

Reconstructability criterion을 통한 granular-based RBF NN의 최적화

박호성\*, 오성권\*\*

수원대학교 산업기술연구소\*, 수원대학교 전기공학과\*\*

Optimization of granular-based RBF NN with the aid of reconstructability criterion

Ho-Sung Park\*, Sung-Kwon Oh\*\*

Industry Administration Institute, University of Suwon\*, Dept. Electrical Engineering, University of Suwon\*\*

**Abstract** - 본 논문에서는 주어진 데이터의 입자화 특성을 효과적으로 모델 구축에 반영하고자 재구성 평가 기준을 통한 새로운 형태의 입자화 기반 RBF 뉴럴 네트워크를 개발한다. 주어진 데이터들의 입자화 특성을 파악하기 위해서 새로운 형태의 FCM 클러스터링(-Context-based fuzzy clustering)을 이용한다. 즉, 출력 공간의 입자화 특성은 K-means clustering 방법을 사용한 것에 반해, 입력 공간에서의 정보들은 Context-based fuzzy clustering 방법을 이용하여 효율적으로 데이터의 특성을 파악하여 모델의 구축에 반영하였으며, 또한 모델의 최적화를 위하여 RBF 뉴럴 네트워크의 은닉층의 수를 재구성 평가 기준을 통하여 모델의 최적화를 꾀하였다. 제안된 모델의 효율적인 특성을 보여주기 위해 저차원 합성 데이터를 이용하여 모델을 평가한다.

1. 서 론

방사형 기저 함수들로 이루어져 있는 RBF 뉴럴 네트워크는 지능형 시스템의 중요한 연구 분야중의 하나로 자리매김하고 있으며 다수의 비교적인 연구와 다양한 적용이 이루어지고 있다[1, 2]. 전형적인 형태의 기저함수로써 가우시안 함수를 이용한 RBF 뉴럴 네트워크가 다양한 분야에서 사용되어져 왔으며, 특히 어떠한 주어진 실험적인 데이터를 기초로 해서 은닉층에서의 기저함수를 클러스터링 기법 사용을 통해 대안적인 방법으로 처리하는 기술도 널리 사용되어지고 있다[3], 또한 최적화 방법으로 경사 하강법 뿐만 아니라 진화론적 최적화를 포함한 전역 극소 기법들도 시도되어지고 있다[4].

본 연구에서는 정보 입자화의 세심한 처리를 통한 실험적 데이터의 효과적인 사용을 기반으로 하는 탁월한 데이터 유도 구조로서의 granular-based RBF 뉴럴 네트워크를 제안한다. 제안된 모델은 퍼지 클러스터링의 절차에 따라 출력 변수에 의해 구현된 퍼지 입자화를 사용하였으며, 재구성 평가 기준을 통한 입력공간의 특성과 신중하게 형성된 정보 입자화를 통한 입력 공간의 균형 잡힌 표현을 나타내도록 하는 각각의 context에 대한 클러스터의 조절 가능한 수를 고려하였다. 또한 구조적 개발의 다양성을 입증하기 위하여 본 논문에서는 제안된 구조의 장점과 한계를 정량화하기 위하여 저차원 합성 데이터를 실험적 데이터로 이용하였다.

2. Receptive field 형성을 위한 granular computing

2.1 K-means clustering

K-means 클러스터링[5]은 영상과 음성 데이터의 압축과 시스템 모델링을 위한 데이터 전처리와 같은 다양한 영역에서 적용되어져 왔으며, 이는 데이터의 구성과 분류뿐만 아니라 데이터의 압축과 모델의 동정에도 널리 사용되어져 왔다.

본 논문에서는 K-means 클러스터링을 이용하여 출력변수의 {target<sub>k</sub>}을 P-2개로 분류하고 다음 식 (1)과 같이 목적함수가 최소가 되도록 각각의 클러스터의 중심값을 찾는다.

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

또한 출력변수의 최소 및 최대값을 첫 번째와 마지막 클러스터의 중심값으로 설정한다. 따라서 K-means 클러스터링 알고리즘을 이용하여 전체적으로 P개의 중심이 생성된다. 이와 같이 얻어진 중심값들을 점침도를 1/2이 되도록 삼각형 멤버쉽 함수를 통하여 멤버쉽 값을 구한다.

2.2 Context-based fuzzy clustering

Context-based FCM 클러스터링 [3, 6, 7]은 출력공간에서 몇 개의 미리 정의된 context라 불리는 퍼지 집합의 분류들을 기준으로 입력 데이터에서의 공간을 정보 입자화의 특성에 맞게 분류하는 알고리즘이다. 이는 입력공간과 출력공간 영역 사이의 관계를 드러나게 함으로써 기준

의 입력만을 고려한 클러스터링 방법에서 벗어나 입력과 출력을 동시에 고려한 특징을 가지고 있다.

Context-based FCM 클러스터링의 초기 조건은 다음과 같다.

$$U(F_j) = \left\{ u_{ik} \in [0, 1], \sum_{i=1}^c u_{ik} = f_{jk} \forall k, 0 < \sum_{k=1}^N u_{ik} < N \forall i \right\} \tag{2}$$

여기서, f<sub>jk</sub>는 K-means 클러스터링을 통해서 구해진 출력 공간에서의 멤버쉽 함수 값을 의미한다.

Context-based FCM 클러스터링의 목적함수는 다음과 같다.

$$V = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \| \mathbf{x}_k - \mathbf{z}_i \|^2 \tag{3}$$

여기서, m>1은 퍼지 계수를 나타내고 z<sub>i</sub>는 입력변수들의 중심값을 의미한다.

Context-based FCM에 의해서 실행되는 목적함수 V의 최소화는 partition matrix와 중심값들의 반복적으로 갱신되는 것을 실현되며, partition matrix는 다음과 같이 실행된다.

$$u_{ik} = \frac{f_{jk}}{\sum_{i=1}^c \left( \frac{\| \mathbf{x}_k - \mathbf{z}_i \|^2}{\| \mathbf{x}_k - \mathbf{z}_i \|^2} \right)^{2/(m-1)}}, \quad i = 1, 2, \dots, c, \quad k = 1, 2, \dots, N \tag{4}$$

3. Receptive field 최적화를 위한 reconstructability criterion

기존의 context-based FCM 클러스터링 방법을 이용한 모델들[14, 23, 30]은 데이터의 서로 다른 구조 및 복잡성에도 불구하고 서로 다른 context에 똑같은 수의 클러스터를 적용하는 단점을 가지고 있다. 본 논문에서는 reconstructability criterion을 기반으로 각각의 context에 대한 클러스터의 수를 조절하여 각각의 context에서 필요할 때 까지 각각의 context에 대해서 클러스터의 수를 증가하는 것을 고려한다.

제안된 reconstructability criterion 알고리즘은 다음과 같다.

먼저 각각의 입력변수 x<sub>k</sub>에 대해서 다음과 같이 클러스터의 인덱스 (i)를 결정한다.

$$i = \arg \max u_{ik} \tag{5}$$

여기서 u<sub>ik</sub>는 식 (6)에서 주어진 모든 중심값들에 기초해서 FCM 클러스터링 방법에서 사용된 것과 같이 계산된다.

또한 식 (6)과 같이 표현하는 것과 같은 관계를 가지는 데이터의 부분 집합을 Ω<sub>j</sub>로 표기한다.

$$\Omega_j = \{ \mathbf{x}_k \mid i = \arg \max u_{ik} \text{ and } i \in L(j) \} \tag{6}$$

모든 x<sub>k</sub> ∈ Ω<sub>j</sub>는 j-번째 context에 의해서 포함된 중심값들을 기반으로 하여 일정한 정확성을 가지고 "재구성"될 수 있다. 즉, 재구성된 데이터 x̂<sub>k</sub>는 다음과 같이 계산된다.

$$\hat{\mathbf{x}}_k = \frac{\sum_{i \in L(j)} g_{ik}^m \mathbf{v}_i}{\sum_{i \in L(j)} g_{ik}^m} \tag{7}$$

여기서, 소속 정도  $g_{jk}$ 은  $j$ -번째 context에 의해 포함된 중심값들을 기초로 결정된다. 즉,

$$g_{jk} = \frac{1}{\sum_{i \in L(j)} \left( \frac{\| \mathbf{x}_k - \mathbf{v}_i \|}{\| \mathbf{x}_k - \mathbf{v}_j \|} \right)^{2/(m-1)}} \quad (8)$$

여기서,  $i \in L(j)$ 이고  $\mathbf{x}_k \in \Omega_j$ .

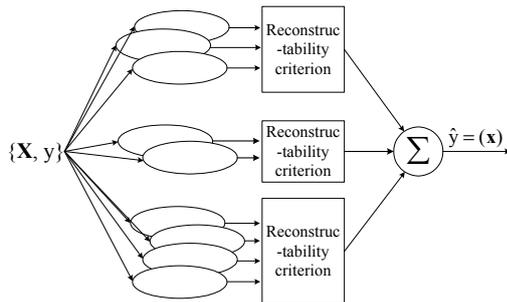
식 (9)에서 보듯이, 원 데이터  $\mathbf{x}_k$ 와 재구성된 데이터  $\hat{\mathbf{x}}_k$ 와의 weighted Euclidean distance 계산을 통해서 재구성의 특징인 평가 기준을 측정한다.

$$V(j) = \sum_{k \in \Omega_j} \| \mathbf{x}_k - \hat{\mathbf{x}}_k \|^2 \quad (9)$$

식 (10)에서 제안된 평가기준의 계산은 모든 context에 대해서 실행되면 이는 결과적으로  $V(1), V(2), \dots, V(P)$ 와 같은 결과값으로 생성된다. 여기서 우리는  $j_0$ -번째 context에서 발생한 가장 작은 값을 가진  $V$ -값을 선택하고 이를 기준으로써 간주한다. 그 후  $j_0$ -번째 context를 제외한 나머지 context에 대해서 클러스터의 수를 증가하여  $V(j)$ 의 값이  $V$ -값보다 작거나 또는 같을 때까지 클러스터의 수를 증가시키면서 context-based FCM 클러스터링 알고리즘을 계산한다. 단, 모델의 복잡성을 고려하여 본 논문에서는 클러스터의 수 증가에 대한 강제 제한 조건으로 '2c'를 한정하였는데 이는 클러스터의 수 증가가 기준인  $V$ -의 값을 가지는 클러스터의 수보다 2배가 초과하지 않도록 하기 위함이다.

#### 4. Granular-based RBF 뉴럴 네트워크

앞에서 언급한 알고리즘 및 구조를 기반으로 본 논문에서 제안한 granular-based RBF 뉴럴 네트워크의 구조는 다음과 같다.



〈그림 1〉 Granular-based RBF 뉴럴 네트워크의 구조

그림 1에서도 볼 수 있듯이 본 논문에서 제안한 모델은 각각의 context에 적당한 수의 클러스터를 조절함으로써 주어진 데이터의 정보 입자화를 모델에 효과적으로 반영하도록 하였다. 은닉층으로부터의 연결은 함수적인 특성을 지니고 있으며, 지역 모델은 식 (10)에서 볼 수 있듯이 입력과 출력에서 형성된 정보 입자화 주위에 선형 형태로 이루어져 있다.

$$h - y_i = \mathbf{a}_i^T (\mathbf{x} - \mathbf{v}_i) \quad (10)$$

제안된 모델의 전체 출력은 다음과 같이 계산된다.

$$y(\mathbf{x}) = \sum_{u_i \in L(1)} u_i(\mathbf{x}) [y_1 + \mathbf{a}_1^T (\mathbf{x} - \mathbf{v}_1)] + \sum_{i \in L(2)} u_i(\mathbf{x}) [y_2 + \mathbf{a}_i^T (\mathbf{x} - \mathbf{v}_i)] + \dots + \sum_{i \in L(P)} u_i(\mathbf{x}) [y_P + \mathbf{a}_i^T (\mathbf{x} - \mathbf{v}_i)] \quad (11)$$

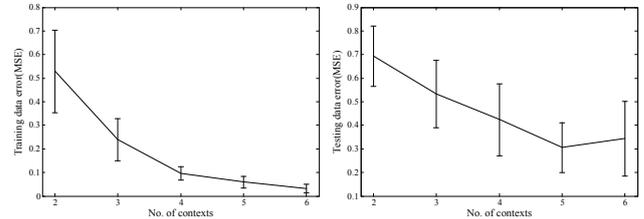
지역 모델들의 파라미터  $\mathbf{a}_i$ 의 최적화는 standard Mean Square Error (MSE)을 이용하여 실행하였다.

#### 5. 시뮬레이션 및 결과 고찰

제안된 설계 방법론의 해석에 대한 탁월한 능력의 우수성을 입증하고자 식 (12)에서 보여준 것과 같이 저차원 합성 데이터를 이용하여 총 200개의 입출력 쌍을 생성하여 학습 데이터와 (60%)과 테스트 데이터 (40%)로 분할하여 모델을 구축하였다. 또한 실험을 통해 발생된 결과들의 타당성 및 좀 더 높은 객관성을 제공하기 위해 10-fold cross validation 모드를 실행하였다.

$$y = (1 + x_1^{-2} + x_2^{-1.5} + x_3^{-1.0})^2, \quad 1 \leq x_1, x_2, x_3 \leq 5 \quad (12)$$

그림 2는 퍼지 계수(m)이 2.0인 경우에 context의 수에 따른 성능 지수 (평균값과 표준 오차) 값을 나타낸 것이다. 그림에서도 알 수 있듯이 학습 데이터의 성능지수는 context의 수가 증가할수록 성능이 개선될 뿐만 아니라 그 표준 오차도 작아지는데 반해 테스트 데이터에 성능지수는 학습 데이터의 성능지수와는 달리 context의 수가 증가한다고 해도 그 성능이 좋아지지 않으며 또한 큰 표준오차를 가짐을 알 수 있다.



〈그림 2〉 Context의 수에 따른 성능지수 : 각 context에 대한 클러스터의 초기 수 : 2

표 1은 제안된 모델의 성능과 linear regression model 및 표준적인 RBF 뉴럴 네트워크의 성능을 비교한 것이다.

〈표 1〉 선택된 모델의 성능과의 비교 분석

Model	No. of nodes	PI	EPI
Linear regression		0.4516 ± 0.0916	0.4954 ± 0.1516
RBFNN I	H=20	2.2125 ± 0.3718	3.1979 ± 1.0636
RBFNN II	H=20	0.5768 ± 0.1091	0.7617 ± 0.2814
Proposed model	m=2.0	H ≤ 20	0.0593 ± 0.0254
	m=2.5	H ≤ 20	0.3054 ± 0.1043
		0.1448 ± 0.0588	0.3502 ± 0.0824

#### 6. 결 론

본 논문에서는 reconstructability criterion을 이용하여 정보 입자화의 세심한 공정을 기반으로 한 Granular-based RBF 뉴럴 네트워크의 새로운 구조를 제안하였다. 본 논문에서 사용한 context-based Fuzzy C-Means는 출력 공간에서 정의된 퍼지 집합들을 입력 공간의 정보 입자화의 형성에 중요한 역할을 수행하고 실험 데이터들의 표현과 관련된 입자화의 효과를 정량화하는 reconstructability criterion을 통하여 각각의 context에 대해서 적합한 클러스터의 수를 선택할 수 있게 도와준다. 본 연구에서는 포괄적인 실험적 예제를 위해서 저차원 합성 데이터를 사용하여 었다. 또한 널리 사용되어진 RBF 구조들과 선형 회귀 방식성을 이용하여 제안된 모델의 성능을 정량적으로 평가하였다.

#### 감사의 글

이 논문은 2008년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임[KRF-2008-359-D00007]. 그리고 2009년도 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임[KRF-2009-0074928].

#### 참고 문헌

- [1] M. Han and J. Xi, "Efficient clustering of radial basis perceptron neural network for pattern recognition," Pattern Recognition, Vol. 37, pp. 2059-2067, 2004.
- [2] J. M. Vilaplana, J. L. P. Molina, and J. L. Coronado, "Hyper RBF model for accurate reaching in redundant robotic systems," Neurocomputing, Vol. 61, pp. 495-501, 2004.
- [3] 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.
- [4] D. Shi, D. S. Yeung, and J. Gao, "Sensitivity analysis applied to the construction of radial basis function networks," Neural Networks, Vol. 18, pp. 951-957, 2005.
- [5] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
- [6] W. Pedrycz, "Conditional fuzzy c-means," Pattern Recognition Letters, Vol. 17, pp. 625-631, 1996.
- [7] W. Pedrycz and K. C. Kwak, "The development of incremental models," IEEE Trans. Fuzzy Systems, Vol. 15, No. 3, pp. 507-518, 2007.