

데이터 입자 기반 유전론적 퍼지 관계 다항식 뉴럴네트워크의 최적 동정

Optimal Identification of Data Granules-based Genetically Optimized Fuzzy Relation Polynomial Neural Networks

이인태, 이영일, 오성권
수원대학교 전기공학과

In-Tae Lee, Young-Il Lee, Sung-Kwun Oh

Dept. of Electrical Engineering, The University of Suwon

E-mail : ohsk@suwon.ac.kr

요 약

본 논문에서는 정보 입자화와 유전자 알고리즘을 기반으로 최적 퍼지 다항식 뉴럴네트워크를 제안하고, 유전자 알고리즘을 사용하여 종합적인 설계방법을 개발한다. 제안된 모델은 기존의 진화론적 퍼지 다항식 뉴럴네트워크의 구조를 정보입자화를 통해 좀 더 빠르게 최적의 해공간에 접근시키는데 그 목적이 있다. 퍼지 관계기반 다항식 뉴럴네트워크는 퍼지 다항식 뉴런이 기초가 되어 가능한 구조적이고 요소적으로 모델의 성능을 향상 시켜준다. 퍼지 다항식 뉴런의 최적 구조를 위해 유전자 알고리즘을 이용하여 입력변수의 수와 후반부 다항식의 차수 입력변수 수에 따른 입력변수 그리고 멤버십 함수의 수를 동조한다. 여기서, 클러스터링의 하나의 방법인 HCM에 의해 퍼지 규칙 각각의 전반부와 후반부에 데이터 중심값을 이용하여 다항식함수의 파라미터값을 결정한다.

제안된 유전론적 퍼지 관계 다항식 뉴럴네트워크의 성능평가는 기존 퍼지 모델링에서 이용된 표준 데이터를 활용하여 평가한다.

1. 서론

모델링 분야에서 신경회로망, 퍼지 집합 그리고 진화론적 알고리즘은 매우 광범위하게 발전하고 있는 것과 동시에, 시스템 모델링에 대한 새로운 방법론적 이슈가 지속적으로 등장하고 이에 대한 우리의 인식 또한 증가하고 있다. 이에 본 논문에서는 복잡한 비선형 문제의 분석과 해결을 위한 효과적인 고급 컴퓨터 지능 모델의 개발을 목표로 한다.

본 논문에서 제안한 모델의 기본 모델이 되는 퍼지규칙 기반 자기구성 퍼지 다항식 뉴럴네트워크(Self-Organizing Fuzzy Polynomial Neural Networks ; SOFPNN)[1]는 FIS와 GMDH[2]의 확장(Extension), 결합(Combination), 및 융합(Fusion)에 의해 설계되며 이는 퍼지 집합 기반

다항식 뉴럴네트워크(Fuzzy Set-based Polynomial Neural Networks; FSPNN)와 퍼지 관계 기반 다항식 뉴럴네트워크(Fuzzy Relation-based Polynomial Neural Networks; FRPNN)로 나뉘어 진다. 이것은 퍼지집합과 퍼지 관계를 기반으로 한 FIS가 GMDH의 확장된 형태인 SOPNN(Self-Organizing Polynomial Neural Networks)[3]과의 융합으로 생성된 구조이다.

기본 모델로 사용된 FRPNN은 유전자 알고리즘(Genetic Algorithms;GAs)[4, 5]을 사용하여 구조를 최적화시킨다. 이는 기존의 FRPNN 모델들이 선호된 네트워크 구조를 찾기 위해서 노드의 입력변수의 수와 규칙 후반부 다항식 차수 입력변수 그리고 선택된 멤버십 함수의 수를 설계

자의 경험 또는 반복적인 학습을 통해서 얻을 수 있어서 최적의 네트워크 구조를 구축하는데 어려움이 내재되어 있었다. 따라서 객관적이고 신뢰성 있는 최적의 모델을 구축하기 위해서 유전자 알고리즘을 통해 모델을 구축하였다. 비록 GAs를 통해 구조를 최적화 하였더라도 데이터의 특성까지는 반영되지는 못한다. 따라서 주어진 데이터의 특성을 해석하기 위하여 HCM 방법을 결합한 Information Granulation[6, 7]기반 진화론적 최적 퍼지 관계 다항식 뉴럴네트워크 (Information Granulation-based gFRPNN; IG_gFRPNN)를 제안한다.

본 논문의 평가를 위해서 가스터빈 발전소의 NOx 배출 공정 데이터[8]를 이용하여 제안된 모델의 타당성, 정확성 및 예측 성능 등을 기존 지능 모델들과 비교를 통해 그 우수성을 보인다.

2. 데이터 입자

데이터 입자는 근접성, 유사성 또는 기능성 등의 기준에 의해 서로 결합된 물체(특히, 데이터 점)의 연결된 모임으로 간주되며, 기존보다 적은 계산적인 복잡성을 위해 잘 정의된 하위문제(모듈)들로 분할된다. 본 논문에서는 입출력 관계의 특성을 이용하여 모델링을 수행한다. 따라서 데이터의 특성을 파악하는 것이 매우 중요하다. 입출력 데이터의 특성을 찾아내기 위해 HCM 클러스터링 알고리즘을 이용한다.

본 논문에서는 데이터들간의 거리를 기준으로 근접한 정도를 측정하여 데이터를 특성별로 분류하는 HCM 클러스터링 알고리즘[4]을 이용한다. HCM 클러스터링에 의한 데이터 분류는 다음과 같다.

- [단계 1] 클러스터 개수 ($2 \leq c \leq n$)를 결정하고, 소속행렬 U를 초기화한다.
- [단계 2] 각각의 클러스터에 대한 중심벡터를 구한다.
- [단계 3] 각각의 클러스터 중심과 데이터와의 거리를 계산하여 새로운 소속행렬을 생성한다.
- [단계 4] 만일 종료 조건을 만족한다면 종료하고, 그렇지 않으면 [단계 2]로 간다.

3. Fuzzy Polynomial Neuron

FPN은 FRPNN의 가장 기본이 되는 처리 단위이다. FPN은 퍼지 집합과 뉴럴 네트워크의 개념에 기초해서 자세히 설명된다. 그림 1과 같이 2개의 기본적인 기능 모듈로 구성되어 있다.

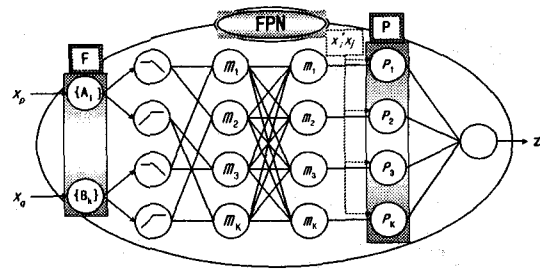


그림 1 FPN 모듈의 일반적인 구조

그림 1은 입력변수가 2개이고 멤버십 함수가 2x2인 경우를 보여주고 있다. "F" 로 표기되어진 것은 퍼지 집합의 집합체로써 입력 수의 변수들과 뉴론에 의해 실현된 처리 부분사이를 상호 연결해 준다. 그림에서 Xp와 Xq를 입력변수들로 표시한다. "P" 로 표기된 두 번째 모듈은 함수 기반 비선형(다항식) 처리에 관한 것이다.

4. IG-based gFRPNN 설계

데이터 입자화 기반 진화론적 퍼지 다항식 뉴럴 네트워크의 최적화 설계과정은 다음과 같다.

- [단계 1] 출력변수 y와 시스템 입력변수 결정.
- [단계 2] 입출력 데이터를 학습용과 평가용 데이터로 분할한다.
- [단계 3] gFRPNN의 퍼지 추론 방법 및 동정을 설계하는 부분과 모델의 구조를 결정하는 초기 정보를 결정한다.
- [단계 4] gFRPNN 모델을 구축하는데 기본 유닛인 FPN을 유전자 알고리즘에서 각각의 염색체에 의한 구조 선택을 통해 최적 모델을 구축
 - [단계 4-1] 입력변수의 수 선택
 - [단계 4-2] 퍼지규칙 후반부 다항식 차수 선택
 - [단계 4-3] 입력변수 선택
 - [단계 4-4] 멤버십 함수의 수 선택
- [단계 4-1]부터 [단계 4-4]까지 수행해서 얻은 선택된 입력변수와 퍼지규칙 후반부 다항식 차수를 이용하여 퍼지 추론을 적용한다.

표 1 퍼지규칙 후반부 구조의 서로 다른 형태

입력수 차수	1	2	3
0(Type 1)	Constant		
1(Type 2)	Linear	Bilinear	Trilinear
2(Type 3)	Quadratic	Biquadratic	Triquadratic
2(Type 4)		Modified Biquadratic	Modified Triquadratic

- [단계 5] 전반부 동정
HCM 클러스터링을 통한 데이터 입자에 의한

전반부 동정은 다음과 같다.

전체 데이터 집합 $U = \{x_1, x_2, \dots, x_l; y\}$ 이고, 여기서 $x_k = [x_{1k}, \dots, x_{mk}]^T$, $y = [y_1, \dots, y_m]^T$, l 은 변수의 수 그리고 m 은 데이터의 수로 가정한다. 데이터 집합 X_k 를 c 개(멤버십 함수의 수)의 클러스터(데이터 입자)로 분류한다.

$$v_{kg} = \{v_{k1}, v_{k2}, \dots, v_{kc}\} \quad (2)$$

여기서, $k=1, 2, \dots, l$, $g = 1, 2, \dots, c$. 중심 벡터 v_{kg} 를 멤버십함수의 정점으로 설정한다.

[단계 6] 후반부 동정

퍼지 모델의 후반부 동정도 전반부와 마찬가지로 구조 동정과 파라미터 동정으로 나뉘어진다. 데이터 입자에 따른 다항식 함수의 초기값을 가지고 구조를 동정한다.

j 번째 규칙의 퍼지 공간에 속한 데이터 집합을 찾는다. j 번째 규칙의 퍼지 공간에 속한 입력 데이터 쌍을 찾는다. 앞서 찾은 입력 데이터의 쌍으로부터 그에 대응하는 출력 데이터를 찾는다.

각 규칙에서 산술 평균에 의한 데이터 집합의 중심 벡터 V_j 를 계산한다.

$$V_j = \{V_{1j}, V_{2j}, \dots, V_{kj}; M_j\} \quad (1)$$

여기서, $k=1, 2, \dots, l$, $j=1, 2, \dots, n$. V_{kj} 와 M_j 는 각각 입력 데이터와 출력 데이터의 중심값이다.

예를 들어, 2차식 퍼지추론은 후반부가 2차식 함수의 다항식 형태를 가지면, 퍼지 집합 다항식 뉴런은 모델은 식 (2)와 같은 구형 규칙으로 구성된다.

$$R^j: \text{If } x_1 \text{ is } A_1 \text{ and } \dots \text{ and } x_k \text{ is } A_n \text{ then } y_j - M_j = f_j(x_1, \dots, x_k) \quad (2)$$

$$f_j(x_1, \dots, x_k) = a_{j0} + a_{j1}(x_1 - V_{1j}) + \dots + a_{jk}(x_k - V_{kj}) + a_{j(k+1)}(x_1 - V_{1j})^2 + \dots + a_{j(2k)}(x_k - V_{kj})^2 + a_{j(2k+1)}(x_1 - V_{1j})(x_2 - V_{2j}) + \dots + a_{j(k+2)(k+1)/2}(x_{k-1} - V_{(k-1)j})(x_k - V_{kj}) \quad (3)$$

퍼지 추론에 의해 추론된 값 y^* 은 식 (4)과 같다.

$$y^* = \frac{\sum_{j=1}^n w_{ji} y_j}{\sum_{j=1}^n w_{ji}} = \sum_{j=1}^n \hat{w}_{ji} (f_j(x_1, \dots, x_k) + M_j) \quad (4)$$

후반부 파라미터는 최소자승법에 의해 결정된다.

[단계 7] 모델 구축, 테스트 그리고 노드의 선택

Step 1) gFRPNN 모델 구축을 위한 유전자의 세대, Population 개수, 돌연변이를, 교배율, 그리고 염색체의 길이와 같은 유전자의 정보를 설정한다.

Step 2) 테스트 데이터를 이용해서 구해진 성능 지수를 유전자 알고리즘의 적합도 함수로 계산하기 위해서 다음 식 (5)와 같이 변환하여 유전자 알고리즘의 적합도 함수로 사용한다.

$$F(\text{Fitness Function}) = 1/(1+EPI) \quad (5)$$

Step 3) 다음 세대의 생성을 위해서 구해진 개체군 정보와 적합도 값들을 이용하여 선택연산, 교배, 돌연변이를 실행한다.

Step 4) 내림 차순으로 정렬된 적합도 값들 중에서 중복된 적합도 값들은 제거한다.

Step 5) 중복된 적합도 값들을 하나로 처리한 후 설계자가 노드의 제한된 개수 W 개만큼 선호되는 적합도 값들을 선택한다.

Step 6) Step 5)에서 선택되어진 노드들을 다시 초기 population 번호 순서에 따른 오름차순으로 재정렬한다.

Step 7) 다음세대에서의 엘리트 전략을 위해서 선택된 노드들 중 최적의 적합도 값을 가지는 노드를 선택한다.

Step 8) Step 3)에서 얻어진 정보를 가지고 다음 세대의 population을 생성한 후 Step 2)에서 Step 7)까지 반복한다. 여기서 생성된 적합도 값들 중에서 가장 낮은 적합도 값을 가진 노드를 이전 세대의 최적의 적합도 값을 가진 노드로 대체하는 엘리트 전략을 실행한다.

Step 10) 이전 세대에서 구해진 적합도 값들과 현재 세대에서 구해진 적합도 값들을 결합하여 다시 Step 4)부터 Step 7)까지 반복하고 정해진 세대까지 반복 실행한다.

[단계 8] [단계 7]의 현재 층에서 얻어진 최적 노드의 최대 적합도 값 F_1 이 다음 부등식을 만족하는 경우에 알고리즘을 종료한다.

$$F_l \leq F_* \quad (6)$$

단, F_* 는 이전 층의 최적 노드인 최대 적합도 값이다. 본 논문에서의 모델의 성능지수식으로 다음 식을 이용하였다.

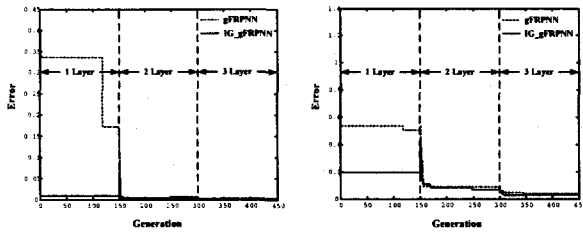
$$E = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (7)$$

[단계 9] 현재 층에서 보존된 노드의 출력들에서 다음 층의 새로운 입력들로, $x_{1j} = z_{1i}$, $x_{2j} = z_{2i}$, ..., $x_{wj} = z_{wi}$ 으로 구성하고, [단계 4]로 간다(여기서, $j=i+1$). 이 후, [단계 4]부터 [단계 9]까지를 반복한다. 알고리즘이 종료할 경우, 마지막 층의 최적 적합도 값 F_* 을 얻은 노드의 다항식에 입력이 되는 전 층의 출력을 대입하고 동일한 조작을 제 3 층까지 반복하여 최종 추정 모델 \hat{y} 를 얻는다.

5. 실험 데이터를 통한 결과 고찰

제한된 모델의 성능을 테스트하기 위해서 가스 터빈 NOx 공정 데이터[8]를 이용하였다. 가스 터빈 발전소의 NOx 배출 공정은 5입력 1출력 [Tamb, COM, LPT, Pcd, Texh; NOx]을 가진다.

출력은 NOx의 배출 양으로 배출 메카니즘을 모델링한다.



(a) Training data (b) Testing data
 그림 2 가우시안 멤버쉽 함수일때 성능지수에 따른 최적화 과정

그림 2는 최대 선택될수 입력변수의 수가 3이며 시스템 입력을 모두 후반부 다항식에 사용하였을 때의 그림이다. 여기에서 정보 입자화를 사용한 모델이 보다 빠르게 최적 해에 수렴함을 볼 수 있다.

표 2 기존모델과의 성능지수 비교

Model				PI	EP	
Multi-FNN [9]				0.7200	2.0250	
GA_PNN				0.0620	0.1070	
gFRPNN	3	Max	M type	T	PI	EP
			T	3	0.0336	0.1311
			G	4	0.0019	0.0837
			T	2(T*)	0.0060	0.0585
			G	3(T*)	0.0015	0.0387
Our model	3		T	3	0.0273	0.1113
			G	3	0.0037	0.0553
			T	4(T)	0.0013	0.0437
			G	4(T)	0.0017	0.0274

T : Triangular , G : Gaussian-like
 T*: 후반부 다항식에 시스템 입력을 사용

표 2은 NOx 배출공정을 사용한 모델들과의 성능지수 비교를 보여주고 있다. 표에서 보여주듯이 기존의 다른 모델들보다 데이터의 특성을 파악하여 모델을 설계하는 것이 최적화 모델을 설계하는데 우수함을 알 수 있다.

6. 결론

본 논문에서는 IG_based gFRPNN의 설계방법을 제안하였다. 기존의 gFRPNN은 구조면에서 설계자의 경험 또는 학습이 아닌 객관적인 모델을 설계하기 위해 유전자 알고리즘을 사용하여 노드내의 입력변수의 수와 그에 따른 입력변수, 퍼지규칙 후반부 다항식의 차수 그리고 멤버쉽 함수의 수를 선택하여 모델을 설계하였다. 하지만 gFRPNN은 데이터의 특성까지는 고려하지 않고 설계를 하기에 약간의 단점이 있었다. 이를

극복하고자 데이터 입자 기반 진화론적 최적 퍼지 관계 다항식 뉴럴네트워크를 제안하였다. 이 모델은 gFRPNN의 단점인 데이터의 특성까지도 고려한 모델로써 다변수 공정에서 보다 빠르게 최적 해를 찾을 뿐만 아니라 성능 또한 기존 gFRPNN보다 우수하였다. 결과적으로 다변수 공정에서 IG_based gFRPNN 모델이 우수한 예측 성능을 갖는다.

감사의 글

본 연구는 산업자원부의 지원에 의하여 기초전력연구원 (R-2004-B-274) 주관으로 수행된 과제임.

7. 참고문헌

- [1] S.-K. Oh and W. Pedrycz, "Fuzzy Polynomial Neuron-Based Self-Organizing Neural Networks", *Int. J. of General Systems*, Vol. 32, No. 3, pp. 237-250, May, 2003.
- [2] A. G. Ivakhnenko, "Polynomial theory of complex systems", *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-1, pp. 364-378, 1971.
- [3] S.-K. Oh and W. Pedrycz, "The design of self-organizing Polynomial Neural Networks", *Information Science*, Vol. 141, pp. 237-258, 2002.
- [4] Holland, J. H., *Adaptation In Natural and Artificial Systems*, The University of Michigan Press, Ann Arbour. 1975.
- [5] K. De Jong. Are genetic algorithms function optimizers? In Proc. of PPSN II (Parallel Problem Solving from Nature), pages 3-13, Amsterdam, North Holland, 1992.
- [6] L. A Zadeh, "Fuzzy logic = Computing with words," *IEEE Trans. Fuzzy Syst.*, Vol. 4, No. 2, pp. 103-111, 1996.
- [7] W. Pderycyca and G. Vukovich, "Granular neural networks," *Neurocomputing*, Vol. 36, pp. 205-224, 2001.
- [8] G. Vachtsevanos, V. Ramani, and T. W. Hwang, "Prediction of Gas Turbine NOx Emissions using Polynomial Neural Network", *Technical Report*, Georgia Institute of Technology, Atlanta, 1995.
- [9] S. K. Oh, W. Pedrycz, and H. S. Park, "Rule-based Multi-FNN Identification with the Aid of Evolutionary Fuzzy Granulation", *Journal of Knowledge-Based Systems*, 2003.