0.190	
Pedrycz's model [20]	0.320	
Leski's model [21]	0.047	
Kang's model [22]	0.161	
Kim's model [23]	0.034	0.244
Lin's model [24]	0.071	0.261
PNN-1 (5th layer) [9]	0.021	0.085
	0.017	0.101
Kim's model [10]	0.013	0.114
Our model	0.012	0.108

4.2 pH 중화 공정

기존의 PNN에서와 같이 제안된 모델을 pH중화 공정에 적용하였다. 본 논문에서 예제로 사용된 pH중화 공정은 [9]에서 사용한 것과 같은 공정이며, 아래에서 보인 공정에 대한 설명은 [9]에서 인용하였다.

pH중화 공정은 비선형성과 불확실성이 강하고, 중화점 근처에서 매우 민감하게 반응하여 제어가 매우 어려운 것으로 알려져 있다. pH는 물에 용해된 산성과 염기성의 농도를 나타낸 것이다. 회석을 위해 수치적으로 수소이온[H+]농도의 식으로 나타내며 (8)로 정의된다.

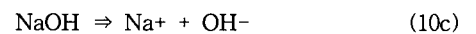
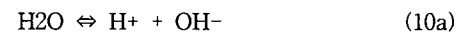
$$pH = -\log_{10}[H^+] \quad (8)$$

pH공정은 유입물질과 반응물질, 그리고 혼합용기의 선택에서 이들의 변동에 매우 민감한 경향이 있다. CSTR (continuously stirred tank reactor)에서 pH공정의 주요특성을 설명하는 잘 알려진 수학적 동적방정식이 있다. 이 모델은 구성물질의 균형과 화학적인 평형상태와 같은 원칙에 기본을 두고 발전되었다. 본 논문에서는 출력 pH에 관하여 다룬다. CSTR은 완전혼합이라는 가정하에 모델링은 수행된다. pH공정은 두 개의 유입량을 가진다고 가정한다. 전체체크기가 $v(l)$ 인 Tank안으로 F_a, F_b 가 흘러 들어간다. F_a 는 농도가 C_1 인 산의 흐름율이고, F_b 는 농도가 C_2 인 염기의 흐름율이다. CSTR의 동적 방정식은 다음과 같다.

$$\frac{Vdw_a}{dt} = F_a C_a - (F_a + F_b) W_a \quad (9a)$$

$$\frac{Vdw_b}{dt} = F_b C_b - (F_a + F_b) W_b \quad (9b)$$

(9)는 입력의 흐름율 F_a, F_b 에 의하여 시간에 따른 w_a, w_b 의 농도변화를 보여준다. 출력으로 pH를 얻기 위해, w_a, w_b 그리고 pH값 사이의 관계를 알아야 한다. 이 관계는 적정곡선 혹은 특성곡선으로 알려진 비선형 대수방정식에 의하여 묘사된다. 반응메카니즘은 다음과 같다.



전기적 중성조건에 따라 이온의 전하합은 용액 내에서 0이 되어야한다.

$$[Na^+] + [H^+] = [OH^-] + [AC^-] \quad (11)$$

[X]; X 이온의 농도

물과 아세트산의 평형관계는 다음과 같다.

$$K_a = \frac{[AC^-][H^+]}{[HAC]} \quad (12 a)$$

$$K_w = [H^+][OH^-] \quad (12 b)$$

$w_a = [HAC] + [AC^-], w_b = [Na^+]$ 을 K_a, K_w 에 대하여 대입하고 정리하면 다음과 같다.

$$[H^+]^3 + [H^+]^2(K_a + w_b) + [H^+](K_a(w_b - w_a) - K_w) - K_a K_w = 0 \quad (13)$$

여기서 (8)을 이용하면, (13)은 다음과 같이 (14)가 된다.

$$w_b + 10^{-pH} - 10^{pH - pK_a} - \frac{w_a}{1 + 10^{pK_a - pH}} = 0 \quad (14)$$

여기서, $pK_a = -\log_{10} K_a$

w_a, w_b 그리고 pH 사이의 관계는 (14)에 의하여 정의된다. 입력변수는 $F_b(t-3), F_b(t-2), F_b(t-1), Y_{pH}(t-3), Y_{pH}(t-2), Y_{pH}(t-1)$ 이고 출력변수는 $Y_{pH}(t)$ 이며 500개의 입력력 데이터 쌍이 사용되었다. 데이터 쌍 (F_b, Y_{pH})는 [9]에서 사용한 것을 이용하였으며 CSTR에서 pH공정의 주요특성을 설명하는 동적방정식을 이용하여 시스템의 입출력 데이터를 추출하는 과정은 [9]에서 상세히 다룬다.

앞 절에서 언급한 제안된 진화론적 설계방법을 다시 적용하여 시뮬레이션 결과를 살펴보면 다음과 같다. 성능지수가 최적으로 수렴해 가는 과정은 그림 10과 같다. 이때 유전연산자 및 설계파라미터는 표 5와 동일하다.

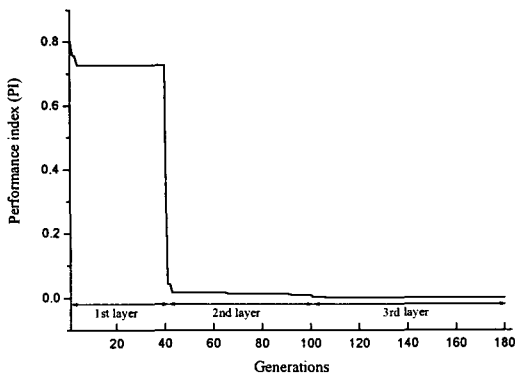


그림 10 유전자 알고리즘에 의한 성능지수의 수렴
Fig. 10 Convergence process of performance index by GA

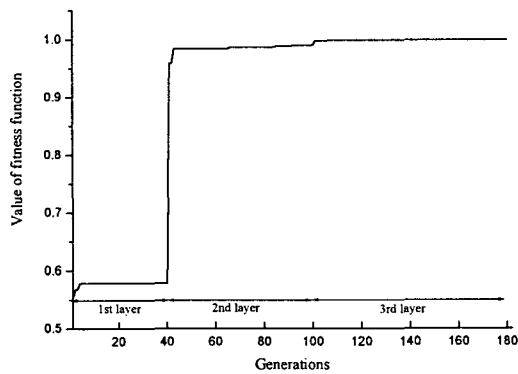


그림 11 유전자 알고리즘에 의한 적합도의 수렴
Fig. 11 Convergence process of fitness value by GA

성능지수에 대한 목적함수의 적합도 수렴과정은 그림 11에서 나타내었다. 그림 12에서 제안된 방법과 기존 PNN의 성능지수를 각 층에 따라 나타내었다. 기존 PNN의 Basic PNN-Case 1은 4입력 Type 2의 결과값으로 15층에서 0.00153이며, Basic PNN-Case 2는 1층에서 4입력 Type 1, 2층 이상에서 4입력 Type 2의 결과값으로 역시 15층에서

0.00520이며 Modified PNN-Case 1은 1층에서 2입력 Type 2, 2층 이상에서 4입력 Type 2의 결과값으로 0.00386을 10층에서 얻었으며 Modified PNN-Case 2는 1층에서 3입력 Type 1, 2층 이상에서 4입력 Type 2로 모델링을 하여 10층에서 0.01242의 결과값 [9]을 얻은 것과 비교한 것이다. GA를 통해 생성된 최종적인 최적 네트워크는 그림 13과 같다.

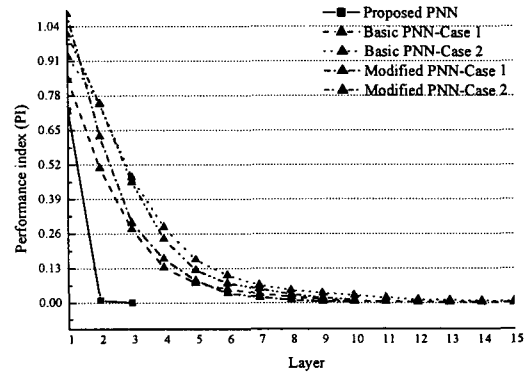


그림 12 제안된 방법과 기존 PNN과의 오차 비교
Fig. 12 Error comparison of the proposed method with conventional PNN

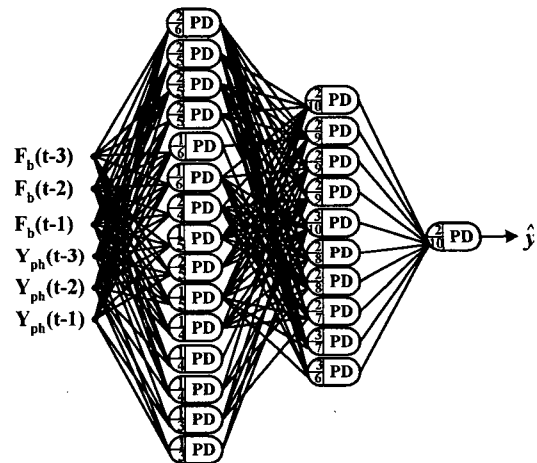


그림 13 유전자 알고리즘을 통해 생성된 최적의 구조
Fig. 13 Optimal structure generated via GA

표 8 기존 모델방법과의 동정 오차비교
Table 8 Comparison of identification error with previous modeling methods

Model		MSE PI
Nie's model [18]	USOCPN	0.230
	SSOCPN	0.012
Basic PNN [9]	Case 1 (15th layer)	0.0015
	Case 2 (15th layer)	0.0052
Modified PNN [9]	Case 1 (10th layer)	0.0039
	Case 2 (10th layer)	0.0124
Our model (3rd layer)		0.0008

기존의 PNN으로 최종적인 네트워크를 보이는 것은 많은 층수로 인해 불가능했다. 하지만 제안된 방법은 3층에서 성능지수 (PI)가 0.00088 이며 비교된 PNN의 결과값보다 상당히 우수한 성능을 보이며 유전자 알고리즘으로 구조를 동조하여 최적화시킴으로 더 적은 수의 층으로 모델을 추정할 수 있었다. 표 8에서 기존 PNN의 결과값과 함께 다양한 퍼지모델과의 결과비교를 나타내었다. Nie's model [18]중 USOCPN은 Counter-Propagation Network (CPN)에 기반을 둔 Unsupervised Self-Organizing CPN 알고리즘이며 31개의 rule을 사용하여 0.230을 얻었고, SSOCPN은 Supervised Self-Organizing CPN 알고리즘으로 34개의 rule을 사용하여 0.012를 얻었다.

5. 결 론

본 논문에서는 복잡한 비선형 요소가 많은 시스템의 체계적이고 효율적인 최적 모델을 얻기 위해 다항식 뉴럴네트워크 구조의 최적 설계와 관련하여 진화론적 설계방법을 제안하였다. 기존의 PNN은 입력변수의 선택과 다항식의 차수, 입출력 데이터의 분할, PD를 정의하여 시스템 방정식을 예측한다. 또한 변수와 차수로 정의된 PD를 계층적으로 조합하여 비선형 모델 추정식을 얻는다. 그러므로 선형 모델에 비해 우수한 시스템 묘사력을 가지고 있다. 하지만 PNN의 파라미터인 입력변수와 차수가 시스템의 성능과 직접적인 연관을 가지고 있음으로 객관적이고 신뢰성 높은 선택방법이 필요하다. 이들 파라미터의 선택문제를 해결하기 위해 진화(유전자) 알고리즘이 사용되었다. 그 결과 PNN의 각 노드에 인가되는 입력변수와 차수의 형태를 자동적으로 검색하여 최적의 해를 얻었다. 시뮬레이션에서 알 수 있듯이 본 논문에서 제안된 모델이 다른 지능형 모델과 기존의 PNN보다 우수한 성능뿐만 아니라 네트워크 사이즈도 작았다. 제안된 모델의 특징은 다음과 같다. 일반적인 다층 퍼셉트론 구조의 경우와 달리 고정되어 있지 않고 진화알고리즘에 의하여 각 층의 노드가 최적화된다. 또한 모델 스스로 노드의 선택과 제거를 통해 네트워크 구조를 생성한다. 근사화 능력과 일반화 능력사이의 적절한 타협점을 찾기 위해 하중값을 가진 목적함수를 정의하였고, 과도한 연산시간과 오퍼피팅을 피하기 위해 탐색체는 3개의 부염색체로 구성된다. 유전자 알고리즘에 의하여 각 노드의 PD의 정보는 최적으로 탐색된다.

감사의 글

GA 기반 GMDH 알고리즘에 대한 Prof. A. G. Ivakhnenko의 보이지 않는 성원에 감사 드립니다.

참 고 문 헌

[1] K. J. Astrom and P. Eykhoff, "System identification-a survey," *Automatica*, Vol. 7, pp. 123-162, 1971.
 [2] A. G. Ivakhnenko, "Polynomial theory of complex

systems", *IEEE Trans. Syst., Man, Cybern.*, Vol. SMC-1, No. 1, pp. 364-378, 1971.
 [3] A. G. Ivakhnenko and N. A. Ivakhnenko, "Long-term prediction by GMDH algorithms using the unbiased criterion and the balance-of-variables criterion," *Sov. Automat. Contr.*, Vol. 7, pp. 40-45, 1974.
 [4] —, "Long-term prediction by GMDH algorithms using the unbiased criterion and the balance-of-variables criterion, part 2," *Sov. Automat. Contr.*, Vol. 8, pp. 24-38, 1975.
 [5] A. G. Ivakhnenko, V. N. Vysotskiy, and N. A. Ivakhnenko, "Principal version of the minimum bias criterion for a model and an investigation of their noise immunity," *Sov. Automat. Contr.*, Vol. 11, pp. 27-45, 1978.
 [6] A. G. Ivakhnenko, G. I. Krotov, and N. A. Ivakhnenko, *Identification of the mathematical model of a complex system by the self-organization method*, in *Theoretical Systems Ecology: Advances and Case Studies*, E. Halfon, Ed. New York: Academic, 1970, ch. 13
 [7] S. J. Farlow, *Self-Organizing Methods in Modeling, GMDH Type-Algorithms*, New York: Marcel Dekker, 1984.
 [8] S. Barada, and H. Singh, "Generating Optimal Adaptive Fuzzy-Neural Models of Dynamical Systems with Applications to Control," *IEEE Trans. Syst., Man, Cybern., part C*, Vol. 28, No. 3, pp. 371-391, 1998.
 [9] 오성권, 김동원, 박병준, "다항식 뉴럴네트워크 구조의 최적 설계에 관한 연구", *Trans. KIEE*, Vol. 49D, No. 3, pp. 145-156, 2000.
 [10] 김동원, "자기구성 다항식 뉴럴네트워크의 진화론적 설계", Master's thesis, Dept. Control Instrum., Wonkwang Univ., 2002.
 [11] J.H. Holland, *Adaptation in Natural and Artificial Systems*, The Univesity of Michigan Press, Ann Arbor, M.I., 1975.
 [12] C.T. Lin, C.P. Jou, and C.J. Lin, "GA-based reinforcement learning for neural networks", *International Journal of System Science*, Vol. 29, No. 3, pp. 233-247, 1998.
 [13] A. Homaifar and E. McCormick, "Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms", *IEEE Trans. Fuzzy Syst.*, Vol. 3, pp. 129-139, 1995.
 [14] G.E.P. Box and F.M. Jenkins, *Time Series Analysis : Forecasting and Control*, 2nd ed. Holden-day, 1976.
 [15] F.G. Shinsky, *pH and pION Control in Process and Waste Streams*, Wiley, New York, 1973.
 [16] 장병탁, "유전 알고리즘이론 및 응용", 전자공학회지,

제 22권, 제 11호, 1995.

- [17] Sung-Kwun Oh, Dong-Won Kim, Byoung-Jun Park, and Hyung-Soo Hwang, "Advanced Polynomial Neural Networks Architecture with New Adaptive Nodes", *Trans. on Control, Automation and Systems Engineering*, Vol. 3, No. 1, pp. 43-50, 2001.
- [18] J. Nie, A.P. Loh, C.C. Hang, "Modeling pH neutralization processes using fuzzy-neural approaches", *Fuzzy Sets Syst*, Vol. 78, pp. 5-22, 1996.
- [19] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling", *IEEE Trans. Fuzzy Syst.*, Vol. 1, No. 1, pp. 7-31, 1993.
- [20] W. Pedrycz, "An identification algorithm in fuzzy relational system", *Fuzzy Sets Syst.*, Vol. 13, pp.153-167, 1984.
- [21] J. Leski, and E. Czogala, "A new artificial neural networks based fuzzy inference system with moving consequents in if-then rules and selected applications", *Fuzzy Sets Syst.*, Vol. 108, 289-297, 1999.
- [22] S.J. Kang, C.H. Woo, H.S. Hwang, and K.B. Woo, Evolutionary Design of Fuzzy Rule Base for Nonlinear System Modeling and Control, *IEEE Trans. Fuzzy Syst.*, Vol. 8, No. 1, Feb., 2000.
- [23] E. Kim, H. Lee, M. Park, M. Park, "A simple identified Sugeno-type fuzzy model via double clustering," *Inf. Sci.*, Vol. 110, pp. 25-39, 1998.
- [24] Y. Lin, G.A. Cunningham III, "A new approach to fuzzy-neural modeling", *IEEE Trans. Fuzzy Syst.*, Vol. 3, No. 2, pp. 190-197, 1995.
- [25] D. W. Kim, G. T. Park, "A Design of EA-based Self-Organizing Polynomial Neural Networks using Evolutionary Algorithm for Nonlinear System Modeling", *IEEE Trans. Syst. Man Cybern. Part B-Cybern.*(submitted)
- [26] S. Oh, and W. Pedrycz, "Identification of fuzzy systems by means of an auto-tuning algorithm and its application to nonlinear systems," *Fuzzy Sets Syst.*, Vol. 115, pp. 205-230, 2000.

저 자 소 개



김 동 원 (金東源)

1974년 03월 13일생. 2000년 원광대학교 제어계측공학과 졸업. 2002년 원광대학교 대학원 제어계측공학과 졸업(석사). 2002년~현재 고려대학교 대학원 전기공학과 박사과정. 관심분야: GMDH 타입 알고리즘, 소프트 컴퓨팅기반 뉴로-퍼지 시스템, 시스템 모델링 및 지능제어.

Tel : 02-929-5185, Fax : 02-929-5185

E-mail : dwkim@elec.korea.ac.kr



박 귀 태 (朴貴泰)

1947년 10월 25일생. 1975년 고려대학교 전기공학과 졸업. 1977년 동 대학원 전기공학과 졸업(석사). 1981년 동 대학원 전기공학과 졸업(공박). 1975~1978 한국 원자력 연구소 연구원. 1978~1981 광운대학교 전기공학과 조교수. 1981년~현재 고려대학교 전기공학과 교수. 2000~현재 (사) IBS KOREA 회장. 관심분야: 퍼지 및 신경망, 컴퓨터 비전, 지능제어, 인텔리전트 빌딩시스템 및 정보통신.

Tel : 02-929-5185, Fax : 02-929-5185

E-mail : gtpark@korea.ac.kr