정보 유사성 기반 입자화 중심 RBF NN의 진화론적 설계

논 문
59-2-35

Genetic Design of Granular-oriented Radial Basis Function Neural Network Based on Information Proximity

박 호 성* · 오 성 권† · 김 현 기**
(Ho-Sung Park · Sung-Kwun Oh · Hyun-Ki Kim)

Abstract - In this study, we introduce and discuss a concept of a granular-oriented radial basis function neural networks (GRBF NNs). In contrast to the typical architectures encountered in radial basis function neural networks(RBF NNs), our main objective is to develop a design strategy of GRBF NNs as follows : (a) The architecture of the network is fully reflective of the structure encountered in the training data which are granulated with the aid of clustering techniques. More specifically, the output space is granulated with use of K-Means clustering while the information granules in the multidimensional input space are formed by using a so-called context-based Fuzzy C-Means which takes into account the structure being already formed in the output space. (b) The innovative development facet of the network involves a dynamic reduction of dimensionality of the input space in which the information granules are formed in the subspace of the overall input space which is formed by selecting a suitable subset of input variables so that the this subspace retains the structure of the entire space. As this search is of combinatorial character, we use the technique of genetic optimization to determine the optimal input subspaces. A series of numeric studies exploiting some nonlinear process data and a dataset coming from the machine learning repository provide a detailed insight into the nature of the algorithm and its parameters as well as offer some comparative analysis.

Key Words : Radial basis function neural network, Proximity, Context-based fuzzy C-means clustering, Information granules, Genetic algorithm

1. 서 론

차원성 문제는 뉴로 컴퓨팅 분야에서 중요한 주제의 한 부분으로써 나타나고 있으며 이는 고차원의 실세계 문제를 처리하는데 있어서 우리의 근본적인 도전으로 주어지고 있을 뿐만 아니라 이를 해결하기 위한 강한 동기를 부여함으로써 많은 연구들이 진행되어져 왔다[1-2]. 이와 같은 연구들이 강조하는 부분은 효율적인 데이터 시각화와 데이터의 표현, 계산의 복잡성에 대한 감소 그리고 정확도의 중요성이며, 더 나아가 모델의 증가된 차원성으로 인한 반복적인 학습 절차를 처리할 때 요구되어지는 많은 계산적 시간과 복잡성을 개선해야하는 어려움에 직접적으로 직면하고 있다.

이러한 차원성의 이슈는 방사형 기저 함수 신경회로망[3-4]와 같은 학습의 문맥에서 연구되어져 왔다. 방사형 기저 함수 신경회로망은 주어진 간단한 형태의 구조와 일반적인 근사화 능력[5] 때문에 방대한 지능형 시스템의 중요한 연구 분야중의 하나로 연구되어져 왔으며 패턴인식, 신호처리, 시계열 예측 그리고 비선형 시스템의 모델링과 제어와 같은 많은 분야에서 적용되어져 왔다[6-8]. 그러나 이러한 문헌들은 주어진 시스템의 입출력 데이터에 대한 정보의 고려 없이 진화론적 알고리즘에 의존하는 경향이 뚜렷하며 따라서 주어진 데이터의 정보를 통한 효율적인 차원성 감소에 대한 연구가 미비한 상태이다. 이에 본 논문에서는 주어진 정보 입자화의 효율적인 처리를 이용하여 정보 유사성 기반 입자화 중심 방사형 기저 함수 신경회로망(GRBR NNs)을 제안한다. 본 연구에서 우리의 목적은 (a) 네트워크의 구조는 학습 데이터에서 고려된 구조의 전체적인 영향으로 퍼지 클러스터링의 도움을 받아 입자화 되고, (b) 은닉층들의 형태는 주어진 전체 데이터들과 선택된 입력변수들과의 유사성을 측정하여 효율적이고 최적의 입력 공간에서의 차원성 감소를 실현하는 전략을 개발하는 데 있다.

즉, 제안된 GRBF NNs의 개발에 있어서, 네트워크의 구조는 정보 입자화의 도구로써 실현된 context-based fuzzy C-means 클러스터링[8-10]을 이용하여 네트워크 구조를 형성한다. 여기서, context-based fuzzy C-means 클러스터링 방법은 입력 공간(입력 변수)에서의 데이터 구조가 출력 공간에서 미리 정의된 몇 개의 언어적인 경계값을 기반으로 조건화 된다. 또한 유전자 알고리즘[11-12]을 이용하여 방사형 기저 함수 신경회로망의 은닉층에 사용될 각각의 노드들에 대한 입력 변수들을 선택하며, 최종적으로 유전자 알고리

* 정 회 원 : 수원대 산업기술연구소 학술연구교수 · 공박
† 교신저자, 정회원 : 수원대 공대 전기공학과 교수 · 공박
E-mail : ohsk@suwon.ac.kr
** 정 회 원 : 수원대 공대 전기공학과 교수 · 공박
접수일자 : 2009년 7월 24일
최종완료 : 2009년 12월 28일

즘을 통해 선택된 입력 변수들은 전체 입력 변수들과의 유사성 측정[13-15]을 통하여 은닉층에서 사용될 입력 공간의 효율적인 차원의 감소를 위한 입력 변수들의 최적 부분 집합을 결정한다.

2. 일반적인 방사형 기저 함수 신경회로망

2.1 방사형 기저 함수 신경회로망의 형태

방사형 기저 함수 신경회로망[3-7, 16-19]는 3개의 층을 가진 신경 회로망으로써 n 차원의 입력 벡터 $\mathbf{x}=[x_1, x_2, ..., x_n]^T$는 은닉층을 통해서 비선형식으로 변환된다. 결과적으로 얻어진 활성 레벨들은 출력층에 위치한 뉴론에 의해서 선형 형태로 변환된다. 공통적으로 은닉층에서의 활성함수는 다음과 같이 가우시안 기저 함수 형태로 표현된다.

$$\Phi_i(\mathbf{x}) = \exp\left(-\frac{\| \mathbf{x}-\mathbf{v_i} \|^2}{2r_i^2}\right) \tag{1}$$

여기서, $\mathbf{v}_i$와 r_i는 i번째 노드의 가우시안 기저 함수의 중심값과 폭을 의미한다. 네트워크의 출력 $y(\mathbf{x})$는 각각의 활성화 레벨의 선형 조합으로써 다음과 같이 계산된다.

$$y(\mathbf{x}) = \sum_{i=1}^{K} w_i \Phi_i(\mathbf{x}) \tag{2}$$

2.2 방사형 기저 함수 신경회로망의 설계

방사형 기저 함수 신경회로망에 적용된 구조로서 모든 방사형 기저 함수 신경회로망의 설계는 (a) 은닉층에서의 개발과 (b) 네트워크의 출력층에서의 뉴론들의 연결 값의 최적화로 구성될 수 있다. 그러나 표준적으로 네트워크의 출력층에서 뉴론들의 연결 값의 최적화 임무보다 더욱 도전적으로 중요한 것은 은닉층에서의 개발에 대한 네트워크의 효율적인 기능이다. 즉, 은닉층에서의 개발이 결정적인 방사형 기저 함수 신경회로망 설계에 있어서 중요한 특징으로 부각된다. 그 이유로는 다음과 같이 들 수 있겠다.

- 은닉층의 형태와 수는 기본적인 설계의 초점이 된다.
- 입력공간에서의 은닉층의 분포는 네트워크의 기능에서 중요한 관련이 있다
- 각각의 은닉층에서의 파라미터의 최적화는 본질적으로 가장 중요하다.

기존의 논문들에서는 이와 같은 중요한 설계 이슈들을 다루기 위해서 여러 가지 방법을 시도하였다. 즉, 방사형 기저 함수 노드 상호간의 연결 가중치 조절과 방사형 기저 함수 노드의 중심 및 폭과 같은 파라미터 변경 등이 연구되었으며[16-17], 은닉층의 수에 대한 결정 역시 다수의 논문을 통해서 시도되었다[17-18]. 또한 은닉층의 형태가 결정되어진 후에, 출력층에서의 노드들의 가중치에 대한 최적화가 직접적으로 이루어진 논문들도 있다[19].

방사형 기저 함수 신경회로망의 설계는 고차원성의 문제가 거론될 때 마다 더욱 어려운 문제에 직면하게 된다. 즉, 이러한 문제들은 은닉층 형태에 대한 한계에 도달하게 되는데 이는 증가된 차원성이 네트워크의 전체적인 근사화 및 일반화 능력의 감소뿐만 아니라 무의미한 입력의 수를 야기하기 때문이다. 본 논문에서는 이와 같은 단점을 극복하는 것이 개발의 목적이다.

3. 입자화 중심 데이터 처리 기법

3.1 K-means clustering을 통한 출력공간의 입자화

K-means 클러스터링은 영상과 음성 데이터 압축[20], 시스템 모델링을 위한 데이터 전처리[21]와 같은 다양한 영역에서 사용되어져 왔다. 이는 K-means 클러스터링 방법들이 데이터 압축과 모델 동정에 사용되어진 것뿐만 아니라 구성과 분류에도 널리 사용되어져 왔다는 것을 강조한다.

본 논문에서는 K-Means 알고리즘을 이용하여 출력 데이터인 $\{target_k\}$를 P-2개로 분류하고 다음과 같은 목적함수 Q가 최소화 되는 각각의 클러스터에 대한 중심값을 구한다.

$$Q = \sum_{i=2k=1}^{P-1} \sum^{N} (target_k - y_i)^2 \tag{3}$$

또한 출력변수의 최소 및 최대값은 클러스터의 첫 번째와 마지막 중심값으로 설정한다. 따라서 K-means 클러스터링 알고리즘을 이용하여 전체적으로 P개의 중심이 생성된다. 이와 같이 얻어진 중심값들은 각각 $y_1, y_2, ..., y_P$로 표기하고 K-means 클러스터링 방법을 통해 얻어진 멤버쉽 함수의 값을 $T_1, T_2, ..., T_P$로 각각 표기한다. 특별히 j-번째 fuzzy set에서의 데이터 $\{target_k\}$의 멤버쉽 정도를 t_{jk}로 표기한다.

3.2 Context-based Fuzzy C-means clustering을 통한 입력공간의 입자화

Context-based FCM 클러스터링은 출력공간에서 미리 정의된 몇 개의 context라 불리는 퍼지 집합의 분류들을 기준으로 입력 데이터에서의 공간을 정보 입자화의 특성에 맞게 분류하는 알고리즘이다[8-10]. 이는 입력공간과 출력 공간 영역에서 사이의 관계를 드러나게 함으로써, 입력과 출력을 동시에 고려한 특징을 가지고 있다.

Context-based FCM 클러스터링 방법으로 입력 데이터 $\{\mathbf{x}_k\}$를 "c"개의 클러스터로 분할한다. $U(T_j)$라 표시되는 j-번째 context에 의해서 유도되는 partition matrix를 구하기 위한 context-based FCM 클러스터링 알고리즘은 다음과 같다.

$$\mathbf{U}(T_j) = \left\{ u_{ik} \in [0,1],\ \sum_{i=1}^{c} u_{ik} = t_{jk}\ \forall k,\ 0 < \sum_{k=1}^{N} u_{ik} < N\ \forall i \right\} \tag{4}$$

여기서, c는 입력 데이터에서의 클러스터의 수, t_{jk}는 삼각형 멤버쉽 함수에 의해서 얻어진 j-번째 context에 속해 있

는 k-번째 데이터들의 멤버쉽 값을 의미한다.

Context-based FCM 클러스터링의 목적함수는 다음과 같다.

$$V = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^{m} \| \mathbf{x_k} - \mathbf{z_i} \|^2 \tag{5}$$

여기서, m>1은 퍼지화 계수를 나타내고, $\mathbf{z}_i$는 입력변수들의 중심값을 의미한다.

목적함수 V는 식 (4)에서 표현된 조건에 만족하도록 최소화되며, 이는 각각의 context들에 의해 분류된 "P"개의 클러스터링에 적용된다. 간단히 다음과 같이 쓸 수 있다.

$$\mathrm{Min}\, V \mathrm{subject\ to}\ \mathbf{U}(T_j), \quad j = 1, 2, ..., P \tag{6}$$

Context-based FCM에 의해서 실행되는 목적 함수 V의 최소화는 partition matrix와 중심값들이 반복적으로 갱신되는 것으로 실현된다.

Partition matrix의 갱신은 다음과 같이 실행된다.

$$u_{ik} = \frac{t_{jk}}{\sum_{l=1}^{c}\left(\frac{\| \mathbf{x_k} - \mathbf{z_i} \|}{\| \mathbf{x_k} - \mathbf{z_l} \|}\right)^{2/(m-1)}}, \quad \begin{matrix} i = 1, 2, ..., c, \\ k = 1, 2, ..., N \end{matrix} \tag{7}$$

중심값 ($\mathbf{z}_1$, $\mathbf{z}_2$, ..., $\mathbf{z}_c$)은 다음과 같은 식을 이용하여 계산된다.

$$\mathbf{z}_i = \frac{\sum_{k=1}^{N} u_{ik}^{m} \mathbf{x}_k}{\sum_{k=1}^{N} u_{ik}^{m}} \tag{8}$$

여기서, i=1, 2, ..., c.

4. 정보 유사성 기반 차원성 축소

본 논문에서는 차원성 축소를 위해서 간단하지만 효율적인 알고리즘인 유사성 기반 메커니즘을 사용한다. 즉, 본 논문에서 추구하고자 하는 본질은 효율적이고 최적의 모델을 구축하는데 있어서 필요 없는 입력 변수 또는 변수들의 집합을 제거함으로써 전체 공간에서 구축된 원래의 구조에 비해서 정보 유사성을 기반으로 한 입력공간의 차원성이 감소될 때의 네트워크 성능이 우수하게 유지되는 것이다.

알고리즘적 관점에서 볼 때, 감소된 공간에서 구조는 클러스터링 데이터에 의해서 정량화될 수 있으며 유사성 기반 partition matrix의 형태는 전체 입력변수들의 partition matrix와 새로이 선택된 입력변수들의 partition matrix의 차이가 어느 정도냐는 비교에 따라서 형성된다. 이와 같은 2개의 partition matrix들의 비교는 상응하는 partition matrix에 의해서 실현된 유사성 matrix에 의해서 계산을 실행한다[13-15].

정보 유사성 matrix에 대한 주요 개념들은, 먼저 k1과 k2로 표현되어진 2개의 패턴을 고려하며 이들은 각각 P_{k1}과 P_{k2}와 같은 partition matrix을 가진다.

Prox(P_{k1}, P_{k2})로 표기되어진 2개의 데이터 사이의 유사성은 다음과 같이 표현된다.

$$\begin{aligned} &\mathrm{Prox}(U_P) = [P_{k1,k2}], \\ &P_{k1,k2} = \sum_{i=1}^{c} \min(u_{ik1}, u_{ik2}), \quad k1, k2 = 1,...,N \end{aligned} \tag{9}$$

여기서 U_p는 p-번째 context의 proximity matrix를 의미하고, Prox(U_p)=[$P_{k1,k2}$]로 나타낸다. u_{ik1}, u_{ik2}는 식 (7)을 이용하여 계산된다. 식 (9)를 살펴보면, 유사성 함수는 대칭이며 같은 데이터 점들(즉, k1=k2)이거나 두 개의 데이터에 대한 멤버쉽 값들이 같을 때 '1'의 값을 나타낸다.

전체 입력변수들 중에서 선택된 입력변수들을 고려했을 때, 선택된 입력변수에 대해서 상응되는 partition matrix (f_{ik})는 다음과 같은 방법으로 계산된다.

$$f_{ik} = \frac{t_{pk}}{\sum_{j=1}^{c}\left(\frac{\| \mathbf{x}_k^l - \mathbf{v}_i \|}{\| \mathbf{x}_k^l - \mathbf{v}_j \|}\right)^{2/(m-1)}} \tag{10}$$

식 (10)에서 t_{pk}는 "3.1 K-means clustering을 통한 출력공간의 입자화"에서 얻어진 partition matrix에서 p-번째 행을 나타내고, $\mathbf{x}^l$은 선택된 입력 변수들의 벡터이다. 여기서, $\mathbf{v}_i$는 선택된 입력 집합 ($\mathbf{x}^l$)의 중심값을 의미하고 이 값들은 식 (8)에서 보여준 것과 같은 방법으로 계산된다. 여기서, 전체 입력변수들 중에서 입력변수들의 선택은 유전자 알고리즘을 통해서 선택하게 된다.

식 (10)을 통해 얻어진 선택된 입력변수들에 대한 partition matrix을 이용하여 다음과 같이 선택된 입력 변수들에 대한 proximity matrix를 계산한다.

$$\begin{aligned} &\mathrm{Prox}(F_p) = [F_{k1,k2}], \\ &F_{k1,k2} = \sum_{i=1}^{c} \min(f_{ik1}, f_{ik2}), \quad k1, k2 = 1,...,N \end{aligned} \tag{11}$$

식 (9)을 통해서 얻어진 전체 입력변수들에 대한 proximity matrix와 식 (11)을 통해서 얻어진 선택된 입력변수들에 대한 proximity matrix를 이용하여 다음과 같이 거리를 계산한다.

$$\begin{aligned} V_p &= \| \mathrm{Prox}(U_p) - \mathrm{Prox}(F_p) \| \\ &= \| P_{k1,k2} - F_{k1,k2} \| = \sum_{k1=1}^{N}\sum_{k2>k1}^{N} |P_{k1,k2} - F_{k1,k2}| \end{aligned} \tag{12}$$

우리의 목적은 감소된 입력 공간에서 모델의 구조를 유지하는 것이고 이와 같은 효과는 V_p의 값에 의해서 정량화된다. 즉, V_p의 값이 작으면 작을수록 감소된 공간에서의 정보 입자화에 대한 구조가 전체 공간에서의 정보 입자화 구조와 비슷하다는 것을 알 수 있다. 형식적인 관점으로부터, 식 (12)의 최소화는 수학적 조합의 특성을 나타낸다. 어떠한 차원의 입력 변수들의 부분 집합을 **I**로 표기하면, s는 다음과

같이 나타낼 수 있다. s<n (여기서, n은 전체 입력 공간에서의 차원의 수를, s는 선택된 입력 공간의 차원의 수를 나타낸다.). V_p의 최소화는 적합하게 선택되어진 입력변수들의 부분 집합 $\mathbf{I}$의 선택에 의해서 실행되며 이는 다시 말해서 다음과 같이 표기할 수 있다. 즉, $\mathbf{I}_{opt} = \arg\min_{\mathbf{I}} V_p$. 조합적인 특성에 기인한 입력변수들의 선택에 대한 문제점은, 유전자 알고리즘과 같은 진화론적 최적화의 기법을 이용하여 해결하였다.

5. 입자화 중심 방사형 기저 함수 신경회로망의 형태

본 장에서는 본 논문에서 제안한 입자화 중심 방사형 기저 함수 신경회로망의 주요한 구조적 특징에 대해서 살펴본다. 제안된 네트워크는 context-based 클러스터링 방법의 개념을 기반으로 하고 있으며 모든 입력 변수들에 대해서 형성된 퍼지 분할의 형태는 그림 1에서 보여준 것과 같은 형태를 이루고 있다.

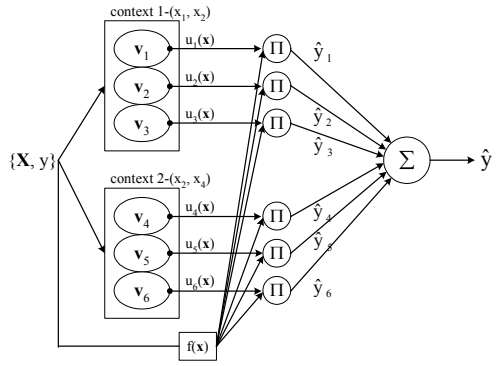


그림 1 입자화 중심 방사형 기저 함수 신경회로망 구조
Fig. 1 Structure of the Granular-oriented RBF neural network

제안된 입자화 중심 방사형 기저 함수 신경회로망은 퍼지 규칙 기반 시스템의 형태로써, 제안된 구조는 다음과 같은 규칙의 형태로 표현 할 수 있다.

$$R^i: \text{If } \mathbf{x} \text{ is } "p-\text{th context}" \text{ and } "c-\text{th cluster}" \quad (13)$$
$$\text{then } \hat{y}_i = \mathbf{a}_0^T + \mathbf{a}_i^T \mathbf{x}^i$$

여기서, $\hat{y}_i$는 i번째 지역 모델 (i=1, 2, ..., p×c), $\mathbf{x}^i$는 p번째 context에 해당되는 선택된 입력 집합, 그리고 R^i는 i번째 규칙을 의미한다.

그림 1에서 표현된 모든 규칙을 고려한다면, 네트워크의 전체 구조는 다음과 같이 표현된다.

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{c} u_i(\mathbf{x})\left[\mathbf{a}_0^T + \mathbf{a}_i^T \mathbf{x}^1\right] + \sum_{i=c+1}^{2c} u_i(\mathbf{x})\left[\mathbf{a}_0^T + \mathbf{a}_i^T \mathbf{x}^2\right] \quad (14)$$
$$+ \ldots + \sum_{i=(p-1)c+1}^{p\times c} u_i(\mathbf{x})\left[\mathbf{a}_0^T + \mathbf{a}_i^T \mathbf{x}^p\right]$$

여기서, 은닉층의 출력값인 $u_i(\mathbf{x})$값들은 식 (7)을 통해 얻어진 값들이다.

제안된 모델의 성능 지수식은 식 (15)에서와 같이 Root Mean Square Error (RMSE)를 사용한다.

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(target_k - \hat{y}_k\right)^2} \quad (15)$$

각각의 지역 모델의 계수는 최소자승법(LSE)을 사용한다.

6. 유전자 알고리즘

최적화 알고리즘으로 사용되었던 유전자 알고리즘[11-12]는 조합적인 문제를 해결하는데 있어서 최적의 도구로써 사용되어져있으며 모든 진화 기반의 탐색 알고리즘 가운데 유연하고 강인한 탐색능력으로 인해 불연속 함수 문제를 포함한 최적화문제를 해결할 수 있는 알고리즘이다.

본 논문에서는 유전자 알고리즘을 시뮬레이션하기 위해 설정된 각각의 개체에 존재하는 염색체의 수와 군집의 크기에 대한 파라미터의 정의로써, 한 개체에 여러 변수의 정보를 담은 직렬 방식을 사용한다. 재생산(Reproduction)연산으로 roulette-wheel 방식, 돌연변이연산은 선택된 비트를 반전시키는 invert 방식, 교차(crossover)연산은 one-crossover 방식을 사용하였다. 그러나 roulette-wheel 방식은 확률적 속성 때문에 선택 과정에서 최적자를 반드시 선택하지 못하는 단점을 가진다. 이를 보완해줄 수 있는 한 방법이 본 논문에서는 엘리트 전략(Elitist strategy)을 사용하였다[12].

본 논문에서 사용된 유전자 알고리즘은 차원성 감소를 위한 객관적이고 효율적인 입력 변수들을 선택하는 사용되어지며, 식 (16)에서 보여준 것과 같은 적합도 함수의 값이 최대가 되도록 사용하였다.

$$\text{Fitness function} = 1/V_p \quad (16)$$

식 (16)에서 V_p는 식 (12)에서 구해진 값이다.

그림 2는 염색체의 배열과 선택된 입력 변수들이 선택되어지는 과정을 보여준다.

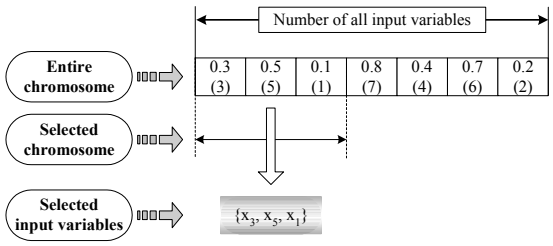


그림 2 유전자 알고리즘을 통한 입력 변수들의 선택
Fig. 2 Selection of input variables through genetic algorithm

7. 시뮬레이션 및 결과고찰

본 논문에서 제안된 네트워크 구조는 몇 개의 수치적 실험을 통해서 모델의 특징과 성능에 대해서 알아본다. 첫 번째 데이터 집합은 하수처리 공정 데이터와 가스 터빈 발전

소의 NOx 배출 메카니즘 데이터이고 두 번째 데이터 집합은 Machine Learning data 집합이다. 모든 데이터는 전체 데이터를 랜덤하게 학습 데이터(60%)와 테스트 데이터(40%)로 나눴으며, 제안된 모델의 정량적인 평가를 위하여 식 (15)에서 표현한 것과 같이 성능 지수식은 Root Mean Square Error (RMSE)를 이용하였다.

또한 제안된 모델의 성능에 있어서 퍼지화 계수의 영향에 따른 성능을 분석하고자, 본 논문에서는 퍼지화 계수 "m"에 대한 다양한 값(1.1, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0)을 적용하였고, 선택될 입력 변수의 수는 시스템 입력이 "n"개라면 1부터 (n-1)개 까지 개별적으로 미리 결정하여 모델을 평가하였다. 또한 입출력 공간에서 형성된 입자화를 제어할 수 있는 2개의 주요한 파라미터 "P"와 "c"를 통해서 실험적 성능을 평가하였다. 여기서 "P"는 context의 수를, "c"는 클러스터의 수를 의미한다. 실험에 사용된 context의 수는 2~6개로, 각각의 context에 대한 클러스터의 수는 2~5개로 한정하였다.

표 1은 수치적 실험적 예제에서 사용된 네트워크의 유전자 알고리즘의 파라미터들과 각각의 실험적 데이터들의 전체 입력변수들의 수를 나타낸 것이다.

표 1 유전자 알고리즘 및 모델 구축을 위한 정보

Table 1 Values of the parameters of the genetic algorithm and the proposed model

	Parameter \ Data	Sewage treatment data	Auto mobile MPG data	NOx process data
GA	Generation size	30		
	Population size	80		
	Crossover rate	0.65		
	Mutation rate	0.1		
	No. of all input variables	4	7	5
RBF NN	No. of contexts(P)	2 ~ 6		
	No. of clusters for each context(c)	2 ~ 5		
	No. of input to be selected(S)	1 ~ 3	1 ~ 6	1 ~ 4

7.1 Sewage treatment process data

본 논문에서는 수도권 하수처리장 중의 하나를 모델로 선정하여 선정된 처리장의 1년 동안의 수질 데이터를 토대로 4입력 1출력(MLSS, WSR, RRSP, DOSP ; ESS)의 활성오니 공정[22]를 제안한 알고리즘을 이용하여 최적의 모델을 구축하고자 한다. 하수처리 공정으로부터 얻어진 52개 데이터를 이용하여 랜덤하게 학습 데이터(60%)와 테스트 데이터(40%)로 나누어 이용하였다.

표 2는 각각의 퍼지계수에 대해서, context의 수에 따른 가장 좋은 성능지수의 값을 가졌을 때의 성능지수 값과 선택된 입력변수 그리고 클러스터의 수를 나타낸 것이다. 표 2에서 'S'는 전체 입력변수 중에서 진화론적으로 최적화된 상태에서 각각의 context에 대해서 가장 낮은 성능지수를 가졌을 때의 선택된 입력 변수의 수를 나타내고, 'T'는 모든 입력 변수들을 사용한 경우를 나타낸다. 또한 'PI'와 'EPI'는 각각 학습 데이터와 테스트 데이터에 대한 성능지수를 나타낸다.

표 2 제안된 모델의 성능지수

Table 2 Performance index of the proposed network

(a) Values of the PI

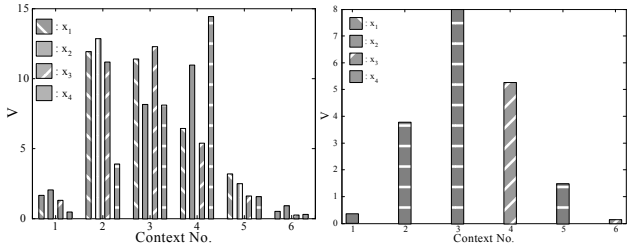
		Number of contexts(P)									
		2		3		4		5		6	
m	S	c	PI	c	PI	c	PI	c	PI	c	PI
1.1	3	4	2.1852	4	1.5844	3	0.8759	2	0.6718	2	0.5992
	T	4	1.8813	3	1.6154	4	0.6211	4	0.4469	5	0.2715
1.5	2	4	1.9732	4	1.3956	5	0.7909	2	0.5988	2	0.4780
	T	5	1.8549	2	1.1238	2	0.5033	5	0.3770	3	0.4657
2.0	1	2	1.8858	3	0.9480	3	0.7414	3	0.4941	2	**0.4013**
	T	5	1.7204	4	0.8355	5	0.2865	3	0.7323	2	0.2677
2.5	1	5	1.9477	3	1.1366	2	0.7642	5	0.8297	2	0.4467
	T	2	1.5928	4	0.8311	5	0.4046	5	0.4617	4	0.3352
3.0	1	4	2.1021	5	1.7774	2	0.7432	5	0.6899	2	0.4526
	T	4	1.7180	2	1.1174	3	0.5566	3	0.2470	5	0.0599
3.5	3	3	1.9692	3	1.1372	2	1.0221	3	0.6640	4	0.8444
	T	3	2.0829	3	1.2763	3	1.0710	3	0.8904	5	0.2224
4.0	2	2	2.2033	5	1.8520	5	1.8678	2	0.9831	5	0.8363
	T	3	1.8976	3	0.9269	4	0.9572	3	0.3397	5	0.2211
4.5	3	2	2.3861	3	1.2757	3	0.9331	2	1.1878	5	0.4540
	T	2	1.9025	5	1.5831	5	0.8625	2	1.1244	4	0.5204
5.0	3	3	2.3451	3	1.0812	5	1.0372	2	0.9273	3	1.6570
	T	2	2.1512	4	1.5265	4	1.4036	4	1.5500	3	1.0201

(b) Values of the EPI

		Number of contexts(P)									
		2		3		4		5		6	
m	S	c	EPI	c	EPI	c	EPI	c	EPI	c	EPI
1.1	3	4	2.4197	4	2.2150	3	1.8333	2	1.8434	2	1.0900
	T	4	2.7922	3	2.0335	4	2.0260	4	2.2640	5	1.6397
1.5	2	4	2.0635	4	1.3619	5	1.3913	2	1.1987	2	0.6564
	T	5	2.1665	2	1.7871	2	2.1206	5	2.4932	3	2.0033
2.0	1	2	1.9440	3	1.1957	3	0.9457	3	0.7084	2	**0.4924**
	T	5	2.2439	4	1.6746	5	1.5987	3	1.0908	2	1.1810
2.5	1	5	2.0639	3	1.3047	2	1.1102	5	1.1123	2	0.4480
	T	2	2.4039	4	1.9267	5	1.4814	5	1.3674	4	1.4803
3.0	1	4	1.5680	5	1.8370	2	1.0531	5	0.7853	2	1.1931
	T	4	3.0871	2	1.6015	3	1.3193	3	2.4600	5	2.3884
3.5	3	3	1.9392	3	1.5190	2	1.7334	3	1.0520	4	2.0935
	T	3	2.1848	3	1.8339	3	1.4879	3	2.4492	5	2.7052
4.0	2	2	1.6976	5	2.3821	5	2.6756	2	1.8036	5	1.2476
	T	3	3.1908	3	2.2887	4	1.1500	3	3.6793	5	2.6467
4.5	3	2	2.7544	3	2.6742	3	1.7795	2	2.0854	5	3.9438
	T	2	2.8910	5	2.3697	5	3.1870	2	2.8577	4	3.7372
5.0	3	3	2.8753	3	1.4810	5	2.1747	2	2.1128	3	2.6190
	T	2	2.6137	4	2.7434	4	3.2698	4	3.8657	3	2.4284

Sewage treatment process 데이터를 이용하였을 때 가장 성능이 좋은 모델은 PI=0.4013, EPI=0.4923으로써, 전체 4개의 입력 중에서 1개의 입력을 사용하고, 퍼지화 계수(m)는 2.0, 6개의 context, 그리고 각각의 context에 2개의 클러스터를 사용한 경우로써 이때 각각의 context에 선택되어진 입력변수는 {4}, {4}, {4}, {3}, {4}, 그리고 {3} 입력이 선택되어졌다. 모든 입력변수가 사용된 경우에서 가장 좋은 성능을 가진 모델은 PI=0.2677, EPI=1.1810으로써 이때 사용된 파라미터들은 퍼지화 계수(m)는 2.0, 6개의 context와 각각의 context에 2개의 클러스터가 존재하는 경우이다. 결과 값에서도 알 수 있듯이, 전체 입력변수를 사용한 경우의 성능은 선택된 입력변수를 사용한 경우에 비해서 근사화 능력은 우수하나 일반화 능력이 현저하게 저조함을 알 수 있다. 따라서 일반화와 근사화 상호 능력을 고려한다면 제안된 알고리즘을 통해서 생성된 모델의 구조가 우수함을 알 수 있다.

그림 3은 표 2에서 보여준 모델의 성능 중에서 최고의 성능을 가진 경우 (PI=0.4013, EPI=0.4923)에서의 유사성 기반 인덱스인 "V"의 값을 나타낸 것이다.

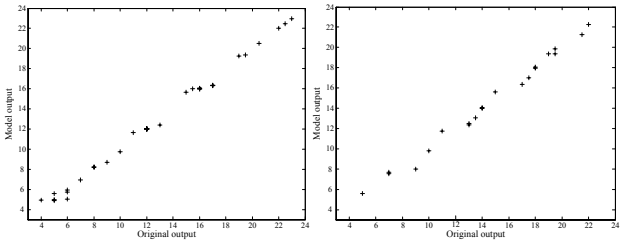


(a) All input variables (b) Selected input variables

그림 3 유사성 기반 "V"의 값

Fig. 3 Values of the proximity-based index "V"

그림 4는 그림 3에서 선택된 모델의 근사화와 일반화 능력을 나타낸 것이다.

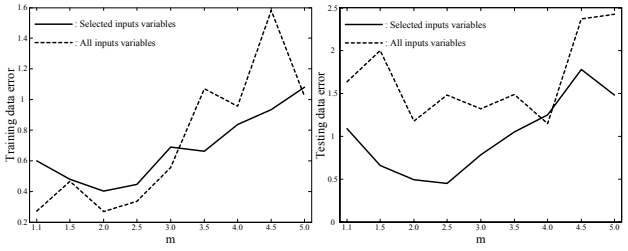


(a) Training data (b) Testing data

그림 4 원 출력과 모델 출력과의 선형성

Fig. 4 Scatter plots of data versus the output of the network

그림 5는 다양한 퍼지화 계수(m)의 값에 따른 제안된 모델의 최고의 성능 지수를 나타낸 것이다.



(a) Training data (b) Testing data

그림 5 퍼지와 계수에 따른 성능지수

Fig. 5 Performance index reported as a function of fuzzification coefficient

표 3은 표준 선형 회귀 모델과 기존의 지능형 모델에서 생성된 결과와 제안된 모델과의 성능을 요약한 것이다. 표 3에서 보이듯이, 표준 회귀 모델은 객관적인 비교를 위해서 10-fold cross validation을 실행하여 평균값과 표준오차를 나타내었다.

7.2 Automobile Miles Per Gallon (MPG) data

본 논문의 성능을 평가하기 위한 두 번째 데이터는 Machine Learning data로써 일반적으로 지능형 모델링에

표 3 기존의 모델들과의 성능 비교

Table 3 Comparative analysis of the performance of various model

Model	m	S	P	c	PI	EPI	R	비고
Linear regression model					3.7739 ±0.218	4.4558 ±0.470		RMSE
Fuzzy model [22]		2	Simple		13.726	16.206	4	MSE
		3			12.625	12.563	8	
		T			7.517	20.944	16	
		2	Linear		6.396	54.233	4	
		3			2.220	1056.9	8	
		T			0.0003	800.81	16	
Hybrid Fuzzy [23]		2	Simple		12.403	12.200	4	
		3			12.047	12.616	8	
		T			6.670	15.410	16	
		2	Linear		7.157	24.658	4	
		3			9.0e-6	212.43	8	
		T			1.0e-6	120.10	16	
Fuzzy set-based FNN[24]		2	Simple		13.401	8.287	4	
					9.971	9.428	6	
			Linear		12.307	9.828	4	
					12.932	7.786	6	
Fuzzy relation-based FNN[25]		2	Simple		12.943	12.176	4	
			Linear		10.584	12.108	4	
Proposed model	1.1	3	6	2	0.5992	1.0900	12	RMSE
		T	6	5	0.2715	1.6397	30	
	1.5	2	6	2	0.4780	0.6564	12	
		T	6	3	0.4657	2.0033	18	
	2.0	1	6	2	0.4013	0.4924	12	
		T	6	2	0.2677	1.1810	12	
	2.5	1	6	2	0.4467	0.4480	12	
		T	6	4	0.3352	1.4803	24	
	3.0	1	5	5	0.6899	0.7853	25	
		T	4	3	0.5566	1.3193	12	
	3.5	3	5	3	0.6640	1.0520	15	
		T	4	3	1.0710	1.4879	12	
	4.0	2	6	5	0.8363	1.2476	30	
		T	4	4	0.9572	1.1500	16	
	4.5	3	4	3	0.9331	1.7795	12	
		T	3	5	1.5831	2.3697	15	
	5.0	3	3	3	1.0812	1.4810	9	
		T	6	3	1.0201	2.4284	18	

R : Number of rules

있어서 벤치마킹되고 있는 MPG 데이터를 이용하여 제안된 모델의 성능을 평가하였다. MPG 데이터는 전체 392개의 입출력 데이터 쌍으로 이루어져 있으며, 출력은 gallon당 mile로 표현되는 자동차 연료 소비이며 입력은 출력과 관련된 7가지 요소(Cylinders, Displacement, Horsepower, Weight, Acceleration, Model year, Origin)들로 구성되어 있다.

표 4는 MPG 데이터를 이용한 퍼지화 계수(m)와 context의 수(P) 그리고 클러스터의 수(c)에 따른 모델의 최고의 성능을 나타낸 것으로 가장 좋은 결과를 얻은 네트워크는 퍼지화 계수(m)이 2.0이고 선택된 입력 변수의 수는 3, context의 수는 6개 그리고 context에 해당되는 클러스터의 수는 4개인 경우이며, 이 때 성능은 PI=0.8314, EPI=0.8546이다. 또한 각각의 context에 해당되는 선택된 입력변수들은 {3, 4, 5}, {3, 5, 6}, {5, 6, 7}, {5, 6, 7}, {5, 6, 7} 그리고 {5, 6, 7}이다. 선택된 입력변수에서도 알 수 있듯이, 전체 입력변수 중에서 5번째 입력(acceleration)은 모든 context에 선택되었으며, 6번째 입력(model year)과 7번째 입력(origin) 역시 최고의 모델을 구축하는데 중요한 입력변수로써 기여되어짐을 알 수 있다.

표 4 학습 데이터와 테스트 데이터에 대한 성능지수 값
Table 4 The values of the performance index for training and testing data

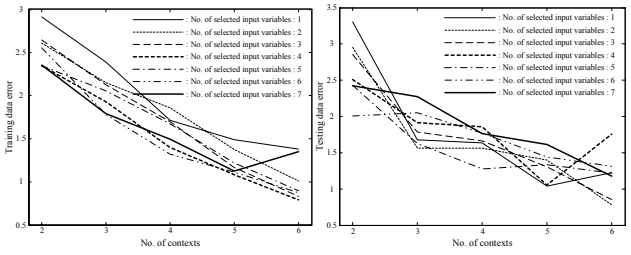
(a) Values of the PI

		Number of contexts(P)									
		2		3		4		5		6	
m	S	c	PI	c	PI	c	PI	c	PI	c	PI
1.1	3	4	2.9799	3	2.4145	4	1.8757	5	1.5114	4	1.3385
	T	4	2.4700	2	1.8801	5	1.5613	4	1.2994	5	1.2665
1.5	6	2	2.2865	2	1.9188	3	1.4279	4	1.2876	3	1.0138
	T	5	2.3781	5	1.7319	4	1.5168	4	1.4786	5	1.0185
2.0	3	2	2.6443	4	2.1254	3	1.6947	3	1.1644	4	**0.8314**
	T	3	2.3508	2	1.7914	5	1.4954	3	1.1215	2	1.3512
2.5	5	2	2.5091	2	2.0320	4	1.8688	5	1.1903	5	0.8331
	T	5	2.5269	3	2.2187	2	1.8992	2	1.6745	5	1.3315
3.0	6	2	2.6331	5	2.1852	3	1.9127	5	1.8373	2	1.0031
	T	4	2.5648	4	2.2611	3	1.8986	3	1.5099	5	1.5891
3.5	1	2	3.4279	4	2.6912	3	2.3796	3	2.4875	5	1.9138
	T	3	2.6743	5	2.1836	3	2.0635	2	1.8483	5	1.3139
4.0	4	4	2.5994	2	2.4869	3	2.4187	3	1.7580	5	1.0632
	T	3	2.6594	2	2.2458	3	2.1875	5	1.6303	3	1.1372
4.5	5	5	2.9132	5	2.6420	2	2.2253	3	2.0782	3	1.1214
	T	5	2.7367	5	2.2564	2	2.0544	4	1.2070	4	2.2664
5.0	5	4	2.6863	4	2.3075	3	1.9904	4	2.0471	2	1.1839
	T	4	2.8052	5	2.3533	3	2.2705	5	1.9253	2	1.0537

(b) Values of the EPI

		Number of contexts(P)									
		2		3		4		5		6	
m	S	c	EPI	c	EPI	c	EPI	c	EPI	c	EPI
1.1	3	4	2.8022	3	2.0260	4	1.8079	5	1.6122	4	1.4178
	T	4	2.3921	2	2.2367	5	1.8943	4	1.8404	5	2.2027
1.5	6	2	2.3308	2	1.7202	3	1.5160	4	1.3656	3	1.2894
	T	5	2.2033	5	2.2233	4	1.5435	4	1.7459	5	1.6778
2.0	3	2	2.8527	4	1.7847	3	1.6618	3	1.3054	4	**0.8546**
	T	3	2.4210	2	2.2719	5	1.7639	3	1.6108	2	1.1779
2.5	5	5	2.5618	2	1.9484	4	1.9141	5	2.1685	5	1.1440
	T	5	2.2900	3	1.9967	2	1.3764	2	1.3148	5	2.0363
3.0	6	2	2.5096	5	2.0699	3	1.8708	5	1.9000	2	1.1999
	T	4	2.6451	4	2.0665	3	1.9924	3	2.0374	5	1.6878
3.5	1	2	3.5317	4	2.7818	3	2.3920	3	1.9224	5	1.6638
	T	3	2.4320	5	2.4184	3	2.3326	2	1.6054	5	2.3847
4.0	4	4	2.8837	2	2.3397	3	1.8335	3	2.6757	5	1.0988
	T	3	2.7412	2	2.6640	3	2.3714	5	2.2116	3	2.8740
4.5	5	5	2.4915	5	2.4892	2	2.1411	3	2.0867	3	1.2377
	T	5	2.7124	5	2.5314	2	2.3759	4	2.4602	4	2.5268
5.0	5	4	2.9412	4	2.8775	3	2.5412	4	2.4538	2	2.1146
	T	4	2.6940	5	2.1600	3	2.4039	5	2.8817	2	1.5970

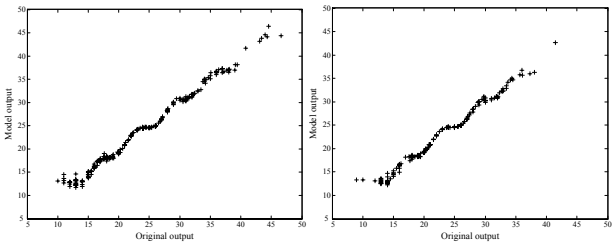
그림 6은 퍼지화 계수(m)가 2.0일 때, 선택된 입력변수들의 context의 수에 최고의 성능 지수만을 나타낸 것이다.



(a) Training data (b) Testing data

그림 6 Context 수의 증가에 따른 성능지수의 변화
Fig. 6 Performance index regarded as a function of the number of contexts

그림 7은 PI=0.8314이고 EPI=0.8546일 때 원 데이터의 출력과 모델의 출력과의 근사화와 일반화 능력의 선형성을 나타낸 그림이다.



(a) Training data (b) Testing data

그림 7 모델의 근사화 및 일반화 능력의 선형성
Fig. 7 Scatter plots of the output of the network vs. data

표 5는 MPG 데이터를 이용한 기존의 지능형 모델들과의 성능을 비교 분석한 것이다.

표 5 기존의 모델들과의 성능 비교
Table 5 Result of comparative analysis

Model	m	S	P	c	PI	EPI	R	비고
Linear regression model					3.383 ±0.194	3.472 ±0.295		RMSE
Incremental model[10]		T	6	6	2.390 ±0.142	3.060 ±0.285	36	
Linguistic model[26]	2.0	T	6	6	2.86 ±0.83	3.14 ±0.98	36	
FRBFNN[27]	2.0	T			2.905 ±0.441	12.024± 1.897	≤18	MSE
	2.75				5.804 ±0.394	7.936 ±1.100	≤33	
Proposed model	1.1	3	6	4	1.3385	1.4178	24	RMSE
		T	5	4	1.2994	1.8404	20	
	1.5	6	6	3	1.0138	1.2894	18	
		T	6	5	1.0185	1.6778	30	
	2.0	3	6	4	0.8314	0.8546	24	
		T	6	2	1.3512	1.1779	12	
	2.5	5	6	5	0.8831	1.1440	30	
		T	5	2	1.6745	1.3148	10	
	3.0	6	6	2	1.0031	1.1999	12	
		T	6	5	1.5891	1.6878	30	
	3.5	1	6	5	1.9138	1.6638	30	
		T	5	2	1.8483	1.6504	10	
	4.0	4	6	5	1.0632	1.0988	30	
		T	5	5	1.6303	2.2116	25	
	4.5	5	6	3	1.1214	1.2377	18	
		T	5	4	1.2070	2.4602	20	
	5.0	5	6	2	1.1839	2.1146	12	
		T	6	2	1.0537	1.5970	12	

7.3 가스 터빈 발전소의 NOx 배출 메커니즘 data

본 장에서는 화력발전소의 가스 터빈에서 발생되는 독성이 강한 NOx의 배출 메카니즘에 제안된 알고리즘을 도입하여 배출 패턴 모델을 구축하고 평가한다. 총 5입력(Tamb, COM, LPT, Pcd, Texh) 1출력(NOx)로 구성된 총 260개의 데이터 중, 학습 데이터와 테스트 데이터를 60%-40%로 랜덤하게 나누어 제안된 모델의 성능을 평가 하였다.

NOx 배출공정 데이터의 경우, 퍼지화 계수(m)가 2.0이고 최적의 성능은 PI=0.0430, EPI=0.0604로써 각각의 context에 선택된 입력변수의 수는 3개이고 이때 context의 수는 6개이고 각각의 context에 해당된 클러스터의 수는 2개이다. 각 context에 선택된 입력변수들은 모두 {1, 2, 3, 4} 입력들이 선택되었으며 이는 각각의 은닉층에서 사용된 감소된 입력변수들이 모든 입력변수들을 사용해서 얻는 경우에 비해서 낮은 성능지수를 갖는데 유익함을 강조할 수 있다.

표 6은 기존의 지능형 모델들의 성능과 본 논문에서 제안한 모델들의 성능을 비교 분석한 것이다.

표 6 선택된 모델들의 비교 분석

Table 6 Comparative analysis of selected models

Model	m	S	P	c	PI	EPI	R	비고
Linear regression model					4.1796 ±0.121	4.4652 ±0.191		RMSE
Fuzzy set-based FNN[24]		T	Simpli		6.269	8.778	30	MSE
			Linear		3.725	5.291	30	
Fuzzy relation-based FNN[25]		T	Simple		0.700	1.649	32	
			Linear		0.080	0.190	32	
Multi-FNN [28]			Linear		0.720	2.910	30	
Proposed model	1.1	3	6	3	1.2838	1.3027	18	RMSE
		T	6	4	0.1385	0.2401	24	
	1.5	4	6	3	0.1727	0.1800	18	
		T	5	4	0.1563	0.2084	20	
	2.0	4	6	2	0.0430	0.0604	12	
		T	6	4	0.0472	0.0617	24	
	2.5	4	6	5	0.0593	0.0636	30	
		T	6	3	0.0645	0.0729	18	
	3.0	4	6	5	0.4273	0.3453	30	
		T	6	2	0.1606	0.2169	12	
	3.5	4	6	3	0.4440	0.6763	18	
		T	6	3	0.4389	0.9800	18	
	4.0	4	6	5	0.5085	0.9358	30	
		T	6	5	0.5573	0.7482	30	
	4.5	4	6	2	0.7828	0.8868	12	
		T	6	3	0.7804	1.0818	18	
	5.0	4	6	5	0.7243	1.5722	30	
		T	6	2	1.0013	1.3033	12	

8. 결 론

본 논문에서는 주어진 정보 입자화의 효율적인 처리를 이용하여 정보 유사성 기반 입자화 중심 방사형 기저 함수 신경회로망에 대해서 새로운 구조와 설계 방법론에 대해서 제안하고 폭넓은 성능평가를 하였다. 제안된 모델의 설계 관점에서 가장 큰 특징은 (a) 정보 입자화에 대한 유사성을 측정하여 조절 가능한 입력변수들은 효과적으로 차원성 감소에 대한 문제점을 해결함으로써 네트워크의 성능을 개선하고 또한 모델의 증가된 차원성으로 인한 반복적인 학습 절차를 처리할 때 요구되어지는 많은 계산적 시간과 복잡성을 개선할 수가 있었다. (b) 모델의 설계에 있어서 K-means 클러스터링을 통한 출력공간의 입자화와 context-based FCM 클러스터링을 통한 입력공간의 입자화를 통하여 효율적인 네트워크를 설계할 수 있었다는 것이다.

감사의 글

이 논문은 2008년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임[KRF-2008-359-D00007] 그리고 본 연구는 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음 [GRRC 수원2009-B2, U-city 보안감시 기술협력센터]

참 고 문 헌

[1] J. A. Cumming and D. A. Wooff, "Dimension reduction via principal variables," Computational Statistics & Data Analysis, Vol. 52, pp. 550-565, 2007.

[2] S. H. Huang, "Dimensionality reduction in automatic knowledge acquisition: a simple greedy search approach," IEEE Trans. Knowledge and Data

[3] O. Buchtala, M. Klimek, and B. Sick, "Evolutionary optimization of radial basis function classifiers for data mining applications," IEEE Trans. SMC-B, Vol. 35, No. 5, pp. 928-947, 2005.

[4] A. Alexandridis, P. Patrinos, H. Sarimveis, and G. Tsekouras, "A two-stage evolutionary algorithm for variable selection in the development of RBF neural network models," Chemometrics Intell. Lab. Syst., Vol. 75, pp. 149-162, 2005.

[5] J. Park and I. Sandberg, "Universal approximation using radial-basis function networks," Neural Comput., Vol. 3, pp. 246-257, 1991.

[6] M. J. Er, W. Chen, and S. Wu, "High-speed face recognition based on discrete cosine transform and RBF neural networks," IEEE Trans. Neural Networks, Vol. 16, No. 3, pp. 679-691, 2005.

[7] N. Xie and H. Leung, "Blind equalization using a predictive radial basis function neural network," IEEE Trans. Neural Networks, Vol. 16, No. 3, pp. 709-720, 2005.

[8] W. Pedrycz, "Conditional fuzzy clustering in the design of radial basis function neural networks," IEEE Trans. Neural Networks, Vol. 9, No. 4, pp. 601-612, 1998.

[9] W. Pedrycz, "Conditional fuzzy c-means," Pattern Recognition Letter, Vol. 17, pp. 625-631, 1996.

[10] W. Pedrycz and K. C. Kwak, "The development of incremental models," IEEE Trans. Fuzzy Systems, Vol. 15, No. 3, pp. 507-518, 2007.

[11] J. H. Holland, Adaptation in Natural and Artificial Systems, The University of Michigan Press, Ann Arbour, 1975.

[12] K. D. Jong, Are genetic algorithms function optimizers?, In Proc. of PPSN II (Parallel Problem Solving from Nature), Amsterdam, North Holland, 1992.

[13] W. Pedrycz, "Fuzzy clustering with a knowledge-based guidance," Pattern Recognition Letter, Vol. 25, pp. 469-480, 2004.

[14] W. Pedrycz, V. Loia, and S. Senatore, "P-FCM: a proximity-based fuzzy clustering," Fuzzy Sets and Systems, Vol. 148, pp. 21-41, 2004.

[15] V. Loia, W. Pedrycz, and S. Senatore, "P-FCM: a proximity-based fuzzy clustering for user-centered web applications," Int. J. Approximate Reasoning, Vol. 34, pp. 121-144, 2003.

[16] P. Singla, K. Subbarao, and J.L. Junkins, "Direction-dependent learning approach for radial basis function networks," IEEE Trans. Neural Networks, Vol. 18, No. 1, pp. 203-222, 2007.
[17] K. Z. Mao and G. B. Huang, "Neuron selection for RBF neural network classifier based on data structure preserving criterion," IEEE Trans. Neural Networks, Vol. 16, No. 6, pp. 1531-1540, 2005.
[18] D. S. Yeung, W. W. Y. Ng, D.Wang, E. C. C. Tsang, and X. Z. Wang, "Localized generalization error model and its application to architecture selection for radial basis function neural network," IEEE Trans. Neural Networks, Vol. 18, No. 5, pp. 1294-1305, 2007.
[19] X. Hong, "A fast identification algorithm for Box - Cox transformation based radial basis function neural network," IEEE Trans. Neural Networks, Vol. 17, No.4, pp. 1064-1069, 2006.
[20] S. P. Lloyd, "Least squares quantization in PCM," IEEE Trans. Information Theory, Vol. 2, pp. 129-137, 1982.
[21] H. S. Park, W. Pedrycz, and S. K. Oh, "Evolutionary design of hybrid self-organizing fuzzy polynomial neural networks with the aid of information granulation," Expert Systems with Applications, Vol. 33, pp. 830-846, 2007.
[22] 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 Systems, Vol. 115, No. 2, pp. 205-230, 2000.
[23] S. K. Oh, W. Pedrycz and B. J. Park, "Hybrid Identification of Fuzzy Rule-Based Models," Int. J. of Intelligent Systems, Vol. 17, No.1, pp. 77-103, Jan. 2002
[24] S. K. Oh, W. Pedrycz and H. S. Park, "Hybrid Identification in Fuzzy-Neural Networks," Fuzzy Sets & Systems, Vol. 138, pp. 399-426, 2003.
[25] H. S. Park and S. K. Oh, "Fuzzy Relation-Based Fuzzy Neural-Networks Using a Hybrid Identification Algorithm," International Journal of Control, Automations, and Systems, Vol. 1, No. 3, pp. 289-300, 2003.
[26] W. Pedrycz and K. C. Kwak, "Linguistic Models as a Framework of User-Centric System Modeling," IEEE Trans. SMC-A, Vol. 36, No. 4, pp. 727-745, 2006.
[27] W. Pedrycz, H. S. Park, and S. K. Oh, "A granular-oriented development of functional radial basis function neural networks," Vol. 72, pp. 420-435, 2008.
[28] H. S. Park and S. K. Oh, "Multi-FNN Identification Based on HCM Clustering and Evolutionary Fuzzy Granulation," International Journal of Control, Automations, and Systems, Vol. 1, No. 2, pp. 194-2002, 2003.

저 자 소 개

박 호 성 (朴 浩 成)

1999년 원광대학교 제어계측공학과 졸업. 동 대학원 석사(2001), 박사(2005). 2005-2006년 원광대학교 전기전자 및 정보공학부 전임강사, 2006-2007년 캐나다 Alberta 대학 전기 및 컴퓨터공학과 Post-Doc. 2008~현재 수원대학교 산업기술연구소 학술연구교수. 관심분야는 퍼지추론시스템, 뉴럴 네트워크, 퍼지-뉴럴네트워크, 진화론적 최적화 알고리즘, Granular computing, Neuro-fuzzy computing, Evolutionary computing등의 Computational Intelligence등.
Tel : 031-222-6544
Fax : 031-220-2667
E-mail : parkhs@suwon.ac.kr

오 성 권 (吳 聖 權)

1981년 연세대학교 전기공학과 졸업, 동 대학원 석사(1983), 박사(1993). 1983-1989년 금성산전연구소(선임연구원). 1996-1997년 캐나다 Manitoba 대학 전기 및 컴퓨터공학과 Post-Doc. 1993-2004년 원광대학교 전기전자 및 정보공학부 교수. 2005~현재 수원대학교 전기공학과 교수, 2002~현재 대한전기학회, 제어로봇시스템학회, 퍼지및지능시스템학회 편집위원. 관심분야는 퍼지 시스템, 퍼지-뉴럴 네트워크, 자동화 시스템, 고급 computational intelligence, 지능 제어등.
Tel : 031-229-8162
Fax : 031-220-2667
E-mail : ohsk@suwon.ac.kr

김 현 기 (金 賢 起)

1977년 연세대학교 전기공학과 졸업, 동 대학원 석사(1985), 박사(1991). 1997~현재 경기지역 산학연 기술지원단 단장, 1998~현재 산업기술정책연구소 심의위원, 2000~현재 사단법인 산학연 컨소시엄 전국협의회 회장, 1989~현재 수원대학교 전기공학과 교수, 관심분야는 시스템 자동화 및 감시제어, 지능형 모델링 및 제어등.
Tel : 031-220-2666
Fax : 031-220-2667
E-mail : hkkim@suwon.ac.kr