

## HCM 및 최적 알고리즘을 이용한 퍼지-뉴럴네트워크구조의 설계

윤기찬, 박병준, 오성권  
원광대학교 전기전자공학부, 전라북도 익산시 신동면 344-2 ☎ 570-749

### Design of Fuzzy-Neural Networks Structure using HCM and Optimization Algorithm

Kichan Yoon, Byoungjun Park and Sungkwan Oh  
Dept. of Electrical and Electronic Engineering, Wonkwang Univ., Iksan, KOREA

**Abstract** - This paper presents an optimal identification method of nonlinear and complex system that is based on fuzzy-neural network(FNN). The FNN used simplified inference as fuzzy inference method and Error Back Propagation Algorithm as learning rule. And we use a HCM Algorithm to find initial parameters of membership function. And then to obtain optimal parameters, we use the genetic algorithm. Genetic algorithm is a random search algorithm which can find the global optimum without converging to local optimum. The parameters such as membership functions, learning rates and momentum coefficients are easily adjusted using the genetic algorithms. Also, the performance index with weighted value is introduced to achieve a meaningful balance between approximation and generalization abilities of the model. To evaluate the performance of the FNN, we use the time series data for gas furnace and the sewage treatment process.

### 1. 서 론

본 연구에서는 효율적이고 체계적인 비선형 시스템의 모델을 구축하기 위해서 퍼지-뉴럴네트워크(FNN) 모델의 최적 동정 알고리즘을 제안한다. 기존 방법은 주어진 데이터에 대하여 퍼지 공간을 균등하게 분할하여 멤버쉽 함수를 정의했다. 이것은 퍼지-뉴珥네트워크의 성능을 저하시켜 최적의 모델을 얻기가 어렵다. 복잡하고 비선형적인 시스템에 대해서는 퍼지 공간을 분할하는데 있어서 주어진 데이터의 특성을 부여한 공간을 설정해야 한다. 그러므로 퍼지-뉴珥네트워크의 멤버쉽 함수 파라미터의 결정은 초기에 들어오는 입력 출력 데이터를 사용하여 HCM 방법으로 구축한다. HCM 방법으로 구한 클러스터의 중심값을 멤버쉽 함수의 정점으로 사용한다. 모델의 성능향상을 위하여 신경망의 학습은 오류 역전파 알고리즘을 사용한다. 하지만 오류 역전파 알고리즘은 최급강화법이라는 수학적인 방법을 사용하므로 극소점이 아닌 지역극소의 빠질 위험이 있다. 그러므로 여기에 자연의 유전학과 진화이론에 바탕을 둔 확률적인 탐색 방법인 유전자 알고리즘을 이용한다. 이 개선된 방안은 가스로 공정, 하수처리공정 등의 비선형공정에서 얻어진 데이터를 이용하여 더 나은 모델을 구축할 수 있음을 보인다. 그리고 학습 및 테스트 데이터의 성능 결과의 상호 균형을 얻기 위한 하중값을 가진 성능지수가 제시된다.

### 2. 본 론

#### 2.1 퍼지-뉴런의 구조

Yamakawa[1,2]에 의해 제안된 퍼지-뉴珥네트워크

모델은 퍼지추론부에 보수적 멤버쉽함수를 가지는 간략 추론법이 사용되고, 신경망의 학습은 오류역전파 알고리즘(Error Back propagation Algorithm)을 합성하여 다른 모델들에 비해 학습속도가 빠르고 수렴능력이 우수하다.

퍼지-뉴珥네트워크의 입력은 각각 퍼지뉴런에 연결되어 있으며 계통의 입력변수들이 하나씩 할당된다. 각 퍼지-뉴런에 입력된 신호는 퍼지-뉴런의 특성을 나타내는 비선형함수  $f_i$ 에 의해 출력  $f_i(x_i)$ 를 생성한다. 일반적인 신경망과는 달리 합수  $f_i$ 내에 합성함수는 포함되지 않는다. 이렇게 각 퍼지-뉴런에서 출력된 값들은 식(1)과 같이 모두 대수적으로 합해져  $\hat{y}$ 를 최종적으로 출력하게 된다

$$\hat{y} = f_1(x_1) + f_2(x_2) + \dots + f_m(x_m) = \sum_{i=1}^m f_i(x_i) \quad (1)$$

입력 신호  $x_i$ 는 각 멤버쉽함수를 거쳐 적합도가 구해지고, 적합도는 연결하중과 곱해지며 이들이 대수적으로 합해져 퍼지뉴런의 출력  $f_i(x_i)$ 가 된다. 퍼지 추론은 아래 식(2)과 같이 표현된다.  $A_{ij}$ 는 전반부의 퍼지변수이고  $\mu_{ij}$ 는 멤버쉽함수로 표현된다. 후반부의  $w_{ij}$ 는 연결하중을 나타내고, 상수이다.  $x_i$ 에서  $f_i(x_i)$ 로의 사상(mapping)은 퍼지추론과 비퍼지화에 의해 이루어진다.

$$\begin{aligned} R^1 &: \text{If } X_i \text{ is } A_{1i}, \text{ then } y_{1i} = w_{1i} \\ R^j &: \text{If } X_i \text{ is } A_{ji}, \text{ then } y_{ji} = w_{ji} \\ R^n &: \text{If } X_i \text{ is } A_{ni}, \text{ then } y_{ni} = w_{ni} \end{aligned} \quad (2)$$

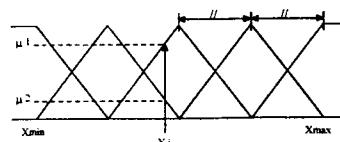


그림1. 기준방법을 이용한 멤버쉽함수 정의

그림1과 같이 전반부의 멤버쉽함수는 삼각형이고 이를 하는 멤버쉽함수와는 보수적(complementary)인 관계를 가진다. 입력변수  $x_i$ 는 동시에 두개의 멤버쉽함수에만 관계되고  $k$ ,  $k+1$ 로 나타나게 되는 두개의 이웃하는 멤버쉽함수에 의한 입력변수의 적합도의 합은 언제나 1이 된다

#### 2.2 HCM 알고리즘에 의한 멤버쉽함수 정의

HCM은 데이터들간의 거리를 기준으로 하여 근접한 정도를 측정하고, 이를 통해 데이터를 분류하는 방법이다.

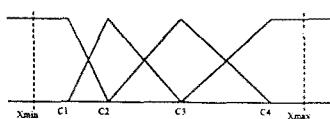


그림 2. HCM을 이용한 멤버쉽함수 정의

기존의 FNN모델의 구조는 주어진 데이터에 대하여 최소값과 최대값을 임의의 개수로 균등하게 분할하여 일률적으로 멤버쉽함수를 정의했다. 그러므로 주어진 데이터의 특성을 살리지 못한다는 점에서 모델의 성능에 좋은 영향을 끼치지 못했다. 따라서 이 특성에 맞은 멤버쉽함수를 정의하는 것이 보다 효율적이다. 본 논문에서는 HCM(Hard C-mean)방법을 이용하여 입력데이터의 특성을 분류하여 클러스터들의 중심을 그림2와 같이 각 멤버쉽함수의 정점으로 사용한다. 그림2에서 \$X\_{\min}\$은 데이터의 최소값, \$X\_{\max}\$는 데이터의 최대값이고 \$c\_n\$은 각 클러스터의 중심이다.

### 3. 유전자 알고리즘

최적화 문제에서 탁월한 성능을 발휘하는 유전자 알고리즘은 자연 세계의 진화 과정(유전자적 특성, 적자생존)을 이용한 탐색 알고리즘이다. 기존의 방법들과는 달리 미분에 대한 제약도 없고, 지역극소가 존재하더라도 문제없이 적용이 가능하므로 최적화 해결에 많이 응용되고 있다. 제안된 모델의 최적화를 위하여 사용된 유전자 알고리즘은 변수를 2진코딩하여 코딩된 문자열을 하나의 개체로 그리고 개개의 비트를 유전자처럼 취급하고, 개개의 점을 탐색하는 것이 아닌 동시에 여러지점을 탐색한다. 또한 미분과 같은 수학적 연산이 아닌 결과의 적합도를 목적함수로 수행하며, 결정적인 방법이 아닌 확률적인 방법을 사용한다.

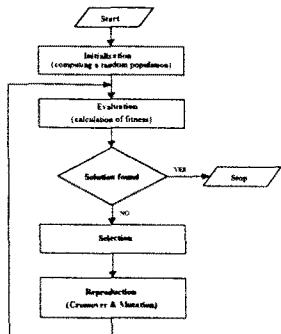


그림3. 유전자 알고리즘의 수행 과정

본 논문에서는 제안된 모델의 최적화를 위하여 HCM 알고리즘에 의해 얻어진 중심값, 신경망의 학습률 및 모멘텀계수를 유전자 알고리즘을 통하여 자동동조를 한다.

### 4. 하중값을 가진 목적함수

본 논문에서는 학습 및 테스트 데이터의 성능 결과의 대하여 위한 하중값을 가진 목적함수가 제시된다.

$$f = \theta \times PI + (1 - \theta) \times E\_PI \quad (3)$$

PI는 학습에 대한 성능지수, \$E\\_PI\$는 테스트에 대한 성능지수. \$\theta\$는 PI와 \$E\\_PI\$에 대한 하중값을 나타낸다. 이 목적함수는 데이터의 성능향상이 최대가 되도록 하기 위해 전반부 멤버쉽함수 즉 삼각형 소속함수의 모든 파라미터들은 변화한다. 하중값 설정에 따라 다음과 특징을

가진다.

1. \$\theta=1\$이면 모델은 학습을 바탕으로 최적화된다.
2. \$\theta=0\$이면 모델은 테스트를 바탕으로 최적화된다.
3. \$\theta=0.5\$이면 학습과 테스트는 모두 같은 비중을 가지고 평가한다.
4. \$\theta \in [0, 1]\$에 대해서 \$\theta = \alpha\$이면 학습과 테스트 모두를 포함하고, \$\alpha\$의 선택은 퍼지모델의 근사화 일반화 사이에서 최적화에 대한 방향을 설정한다.

### 5. 시뮬레이션

본 논문에서는 예측을 위한 모델을 위해서 전체데이터의 절반은 학습을 위하여, 나머지 절반은 모델의 예측성능을 평가하기 위하여 사용한다.

#### 5.1 가스로공정

가스로 공정은 많은 연구자들이 제안하는 모델링 기법을 성능 평가하는데 널리 사용되는 데이터이다. 가스로 공정 데이터는 1개의 입력과 1개의 출력으로 구성된 296쌍의 입력출력 데이터집합이다. 입력(u)은 GAS의 흐름을 나타내고 출력(y)은 이산화탄소(\$CO\_2\$)의 농도를 나타낸다. 본 논문에서는 제안된 모델의 입력으로는 \$u(t-3)\$과 \$y(t-1)\$을, 출력은 \$y(t)\$인 2입력 1출력 구조를 사용한다.

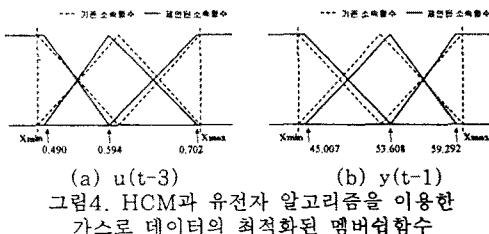


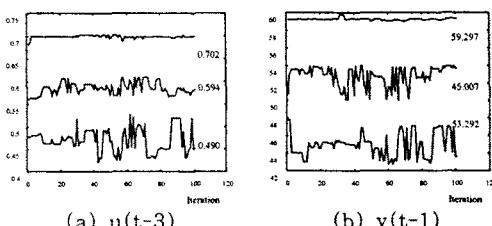
그림4. HCM과 유전자 알고리즘을 이용한 가스로 데이터의 최적화된 멤버쉽함수

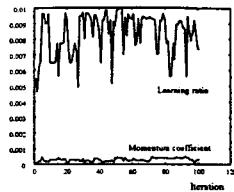
표1. 성능지수 비교

압력(1)	입력(2)	기존 모델(6)		Our Model	
		PI	E_PI	PI	E_PI
\$u(t-3)\$	\$y(t-1)\$	0.029184	0.296210	0.038492	0.283932

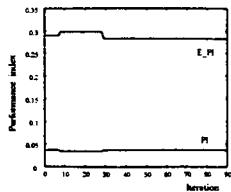
표2. 하중값을 가진 목적함수에 의한 성능지수

압력(1)	입력(2)	Weight (\$\theta\$)	Our Model	
			PI	E_PI
\$u(t-3)\$	\$y(t-1)\$	0	0.186193	0.252302
		0.2	0.064855	0.255015
		0.4	0.052280	0.262260
		0.6	0.038492	0.283932
		0.75	0.037247	0.281824
		1	0.029003	0.332282





(c) 학습률과 모멘텀 계수  
그림5. 유전자 알고리즘에 의한 최적화 과정 ( $\theta = 0.75$ )

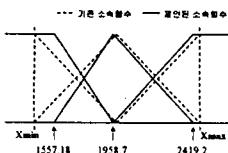


(d) PI와 E\_PI  
그림5. 유전자 알고리즘에 의한 최적화 과정 ( $\theta = 0.75$ )

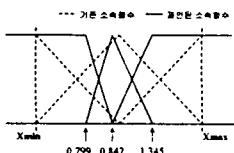
그림5는 가스로공정 데이터를 HCM 알고리즘에 의해 멤버쉽함수를 정의하고 유전자 알고리즘에 의해 자동 동조되는 과정을 나타낸 것이다.

## 5.2 하수처리 공정

하수처리를 위한 다양한 방법 중 가장 많이 사용되는 공정이 활성오너니를 이용한 방법으로, 침사지, 최초침전지, 폭기조 및 최종침전지로 구성된다. 하수처리 공정 데이터는 4개의 입력과 1개의 출력구조를 가진다. 본 논문에서는 제안된 방법을 통해 수도권 하수처리장 중의 하나를 모델로 선정하여 1년분 수질 데이터로부터 모델링하였다. 혼합액 부유물(MLSS), 잉여오나흐름을 (WSR), 반송율 설정치(RRSP), 용존산소 설정치(DOSP)을 입력으로 하고, 부유물의 농도(ESS)를 출력으로 하는 2입력 1출력 구조를 가진다.



(a) MLSS



(b) WSR

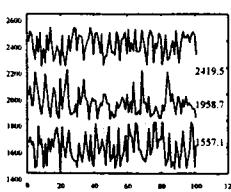
그림6. HCM과 유전자 알고리즘을 이용한 수처리 데이터의 최적화된 멤버쉽함수

표3. 성능지수 비교

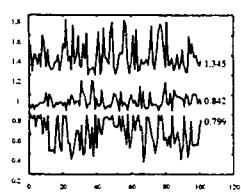
압력(1)	입력(2)	기존 모델(6)		Our Model	
		PI	E_PI	PI	E_PI
MLSS	WSR	11.345088	12.738689	9.955685	11.155324

표4. 하중값을 가진 목적함수에 의한 성능지수

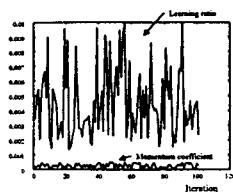
압력(1)	입력(2)	Weight ( $\theta$ )	Our Model	
			PI	E_PI
MLSS	WSR	0	17.234499	5.902997
		0.2	15.589808	6.434089
		0.4	13.021000	7.436243
		0.6	9.955685	11.155324
		0.75	8.958161	14.328180
		1	8.791540	15.582921



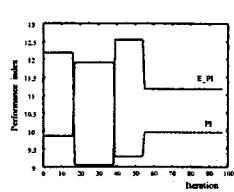
(a) MLSS



(b) WSR



(c) 학습률과 모멘텀 계수  
그림7. 유전자 알고리즘에 의한 최적화 과정 ( $\theta = 0.6$ )



(d) PI와 E\_PI  
그림7. 유전자 알고리즘에 의한 최적화 과정 ( $\theta = 0.6$ )

그림7은 수처리공정 데이터를 HCM 알고리즘에 의해 멤버쉽함수를 정의하고 유전자 알고리즘에 의해 자동 동조되는 과정을 나타낸 것이다.

## 5. 결 론

본 논문에서는 퍼지추론과 신경망이 결합된 퍼지-뉴럴네트워크를 사용하여, 비선형공정에 대한 모델링을 행하였다. 체계적이고 효율적인 모델링을 얻기 위하여 클러스터링 알고리즘과 유전자 알고리즘을 이용한 초기값 설정이 시도되었다. 비선형성이 상대적으로 약한 가스로에서는 주어진 데이터가 비교적 균일하게 분포되어 있으므로 기존 모델에 비해 학습성능이 약간 저하 되지만 예측 성능을 평가하기 위한 테스트 성능은 향상됨을 알 수 있다. 그리고 비선형성이 강한 수처리에서는 학습성능, 테스트 성능 모든 측면에서 우수한 결과를 얻을 수 있다. 그리고 하중값을 가진 목적함수에서 하중값의 설정은 근사화·일반화에 따라 학습과 테스트 사이에 상호 연계를 통한 최적화 향상을 위한 방향을 제시하였다.

## 참 고 문 헌

- [1] Takeshi Yamakawa, "A Neo Fuzzy Neuron and Its Applications to System Identification and Prediction of the System Behavior", *Proceedings of the 2nd International Conference on Fuzzy logic & Neural Networks*, pp.477-483, 1992.
- [2] Takeshi Yamakawa, "A New Effective Learning Algorithm for a Neo Fuzzy Neuron Model", *5th IFS World Conference*, pp.1017-1020, 1993.
- [3] David E. Goldberg, "Genetic Algorithms in search, Optimization & Machine Learning", Addison-wesley.
- [4] Zbigniew Michalewicz, "Genetic Algorithms + Data Structure = Evolution Programs", Springer-Verlag.
- [5] 오성권, 노석범, 남궁문, "퍼지-뉴럴 네트워크 구조에 의한 비선형 공정시스템의 지능형 모델링", 한국퍼지 및 지능시스템학회 논문집 제5권 제4호, pp.41-55, 1995.
- [6] 최재호, 오성권, 안태천, 황형수, "유전자 알고리즘을 사용한 퍼지-뉴럴네트워크 구조의 최적모델과 비선형공정시스템으로의 응용", 한국퍼지 및 지능 시스템학회 '96 추계학술대회 논문집 Vol.6, No.2, pp302-305, 1996.
- [7] 최재호, 박현성, 오성권, 안태천, "FNN성능 개선을 위한 클러스터링기법의 적용", 한국퍼지 및 지능 시스템학회 '97 추계학술대회 학술발표논문집 Vol. 7, No. 2, pp135-138, 1997.