

HCM 클러스터링 기반 FNN 구조 설계

박호성, 오성권

원광대학교 공과대학 전기전자및정보공학부, 전라북도 익산시 신용동 344-2 ☎670-749

Design of FNN architecture based on HCM Clustering Method

Ho-Sung Park and Sung-Kwun Oh

School of Electrical Electronic & Information Engineering, Wonkwang Univ., Iksan, KOREA

Abstract - In this paper, we propose the Multi-FNN(Fuzzy-Neural Networks) for optimal identification modeling of complex system. The proposed Multi-FNNs is based on a concept of FNNs and exploit linear inference being treated as generic inference mechanisms. In the networks learning, backpropagation(BP) algorithm of neural networks is used to update the parameters of the network. In order to control of nonlinear process with complexity and uncertainty of data, proposed model use a HCM(Hard C-Means)clustering algorithm which carry out the input-output data preprocessing function and Genetic Algorithm which carry out optimization of model. The HCM clustering method is utilized to determine the structure of Multi-FNNs. The parameters of Multi-FNN model such as apexes of membership function, learning rates, and momentum coefficients are adjusted using genetic algorithms. An aggregate performance index with a weighting factor is proposed in order to achieve a sound balance between approximation and generalization abilities of the model. NOx emission process data of gas turbine power plant is simulated in order to confirm the efficiency and feasibility of the proposed approach in this paper.

1. 서 론

퍼지 집합과 신경망 이론들은 비선형 공정 처리 영역에서 중요한 역할을 한다. 퍼지 논리와 신경망들은 비선형 공정 처리 시스템을 위해서 서로 다른 능력을 제공함으로써 다른 계산적인 모델들이다. 최근에 이러한 두 가지 기술이 하나의 통합된 구조로 집적되기 위해 많은 노력이 있어왔다. 이를 퍼지 뉴럴 네트워크라 부른다. 이는 비선형 공정에 대한 복잡성과 불확실성을 제어하기 위해서 퍼지 집합 이론[1]을 이용하여 비선형적이고 복잡한 시스템의 특성을 해석하는데 적용함으로써 시스템의 성능 및 기능면에 있어서 기존의 수학적 모델보다 좋은 결과를 얻을 수 있으며, 신경망(Neural networks)은 학습을 통하여 모델이 지능적으로 행동할 수 있도록 하였고 계산처리에 있어서 병렬 분산처리를 함으로써 계산 처리 능력을 증가시키는 장점을 가졌다.

이에 본 논문에서는 지능형 모델의 하나로 다른 퍼지-뉴럴 네트워크모델에 비해서 학습속도가 빠르고, 수렴특성이 매우 우수하며 후반부 추론은 기존의 간략 추론을 사용한 퍼지-뉴럴 네트워크(FNN) 모델[2,3]을 기반으로 하여 후반부 추론 방법을 선형 추론 방법으로 확장한 퍼지-뉴럴 네트워크 모델을 제안한다. 이는 보수적 멤버십 함수를 가지는 퍼지규칙과 오류 역전파 알고리즘을 사용한 신경망의 학습을 결합한 형태를 사용하고, 각 입

력변수에 의한 퍼지 입력공간분할에 의해 생성되는 네트워크 구조이다. 또한 데이터의 복잡성과 불확실성에 따른 비선형 공정을 제어하기 위해 HCM 클러스터링[4]을 이용하여 입출력 데이터의 전처리 기능을 수행함으로써 전체 FNN의 구조를 결정하고 모델의 멤버십함수의 정점, 학습율, 모멘텀 계수와 같은 각각의 FNN 모델 파라미터를 동조, 최적화를 위해 유전자 알고리즘[5,6]을 사용한다. 더욱이 모델의 근사화와 일반화 능력 사이에 합리적 균형을 얻기 위해 하중계수를 가진 성능지수[7]를 사용한다. 모델의 성능지수는 데이터의 개수, 비선형성의 정도(입출력 데이터 분포)등에 의해 상호 의존성을 가진 근사화 및 예측의 정확도에 따라 차이가 나므로 하중계수의 선택과 조절을 통하여 공정시스템의 최적 모델을 구축할 수 있다. 제안된 모델은 가스터빈 발전소의 NOx 방출 공정 데이터[8]를 사용하여 성능을 평가한다.

2. 본 론

2.1 퍼지-뉴럴 네트워크

본 논문에서 제안한 FNN의 기본 모델은 그림 1에서 보여준 바와 같이 다른 모델에 비해서 학습속도가 빠르고, 수렴특성이 매우 우수하며 후반부 추론 방법으로 선형 추론을 사용한다.

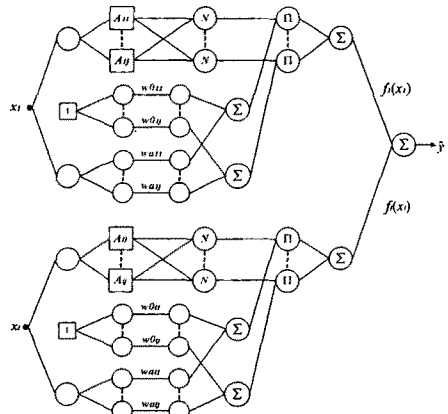


그림 1. 퍼지-뉴럴 네트워크의 구조

각 퍼지-뉴런에서 출력된 값들은 식 (1)과 같이 모두 대수적으로 합쳐져 \hat{y} 를 최종적으로 출력하게 된다.

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

여기서, m 은 퍼지-뉴런의 수이다.

각각의 퍼지-뉴런 (f_i)의 내부구조는 다음과 같이 퍼지 규칙을 이룬다.

$$R^i: IF \ x_i \text{ is } A_{ij} \text{ Then } \ y_{ij} = w_{0ij} + x_i w_{a_{ij}} \quad (2)$$

식(2)의 최종 추론 결과는 식 (3)과 같이 표현된다.

$$f_i(x_i) = \frac{\sum_{j=1}^m \mu_{ij}(x_i) \cdot (w_{0ij} + x_i w_{a_{ij}})}{\sum_{j=1}^m \mu_{ij}(x_i)} \quad (3)$$

식(3)은 식 (4)와 같이 다시 쓸 수 있다.

$$f_i(x_i) = \mu_{ik}(x_i) \cdot (w_{0ik} + x_i w_{a_{ik}}) + \mu_{i,k+1}(x_i) \cdot (w_{0i,k+1} + x_i w_{a_{i,k+1}}) \quad (4)$$

m 개의 입출력 데이터 쌍이 주어진 경우, 퍼지-뉴럴 네트워크의 최종 출력 오차는 식 (5)와 같이 된다.

$$E = \frac{1}{m} \sum_{j=1}^m (y_b - \hat{y}_b)^2 \quad (5)$$

여기서, y_b 는 주어진 출력 데이터, \hat{y}_b 는 모델의 출력 데이터를 나타낸다.

본 논문에서는 HCM 클러스터링을 이용하여 데이터 각각의 거리를 기준으로 서로 유사한 특성을 가진 그룹으로 데이터를 분류하여 특성에 맞는 각각의 퍼지-뉴럴 네트워크를 구축하였다.

2.2 HCM 클러스터링 방법

클러스터링 알고리즘이란 데이터의 분류를 위해서 사용되는 것으로 데이터의 내부가 비슷한 패턴, 속성, 형태 등의 기준을 통해 데이터를 분류하여 내부의 구조를 찾아내는 것이다. 본 논문에서는 클러스터링 알고리즘중 데이터들간의 거리를 기준으로 근접한 정도를 측정하고, 이를 바탕으로 데이터를 특성별로 분류하는 HCM 클러스터링을 이용하여 데이터들의 특성을 파악한다. 주어진 데이터 분류는 먼저 HCM 클러스터링에 의해 학습 데이터를 분류하고, 분류된 학습 데이터의 중심에 의해 테스트 데이터를 분류한다.

HCM 클러스터링에 의한 학습 데이터 분류는 다음과 같다.

[단계 1] 클러스터의 개수 ($2 \leq c \leq n$)를 결정하고, 소속행렬 U 를 $U^{(0)} \in M_c$ 으로 초기화한다.

$$M_c = \{U \mid u_{ij} \in \{0, 1\}, \sum_{i=1}^n u_{ik} = 1, 0 < \sum_{k=1}^c u_{ik} < n\} \quad (6)$$

여기서, u_{ij} 는 소속행렬의 파라미터.

[단계 2] 각각의 클러스터에 대한 중심벡터 V 를 구한다.

$$V_i = \{v_{i1}, v_{i2}, \dots, v_{ij}, \dots, v_{im}\} \quad (7)$$

$$v_{ij} = \frac{\sum_{k=1}^n u_{ik} \cdot x_{kj}}{\sum_{k=1}^n u_{ik}}$$

여기서, i 는 i 번째 클러스터, j 는 j 번째 입력 변수($j = 1, \dots, m$), n 은 data 수.

[단계 3] 각각의 클러스터 중심과 데이터와의 거리를 계산하여 새로운 소속행렬 $U^{(r)}$ 을 생성한다.

$$d_{ik} = d(x_k - V_i) = \|x_k - V_i\| = \left[\sum_{j=1}^m (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (8)$$

$$u_{ik}^{(r+1)} = \begin{cases} 1 & d_{ik}^{(r)} = \min\{d_{jk}^{(r)}\} \text{ for all } j \in c \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

여기서, d_{ik} 는 k 번째 데이터 sample x_k 와 i 번째 클러스터 중심 V_i 의 기하학적 거리

[단계 4] 만일 식 (10)을 만족한다면 종료하고, 그렇지 않으면 $r = r+1$ 로 놓고 [단계 2]로 간다.

$$\|U^{(r+1)} - U^{(r)}\| \leq \varepsilon \text{ (tolerance level)} \quad (10)$$

2.3 유전자 알고리즘

유전자 알고리즘은 자연 선택과 유전학에 기반으로 하는 확률적인 탐색방법으로써 탐색과 해의 가능영역들을

균형있게 이용하기 위하여 생산, 교차, 돌연변이의 과정을 수행하는 일반성 있는 탐색법으로 비선형 최적화 이론에 탁월한 성능을 발휘하고 있다[5,6]. 기존의 다른 탐색방법들과 달리 유전자 알고리즘은 해가 될 가능성이 있는 개체집단을 유지하면서 그들 모두가 동시에 최적점을 찾아나가기 때문에 지역 극소에 빠질 위험을 어느 정도 해결할 수 있다는 점과 더불어 모델의 성능지수가 최소가 되는 전역 극소 영역을 찾는 능력을 갖고 있으며, 또한 기존의 방법들과는 달리 선형, 연속, 미분가능 등의 제한이 없기 때문에 다양한 분야에 별다른 제한 없이 적용할 수 있다는 장점을 가진다는 것이 중요한 특징이다.

본 논문에서는 퍼지-뉴럴 네트워크 각각의 모델에 대한 멤버쉽함수의 정점과 학습을, 모멘텀계수를 최적으로 설정해 주기 위하여 유전자 알고리즘을 이용하여 모델의 최적화 과정을 수행하게 된다.

2.4 합성 목적함수

입출력 데이터로 시스템을 모델링 할 때 가장 문제가 되는 것은 구축된 모델이 학습 데이터만을 사용한 모델로써 시스템을 구축하는 경우이다. 이에 본 논문에서는 주어진 학습 데이터를 이용하여 모델을 구축하고 테스트 데이터를 사용하여 모델을 평가하기 위하여 근사화 및 일반화의 상호 균형과 의존 능력을 가진 합성 목적 함수(성능지수)를 이용하여 최적 모델 구조를 생성한다. 이를 위해 하중계수를 가진 목적함수인 합성 성능지수로써 식 (11)과 같이 정의한다.

$$f(PI, E_PI) = \theta \times PI + (1 - \theta) \times E_PI \quad (11)$$

θ 는 PI 와 E_PI 에 대한 하중계수를 나타내고, PI 는 학습 데이터 성능지수, E_PI 는 테스트 데이터 성능지수를 나타낸다. 식 (11)과 같이 정의한 목적함수는 하중계수의 설정에 따라 다음과 같은 특징을 가진다.

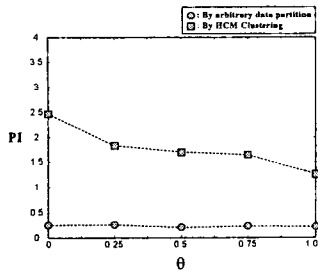
$\alpha \in [0, 1]$ 에 대해서 $\theta = \alpha$ 이면 학습과 테스트 모두를 포함하고, α 의 선택은 모델의 근사화와 일반화 사이에서 상호 균형에 의한 최적화를 시도한다. 따라서 합성 성능지수는 모델 자체의 근사화 능력과 일반화 능력을 상호 고려한 목적함수의 형태로 적절한 하중계수 선정에 따라 최적 퍼지-뉴럴 네트워크모델을 구축할 수 있다.

3. 시뮬레이션 및 결과고찰

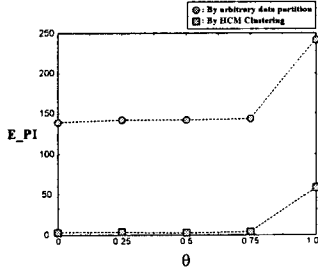
대기상 질소산화물(NO_x)은 주로 연소공정을 통해 배출되며 산업공정, 발전소, 자동차 배기가스 그리고 난방에 의해서 발생하는데, 그 중 대기에서 문제가 될 만큼 존재하는 것들은 NO_2 , NO 및 NO_2 이고 이들 물질을 대기오염 측면에서 NO_x (질소 산화물)로 일컫는다.

NO_x 의 저감방법으로는 발생원으로부터 NO_x 의 배출을 억제하는 방법과 배출가스의 NO_x 를 제거하는 방법들을 수 있는데, 최근에는 Gas Turbine에 사용하는 입출력 변수 중에 입력으로 T_{amb} (The ambient temperature at site, in degrees F), COM (The compressor speed, in rpm), LPT (The low pressure turbine speed, in rpm), Pcd (The compressor discharge pressure, in psia), $Texh$ (The turbine exhaust temperature, in degrees F)등의 주변 환경변수를 조절하며 출력으로는 Parts per million-volume dry 단위의 NO_x 양으로 배출 메카니즘을 모델링하고, 최적 모델을 구축하여 시뮬레이터를 구성함으로써 NO_x 를 효과적으로 관리하는 방법이 연구되고 있다[8].

그림 2는 HCM 클러스터링 방법의 유효함을 나타내기 위하여 가스터빈 발전소의 NO_x 배출 공정 전체 데이터를 임의로 분할한 것과 HCM 클러스터링 방법을 이용하여 분할한 데이터로 본 논문에서 제안한 방법으로 시뮬레이션한 결과이다.



(a) 학습 데이터



(b) 테스트 데이터

그림 2. FNN의 성능지수

그림 2에서 볼 수 있듯이 클러스터링 방법을 이용하여 주어진 데이터 특성에 맞게 데이터를 분할한 후 모델링한 결과가 그렇지 않고 임의로 데이터를 분할 한 후 모델링한 결과보다 아주 우수함을 알 수 있다.

표 1에서는 HCM 클러스터링을 통해 얻은 클러스터 개수와 하중계수 θ 에 따른 퍼지-뉴럴 네트워크 모델을 선형 추론방법을 각각 이용했을 때의 성능지수를 나타낸 것이다. 하중계수 θ 의 설정에 따라 FNN모델의 멤버십 함수의 파라미터들과 학습율, 모멘텀 계수를 유전자 알고리즘에 의해 자동 동조함으로써 근사화 능력과 일반화 능력을 가진 최적 FNN모델의 성능지수를 구한 것이다.

표 1. 클러스터 개수에 따른 성능지수

θ	FNN							
	클러스터 1		클러스터 2		클러스터 3		클러스터 4	
	PI	E_PI	PI	E_PI	PI	E_PI	PI	E_PI
0.0	4.46	5.49	2.45	3.01	1.34	2.12	1.09	2.21
0.25	3.91	5.77	1.83	3.01	1.16	2.04	1.43	2.15
0.5	4.03	6.02	1.69	2.84	1.14	2.65	1.07	2.25
0.75	3.62	5.56	1.64	3.72	1.05	3.07	0.72	2.02
1.0	3.79	6.81	1.25	58.0	0.84	3.18	0.79	2.91

표 1에서 볼 수 있듯이 데이터가 균일하지 않고 비선형성이 강한 데이터일수록 본 논문에서 제안한 퍼지-뉴럴 네트워크 모델이 효과적임을 알 수 있다.

4. 결 론

논문에서는 비선형적이고 복잡한 실 시스템의 데이터 특성을 해석하고 데이터의 특성에 따른 체계적이고 효율적인 모델을 구축하기 위하여 HCM 클러스터링 기반 FNN 구조를 설계하였다. 즉, 제안된 모델은 기존 모델들의 단점인 데이터의 특성을 고려하지 않고 모델을 구축하는 것에서 벗어나 HCM 클러스터링 알고리즘을 이용하여 데이터의 특성에 맞는 최적의 FNN 모델을 구축하였다. HCM 클러스터링을 이용하여 입출력 데이터의

분포 특성에 따른 FNN 모델 구축 후에 유전자 알고리즘을 이용하여 각각의 FNN의 멤버십 함수의 정점, 학습율, 모멘텀 계수를 최적을 동조하였다. 더불어 하중계수를 가진 합성 목적함수에 의해 모델의 근사화 및 일반화를 고려한 최적의 FNN 모델을 생성할 수 있음을 보였다.

시뮬레이션 및 결과 고찰에서 볼 수 있듯이 본 논문에서 제안한 알고리즘은 비선형성이 강하며 균일하지 않고 분산된 구조와 입출력 데이터 수가 적은 경우에는 학습 및 테스트 성능이 향상된 모델을 구축할 수 있었다.

감사의 글

본 논문은 2001년 에너지 관리공단의 에너지기술 학술 진흥사업 연구비에 의하여 지원되었음.

(참 고 문 헌)

- [1] L. A Zadeh, "Fuzzy sets", Inf. Control 8, 338-353, 1965.
- [2] 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.
- [3] T. Yamakawa, "A New Effective Learning Algorithm for a Neo Fuzzy Neuron Model", 5th IFSA World Congress, pp. 1017-1020, 1993.
- [4] BEZDEK, J.C. "Pattern Recognition with Fuzzy Objective Function Algorithms. New York", Plenum, 1981
- [5] David E. Goldberg, "Genetic Algorithm in search. Optimization & Machine Learning", Addison wesley, 1989.
- [6] Zbigniew Michalewicz, "Genetic Algorithms + Data Structures = Evolution Programs", Springer-Verlag, 1992.
- [7] S. K. Oh and W. Pedrycz, "Identification of Fuzzy Systems by means of an Auto-Tuning Algorithm and Its Application to Nonlinear Systems", Fuzzy sets and Syst., Vol. 115, No. 2, pp. 205-230, 2000.
- [8] G. Vachtsevanos, V. Ramani, T. W. Hwang, "Prediction of Gas Turbine NOx Emissions using Polynomial Neural Network", Technical Report, Georgia Institute of Technology, Atlanta, 1995.
- [9] 오성권, "C 프로그래밍에 의한 퍼지모델 및 제어시스템," 내출판사, 2002.