공간 탐색 최적화 알고리즘을 이용한 K-Means 클러스터링 기반 다항식 방사형 기저 함수 신경회로망: 설계 및 비교 해석

K-Means-Based Polynomial-Radial Basis Function Neural Network Using Space Search Algorithm: Design and Comparative Studies

김 옥 동, 오 성 권*

(Wook-Dong Kim¹ and Sung-Kwun Oh¹)

¹The University of Suwon

Abstract: In this paper, we introduce an advanced architecture of K-Means clustering-based polynomial Radial Basis Function Neural Networks (p-RBFNNs) designed with the aid of SSOA (Space Search Optimization Algorithm) and develop a comprehensive design methodology supporting their construction. In order to design the optimized p-RBFNNs, a center value of each receptive field is determined by running the K-Means clustering algorithm and then the center value and the width of the corresponding receptive field are optimized through SSOA. The connections (weights) of the proposed p-RBFNNs are of functional character and are realized by considering three types of polynomials. In addition, a WLSE (Weighted Least Square Estimation) is used to estimate the coefficients of polynomials (serving as functional connections of the network) of each node from output node. Therefore, a local learning capability and an interpretability of the proposed model are improved. The proposed model is illustrated with the use of nonlinear function, NOx called Machine Learning dataset. A comparative analysis reveals that the proposed model exhibits higher accuracy and superb predictive capability in comparison to some previous models available in the literature.

Keywords: radial basis function neural networks, K-Means clustering algorithm, weighted least square estimation, space search optimization algorithm, polynomial function

1. 서론

최근 시스템 모델링의 문제는 최적화 기술과 함께 다양한 분야에서 폭넓게 연구되어지고 있다[1,2]. 이러한 연구 중에서 선형 또는 비선형 미분 방정식에 기반을 둔 정량적인 수학적 모델과 인공지능의 영역에서 나타나는 정성적 특성을 가진 지능형 모델은 시스템 모델링을 위한 대표적인 기술이다. 시스템 모델링과 이를 위한 동정의 근본적인 목적은 물리적 현상의 특성을 적절하게 반영할 수 있는 모델을 개발하는 것이다. 기존 모델링의 대부분은 미분 또는 차분 방정식의 언어에서 표현되어진 선형 또는 비선형 함수의 형태로 수학적 모델을 이끌어 냈다. 그러나 정교한 수학적 모델이 거의 완벽하게 시스템을 표현할 수 있다하더라도 그것은 시스템 동작과 변수 사이의 종속관계에 대한 원하는 정보를 제공하지 못한다. 또한 모델링을 하고자 하는 시스템이 복잡하고, 대규모 구조인 경우는 함수식으로 모델의 표현이 한정되지 않기 때문에 적용하기가 어렵다.

이러한 문제와 어려움을 해결하기 위해 퍼지 모델[3], 신경 회로망[4]과 같은 지능형 모델에 관한 연구가 진행되었으며, 최근에는 이러한 모델들의 최적화를 통해 모델이 가지고 있는 독특한 특성을 살리면서 지능 모델의 효율과 성능 향상을 최대한 이끌어 내는 연구가 활발히 진행되어지고 있다[5,6].

방사형 기저 함수 신경회로망(radial basis function neural networks; RBFNNs)은 기존의 다중(multilayer) 신경 회로망보다 구조가 간단하고 학습능력이 빠른 특징을 갖고 있으며, 퍼지 모델에 비해 비선형적이고 다변수인 시스템을 대상으로 한 모델링에서 또한 우수한 성능을 보인다. 따라서 function approximation, regularization, noisy interpolation, 패턴인식과 같은 다양한 응용분야에서 연구되고 있다[7].

현재 컴퓨터 지능분야에서는 두 개 이상의 모델을 합성한 퍼지 신경 회로망[8], 퍼지 다항식 신경 회로망[9]과 같은 다양한 형태의 모델이 제안되고 있으며, 주어진 입력과 출력데이터를 분석하고 유사성을 판단하기 위해 클러스터링 기술을 결합한 모델이 연구되고 있다. 클러스터링 방법으로는 FCM (Fuzzy C-Means) [10], K-Means [11], Mountain [12], Subtractive [13] 등이 대표적으로 퍼지 모델 및 신경회로망에 적용되고 있다. 또한 모델의 전체적인 구조 및 모델을 구성하는데 필요한 파라미터들을 최적화시키기 위해 많은 연구가 이루어지고 있으며 유전자 알고리즘(GA: Genetic

* 책임저자(Corresponding Author)

논문접수: 2011. 5. 20., 수정: 2011. 6. 5., 채택확정: 2011. 6. 20.
김옥동, 오성권: 수원대학교 전기공학과

(wdkim@suwon.ac.kr/ohsk@suwon.ac.kr)

※ 이 논문은 2010년도 정부(과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2010-D00065). 그리고 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음 [GRRC 수원2011-B2, U-city 보안감시 기술협력센터].

Algorithms)과 입자군집 최적화 알고리즘(PSO: Particle Swarm Optimization) 그리고 두 알고리즘의 개선된 방법들이 모델을 최적화하는데 주로 사용되고 있다. 또한 산업 및 공학[21,22] 그리고 의학 분야에서 지능형 알고리즘을 이용하여 추가전망[14], 오일수요[15], 질병[16], 응용지수[17]에 측 등 다양한 적용방법이 연구되고 있다.

본 논문에서는 새로운 공간탐색 최적화알고리즘(SSOA: Space Search Optimization Algorithm)로 최적화된 K-Means 클러스터링 기반 방사형 기저 함수 신경회로망을 제안한다. 제안된 모델은 기존의 방사형 기저 함수 신경회로망의 구조로 이루어져있다. 그러나 기존의 모델과의 차이점은 다음 표이 크게 세 가지로 나눌 수 있다.

첫째, 은닉층 활성화함수의 중심점 및 분포상수를 정의하는데 있어 클러스터링 알고리즘을 사용하였다. K-Means 클러스터링 알고리즘을 사용하여 입력 데이터를 은닉층의 노드 수만큼 클러스터링하여 각 클러스터의 중심점을 활성화함수의 중심점으로 사용하였다. 또한 활성화함수의 분포상수는 주어진 입력 데이터와 클러스터 중심점간의 표준편차를 구해 적용하였다.

둘째, 연결가중치를 다항식으로 구성하였다. 기존의 방사형 기저 함수 신경회로망은 단일 상수항을 사용하여 출력을 표현하였지만, 제안된 모델의 연결가중치는 입력이 고려된 1차식 이상의 다항식으로 연결가중치를 구성하였다. 이는 모델의 최종 출력을 구할 때 입력을 고려하여 좀 더 정확하고 안정적인 성능을 나타낸다.

마지막으로 공간탐색 최적화 알고리즘을 사용하여 기존의 학습방법과 더불어 모델의 파라미터를 동조하였다. 제안된 모델의 연결가중치는 가중최소자승법(WLSE: Weighted Least Square Estimation) [18]을 사용하며, 활성화함수의 중심점 및 분포상수는 공간탐색 최적화 알고리즘(SSOA)을 사용하여 최적의 파라미터를 동조한다. 일반적인 모델의 경우 연결가중치학습을 위해 최소자승법(LSE: Least Square Estimation)이나 오류역전파(BP: Back Propagation)를 사용한다. 최소자승법은 오직 연결가중치만을 학습하며 활성화함수의 중심점 및 분포상수를 학습할 수 없고, 오류역전파는 경사하강법을 주로 사용하여 연결가중치뿐만 아니라 활성화함수의 중심점 및 분포상수를 학습한다. 하지만 지역 극소에 빠질 가능성이 많으며 학습률을 비롯한 초기 파라미터 설정값에 따른 영향으로 최적의 파라미터를 찾아내는데 많은 연산시간과 시행착오를 겪어야하는 문제점이 있다.

본 논문은 다음과 같은 구성으로 이루어져있다. 먼저 II 장에서는 제안된 모델인 K-Means 클러스터링 알고리즘이 결합된 방사형 기저 함수 신경회로망에 대해서 알아보고, III 장에서는 제안된 모델의 학습 방법인 공간탐색 알고리즘에 대해서 설명하였다. IV 장에서는 제안된 모델의 성능을 평가하기 위해 비선형 함수에서 데이터를 추출한 비선형 데이터와 가스 터빈에서 발생하는 NOx 배출량을 기록한 NOx 데이터[19]를 이용하여 구조적 개발의 다양성을 입증하고 모델의 성능을 확인하였다. 마지막으로 V 장에서는 본 논문에서 제안한 모델에 대한 결론을 서술하였다.

II. K-Means 클러스터링 기반 다항식 방사형

기저 함수 신경 회로망

본 장에서는 제안된 K-Means 클러스터링[11] 기반 다항식 방사형 기저 함수 신경 회로망의 구조에 대해 알아보고, 활성화함수의 초기 중심점을 결정하는 K-Means 클러스터링 방법에 대해 설명하였다. 또한 상수항이외의 1차 선형 및 2차식으로 표현된 다항식 형태의 연결가중치를 소개하였으며, 연결가중치의 학습 알고리즘인 가중최소자승법에 대해 기술하였다.

1. K-Means 클러스터링 기반 다항식 방사형 기저 함수 신경 회로망 구조

제안된 K-Means 클러스터링 기반 방사형 기저 함수 신경회로망은 기존의 방사형 기저 함수 신경회로망의 기본틀을 두고 있다. 은닉층 활성화함수에는 가우시안 함수를 사용하였으며, 가우시안 함수의 중심점은 입력 데이터를 K-Means 알고리즘으로 클러스터링하여 결정된 클러스터의 중심점을 사용한다. 또한 가우시안 함수의 분포상수는 입력과 중심점간의 표준편차를 사용하였다. 그림 1은 본 논문에서 공간탐색 알고리즘을 이용한 K-Means 클러스터링 기반 다항식 방사형 기저 함수 신경회로망의 구조를 나타낸 것이다.

기존 방사형 기저 함수 신경회로망의 은닉층 활성화함수(receptive field)는 가우시안 함수를 사용하였으며, 식 (1)과 같이 표현 할 수 있다.

$$R_i(x) = e^{-\frac{1}{2} \left(\frac{(x_1 - v_{i1})^2 + (x_2 - v_{i2})^2 + \dots + (x_j - v_{ij})^2}{\sigma_i^2} \right)} \quad (1)$$

제안된 모델은 식 (1)을 개선하여 사용하였으며 식 (2)와 같이 표현할 수 있다.

$$R_i(x) = e^{-\frac{1}{2} \left(\frac{(x_1 - v_{i1})^2}{\sigma_{i1}^2} + \frac{(x_2 - v_{i2})^2}{\sigma_{i2}^2} + \dots + \frac{(x_j - v_{ij})^2}{\sigma_{ij}^2} \right)} \quad (2)$$

여기서 $i(i=1, \dots, c)$ 은 은닉층 노드수를 나타내며, $j(j=1, \dots, m)$ 은 입력변수의 수를 나타낸다. 식 (2)의 특징은 각 분포

그림 1. K-Means 클러스터링 기반 다항식 방사형 기저 함수 신경회로망의 구조.

Fig. 1. Architecture of the K-Means clustering-based polynomial-Radial Basis Function Neural Networks.

상수가 입력변수와 중심점의 표준편차를 사용한다. 이는 기존 활성화함수의 경우 각 노드의 활성화함수는 입력변수의 크기와 상관없이 동일한 분포상수를 갖고 있다. 그로인해 각 입력변수간의 범위가 큰 차이를 보이면 입력변수가 큰 쪽의 정보만으로 모델이 학습되기 때문에 상대적으로 다른 입력변수들의 특징이 배제되어 정확한 은닉층 출력(적합도)을 구할 수 없다. 흔히 이런 문제점을 해결하기 위해 모든 입력변수를 일정범위로 정규화 시키는 방법을 사용한다. 하지만 제안된 모델의 경우는 분포상수가 입력변수마다 독립적으로 존재함으로써 입력변수가 중심점과 분포상수의 의해 정규화가 되어 각 입력변수의 특징을 고려한 은닉층의 출력을 구할 수 있기 때문에 별도의 정규화 과정이 필요하지 않다.

모델의 연결가중치는 단일 상수항 이외에 다항식 형태의 연결가중치를 사용하였으며, 크게 4개의 형태 즉 상수, 선형식, 이차식, 수정된 이차식으로 나눌 수 있다.

적합도와 연결가중치로 표현되는 모델의 최종출력 식은 식 (3)과 같이 표현한다. 여기서, $i(i=1, \dots, c)$ 는 은닉층 노드 수를 나타낸다.

$$\hat{y} = \frac{\sum_{i=1}^c R_i(\mathbf{x}) f_i(\mathbf{x})}{\sum_{k=1}^c R_k(\mathbf{x})} = \sum_{i=1}^c \hat{R}_i(\mathbf{x}) f_i(\mathbf{x}) \quad (3)$$

다항식 연결가중치의 파라미터 계수를 동정하기 위해 가중 최소 자승법(Weighted Least Squares Estimator: WLSE)을 사용하여 계수를 추정하였다. 가중최소자승법은 오차 제곱의 합이 최소가 되도록 계수를 추정하는 표준최소자승법(LSE: Standard Least Squares Estimation)과는 달리 오차 제곱에 가중치를 곱해서 계수를 추정한다. 기존 컴퓨터 지능 분야에서 모델링의 계수 추정에 사용되어진 표준최소자승법은 후반부 다항식들의 계수를 한 번에 구하기 때문에 전역 모델의 학습을 수행하게 되고, 은닉층의 각 노드와 출력층 노드 사이의 지역 모델에 대한 해석력(interpretability)이 사라지는 단점을 가지고 있다. 이러한 단점을 보완하고자, 가중최소자승법을 적용하여 다항식 연결가중치의 계수를 각각의 독립적으로 구하여 전역 학습을 수행한다.

가중최소자승법의 성능 평가함수를 행렬식으로 표현하면 식 (4)과 같다.

$$Q = \sum_{i=1}^c (\mathbf{Y} - \mathbf{X}_i \mathbf{a}_i)^T \mathbf{W}_i (\mathbf{Y} - \mathbf{X}_i \mathbf{a}_i) \quad (4)$$

여기서 $\mathbf{a}_i$ 는 추정하고자 하는 j 번째 다항식의 계수, $\mathbf{Y}$ 는 출력데이터, $\mathbf{W}_i$ 는 i 번째 입력공간에 대한 입력 데이터들의 적합도 값을 의미한다. $\mathbf{X}_i$ 는 i 번째 지역 모델의 계수를 추정하기 위한 입력 데이터 행렬을 의미하며 지역 모델이 선형식일 경우 다음처럼 정의된다.

$$\mathbf{X}_j = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ 1 & x_{21} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \cdots & x_{Nm} \end{bmatrix} \quad \mathbf{W}_i = \begin{bmatrix} \hat{R}_{1i}(\mathbf{x}) & 0 & \cdots & 0 \\ 0 & \hat{R}_{2i}(\mathbf{x}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{R}_{Ni}(\mathbf{x}) \end{bmatrix} \quad (5)$$

여기서 N 은 전체 데이터의 수를 나타내며 i 번째 규칙에 대한 지역 모델인 다항식 연결가중치의 계수는 식 (6)에 의해 구해진다.

$$\mathbf{a}_i = (\mathbf{X}_i^T \mathbf{W}_i \mathbf{X}_i)^{-1} \mathbf{X}_i^T \mathbf{W}_i \mathbf{Y} \quad (6)$$

제안된 모델의 성능 평가를 위한 성능지수 식은 식 (7)의 RMSE (Root Mean Square Error)과 식 (8)의 MSE(Mean Square Error)를 사용하였다.

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

$$E = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (8)$$

여기서 N 은 전체 데이터 개수, y_i 는 원 데이터 출력 그리고 $\hat{y}_i$ 는 모델의 최종출력을 의미한다.

2. K-Means 클러스터링 알고리즘

K-Means 클러스터링 알고리즘[11]은 데이터를 분류하는 대표적인 알고리즘이다. 클러스터의 수가 정해지면 각 클러스터마다 임의의 중심점을 생성하여 데이터간의 유클리드 거리를 계산한다. 데이터와 중심점간의 거리를 바탕으로 유사성을 판단하고, 유사한 데이터를 그룹화하는 알고리즘이다. 제안된 모델에서는 은닉층 노드의 수 만큼의 클러스터가 생성되고 학습을 통해 각 클러스터의 중심점이 결정되면 그 중심점을 그대로 활성화함수의 중심점으로 사용한다. 따라서 클러스터의 수와 은닉층의 노드 수는 항상 같게 되며, 데이터의 유사성을 바탕으로 활성화함수를 생성하여 입력 데이터의 특성을 최대한 고려하였다. K-Means 클러스터링 알고리즘 과정을 다음과 같다.

[단계 1] 클러스터 개수를 결정하고, 소속행렬 $U^{(0)}$ 을 0과 1로만 이루어진 값으로 초기화한다.

$$M = \left\{ U \mid u_{ik} \in 0 \text{ or } 1, \sum_{i=1}^c u_{ik} = 1, 0 < \sum_{k=1}^N u_{ik} < N \right\} \quad (9)$$

여기서 u_{ik} 은 소속행렬의 파라미터를 나타내며, c 는 클러스터의 개수를 나타내고, N 은 주어진 데이터 수를 나타낸다.

[단계 2] 소속행렬과 입력 데이터를 가지고 클러스터 개수만큼의 중심점 $\mathbf{v}$ 를 계산한다.

$$v_{ij} = \frac{\sum_{k=1}^N u_{ik} \times x_{kj}}{\sum_{k=1}^N u_{ik}} \quad (10)$$

여기서 $i(i=1, \dots, c)$ 는 클러스터의 수를 나타내며 $j(j=1, \dots, m)$ 는 입력변수의 수를 나타내고 $k(k=1, \dots, N)$ 는 데이터의 수를 나타낸다.

[단계 3] 각 클러스터의 중심점과 데이터와의 거리를 계산하여 거리가 가까운 클러스터의 소속행렬 파라미터 u_{ik} 을 1로 바꿔주면서 새로운 소속행렬 $U^{(r+1)}$ 을 생성한다.

$$d_{ij} = d(\mathbf{x} - \mathbf{v}) = \|\mathbf{x} - \mathbf{v}\| = \sqrt{\sum_{j=1}^m (x_j - v_{ij})^2} \quad (11)$$

$$u_{ik}^{(r+1)} = \begin{cases} 1 & d_{ij}^{(r)} = \min d_{ij}^{(r)} \text{ for all } i \in c \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

여기서 r 은 반복횟수를 나타낸다.

[단계 4] 만일 식 (13)을 만족하면 종료하고, 그렇지 않으면 $r=r+1$ 로 놓고 [단계 2]로 이동한다.

$$\|U^{(r+1)} - U^{(r)}\| \leq \varepsilon \quad (13)$$

또한 활성함수를 정의하기 위해 필요한 분포상수는 입력 데이터와 중심점간의 표준편차를 사용하며 식 (14)과 같다.

$$\sigma_{ij} = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (x_j - v_{ij})^2} \quad (14)$$

모델에 적용된 최적화 알고리즘에 대한 자세한 설명은 3장에 서술하였다.

III. 공간탐색 최적화 알고리즘을 이용한 모델 최적화 방법

본 장에서는 공간탐색 최적화 알고리즘에 대해 설명하고 모델의 파라미터를 최적화시키기 위한 적용 방법에 대해 자세히 기술한다.

1. 공간탐색 최적화 알고리즘

공간 탐색 최적화 알고리즘(SSOA: Space Search Optimization Algorithm)은 탐색 공간을 조절하여 최적의 해를 찾아가는 알고리즘이다. 연산자는 교배연산과 돌연변이 연산으로 이루어져 있으며 공간 탐색 알고리즘의 학습 과정은 다음과 같다.

[단계 1] 염색체 $X_P(t) = \{X_1^P, X_2^P, \dots, X_D^P\}$ 을 주어진 탐색 공간에서 랜덤하게 생성한다. 여기서, P 는 염색체의 인구가 되며 D 는 파라미터의 차원 수가 된다.

[단계 2] 생성된 염색체를 가지고 목적함수의 적합도 $F(t)$ 를 평가한다.

[단계 3] 적합도 중에서 가장 좋은 적합도 $F_{\text{best}}(t)$ 와 가장 나쁜 적합도 $F_{\text{worst}}(t)$ 를 판별한다.

[단계 4] 전체 인구수(population)에서 랜덤하게 M 개의 염색체를 선택한다.

[단계 5] 선택된 염색체를 교배시켜 새로운 염색체 $X_P(t+1)$ 을 생성한다.

$$X_P(t+1) = \{X_1^P, X_2^P, \dots, X_D^P\} \quad (15)$$

$$X_q^P = \sum_{k=1}^M \gamma_k X_q^k, \text{ [where } \sum_{k=1}^M \gamma_k = 1 \text{ } (-0.5 \leq \gamma_k \leq 1.5)]$$

여기서 $q(q=1, \dots, D)$ 이며, $k(k=1, \dots, M)$ 를 나타낸다. γ 는 교배가중치로써 교배연산을 통해 새로운 염색체가 생성되는 탐색 공간의 크기를 결정한다. γ 의 합은 항상 1이 되어야 한다.

[단계 6] 교배 연산된 염색체에서 랜덤하게 개체를 선택하여 돌연변이 연산을 통해 새로운 개체를 생성한다.

$$X_q^P = \text{rand} \times (\text{space}_q^{\text{max}} - \text{space}_q^{\text{min}}) + \text{space}_q^{\text{min}} \quad (16)$$

여기서 space 는 초기에 주어진 탐색 공간이 된다.

[단계 7] 새로 생성된 염색체 $X_P(t+1)$ 을 가지고 적합

도를 평가한다.

[단계 8] 적합도 중에서 가장 좋은 적합도 $F_{\text{best}}(t+1)$

을 찾는다.

[단계 9] 새로운 염색체의 적합도 $F_{\text{best}}(t+1)$ 을 $F_{\text{worst}}(t)$ 와 $F_{\text{best}}(t)$ 을 서로 비교한다.

$$\text{If } (F_{\text{worst}}(t) > F_{\text{best}}(t+1)) \\ X_{\text{worst}}(t) = X_{\text{best}}(t+1) \quad (17)$$

$$\text{If } (F_{\text{best}}(t) > F_{\text{best}}(t+1)) \\ X_{\text{best}}(t) = X_{\text{best}}(t+1) \quad (18)$$

[단계 10] 만약 염색체가 교체되었으면, 다시 적합도의 $F_{\text{worst}}(t)$ 와 $F_{\text{best}}(t)$ 을 판별한다.

[단계 11] 종료조건을 만족하면 프로그램을 종료하며, 그렇지 않으면 [step 4] 이동하여 반복 수행한다.

2. 공간탐색 최적화 알고리즘을 이용한 모델 최적화방법

공간탐색 최적화 알고리즘을 이용하여 은닉층 활성함수의 중심점 및 분포상수를 동조한다. 제안된 모델의 중심점 및 분포상수는 K-Means 클러스터링 알고리즘에 의해 초기값이 결정된다. 이 초기값을 바탕으로 탐색공간을 설정하여 더욱 성능이 개선될 수 있는 최적의 파라미터를 찾고자 한다.

그림 2는 최종 제안된 모델을 나타낸다. 초기 활성함수의 중심점 및 분포상수는 K-Means 클러스터링 알고리즘으로 설정된 후 그 값을 바탕으로 최적화 알고리즘을 수행한다. 결과적으로 최적화 알고리즘을 통해 최적의 중심점 및 분포상수를 찾고 이는 더 우수한 모델의 성능으로 나타난다. 공간탐색 알고리즘의 염색체 구조는 그림 5와 같이 묘사할 수 있다. 염색체의 구조는 중심점과 분포상수의 파라미터 이루어지며, 은닉층의 수와 입력변수의 수에 따라 파라미터의 차원수가 결정된다.

그림 2. 공간탐색 최적화 알고리즘을 이용한 K-Means 클러스터링 기반 방사형 기저 함수 신경회로망.

Fig. 2. K-Means clustering-based radial basis function neural network using space search optimization algorithm.

그림 3. 공간탐색 최적화 알고리즘의 염색체 구조.

Fig. 3. Structure of chromosome of space search optimization algorithm.

위 그림은 그림 3의 입력데이터와 은닉층의 노드 수가 3개 일 때의 염색체 구조를 나타내었다. 총 12개의 파라미터로 구성되어 있으며, 초기 정해진 탐색공간에서 교배연산과 돌연변이연산을 통해 파라미터가 생성되어 모델에 적용된다. 공간탐색 최적화 알고리즘의 목적함수는 식 (7) 또는 식 (8)과 같이 RMSE 또는 MSE를 사용하며, 목적함수가 최소가 될 때의 파라미터가 최적의 파라미터가 되며 모델의 최종 값이 된다.

IV. 시뮬레이션 및 결과 고찰

이번 장에서는 컴퓨터 지능분야에서 주로 쓰이는 대표적인 실험 데이터를 바탕으로 모델의 성능을 평가하였다. 실험 데이터는 총 3개로 첫 번째 데이터는 2입력-1출력의 synthetic 데이터를 사용하였으며, 두 번째 데이터는 가스터빈 발전기의 NOx 배출량을 기록한 데이터를 사용하였다. 모든 실험의 결과들은 좀 더 높은 객관성을 제공하기 위해 5-fold resampling cross validation 모드를 실행하였다. 또한 SSOA로 최적화된 모델과 GA로 최적화된 모델의 수렴속도 및 성능을 비교하였다.

표 1은 모델의 최적화를 위해 사용된 SSOA 알고리즘과 GA 알고리즘의 초기설정 파라미터들을 나타낸 것이다.

1. 비선형 함수 데이터

모델을 평가하기 위하여 비선형 데이터를 이용하였으며, 그림 4와 같은 비선형 함수에서 임의로 500개의 데이터를 추출하였다. 이 데이터에서 학습 데이터 300(60%)개와 테스트 데이터 200(40%)개가 된다.

표 2에 GA로 중심점 및 분포상수를 최적화시킨 결과와 함께 나타냈으며, 성능지수는 식 (8)에서 설명된 MSE (Mean Square Error)로 나타내었다.

표 2에서 최적화가 적용되지 않은 모델은 K-Means 클러스터링만 사용했을 때의 성능지수를 나타내었다. 표에서 보

표 1. 최적화 알고리즘의 초기 설정 파라미터.

Table 1. Initial setup parameters of optimization algorithm.

SSOA		GA	
Parameters	Values	Parameters	Values
Generation size	100	Generation size	100
Population size	50	Population size	50
No. of chromosome for crossover	8	Selection	Roulette wheel
weight of crossover	[-0.5 1.5]	Crossover	One point
Mutation	Uniform	Crossover rate	1
Mutation rate	0.01	Mutation	Uniform
-	-	Mutation rate	0.05
-	-	Elitism	Used

듯이 최적화된 모델의 성능이 우수함을 알 수 있다. 또한 연결가중치로 단일 상수항(constant)보다 1차 선형 및 2차식을 사용했을 때의 성능이 더 우수하게 나타난다. 최적화된 모델 중에서는 GA로 최적화된 모델보다 SSOA로 최적화된 모델의 성능이 전체적으로 우수하다. 가장 우수한 성능지수는 은닉층의 노드 수가 4개 일 때로 최적화를 적용하지 않은 모델의 근사화 지수(학습데이터의 성능지수; PI)는 0.174 ± 0.003 이며, 일반화 지수(테스트 데이터의 성능지수; EPI)는 0.178 ± 0.007 로 나타난다. 그리고 최적화된 모델에서 SSOA의 경우 근사화 지수는 0.061 ± 0.001 , 일반화 지수는 0.068 ± 0.006 이며, GA로 최적화된 모델의 근사화 지수는 0.084 ± 0.007 , 일반화 지수는 0.092 ± 0.010 이다.

(a) A nonlinear function. (b) Data set for experiment.

그림 4. 비선형 함수와 관련된 데이터.

Fig. 4. Nonlinear function and associated data set.

표 2. 비선형 데이터에 SSOA를 적용한 모델과 GA를 적용한 모델의 성능지수 비교.

Table 2. Comparison of performance index of SSOA-based model and GA-based model for nonlinear function.

N	T	Without optimization		With optimization			
		PI	EPI	SSOA-based model	GA-based model	PI	EPI
2	C	0.293 ± 0.004	0.287 ± 0.012	0.226 ± 0.007	0.223 ± 0.153	0.262 ± 0.007	0.259 ± 0.010
	L	0.243 ± 0.006	0.237 ± 0.011	0.185 ± 0.005	0.188 ± 0.004	0.215 ± 0.006	0.213 ± 0.007
	Q	0.192 ± 0.004	0.197 ± 0.011	0.111 ± 0.007	0.120 ± 0.011	0.141 ± 0.006	0.152 ± 0.007
	M	0.241 ± 0.006	0.236 ± 0.011	0.151 ± 0.017	0.157 ± 0.015	0.206 ± 0.005	0.205 ± 0.007
4	C	0.283 ± 0.005	0.275 ± 0.010	0.204 ± 0.008	0.202 ± 0.005	0.222 ± 0.006	0.218 ± 0.008
	L	0.236 ± 0.005	0.231 ± 0.011	0.146 ± 0.003	0.149 ± 0.011	0.171 ± 0.005	0.172 ± 0.007
	Q	0.174 ± 0.003	0.178 ± 0.007	0.061 ± 0.001	0.068 ± 0.006	0.084 ± 0.007	0.092 ± 0.010
	M	0.233 ± 0.006	0.229 ± 0.012	0.135 ± 0.003	0.138 ± 0.010	0.166 ± 0.008	0.169 ± 0.014
6	C	0.284 ± 0.005	0.277 ± 0.010	0.210 ± 0.006	0.207 ± 0.009	0.232 ± 0.004	0.227 ± 0.010
	L	0.235 ± 0.006	0.230 ± 0.011	0.137 ± 0.009	0.142 ± 0.010	0.164 ± 0.010	0.166 ± 0.010
	Q	0.176 ± 0.003	0.181 ± 0.007	0.066 ± 0.004	0.073 ± 0.010	0.095 ± 0.009	0.102 ± 0.011
	M	0.233 ± 0.006	0.228 ± 0.012	0.129 ± 0.006	0.133 ± 0.009	0.164 ± 0.008	0.167 ± 0.014

N: No. of nodes T: Polynomial Type

C: Constant L: Linear Q: Quadratic M: Modified Quadratic

그림 5. SSOA로 최적화된 모델의 다항식 형태별 모델 출력.

Fig. 5. Shape of the model output according to polynomial type of SSOA-based model.

그림 5는 은닉층 노드의 수가 4개 일 때 SSOA로 최적화된 모델을 다항식 형태 별로 모델의 최종출력을 나타내었다. 그림 5(a)와 가장 유사한 모델은 다항식 형태가 quadratic일 때이며, 그다음으로는 modified quadratic, linear 순으로 나타난다.

그림 6. SSOA와 GA의 학습 데이터 수렴과정.

Fig. 6. Convergence process for the training dataset in each case of SSOA and GA.

그림 6은 각 은닉층 노드별 다항식 형태가 quadratic일 때의 SSOA와 GA의 목적함수가 최적화되는 과정을 수렴곡선으로 나타내었다. 앞서 말했다시피, 최적화 알고리즘의 목적함수는 MSE를 사용하였으며, 학습데이터의 MSE가 낮아지도록 최적화가 수행된다. 그림에서 보듯이 SSOA가 더 빠르게 수렴하는 것을 확인할 수 있다.

2. NOx 데이터

NOx 데이터는 가스터빈에서 발생하는 NOx의 배출량을 기록한 데이터로 총 5입력-1출력 260개의 데이터로 이루어져 있다. 주변 환경변수들을 입력변수로 사용하며 각각 Tamb (The ambient temperature at site), COM (The compressor speed), LPT (The low pressure turbine speed), Pcd (The compressor discharge pressure), Texh (The turbine exhaust temperature)등이 있다. 실험을 위해 학습데이터 130(50%)개와 테스트 데이터 130(50%)개로 나누었으며, 모델의 성능평가는 MSE를 사용하였다.

표 3에 최적화를 적용하지 않은 모델과 최적화를 적용한 모델의 성능을 나타내었다. 또한 GA로 최적화된 모델과 SSOA로 최적화된 모델을 비교하였다.

NOx 데이터의 경우 연결가중치를 단일상수항으로 사용한 모델과 다항식으로 사용한 모델과의 성능차이가 확연히 나타난다. 또한 최적화되지 않은 모델에서 가장 좋은 성능을 나타내는 다항식은 은닉층 노드가 6개 일 때의 quadratic

표 3. NOx 데이터에 SSOA를 적용한 모델과 GA를 적용한 모델의 성능지수 비교.

Table 3. Comparison of performance index of SSOA-based model and GA-based model for NOx dataset.

N	T	Without optimization		With optimization			
		PI		SSOA-based model		GA-based model	
		PI	EPI	PI	EPI	PI	EPI
4	C	442.01 ±18.86	491.40 ±59.04	64.19 ±8.19	68.85 ±9.60	51.41 ±14.38	60.47 ±17.89
	L	15.42 ±0.84	19.22 ±1.73	0.87 ±0.19	1.52 ±0.59	0.93 ±0.30	1.76 ±0.26
	Q	2.05 ±0.16	3.06 ±0.50	0.03 ±0.01	0.32 ±0.28	0.05 ±0.01	0.27 ±0.01
	M	7.71 ±0.45	10.15 ±1.07	0.33 ±0.11	1.17 ±0.45	0.35 ±0.09	0.95 ±0.22
6	C	421.74 ±25.93	461.00 ±56.54	51.75 ±2.45	60.98 ±9.57	47.75 ±10.20	50.81 ±9.34
	L	14.99 ±0.98	19.17 ±1.79	0.60 ±0.07	1.57 ±0.41	0.70 ±0.12	1.33 ±0.38
	Q	1.91 ±0.24	3.22 ±1.21	0.02 ±0.01	0.43 ±0.36	0.05 ±0.01	0.24 ±0.06
	M	7.19 ±0.57	9.82 ±0.87	0.33 ±0.03	1.19 ±0.35	0.37 ±0.06	0.90 ±0.27
8	C	351.64 ±37.08	387.24 ±105.5	55.08 ±5.44	65.68 ±11.44	40.21 ±8.55	44.86 ±8.93
	L	13.37 ±1.22	17.17 ±3.17	0.62 ±0.09	1.08 ±0.22	0.70 ±0.11	1.21 ±0.26
	Q	2.16 ±0.38	3.39 ±0.89	0.02 ±0.01	0.43 ±0.47	0.13 ±0.19	0.20 ±0.14
	M	6.74 ±0.22	9.36 ±1.21	0.32 ±0.09	3.54 ±4.66	0.40 ±0.05	1.58 ±0.41

N: No. of nodes T: Polynomial Type.

C: Constant L: Linear Q: Quadratic M: Modified Quadratic.

표 4. 기존 모델과의 성능지수 비교 분석.

Table 4. Comparative analysis of performance of several selected models for the NOx emission process.

Model		Type	N	PI	EPI
Regression model		-	-	17.68	19.23
Standard Neural Networks		-	25	0.02	3.85
Polynomial networks		-	21	0.94	1.09
FNN(GAs+complex) [19]	Constant		30	6.27	8.78
	Linear		30	3.73	5.29
Multi-FNN [20]	Constant		120	2.81	5.16
	Linear		120	0.720	2.03
Proposed model	Without Optimization	Quadratic	4	2.05±0.16	3.06±0.50
			6	1.91±0.24	3.22±1.21
			8	2.16±0.38	3.39±0.89
	With Optimization	Quadratic	6	0.02±0.01	0.43±0.36
			8	0.02±0.01	0.43±0.47

N: No. of rules(nodes)

이다. 하지만 성능이 1.91 ± 0.24 로 나옴에 비해, 최적화된 모델의 경우 SSOA는 0.02 ± 0.01 , GA는 0.05 ± 0.01 로 크게 개선되었음을 알 수 있다. 이 데이터 역시 앞서 실험한 비선형 데이터와 마찬가지로 SSOA로 최적화된 모델이 GA로 최적화된 모델보다 전체적으로 성능이 우수하게 나타난다.

표 4는 제안된 모델의 성능을 기존의 모델과 평가하기 위해 실험결과를 정리하였다.

FNN은 퍼지셋(fuzzy set)을 기반으로 모델의 전반부를 구성하며, 후반부의 연결가중치는 BP(back propagation)으로 학습한다. 또한 멤버십 함수의 정점은 유전자 알고리즘과 (GA)과 complex 방법을 동시에 사용하여 최적화한다. Multi-FNN은 HCM으로 입력 데이터를 클러스터링 하고 각 클러스터에 대해 모델을 설계하여 유전자 알고리즘으로 최적화된 모델을 구축한다. 표에서 보듯이 SSOA로 최적화된 모델의 근사화지수(PI) 및 일반화지수(EPI) 모두 우수함을 알 수 있다. 또한 다른 모델에 비해 적은 노드로 모델이 구성되었음을 알 수 있다.

V. 결론

본 논문에서는 SSOA를 이용한 K-Means 클러스터링 기반 다항식 방사형 기저 함수 신경회로망을 제안하였다. 제안된 모델의 특징은 은닉층 활성화함수의 중심점 및 분포상수를 K-Means 클러스터링 알고리즘을 이용하여 초기값을 설정하고, SSOA로 한번 더 학습을 통해 모델의 성능을 개선시킨다. 또한 연결가중치를 기존의 단일상수항 대신 다항식 형태의 연결가중치를 이용하여 모델의 출력을 표현하였다. 연결가중치의 파라미터 계수는 가중최소자승법(WLSE)을 사용하였다. SSOA는 공간탐색을 조절하여 최적해를 찾아가는 알고리즘으로 교배연산과 돌연변이연산을 갖고 있다. 실험에서 K-Means clustering을 사용한 모델의 결과보다 SSOA로 최적화된 모델의 성능이 더 우수하게 나타났으며, 유전자 알고리즘에 비해 빠른 수렴속도로 모델의 최적 값을 찾아간다.

참고문헌

- [1] L. Zhao, Y. Yang, and Y. Zeng, "Eliciting compact T-S fuzzy models using subtractive clustering and coevolutionary particle swarm optimization," *Neurocomputing*, vol. 72, no. 10-12, pp. 2569-2575, June 2009.
- [2] B. Niu, Y. Zhu, X. He, and H. Shen, "A multi-swarm optimizer based fuzzy modeling approach for dynamic systems processing," *Neurocomputing*, vol. 71, no. 7-9, pp. 1436-1448, March 2008.
- [3] A. Kandal, L. Li, and Z. Cao, "Fuzzy inference and its application to control systems," *Fuzzy Sets and Systems*, vol. 48, no. 1, pp. 99-111, 1992.
- [4] Z.-B. Xu, H. Q. J. Peng, and B. Zhang, "A comparative study of two modeling approaches in neural network," *Neural Networks*, vol. 17, no. 1, pp. 73-85, Jan. 2004.
- [5] L. Sanchez, I. Couso, and J. Casillas, "Genetic learning of fuzzy rules based on low quality data," *Fuzzy Sets and Systems*, vol. 160, no. 17, pp. 2524-2552, Sep. 2009.
- [6] S. Kiranyaz, T. Ince, A. Yildirim, and M. Gabbouj, "Evolutionary artificial neural networks by multi-dimensional particle swarm optimization," *Neural Networks*, vol. 22, no. 10, pp. 1448-1462, Dec. 2009.
- [7] A. A. Frolov, D. Husek, I. P. Muraviev, and P. Y. Polyakov, "A boolean factor analysis by attractor neural network," *IEEE Transactions Neural Networks*, vol. 3, pp. 698-707, May 2007.
- [8] R. A. Aliev, B. G. Guirimov, B. Fazlollahi, and R. R. Aliev, "Evolutionary algorithm-based learning of fuzzy neural networks. part 2: recurrent fuzzy neural networks," *Fuzzy Sets and Systems*, vol. 160, no. 17, pp. 2553-2566, Sep. 2009.
- [9] S. B. Roh, S. K. Oh, and W. Pedrycz, "A fuzzy ensemble of parallel polynomial neural networks with information granules formed by fuzzy clustering," *Knowledge-Based Systems*, vol. 23, no. 3, pp. 202-219, April 2010.
- [10] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, New York, 1981.
- [11] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129-137, Mar. 1982.
- [12] J. Wang, J. Liu, and L. Liu, "A mountain means clustering algorithm," *7th World Congress on Intelligent Control and Automation (WCICA 2008)*, pp. 5045-5049, 2008.
- [13] N. R. Pal and D. Chakraborty, "Mountain and subtractive clustering method: Improvements and generalizations," *International Journal of Intelligent Systems*, vol. 15, no. 4, pp. 329-341, April 2000.
- [14] W. Shen, X. Guo, C. Wu, and D. Wu, "Forecasting

- stock indices using radial basis function neural networks optimized by artificial fish swarm optimization,” *Knowledge-Based Systems*, vol. 24, no. 3 pp. 378-385, April 2011.
- [15] A. Azadeh, M. Saberi, and S. M. Asadzadeh, “An adaptive network based fuzzy inference system-auto regression-analysis of variance algorithm for improvement of oil consumption estimation and policy making: The cases of Canada United Kingdom, and South Korea,” *Applied Mathematical Modelling*, vol. 35, no. 2, pp. 581-593, Feb. 2011.
- [16] D. Wu, K. Warwick, Z. Ma, J. G. Burgess, S. Pan, and T. Z. Aziz, “Prediction of Parkinson’s disease tremor onset using radial basis function neural networks,” *Expert Systems with Applications*, vol. 37, no. 4, pp. 2923-2928, April 2010.
- [17] J. Li and X. Liu, “Melt index prediction by RBF neural network optimized with an MPSO-SA hybrid algorithm,” *Neurocomputing*, vol. 74, no. 5, pp. 735-740, Feb. 2011.
- [18] O. Nelles and R. Isermann, “Basis function networks for interpolation of local linear models,” *Proceedings of IEEE Conference on Decision and Control*, vol. 1, pp. 470-475, 1996.
- [19] S. K. Oh, W. Pedrycz, and H. S. Park, “Hybrid identification in fuzzy-neural networks,” *Fuzzy Set and Systems*, vol. 138, no. 2, pp. 399-426, Sep. 2003.
- [20] S. K. Oh, W. Pedrycz, and H. S. Park, “Rule-based multi-FNN identification with the aid of evolutionary fuzzy granulation,” *Knowledge-Based systems*, vol. 24, no. 1, pp. 1-13, Jan. 2004.
- [21] B. S Kim, S. K. Park, B. H. Choi, E. T. Kim, H. J. Lee, and H. J. Kang, “Collision risk assessment for pedestrians’ safety using neural network,” *Journal of Institute of Control Robotics and Systems (in Korean)*, vol. 17, no. 1, pp. 1-11, Jan. 2011.
- [22] K. S. Seo and B. Y. Hyun, “GPU implementation techniques of genetic algorithm and comparative studies,” *Journal of Institute of Control Robotics and Systems (in Korean)*, vol. 17, no. 32, pp. 328-335, Jan. 2011.

김 욱 동

2009년 수원대학교 전기공학과 졸업.
2011년 동 대학원 석사. 2011년~현재
동 대학원 박사과정. 관심분야는 퍼지
추론 시스템, 뉴럴 네트워크, Granular
Computing, 진화론적 최적화 알고리즘
Computational Intelligence 등.

오 성 권

1981년 연세대학교 전기공학과 졸업.
1983년 동 대학원 석사, 1993년 박사.
1983년~1989년 금성산전연구소(선임연
구원). 1996년~1997년 캐나다 Manitoba
대학 전기 및 컴퓨터공학과 Post-Doc.
1993년~2004년 원광대학교 전기전자
및 정보공학부 교수. 2005년~현재 수원대학교 전기공학과
교수. 2002년~현재 대한전기학회, 제어·로봇·시스템학회,
퍼지 및 지능시스템학회 편집위원. 관심분야는 퍼지 시스
템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 computational
intelligence, 지능 제어 등.