Journal of the Korean Institute of Industrial Engineers
Vol. 29, No. 2, pp. 172-178, June 2003.

보정된 K-medoids 군집화 기법과 이분 탐색기법을 이용한 RBF 네트워크의 중심 개수와 위치의 통합 결정

이대원 · 이재욱†

포항공과대학교 산업공학과

Determining the Number and the Locations of RBF Centers Using Enhanced K-Medoids Clustering and Bi-Section Search Method

Daewon Lee · Jaewook Lee

Department of Industrial Engineering, Pohang University of Science and Technology, Pohang, 790-784

In the recent researches, a variety of ways for determining the locations of RBF centers have been proposed assuming that the number of RBF centers is known. But they have also many numerical drawbacks. We propose a new method to overcome such drawbacks. The strength of our method is to determine the locations and the number of RBF centers at the same time without any assumption about the number of RBF centers. The proposed method consists of two phases. The first phase is to determine the number and the locations of RBF centers using bi-section search method and enhanced k-medoids clustering which overcomes drawbacks of clustering algorithm. In the second phase, network weights are computed and the design of RBF network is completed. This new method is applied to several benchmark data sets. Benchmark results show that the proposed method is competitive with the previously reported approaches for center selection.

Keywords: generalized RBF network, center selection, data partition, bi-section, k-medoids clustering

1. 연구 배경

RBF(Radial Basis Function) 네트워크는 하나의 은닉층(hidden layer)만을 가지는 네트워크의 구조적 간단성으로 인해 비선형 함수의 추정과 패턴 분류의 분야에서 널리 쓰이고 있다.

RBF 네트워크는 입력 데이터(training examples)를 네트워크와 연결하는 입력층 (input layer), 입력 데이터를 radial basis function을 통해 고차원의 공간으로 비선형 변환을 시키는 은닉층(hidden layer), 은닉층의 결과의 선형 조합을 통해 네트워크의 최종 출력값을 계산하는 출력층(output layer) 등의 세 개의 서로 다른 층으로 구성되어 있다.

이러한 RBF 네트워크의 학습은 은닉층의 뉴런(커널 함수 또는 RBF)의 중심의 결정과 은닉층과 출력층을 연결하는 weight의 추정의 두 가지로 구성되어 있다. 학습과정에서 일단 RBF 네트워크의 중심이 결정되어 고정된다면 네트워크의 weight는 바로 구할 수 있다. 이는 RBF 네트워크가 MLP(다중 퍼셉트론)와는 달리 은닉층이 하나이기 때문에 linear least square 알고리듬이나 pseudo-inverse를 이용할 수 있기 때문이다.

따라서 RBF 네트워크의 설계에 있어서 핵심적인 문제는 은닉층 뉴우런, 즉 RBF의 중심의 개수와 위치를 결정하는 것이다. 기존의 연구에서는 중심의 개수를 알고 있다는 가정하에 그 위치만을 결정하는 알고리듬들이 다양하게 제안되었는데, 대표적인 것으로는 입력 데이터 중에서 임의로 중심을 선택, 네트워크의 입력 공간에서 임의로 중심 추출 (Micchelli, 1986), 군집화 기법(clustering algorithm)을 이용한 중심 결정 (Powell, 1985), (Broomhead and Lowe, 1988) 등이 있다. 하지만, 이러한

본 연구는 포항공과대학교 기초과학연구(POSTECH BSRI research-1RB0311501), BK21과제의 지원에 의하여 수행되었음.
† 연락저자 : 이재욱 교수, 790-784 경북 포항시 남구 효자동 산 31번지 포항공과대학교 산업공학과, Fax : 054-279-2870,
e-mail : jaewookl@postech.ac.kr
2003년 3월 접수; 2003년 5월 수정본 접수; 2003년 5월 게재 확정.

기존의 중심 선택 방법론들은 중심의 개수와 위치를 통합적으로 고려하지 않고 있으며, 선택된 중심의 위치가 각 class별 데이터 각각의 분포를 제대로 반영하지 못하는 등 다양한 문제점을 지니고 있다.

본 논문에서는 입력 데이터를 목적 에러율(goal error rate) 이하로 분류할 수 있는 RBF 중심의 위치와 개수를 동시에 결정하는 알고리듬을 제안한다. 제안된 알고리듬의 기본적인 아이디어는 세 가지로 이루어졌다. 첫째, 입력 데이터의 분류 성능을 높이기 위해, 전체 입력 데이터의 분포를 반영하는 RBF 중심 대신에, 입력 데이터를 각각이 속한 class별로 분할하고 각각에 대해 class별 데이터의 분포를 반영하는 중심을 구하였다. 둘째, 기존의 알고리듬은 중심의 개수를 안다는 가정하에서 출발하였으나, 제안된 알고리듬은 이러한 가정이 필요 없이 Bi-Section 알고리듬을 이용해 중심 개수에 대한 검색범위(1~입력데이터 개수)를 줄여가며 목적 에러율에 도달하는 최소의 RBF 중심 개수를 효율적이면서도 자동적으로 결정하여 준다. 셋째, 기존의 군집분석 기법이 지닌 문제점(계산 시간, 입력 데이터의 class 정보의 미사용 등)들을 보완하여 최적의 중심의 위치를 결정하였다.

본 논문의 구성은 다음과 같다. 먼저 2절에서는 본 논문에서 다루게 될 Generalized RBF 네트워크의 기본적인 구조에 대해 알아보고 네트워크 설계에 있어서 핵심적인 문제 중에 하나인 RBF 중심 결정에 관한 기존의 연구들과 각각이 지닌 한계점에 대해 살펴본다. 3절에서는 RBF 중심의 개수와 위치의 통합 결정에 쓰인 기법들의 기본적인 아이디어에 대해 서술하고, 4절에서는 목적 에러율에 도달하는 RBF 네트워크의 최소의 중심 개수와 위치를 결정하는 새로운 알고리듬을 제안한다. 5절에서는 수치 예제들을 통해 본 연구 결과의 성능을 살펴보고 6절에서 결론을 맺는다.

2. RBF 네트워크의 기본구조와 중심 결정에 관한 기존의 연구

RBF 네트워크는 은닉층 뉴런의 개수, 즉 RBF 중심의 개수에 따라 interpolation에 이용되는 RBF 네트워크와 Generalized RBF 네트워크로 나누어 진다. 본 논문에서 제안하는 알고리듬은 overfitting을 방지하고 네트워크의 복잡성을 극복한 Generalized RBF 네트워크에 대해 다룬다.

2.1 Generalized RBF 네트워크의 구조

전통적인 RBF 네트워크는 입력 데이터 모두를 RBF의 중심으로 이용하는 interpolation 문제에 이용되어 왔다. 이러한 RBF 네트워크의 weight를 계산하기 위해서는 $N\times N$ Matrix의 역행렬을 구해야 한다(여기서 N은 입력 데이터의 개수). 이는 N의 값이 커짐에 따라 계산은 N^3에 비례해 복잡해짐을 알 수 있다. 따라서 입력 데이터 모두를 RBF 중심으로 이용함으로써 유발되는 네트워크의 복잡성을 극복하기 위해 입력 데이터보다 더 적은 개수의 중심을 사용하는 Generalized RBF 네트워크가 제안되었다. Generalized RBF 네트워크의 구조는 <그림 1>과 같다.

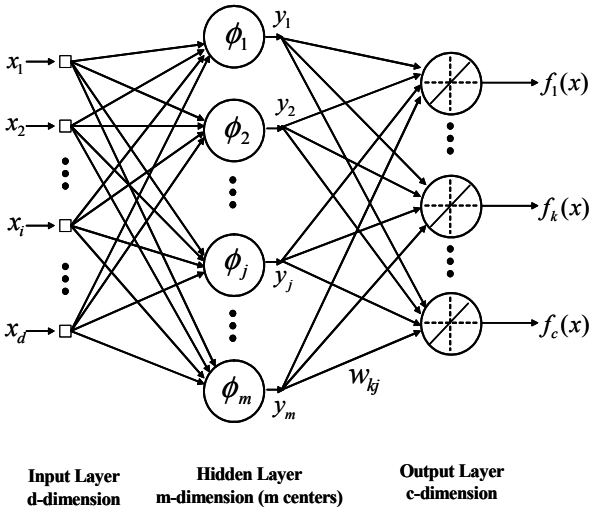

그림 1. Generalized RBF 네트워크의 구조.

이 방법은 기존의 interpolation에 이용되던 네트워크보다 더 낮은 차원의 공간에서 interpolation 문제의 해를 추정한 suboptimal 해를 구하는 방법이다. 추정해인 $F^*(\mathbf{x})$는 다음과 같이 나타낼 수 있다(Poggio and Girosi, 1990).

$$F^*(\mathbf{x}) = \sum_{i=1}^{m} w_i \phi\left(\left\|\mathbf{x}-\mathbf{t}_i\right\|\right) \tag{1}$$

여기서 $\phi(\cdot)$는 RBF를 의미하며 $\mathbf{t}_i$는 RBF의 중심을 나타내고 $\mathbf{t}_i$의 개수는 입력 데이터의 개수 N보다 작은 m개이다. 본 논문에서는 RBF로 아래와 같은 다변량 가우스 함수를 이용하였다.

$$\phi(\mathbf{x},\mathbf{x}_i) = \exp\left(-\frac{1}{2\sigma_i^2}\left\|\mathbf{x}-\mathbf{x}_i\right\|^2\right)$$

여기서 $\mathbf{x}_i$는 함수의 중심이고 σ_i는 함수의 분산이다. 이를 통해 새로운 cost function인 식 (1)을 최소화하는 weight 벡터 $\mathbf{w}$는 식 (2)와 같다(Broomhead and Lowe, 1988).

$$\mathbf{w} = (\mathbf{\Phi}^T\mathbf{\Phi} + \lambda\mathbf{\Phi}_0)^{-1}\mathbf{\Phi}^T\mathbf{d} \tag{2}$$

여기서

$$\mathbf{d} = [d_1, d_2, \cdots, d_N]^T$$

$$\mathbf{w} = [w_1, w_2, \cdots, w_{m_1}]^T$$

$$\mathbf{\Phi} = \begin{bmatrix} \phi(\mathbf{x}_1,\mathbf{t}_1) & \phi(\mathbf{x}_1,\mathbf{t}_2) & \cdots & \phi(\mathbf{x}_1,\mathbf{t}_m) \\ \phi(\mathbf{x}_2,\mathbf{t}_1) & \phi(\mathbf{x}_2,\mathbf{t}_2) & \cdots & \phi(\mathbf{x}_2,\mathbf{t}_m) \\ \vdots & \vdots & & \vdots \\ \phi(\mathbf{x}_N,\mathbf{t}_1) & \phi(\mathbf{x}_N,\mathbf{t}_2) & \cdots & \phi(\mathbf{x}_N,\mathbf{t}_m) \end{bmatrix}$$

$$\mathbf{\Phi}_0 = \begin{bmatrix} \phi(\mathbf{t}_1,\mathbf{t}_1) & \phi(\mathbf{t}_1,\mathbf{t}_2) & \cdots & \phi(\mathbf{t}_1,\mathbf{t}_m) \\ \phi(\mathbf{t}_2,\mathbf{t}_1) & \phi(\mathbf{t}_2,\mathbf{t}_2) & \cdots & \phi(\mathbf{t}_2,\mathbf{t}_m) \\ \vdots & \vdots & & \vdots \\ \phi(\mathbf{t}_m,\mathbf{t}_1) & \phi(\mathbf{t}_m,\mathbf{t}_2) & \cdots & \phi(\mathbf{t}_m,\mathbf{t}_m) \end{bmatrix}$$

이다. 즉, Generalized RBF 네트워크에서 $\phi(\mathbf{x},\mathbf{x}_i)$ 를 성분으로 갖는 행렬 $\mathbf{\Phi}$는 더 이상 대칭이 아닌 $N\times m$ 행렬임을 알 수 있다.

2.2 RBF 중심 결정에 관한 기존의 연구 및 한계점

Generalized RBF 네트워크는 입력 데이터보다 더 작은 개수의 중심을 가지기 때문에 RBF의 중심의 개수($m<N$)와 위치를 결정하는 것이 네트워크의 설계에 있어서 중요한 요소 중에 하나이다. 일단, RBF의 중심이 결정되면 행렬 $\mathbf{\Phi}$ 가 구해지고 최적의 weight는 식 (2)와 같이 이를 pseudo-inverse를 취함으로써 바로 구할 수 있게 된다. RBF의 중심 개수 결정에 관한 기존의 연구는 미비한 실정이고 개수가 주어졌다는 가정하에 중심의 위치를 결정하는 연구는 크게 아래와 같은 3가지의 방법이 있다.

1) **임의로 중심 선택:** 입력 공간에서 임의로 중심을 추출하거나 입력 데이터 중에 임의로 중심을 선택하는 것으로 가장 간단한 방법 중에 하나이다(Lowe, 1989). 그러나 임의로 선택한 중심의 위치는 최적이 아니므로 목적 에러율을 가지기 위한 중심의 개수가 실제 필요한 것보다 많이 쓰이게 되는 단점이 있다. 또한 중심의 개수에 대해 알고 있다는 가정을 하고 있으므로 적절한 수준의 중심의 개수를 찾는 것 또한 어려움이 있다.
2) **군집화 기법을 이용한 중심 선택:** 이 방법은 입력 데이터들을 K-means, K-medoids, SOM, VQ 등과 같은 군집화 기법으로 나누고 이들 군집의 중심을 RBF의 중심으로 이용하는 방법이다 (Duda and Hart, 1973). 이 방법은 군집의 수, 즉 RBF 중심의 수를 알고 있다는 가정을 하지만 실제로 군집의 개수를 결정하는 것이 어렵다. 또한 supervised classification 문제에서 각각의 입력 데이터는 자신이 속한 class가 어디인지에 대한 정보를 알고 있으나 실제 군집화 과정에서는 이에 대한 정보를 반영하지 않으므로 class 별 데이터의 분포를 잘 반영하지 못하는 단점이 있다.
3) **그래디언트(gradient)에 기반을 둔 중심 선택:** 이 방법은 RBF 네트워크에 관련된 모든 매개 변수(가우시안 커널의 분산, 중심의 위치, 네트워크의 weight)를 그래디언트 디센트(gradient-descent) 알고리듬을 이용하여 학습시키는 것이다 (Wettschereck and Dietterich, 1992). 중심의 위치에 관한 학습 식은 아래와 같다.

$$\mathbf{t}_i(n+1) = \mathbf{t}_i(n) - \eta \frac{\partial e(n)}{\partial t_i(n)} \qquad i = 1,2,\ldots,m$$

그러나 이 방법 또한 중심의 개수를 알고 있다는 가정하에 있고 $\frac{\partial e(n)}{\partial t_i(n)}$ 또한 구현하기가 어려울 뿐만 아니라 많은 local minima가 존재해 최적의 위치를 찾기가 어렵다.

3. 새로운 RBF 중심 선택을 위한 기본적인 아이디어

앞 절에서 살펴본 바와 같이 RBF 네트워크는 중심의 위치와 개수만 결정되면 네트워크의 weight를 구할 수 있다. 본 논문에서는 중심의 개수가 결정되어 있다는 가정 없이 중심의 개수와 위치를 통합적으로 결정하는 알고리듬을 제안한다. 이 알고리듬의 기본적인 아이디어는 다음의 세 가지로 나눌 수 있다.

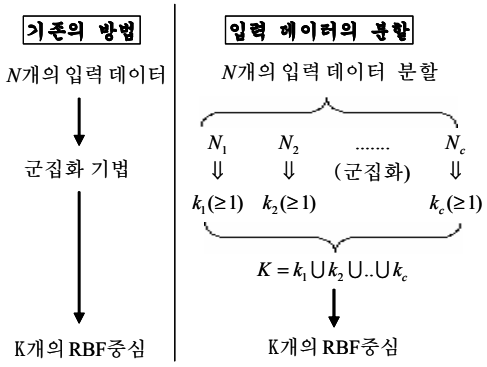

그림 2. 입력 데이터 분할 후 군집화.

3.1 입력 데이터의 분할

기존의 RBF 중심의 위치 결정에 있어서 가장 많이 이용되는 것이 군집화 기법이다. 본 논문 역시 기본적으로 군집화 기법에 기반을 두어 입력 데이터의 분포를 가장 잘 반영하는 RBF의 중심을 구한다. 그러나 분류문제(classification problem)에 있어서 군집화 기법을 이용하여 RBF 중심의 위치를 구하는 기존의 방법들은 각 입력 데이터의 class에 대한 정보를 이용하지 않고 전체 데이터를 unsupervised 데이터로 보고 분석한다. 따라서 기존의 방법은 데이터 전체의 분포를 반영하는 중심을 구하고 있다. 하지만, RBF 네트워크는 supervised 분석

(출력값의 답을 알고 있는 문제)이므로 입력 데이터가 어느 class에 속했는지에 대한 정보를 가지고 있고, 이를 이용하는 것이 중요하다. 본 연구에서는 먼저 N개의 입력 데이터를 class (output label 정보이용) 별로 분할한 뒤, 그 이후의 분할된 데이터 각각에 군집화 기법을 적용하여, 각각의 class별 데이터의 분포를 가장 잘 반영하는 군집의 중심들(1개 이상)의 집합을 c회(class 개수)에 걸쳐 따로 구한다. 이렇게 구한 군집의 중심들을 RBF의 중심으로 이용한다. <그림 2>는 기존의 방법과 본 논문에서 제안하는 입력 데이터 분할 방법의 간략한 진행 과정이다. 최적의 중심 위치 결정은 같은 중심의 개수라면 좀더 낮은 에러율을 가질 수 있고, 동일한 에러율을 가지기 위해서는 기존의 방법보다 좀더 작은 개수의 중심만 있으면 될 것이다. 따라서 class를 반영한 중심의 위치 결정은 최적 중심 위치 결정을 위한 중요한 사항이라 할 수 있다.

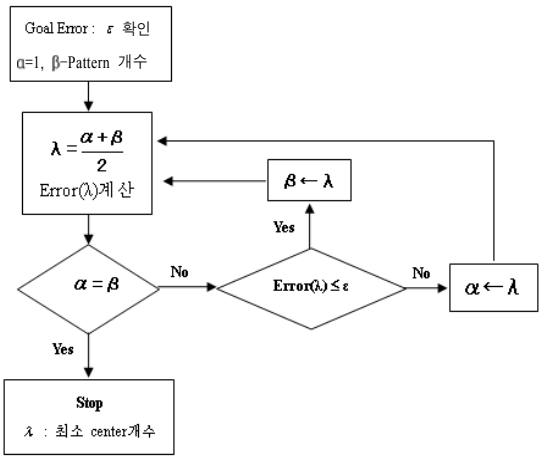

그림 3. Bi-section 알고리듬.

3.2 Bi-Section 알고리듬을 이용한 최소 중심 개수 결정

본 논문에서 제안하는 RBF 네트워크의 중심 개수 결정 방법의 아이디어는 다음과 같다. 은닉층의 차원인 RBF 중심의 개수를 크게 할수록 입력 데이터에 대한 에러율(training error)은 줄어든다(Cover, 1965). 따라서 이러한 에러 곡선은 중심 개수에 대해 단조감소함수이며 이러한 함수의 성질을 이용해 중심 개수의 검색 범위 (1개 ~ 입력 데이터 개수)를 줄여가며 목표 에러율에 도달하는 최소의 중심 개수를 찾을 수 있다. 본 논문에서 목표 에러율에 도달하는 최소 중심 개수를 찾기 위해 제안한 Bi-section 방법은 기존의 이진검색(binary search)과 개념이 비슷하다. 즉 중심 개수가 증가하면 에러율이 작아진다는 가정하에서 검색 공간을 반으로 줄여가며 목표 에러율 이하로 떨어지는 최소의 중심 개수를 찾아내는 것이다. 본 논문에서 제안한 Bi-section 알고리듬은 <그림 3>과 같다. Bi-section은 중심 개수의 결정 시간이 $O(\log_2 N)$에 비례하기 때문에 학습 데이터의 개수가 크더라도 비교적 안정적인 시간 안에 최소 중심 개수를 구할 수 있다는 장점이 있다. <그림 4>는 5절에서 다룰 Vowel 문제의 Bi-section 진행 과정을 나타낸다. <그림 4>를 보면 [1~528]의 검색 범위를 1/2씩 줄여가며 5번의 iteration 만에 목적 에러율(ε) 이하로 떨어지는 최소의 중심 개수에 수렴함을 알 수 있다. 따라서 Bi-section은 효율적($O(\log_2 N)$)으로 그리고 정확하게 검색 범위를 줄여나가 상대적으로 작은 trial만으로 목적 에러율(ε) 이하에 도달하는 최소의 중심 개수를 구할 수 있다.

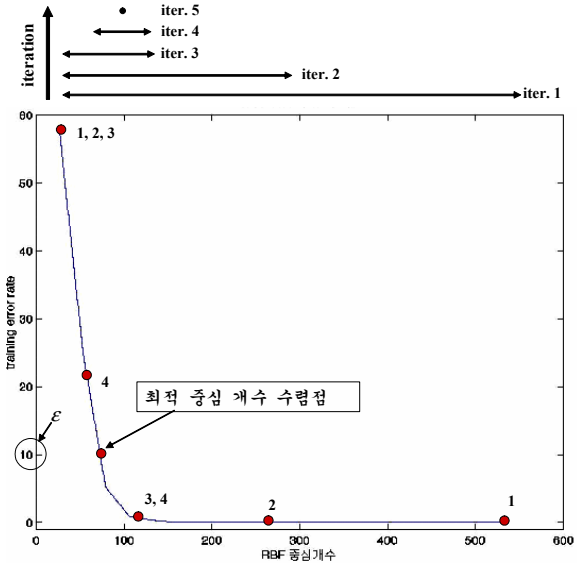

그림 4. Vowel 데이터에 대한 Bi-section 진행 과정.

3.3 중심 위치 결정을 위한 보정된 K-medoids 군집화 기법

본 논문에서는 기본적으로 군집화 기법을 이용하여 RBF 중심의 위치를 결정한다. 먼저 식 (3)과 같은 비용함수를 최소화하는 k개의 군집의 중심을 입력 공간에서 추출한다.

$$W(C) = \min_{C,\{m_k\}_1^K} \sum_{k=1}^{K} \sum_{C(i)=k} \|x_i - m_k\|^2 \qquad (3)$$

여기서 군집의 개수인 k는 3.2절에서 다룬 Bi-section 알고리듬을 이용하여 결정한다. 하지만 데이터의 분포가 Non-convex인 경우는 입력 데이터를 class별로 분할하였더라도 군집의 중심이 다른 class의 영역에 위치하는 경우가 생기게 된다. 따라서 입력 공간에서 추출한 군집의 중심 위치를 분할된 자신의 class의 데이터들 중 가장 가까운 값으로 보정한다. 이처럼 보정된 K-medoids 군집화 알고리듬은 다음과 같다.

알고리듬 I. 보정된 K-medoids 군집화 알고리듬

---

1. 입력 공간에서 임의로 초기 군집의 중심 $\{m_1, \cdots, m_k\}$를 추출한다.

2. 군집의 중심에 대해 식 (3)을 최소화하도록 입력 데이터를 군집화하고 각 군집의 데이터들의 평균값을 구한다.
3. 2에서 구한 평균값들을 새로운 군집의 중심으로 update하고 식 (4)와 같이 입력 데이터들을 가장 가까운 중심이 속한 군집으로 다시 할당한다.

$$C(i) = \arg\min_{1<k<K} \|x_i - m_k\|^2 \qquad (4)$$

4. 입력 데이터들의 군집이 변화가 없을 때까지 2~3단계를 반복한다.
5. 최종 추출된 k개의 군집 중심을 입력 데이터들 중 가장 가까운 값들로 수정한다.

4. 알고리듬

3절에서 제안한 기법들을 이용하여 목표 에러율 이하로 입력 데이터를 분류(classify)하는 Generalized RBF 네트워크의 최소 중심 개수와 위치를 구하는 알고리듬을 아래와 같이 제안한다. 알고리듬 II.의 순서도는<그림 5>와 같다.

알고리듬 II. RBF 네트워크의 중심 개수와 위치의 통합 결정 알고리듬

1. i=1에서 c까지 아래 과정 반복 (여기서 c는 출력 class의 종류)

   PHASE I : 각 class별 최적의 중심 개수와 위치 결정

   (1) 전체 입력 데이터를 i번째 class에 속한 것과 속하지 않은 것으로 두 부분으로 분할한다.

$$\text{training sample} : \{(x_1,d_1),(x_2,d_2),....,(x_N,d_N)\}$$
$$\Rightarrow \begin{cases} D_i \ : \{(x_1,d_1),...,(x_{n_1},d_{n_1})\} & \text{for } d \in \text{ class i} \\ D_{i^\neg} \ : \{(x_1,d_1),...,(x_{n_2},d_{n_2})\} & \text{for } d \notin \text{ class i} \end{cases}$$

   (2) 아래와 같이 sub-sample D_i 중에서 class i를 대표하는 RBF 중심의 위치와 개수를 결정한다[Bi-section method와 알고리듬 I의 군집기법 이용].

   ① i번째 class 데이터에 대한 목표에러 ε과 검색 범위 (search space) 범위 $[\alpha_0,\ \beta_0]$결정.

$$(\alpha_0 = 0,\ \beta_0 = n_1)$$

   ② 알고리듬 I의 군집기법을 이용하여 subsample D_i중 $\lambda_0\left(=\frac{\alpha_0+\beta_0}{2}\right)$개의 군집의 중심을 구한다.

   ③ λ_0개의 중심을 sub-sample D_i중 가장 가까운 점들로 보정한다. 이 점들을 C_{i0}라 한다.

   ④ C_{i0}를 RBFN의 중심으로 이용하여 전체 training sample에 대한 weight를 구한다.

$$\mathbf{W} = \left(\Phi^T\Phi\right)^{-1}\Phi^T\mathbf{d}$$

   ⑤ 위 weight를 이용하여 sub-sample D_i에 대한 에러 ε_0를 구한다(즉 class i에 속한 sample을 속하지 않았다고 classify할 비율).

   ⑥ 알고리듬 I의 군집기법은 군집 중심의 초기 위치에 따라 수렴 위치가 차이가 많이 나므로 1.2.2~1.2.5의 과정을 임의로 n회 반복(사용자가 지정)하여, ε_0가 가장 작게 나오는 C_{i0}를 최종 i번째 class의 λ_0개의 RBF 중심의 위치로 정한다.

   ⑦ 검색범위를 아래와 같이 반으로 줄인다.

$$[\alpha_0,\beta_0] \Rightarrow [\alpha_1,\beta_1] = \begin{cases} \alpha_1 = \alpha_0,\ \beta_1 = \lambda_0 & \text{if } \varepsilon_0 \le \varepsilon \\ \alpha_1 = \lambda_0,\ \beta_1 = \beta_0 & \text{if } \varepsilon_0 > \varepsilon \end{cases}$$

   ⑧ 검색범위가 수렴할 때까지 1.2.1~1.2.7의 과정을 반복

   ⑨ 이를 통해 목표 에러 ε에 도달하는 최소의 RBF 중심인 C_i^*를 구한다(즉 class i의 중심의 위치와 개수가 동시에 결정).

2. 1의 과정을 통해 각 class별로 RBF의 최적의 중심의 개수와 위치를 아래와 같이 구하였다.

   RBF 네트워크의 중심: $C = \{C_1^* \cup\} \ldots C_i^* \cdots \cup C_C^*$

   PHASE II: RBF 중심을 이용해 최적 네트워크 구성

3. 최종 RBF center C를 이용하여 전체 입력 데이터에 대한 Generalized RBF 네트워크를 구성하고 weight를 구한다(식 (2) 이용).

5. 수치예제 실험 결과

4절에서 제안한 RBF 네트워크의 중심 개수와 위치의 통합 결정 알고리듬의 성능을 평가하기 위해서 2-spiral, sonar, heart, vowel 등의 잘 알려진 4종류의 벤치마킹 분류 문제(classification problem)에 적용하였다. 수치예제에 대한 설명은 다음과 같다.

5.1 수치예제 설명

2-Spiral 데이터는 2차원 평면상에 놓인 2개의 서로 다른 소용돌이(spiral) 모양의 입력 데이터를 분류하는 문제이다. 2개

의 소용돌이는 원점을 출발점으로 하여 서로 엉켜있는 형태로 매우 어려운 분류문제 중 하나이다. **Sonar** 데이터는 신경망을 이용하여 sonar 신호를 분류한 Gorman과 Sejnowski의 연구에서 인용하였다. **Heart** 데이터는 심장병의 발생률과 나이, 성별 등의 13 종류의 사람의 특성 간의 관계를 보여준다. **Vowel** 데이터는 log area ratios를 통해 기록된 lpc의 입력값을 통해 British English의 11개의 모음(vowel)을 분류하는 문제이다. 위 수치예제의 속성은 <표 1>에서 정리하였다.

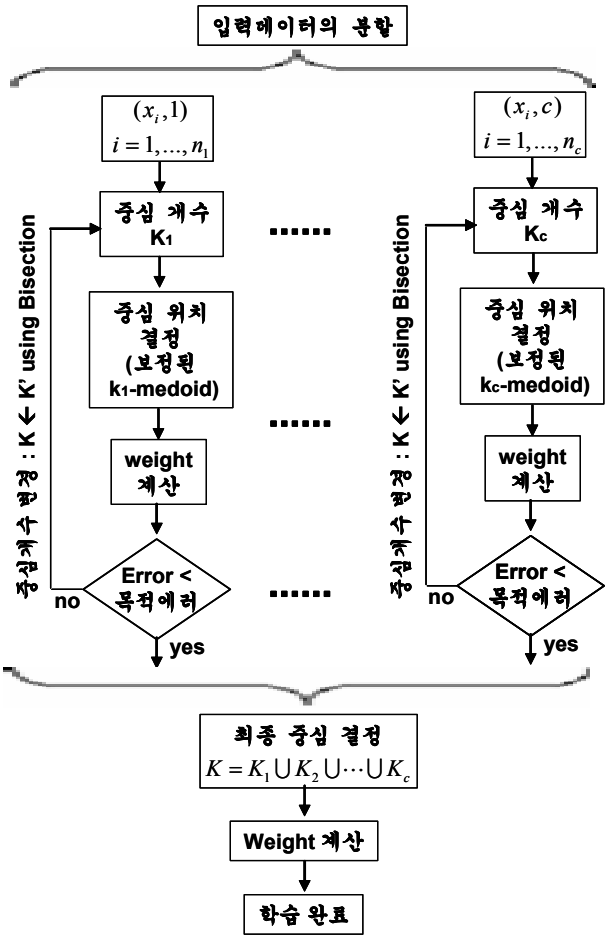

그림 5. 알고리듬 II의 순서도.

표 1. 수치예제 설명

	입력 차원	class 개수	입력 데이터 개수
2-Spirals	2	2	388 (194,194)
Sonar	60	2	104 (55,49)
Heart	13	2	180 (98,82)
Vowel	10	11	528 (48,48,48,48,48,48, 48,48,48,48,48)

*( )안은 각 class 별 입력 데이터의 개수

5.2 실험 결과

5.1절에서 설명한 4종류의 수치예제를 3종류의 알고리듬을 적용하여 실험하였다. Method I은 본 논문에서 제안한 RBF 네트워크의 중심 개수와 위치의 통합 결정 알고리듬이며 Method II와 III은 제안된 알고리듬의 성능과 비교하기 위한 기존의 알고리듬이다. 기존의 방법들은 중심의 개수를 알고 있다는 가정하에 중심의 위치를 결정하지만, 실제로 RBF를 classification의 문제에 적용할 시에 중심의 개수를 구하기란 매우 어렵기 때문에 그 개수를 구하는 과정도 RBF 네트워크의 학습(중심 개수, 중심 위치, Weight의 결정) 시간에 포함하여 비교하였다. Method II는 K-means 군집화 기법을 이용하여 중심의 위치를 결정하는 알고리듬으로 중심의 개수인, k는 한 개씩 증가시켜가며 목표 에러율에 도달하는 최소의 개수로 결정하였다. Method III는 입력 데이터 중에 임의로 중심의 위치를 결정하는 알고리듬으로 중심의 개수는 Method II와 같이 한 개씩 증가시켜가며 목표 에러율에 도달하는 최소의 개수로 결정하였다. 본 논문에서의 RBF인 다변량 가우스 함수의 분산은 1로 고정하였고 목표 에러율은 임의로 0.1로 설정하였다. 결과는 <표 2>에 정리하였다. <그림 6>은 제안된 알고리듬인 Method I을 이용하여 2-Spiral 문제의 분류 결정 영역(decision region)을 나타낸 그림이다.

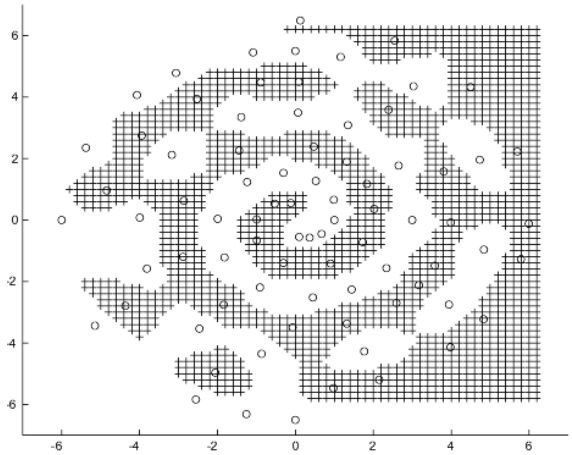

그림 6. 2-Spiral 문제의 결정 영역.

6. 결론 및 토의

본 논문에서는 사용자가 임의로 설정한 목표 에러율 이하에 도달하는 Generalized RBF 네트워크의 최소 중심 개수와 그 위치를 동시에 결정하는 새로운 알고리듬을 제안하였다. 이 알고리듬을 4 종류의 수치예제에 적용한 결과, <표 2>에서 볼 수 있듯이 중심 개수를 결정하는 데 걸리는 시간이 최소 15배에서 최대 85배까지 단축됨을 알 수 있었다. 구해진 중심 개수는 중심을 하나씩 증가시켜 구한 Method II나 III와 거의 유사하거나 더 작게 결정됨을 알 수 있다. 이는 Bi-section 알고리듬이 안정적으로 중심의 개수를 비교적 정확하게 결정하고 있음을

표 2. 수치예제 실험 결과

알고리듬	Method I (제안된 알고리듬)			Method II			Method III		
	중심 개수	계산 시간	에러율	중심 개수	계산 시간	에러율	중심 개수	계산 시간	에러율
2-Spirals	78 (36,42)	351	0.067	84	5650	0.069	88	5925	0.080
Sonar	34 (17,17)	34	0.096	37	1570	0.096	37	1112	0.096
Heart	126 (67,59)	159	0.106	126	13956	0.106	126	13094	0.106
Vowel	61 (2,5,5,2,8, 8,2,2,11,2,14)	256	0.097	65	4000	0.099	65	3688	0.093

* Method I에서 ( )안은 각 class 데이터 별 RBF 중심의 개수

입증한다. 또한 같은 에러율을 가지면서 오히려 더 적은 수의 RBF 중심만이 필요하다는 것은 입력 데이터를 class별로 분할하고 분할된 데이터 각각에 대해 군집화하여 구한 RBF의 중심이 기존의 기법보다 분류문제를 위한 데이터의 분포를 더 잘 반영함을 나타낸다. 즉, 좀더 분류문제에 적합한 중심의 위치를 결정해 주는 것이다. 따라서 본 논문에서 제안한 알고리듬은 입력 데이터의 분할로 기존의 중심 위치 결정 알고리듬보다 좀더 분류문제에 적합한 중심의 위치를 구해주며 최적 중심 개수 또한 Bi-section을 통해 빠른 시간 안에 결정해 준다는 것을 알 수 있다.

본 논문에서 제안한 알고리듬의 의의는 기존의 RBF 중심 결정 알고리듬들이 대부분 중심의 위치만을 고려한 것에 비해, 제안된 알고리듬에서는 RBF 네트워크 구성에 가장 큰 변수 중 하나인 중심의 개수를 위치와 통합적으로 결정한다는 것이다.

참고 문헌

A. C. Micchelli (1986), Interpolation of scattered data: Distance matrices and conditionally positive definite functions, *Construct. Approx,* vol. **2**, pp. 11-22.

C. M. Bishop (1991), Improving the generalization properties of radial basis function neural networks, Neural Computation, vol. **3**(4), pp. 579-588.

Cover, T.M. (1965), Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition, *IEEE Trans. Electronic Computers,* vol. EC-**14**, pp. 326-334.

Duda, R.O., and P.E. Hart (1973), Pattern classification and Scene Analysis, New York : Wiley.

D. S. Broomhead and D. Lowe (1988), Multivariable functional interpolation and adaptive networks, *Complex Syst.,* vol. **2**, pp. 321-355.

J. Barry Gomm and Ding Li Yu (2000), Selecting Radial Basis Function Network Centers with Recursive Orthogonal Least Squares Training, *IEEE Trans. Neural Networks,* vol. **11**(2), pp. 306-314.

K. Z. Mao (2002), RBF Neural Network Center Selection Based on Fisher Ratio Class Separability Measure, *IEEE Trans. Neural Network,* vol. **13**(5), pp. 1211-1217.

L. Bruzzone and D. F. Prieto (1999), A technique for the selection of kernelfunction parameters in RBF neural networks for classification of remotesensing images, IEEE Trans. Geosci., vol. **37**(2), pp. 1179-1184.

Lowe (1989), Adaptive radial basis function nonlinearities, and the problem of generalisation, First IEE International Conference on Artificial Neural Networks, pp.171-175, London.

M. J. Orr (1995), Regularization in the selection of RBF centers, Neural Computation, vol. **7**(3), pp. 606-623.

M. J. D. Powell (1985), Radial basis functions for multivariable interpolation : A review, *in Proc. IMA Conf. Algorithms for the Approximation of Functions and Data,* Shrivenham, U.K..

Poggio and Girosi, (1990), Networks for approximation and learning, Proceedings of the IEEE, vol.**78**, pp. 1481-1497.

S. Chen (1995), Nonlinear time series modeling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning, Inst. Elect Eng. Electron. Lett., vol. **31**(2), pp. 117-118.

Wettschereck and Dietterich (1992), Improving the performance of radial basis function networks by learning center locations, Advances in Neural Information Processing Systems, vol. **4**, pp.1133-1140, San Mateo, CA : Morgan Kaufmann.

Zheng ou Wang and Tao Zhu (2000), Al efficient learining algorithm for improving generalization performance of radial basis function neural networks, Neural Networks, vol. **13**(4,5), 2000.

Z. Uykan, C. Guzelis, M. E. Celebi and H. N. Koivo (2000), Analysis of inputoutput clustering for determining centers of RBFN, *IEEE Trans. Neural Networks,* vol. **11**, pp. 851-858.