

PSO를 이용한 FCM 기반 RBF 뉴럴네트워크의 최적화

최정내\*, 오성권\*\*, 김현기\*\*

\*대림대학 전기과, \*\*수원대학교 전기공학과

Optimization of FCM-based Radial Basis Function Neural Network using PSO

Jeoung-Nae Choi\*, Sung-Kwun Oh\*\*, Hyun-Ki Kim\*\*

\* Department of Electrical Engineering, Daelim college, \*\* Department of Electrical Engineering, The University of Suwon

**Abstract** - 본 논문에서는 FCM 기반 RBF 뉴럴네트워크(FCM-RBFNN) 구조를 제안하고 PSO를 이용한 FCM-RBFNN의 구조 및 파라미터의 최적화 방법을 제시한다. 클러스터링 알고리즘은 퍼지 뉴럴 네트워크에서 멤버쉽함수의 중심점과 반경등을 결정하는 학습에 일반적으로 사용된다. 제안된 FCM-RBFNN에서는 방사기저함수로서 가우시안, 삼각형 타입 등의 정해진 형태를 사용하지 않고 데이터들 사이의 거리에 관계된 계산을 수행하는 FCM에 의해 결정된다. 기존의 RBFNN에서 후반부는 상수형태로써 방사기저함수의 선형결합으로써 표현되는 반면에 제안된 FCM-RBFNN의 후반부는 상수형, 선형, 2차식 등의 다양한 형태의 다항식으로 표현될 수 있으며 다항식의 계수는 WLSE를 이용하여 추정한다.

FCM 기반 RBF 뉴럴 네트워크의 성능은 퍼지규칙의 수, 후반부 다항식의 차수 FCM의 퍼지화 계수에 의하여 결정기 때문에 FCM-RBFNN의 구조와 파라미터의 최적화가 요구된다. 본 논문에서는 PSO를 이용하여 FCM-RBFNN의 구조에 관련된 퍼지규칙의 수, 후반부 다항식의 차수와 파라미터에 관련된 퍼지화 계수를 최적화 한다. 또한 후반부 다항식의 계수는 WLSE를 사용하여 추정한다.

1. 서 론

최근에 강한 비선형 특성을 가진 복잡한 시스템의 모델링에 대한 많은 연구가 수행되고 있다. 비선형적이고 다변수인 시스템을 대상으로 한 퍼지 모델링 기법의 유용성은 이미 잘 알려져 있으며 이들은 퍼지추론에 기초하여 퍼지추론시스템[1-4], 퍼지뉴럴네트워크[6-7], 그리고 RBF 뉴럴네트워크[8]등 다양한 구조로 연구되어 지고 있다.

Clustering 알고리즘은 데이터를 분석하기 위해 사용되는 기법으로써 많은 분야에서 사용되고 있다. 특히 Fuzzy C-Means clustering 알고리즘(FCM)은 각각의 데이터와 각각의 클러스터 중심점과의 거리에 대한 유사도 측정을 기반으로 목적함수를 최적화시키기 위한 알고리즘으로써 많은 분야에서 적용되고 있다[9].

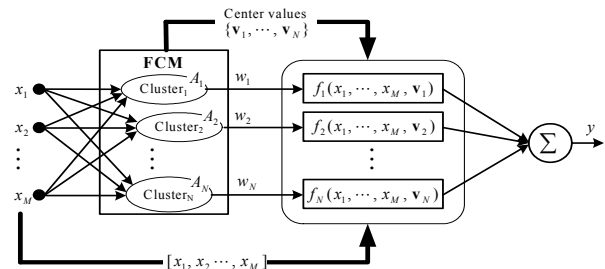
본 논문에서는 FCM 기반 RBF 뉴럴 네트워크 구조(FCM-RBFNN)와 학습 및 최적화 방법을 제안한다. FCM-RBFNN 구조는 Radial basis function neural network(RBFNN)의 변형 및 확장된 구조로써 전반부에 FCM을 사용하고 후반부는 로컬모델로써 다양한 형태의 다항식으로 구성될 수 있다. 퍼지공간 분할과 멤버쉽함수는 FCM에 의해 결정된다. 특히 기존 퍼지모델과 달리 멤버쉽함수는 가우시안이나 삼각형 등의 특정 함수를 사용하지 않고 데이터들의 거리에 기반을 둔다. 각 퍼지공간에 대한 로컬모델로써 상수형, 선형 또는 2차식 등의 다항식을 사용할 수 있으며 다항식의 계수는 Weighted Least square estimator(WLSE)를 사용하여 추정한다.

FCM-RBFNN의 학습은 FCM과 WLSE를 이용하여 학습하지만, 퍼지공간(또는 퍼지규칙, 클러스터)의 수, 후반부 다항식의 차수 및 퍼지화 계수 등을 최적화함으로써 성능을 극대화시킬 수 있다. 본 논문에서는 위와 같은 요소들을 PSO를 이용하여 최적화 한다.

2장에서는 FCM-RBFNN의 구조와 WLSE를 기반으로 한 학습방법에 대하여 설명한다. 3장에서는 PSO를 이용한 FCM-RBFNN의 최적화 방법을 설명하고, 4장과 5장에는 각각 실험과 결론을 설명한다.

2. FCM-RBFNN과 학습

제안된 FCM-RBFNN의 구조는 그림 1과 같다. 이 구조는 RBFNN의 변형 및 확장된 구조로써 가우시안이나 타원형등의 특정 RBF 커널을 사용하지 않고 FCM을 사용한다. FCM은 입력공간을 분할하고, 소속값을 계산한다.



<그림 1> FCM-RBFNN 구조

FCM-RBFNN은 전반부와 후반부로 구성이 되어 있다. FCM은 전반부에 해당한다. 클러스터의 수는 퍼지규칙의 수가 되며 각 데이터의 소속 정도는 각 규칙에 대한 적합도로써 사용된다. 후반부는 각 퍼지공간에서의 로컬모델을 의미하며 다항식 구조를 갖는다. FCM-RBFNN은 식(1)처럼 표현되며 후반부 다항식으로써 식(2)~(5)처럼 4가지 형태를 고려한다. RBFNN은 후반부가 식(2)처럼 상수값을 갖는 형태이다.

$$R^j: IF x_1, \dots, x_n \text{ is } A_j \text{ THEN } y_j = f_j(x_1, \dots, x_n) \quad (1)$$

Type 1 : 간략식(상수형)

$$f_j(x_1, x_2, \dots, x_n) = a_{j0} \quad (2)$$

Type 2 : 선형식

$$f_j(x_1, \dots, x_n) = a_{j0} + a_{j1}x_1 + \dots + a_{jn}x_n \quad (3)$$

Type 3 : 2차식

$$f_j(x_1, \dots, x_n) = a_{j0} + a_{j1}x_1 + \dots + a_{jn}x_n + a_{j(n+1)}x_1^2 + \dots + a_{j(2n)}x_n^2 + a_{j(2n+1)}x_1x_2 + \dots + a_{j(n+2)(n+1)/2}x_{n-1}x_n \quad (4)$$

Type 4 : 변형된 2차식

$$f_j(x_1, \dots, x_n) = a_{j0} + a_{j1}x_1 + \dots + a_{jn}x_n + a_{j(2n+1)}x_1x_2 + \dots + a_{j(n+2)(n+1)/2}x_{n-1}x_n \quad (5)$$

FCM-FNN의 학습은 크게 전반부 멤버쉽함수의 학습과 후반부 다항식의 학습부분으로 구성된다. 전반부 멤버쉽 함수는 삼각형이나 가우시안 형태 등의 특정 형태를 사용하지 않고 FCM 알고리즘을 통하여 멤버쉽 값이 결정된다.

식(1)에서  $A_j$ 는 j번째 클러스터(규칙)를 의미하며, 이때 소속정도는 FCM에서 소속정도를 결정해주는 식(6)으로 부터 구해진다.

$$\mu(A_j) = \omega_j = \frac{1}{\sum_{k=1}^c \left( \frac{\|x - v_j\|}{\|x - v_k\|} \right)^{\frac{2}{p-1}}} \quad (6)$$

여기에서  $j = 1, \dots, c$ ,  $c$ 는 규칙수이며,  $x = [x_1 \dots x_n]$ ,  $n$ 은 입력 변수의 수,  $v_j$ 는 j번째 클러스터의 중심벡터,  $p$ 는 퍼지화 계수이다.

FCM-FNN 모델은 출력은 식(7)처럼 구해진다.

$$\hat{y} = \sum_{j=1}^c \omega_j f_j(x_1, \dots, x_n) \quad (7)$$

후반부 다항식의 계수는 WLSE에 의하여 결정된다. WLSE는 로컬모델의 학습을 수행하기 위한 방법으로써 식(8)처럼 정의 되는 성능평가함수가 최소가 되도록 각각의 로컬모델인 다항식의 계수를 추정하는 방법으로써 각각의 로컬모델을 독립적으로 추

정할 수 있다는 장점이 있다. 이 방법은 컴퓨터의 계산부하를 줄일 수 있고, 각각의 로컬모델로써 서로 다른 차수의 다항식들을 사용할 수 있다는 장점이 있다.

$$J_L = \sum_{j=1}^R (Y - X_j a_j)^T W_j (Y - X_j a_j) \quad (8)$$

여기서,  $W_j$ 는  $j$ 입력공간에 대한 입력데이터의 활성레벨(소속값)을 의미하고,  $X_j$ 는 로컬모델의 계수를 추정하기 위한 입력데이터 행렬을 의미하며 로컬모델이 선형일 경우 다음처럼 정의된다.

$$W_j = \begin{bmatrix} w_{j1} & 0 & \cdots & 0 \\ 0 & w_{j2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{jm} \end{bmatrix}, \quad X_j = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1n} \\ 1 & x_{21} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m1} & \cdots & x_{mn} \end{bmatrix}$$

여기서,  $m$ 은 데이터의 수이다.

$j$ 번째 규칙에 대한 로컬모델인 다항식의 계수는 식(9)에 의해서 구해진다.

$$\hat{a}_j = (X_j^T W_j X_j)^{-1} X_j^T W_j Y \quad (9)$$

### 3. POS을 이용한 FCM-RBFNN의 최적화

PSO는 새, 벌 등의 군집 생활을 하는 생물체들의 이동 모습과 그 원리를 모방한 최적화 알고리즘이다. 집단은 여러 개의 Particle로 이루어져 있으며, 각각의 Particle의 위치는 풀고자 하는 문제에 대한 가능한 각각의 해가 된다. 각각의 Particle들은 매번 반복하여 자신의 위치, 즉 해를 변화시키게 되는데, 이 과정에서 모든 Particle 위치 중에 가장 좋았던, 즉 적합도가 높았던 위치와 각 Particle들의 위치변화 중 가장 적합도가 높았던 위치를 고려하여 위치를 변화시킨다[5].

FCM-RBFNN의 최적화 요소로서 퍼지화 계수, 퍼지규칙의 수 (또는 로컬 모델의 수), 로컬모델의 다항식의 차수를 최적화 하였다.

### 4. 실험 및 고찰

제안된 FCM-RBFNN 및 최적화 방법을 Box와 Jenkins가 사용한 가스로 시계열 데이터를 적용하였다. 입출력 데이터인 가스 흐름율과 연소된 이산화탄소 농도의 가스로 공정에 적용한다.  $u(t)$ 와  $y(t)$ 의 296쌍의 입출력 데이터로부터  $u(t-3)$ ,  $y(t-1)$ 의 2개의 입력변수와 한 개의 출력변수  $y(t)$ 를 설정하고 학습 데이터로 145개 테스트 데이터로 145개의 데이터 쌍을 사용한다.

PSO에서는 최적해 탐색영역으로써 퍼지화계수는  $1.01^5$ , 퍼지규칙의 수는 2~6개, 후반부다항식 차수는 상수항, 선형식, 2차식, 변형된 2차식 4가지 중에서 1가지를 찾도록 설정하였고, Particle의 수(Swarm의 크기)는 30개로 하여 150세대동안 탐색을 하였다. 모델의 평가를 위한 성능지수로서 MSE를 사용하였다.

시뮬레이션 결과 퍼지화 계수(p)는 1.3902, 퍼지규칙의 수는 5개, 후반부 다항식은 2차식일 때 최적값으로 찾아졌다. 이 경우에 학습데이터에 대한 성능지수는 0.0172이고, 테스트 데이터에 대한 성능지수는 0.2997이다.

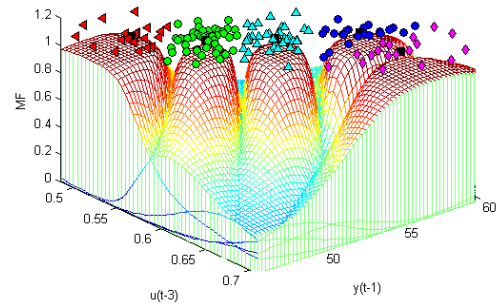
그림 2는 FCM에 의해 클러스터링된 입력데이터와, 각 분할된 공간에서 모델의 활성레벨인 적합도의 분포를 도시한 것이다. 그림 3은 학습데이터와 테스트 데이터에 대한 실제 데이터와 모델의 출력력을 비교를 보인다.

### 5. 결 론

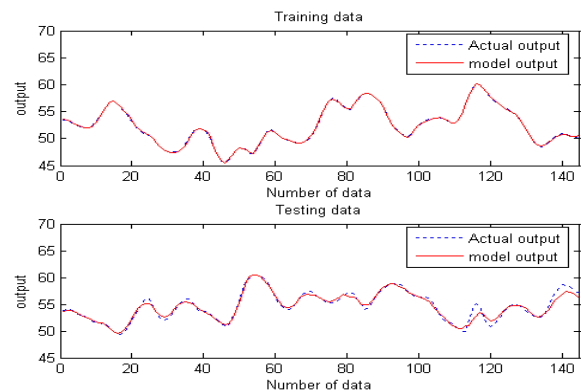
본 논문에서는 FCM-RBFNN 구조를 제안하고 WLSE를 이용한 로컬모델의 학습 및 PSO를 이용한 FCM-RBFNN의 최적화 방법을 제시한다. FCM-RBFNN은 RBFNN의 변형 및 확장된 구조로써 전반부에 FCM을 사용하고 후반부는 로컬모델로써 다양한 형태의 다항식으로 구성될 수 있다. FCM은 퍼지공간을 분할하며 멤버십함수를 결정하고 WLSE를 사용하여 로컬모델인 후반부 다항식의 계수를 추정한다. 또한 PSO를 이용하여 퍼지공간의 수, 후반부 다항식차수 그리고 멤버십함수의 형태를 결정하는 퍼지화 계수를 최적화 한다.

제안된 FCM-RBFNN은 FCM에 의하여 입력이 분할 되기 때문에 격자분할(Grid partition)에 기반하는 퍼지모델에 발생하는 다차원 문제를 해결하고 빠른 학습을 수행할 수 있다. 또한 로컬모델로써 퍼지규칙의 후반부에 고차 다항식을 사용하여 적은 규칙을 가지고 정교한 모델을 구축할 수 있기 때문에 모델의 해석력(interpretability)을 향상시킬 수 있다는 장점이 있다. WLSE를 이용한 로컬모델의 학습방법을 통하여 각각의 로컬모델의 계수를 독립적으로 추정함으로써 컴퓨터의 계산부하를 줄일 수 있을 뿐만 아니라 로컬모델이 서로 다른 형태의 다항식을 가짐으로써 유연성있는 네트워크 구조를

생성할 수 있다는 장점이 있다. 또한 PSO를 이용하여 모델의 구조 및 파라미터를 최적화 함으로써 FCM-RBFNN의 근사화 능력 및 예측 능력을 극대화 할 수 있다 제안된 방법은 시뮬레이션을 통하여 타당성과 성능을 입증하였다.



<그림 2> FCM에 의한 데이터 분할, 공간 분할 및 멤버십 값



<그림 3> 실제출력과 모델출력의 비교

감사의 글  
이 논문은 2007년도 정부(교육인적자원부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임 (F01-2007-000-10058-0)

### [참고 문헌]

- [1] T. Tagaki and M. sugeno, "Fuzzy identification of system and its applications to modeling and control", IEEE Trans. Syst. Cybern., Vol. SMC-15, No. 1, pp. 116-132, 1985
- [2] W. Pedryca and G. Vukovich, "Granular neural networks," Neurocomputing, Vol. 36, pp. 205-224, 2001.
- [3] P. R. Krishnaiah and L. N. Kanal, editors. Classification, pattern recognition, and reduction of dimensionality, Vol. 2 of Handbook of Statistics. North-Holland, Amsterdam, 1982.
- [4] 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.
- [5] J. Kennedy and R. Eberhart, "Particle Swarm Optimization", Proc. of IEEE International Conference on Neural Networks, Vol. IV, pp.1942-1948, Perth, Australia, 1995.
- [6] L. X. Wang, J. M. Mendel, "Generating fuzzy rules from numerical data with applications," IEEE Trans. Systems, Man, Cybern., Vol. 22, No. 6, pp. 1414-1427, 1992.
- [7] J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," IEEE Trans. System, Man, and Cybern., Vol. 23, No. 3, pp. 665-685, 1993.
- [8] A. Staiano, J. Tagliaferri, W. Pedrycz, "Improving RBF networks performance in regression tasks by means of a supervised fuzzy clusering" Automatic structure and parameter," Neurocomputing, Vol. 69, pp. 1570-1581, 2006
- [9] J.C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," Plenum, New York, 1981