

퍼지추론 기반 다항식 RBF 뉴럴 네트워크의 설계 및 최적화

論 文

58-2-28

The Design of Polynomial RBF Neural Network by Means of Fuzzy Inference System and Its Optimization

白 珍 烈* · 朴 炳 俊** · 吳 聖 權†

(Jin-Yeol Baek · Byoung-Jun Park · Sung-Kwun Oh)

Abstract - In this study, Polynomial Radial Basis Function Neural Network(pRBFNN) based on Fuzzy Inference System is designed and its parameters such as learning rate, momentum coefficient, and distributed weight (width of RBF) are optimized by means of Particle Swarm Optimization. The proposed model can be expressed as three functional module that consists of condition part, conclusion part, and inference part in the viewpoint of fuzzy rule formed in 'If-then'. In the condition part of pRBFNN as a fuzzy rule, input space is partitioned by defining kernel functions (RBFs). Here, the structure of kernel functions, namely, RBF is generated from HCM clustering algorithm. We use Gaussian type and Inverse multiquadratic type as a RBF. Besides these types of RBF, Conic RBF is also proposed and used as a kernel function. Also, in order to reflect the characteristic of dataset when partitioning input space, we consider the width of RBF defined by standard deviation of dataset. In the conclusion part, the connection weights of pRBFNN are represented as a polynomial which is the extended structure of the general RBF neural network with constant as a connection weights. Finally, the output of model is decided by the fuzzy inference of the inference part of pRBFNN. In order to evaluate the proposed model, nonlinear function with 2 inputs, waster water dataset and gas furnace time series dataset are used and the results of pRBFNN are compared with some previous models. Approximation as well as generalization abilities are discussed with these results.

Key Words : RBF Neural Network, Fuzzy Inference System, 입자군집최적화, 정보 입자화, HCM 클러스터링

1. 서 론

많은 분야에서 행해지고 있는 시스템 모델링에 관한 연구의 궁극적인 목표는 자연현상을 이해하고, 분석함으로써 원하는 정보를 이용하는 데 있다. 그러나 대부분의 현상들은 비선형을 내포하고 있기 때문에 선형성을 바탕으로 개발된 기존의 방법들을 이용해 이들을 해석하기에는 많은 어려움에 부딪히게 된다. 따라서 비선형 시스템의 복잡성과 불확실성을 해석하기 위해 다양한 연구가 지속되고 있으며, 그 중에서도 인간의 언어, 판단, 행동 등과 같은 양식을 모방한 지능형 모델을 이용하여 기존까지 해결하지 못했던 문제점들을 점진적으로 개선하는 추세이다. 하지만 현재까지 개발된 모델들은 입력변수가 증가하면 시스템이 복잡해지고 함수식으로 표현하는데 한계를 갖고 있다.

RBF(Radial Basis Function) 뉴럴네트워크는 지능형 구조 중 퍼지추론 모델과 함께 가장 많이 사용되는 대표적인 모델 중 하나로, 하나의 은닉층(hidden layer)만으로 구성된 네트워크 구조이다.

RBF 뉴럴네트워크는 빠른 학습시간, 일반화(generality) 능력, 단순화(simplicity)의 특징으로 인해 학습 데이터를 분류하는 작업과 비선형 시스템 모델링 등에 적용되고 있다 [1][3][8][10]. RBF 뉴럴네트워크의 구성은 입력데이터를 네트워크와 연결하는 입력층(input layer), 입력데이터를 RBF에 의해 고차원 공간으로 비선형 변환하는 은닉층(hidden layer), 은닉층의 결과를 선형 조합하여 네트워크의 최종 출력값을 계산하는 출력층(output layer)등의 세 개의 서로 다른 층으로 이루어져 있다. 세 층의 네트워크로 구성된 RBF 뉴럴네트워크는 퍼지추론[15][16]의 추론 메커니즘과 유사한 기능을 수행할 수 있으며, 최근 Jang등에 의해 증명되었다 [2][13]. 퍼지추론에 기반을 둔 모델은 퍼지규칙들로 구성되며, 최적구조 형성을 위해 전반부 동정과 후반부 동정으로 나뉘어 다루어진다. 각 동정은 다시 구조 동정과 파라미터 동정으로 나누어진다. 전반부의 구조 동정은 멤버쉽 함수의 형태와 수를 선택하는 문제로 퍼지 모델이 대상 시스템(object system)을 적절히 표현하기 위해 입력공간을 가장 적합한 퍼지 부분공간(fuzzy subspace)으로 분할하는 문제이며 파라미터 동정은 구조로서 선택된 멤버쉽 함수의 파라미터를 결정하는 문제이다. 이는 시스템의 특성을 얼마나 잘 나타낼 수 있는가 하는 중요한 문제로 비선형 시스템의 성능을 결정하는데 많은 영향을 미친다. 후반부 동정은 퍼지 부분공간을 어떻게 표현할 것인가에 대한 문제를 다룬다. 이와 같이 퍼지추론 메커니즘 관점에서의 RBF 뉴럴네트워크 설계를 위한 전반부 동정의 핵심적인 문제는 RBF의

† 교신저자, 正會員 : 水原大 工大 電氣工學科 教授 · 工博

E-mail : ohsk@suwon.ac.kr

* 準 會 員 : 水原大 工大 電氣工學科 碩士課程

** 正 會 員 : 韓國電子通信研究院 先任研究員

接受日字 : 2008年 9月 18日

最終完了 : 2008年 12月 3日

형태, RBF 중심의 위치 그리고 분포상수를 결정하는 것이다. 기존의 연구에서 중심의 위치를 결정하는 알고리즘들이 다양하게 제안되었으며, 입력 데이터 중 임의의 중심 선택법 [4], 입력 공간에서의 임의의 중심 추출법[5], 군집화 기법 (clustering algorithms)을 이용한 중심 결정법[6][7] 등이 있다.

본 연구에서는 퍼지추론 메커니즘에 기반한 다항식 RBF 뉴럴네트워크(Polynomial Radial Basis Function Neural Network; pRBFNN)를 설계하고 최적 구조와 파라미터 결정을 위한 문제를 다룬다. 제안된 모델은 RBF 뉴럴네트워크의 입력층, 은닉층, 출력층 세 층이 퍼지추론 메커니즘에 의해 'If-then' 형식으로 기술되는 조건부, 결론부, 추론부 세 가지의 기능적 모듈로 표현된다.

제안된 모델은 모의 데이터집합(synthetic dataset), 하수처리 데이터, 가스로 시계열 데이터에 적용되어 평가되며, 기존 모델과의 성능 비교 및 분석을 통해 일반화 측면에서 모델의 타당성과 유용성을 입증한다.

2. 다항식 RBF 뉴럴네트워크(pRBFNN)의 구조

2.1 일반적인 RBF 뉴럴네트워크의 구성

RBF 뉴럴네트워크는의 구성은 입력데이터(training data)를 네트워크와 연결하는 입력층(input layer), 입력데이터를 RBF에 의해 비선형 변환을 시키는 은닉층(hidden layer), 은닉층의 결과를 선형 조합하여 최종 출력값을 계산하는 출력층(output layer) 등의 세 개의 다른 층으로 이루어져 있으며 그림 1과 같다.

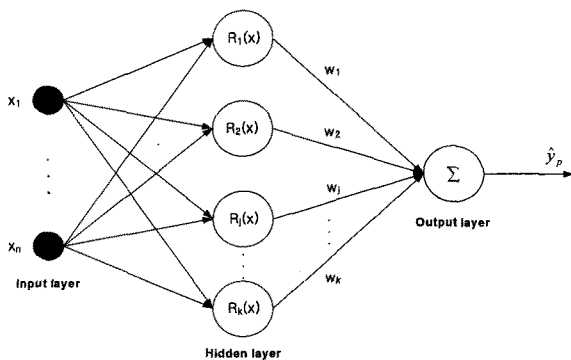


그림 1 일반적인 RBF 뉴럴네트워크
Fig. 1 General Radial Basis Function Neural Network

은닉층의 노드 수, 즉 커널함수인 RBF의 수는 보통 사용자에 의해 결정되고, 활성화함수로 정규분포 형태의 방사형 기저함수를 사용한다. 방사형 기저 함수로는 가우시안 (Gaussian) 함수, Thin plate spline 함수, 역 다항 2차 (Inverse Multiquadratic) 함수, 다항 2차(Multiquadratic) 함수, 3차(Cubic) 함수, 및 선형(Linear) 함수 등이 사용되고 있다. 표 1은 일반적인 RBF 활성화함수를 나타내고 있으며 주로 가우시안 형태가 사용된다.

표 1 일반적인 Radial Basis Function

Table 1 Type of general Radial Basis Function

Radial Basis Function	$R(r) = R(x - d /\sigma)$
Gaussian	$\exp(-r^2/2)$
Thin Plate Spline	$r^2 \log r$
Inverse Multiquadratic	$(r^2 + c^2)^{-1/2}$
Multiquadratic	$(r^2 + c^2)^{1/2}$
Cubic	r^3
Linear	r

RBF의 중심과 폭이 결정된 후 네트워크의 연결자중치는 최소자승법(least square estimation) 또는 경사하강법 (gradient descent method) 등의 학습방법에 의해 결정된다.

2.2 다항식 RBF 뉴럴네트워크(pRBFNN)의 구조

본 논문에서 제안한 다항식 RBF 뉴럴네트워크(Polynomial Radial Basis Function Neural Network; pRBFNN)는 입력층, 은닉층, 출력층으로 구성되는 그림 1의 일반적인 구조를 기반으로 한다. 하지만 기존의 기법과 달리, 퍼지 추론의 추론부에서 필요한 정규화 부분을 그림 2와 같이 은닉층에 추가하여 퍼지 추론 메커니즘의 기능적 모듈 동작 특성을 갖게 하였으며, 입력공간의 특성을 표현하기 위한 활성화함수의 구조 결정에 HCM 클러스터링을 사용하였다.

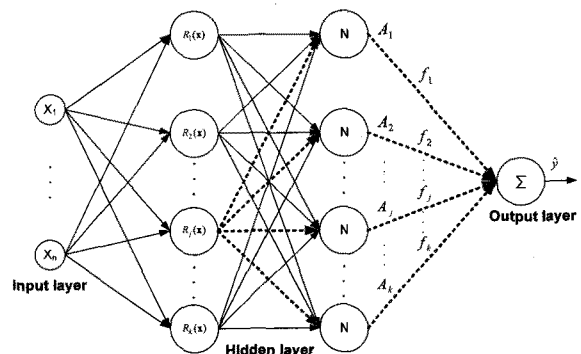


그림 2 퍼지 추론 기반 pRBFNN의 구조
Fig. 2 Architecture of pRBFNN based on the fuzzy inference

본 논문에서 제안한 pRBFNN은 HCM 클러스터링 알고리즘에 의한 활성화 함수의 구조 결정과 다항식 형태의 가중치를 적용하고 모델의 은닉층에 정규화 부분을 적용함으로써 퍼지 규칙과 같이 언어적인 관점에서 해석할 수 있으며 그림 2의 구조로 표현된다.

2.2.1 다항식 RBF 뉴럴네트워크(pRBFNN)의 조건부

2입력 1출력의 입출력 데이터를 2차원 가우시안 RBF에 적용하면 네트워크의 커널함수 출력은 식 (1)과 같이 표현할 수 있다.

$$R_j(x) = e^{-\frac{\|x - v_j\|^2}{2\sigma_j^2}} \quad (1)$$

여기서, $\mathbf{x} = [x, y]$ 는 입력행렬, $\mathbf{v}_j = [v_x, v_y]$ 는 $j(j=1, \dots, k)$ 번째 RBF의 중심벡터, 그리고 σ 는 RBF의 활성화 영역을 결정하는 분포상수이다. $\|\cdot\|$ 은 Euclidean distance를 나타낸다.

$$R(x, y) = e^{-\frac{(x-v_x)^2 + (y-v_y)^2}{2\sigma^2}} = e^{-\frac{(x-v_x)^2}{2\sigma^2}} \cdot e^{-\frac{(y-v_y)^2}{2\sigma^2}} \quad (2)$$

식 (1)을 전개하면 식 (2)와 같이 쓸 수 있다. 이는 2개의 1차원 가우시안 함수(혹은 2개의 멤버십 값)를 AND (product)연산자로 혼합 표현한 식 (3)과 같은 형태를 갖는다.

$$R(x, y) = \mu_A(x) \cdot \mu_B(y) \Leftrightarrow \text{If } x \text{ is } A \text{ and } y \text{ is } B \quad (3)$$

결국 RBF 뉴럴네트워크에서 은닉층의 각 노드는 퍼지추론에서의 퍼지규칙 조건부와 동일하게 표현됨을 알 수 있다 [2][13]. 따라서 본 논문에서 제안한 pRBFNN 구조는 은닉층을 퍼지규칙의 조건부로 대치하고, 상수항의 연결가중치를 다항식의 연결가중치와 퍼지규칙의 결론부로 표현하게 되면 식 (4)와 같은 퍼지규칙으로 기술할 수 있다.

$$R_j : \text{If } x_1 \text{ is } A_{j1}, \dots, \text{and } x_2 \text{ is } A_{j2}, \dots, \text{and } x_n \text{ is } A_{jn}, \text{ then } y = f_j \quad (4)$$

여기서 $R_j : j(j=1, \dots, k)$ 번째 규칙, $A_{jn} : \text{RBF}$ (또는 멤버십 함수), $n : \text{입력변수 수}$, f_j 는 다항식 함수이다. 식 (2)와 (3)을 고려하면 식 (4)는 식 (5) 같이 표현된다.

$$R_j : \text{If } x_1, x_2, \dots, \text{and } x_n \text{ are } A_j, \text{ then } y = f_j \quad (5)$$

이와 같이 커널함수인 RBF는 퍼지추론 메커니즘을 기반으로 하여 제안된 pRBFNN의 조건부(멤버십함수)의 역할을 수행한다.

2.2.2 다항식 RBF 뉴럴네트워크(pRBFNN)의 결론부

pRBFNN의 결론부는 조건부에서 분할한 입력공간을 각 로컬 영역의 다항식 함수로 표현하는 회귀모델로서 퍼지 규칙 식 (4)의 "then" 이후 부분을 형성한다. 본 논문에서는, 기존에 상수항으로 표현된 연결가중치가 아닌 다항식 형태의 연결가중치를 이용한다. 즉, 퍼지규칙 결론부를 나타내는 로컬 회귀 모델은 식 (6)으로 나타내어지는 상수항이외에 식 (7)의 1차 선형식 그리고 식 (8)의 2차식으로 확장됨으로써 로컬영역의 비선형 특징을 반영한다.

$$f_j(x) = a_{j0} \quad (6)$$

$$f_j(x) = a_{j0} + \sum_{i=1}^n a_{ji} \cdot x_i \quad (7)$$

$$f_j(x) = a_{j0} + \sum_{i=1}^n a_{ji} \cdot x_i + \sum_{i=1}^n \left(\sum_{k=1}^n a_{jik} \cdot x_i \cdot x_k \right) \quad (8)$$

식 (6)-(8)의 다항식 함수들은 조건부의 활성화(적합도)에 의해 활성화되며 각 퍼지 규칙의 로컬 회귀 모델로 동작한다.

2.2.3 다항식 RBF 뉴럴네트워크(pRBFNN)의 추론부

그림 2의 $A_j(j=1, 2, \dots, k)$ 는 은닉층 j 번째 뉴런의 출력신호이며, 이 출력신호는 은닉층과 출력층 사이의 j 번째 다항식 로컬모델과의 곱을 통하여 로컬출력을 얻게 되며, " Σ "기호로 표기되는 뉴런은 출력층에서의 입력신호를 합하여 그 결과를 출력층의 최종 출력으로 내보낸다. 이와 같은 일련의 과정은 퍼지 추론과 동일한 절차를 갖으며, 식 (9)와 같이 퍼

지추론의 추론식과 동일한 식으로 표현된다.

$$\hat{y}_p = \frac{\sum_{j=1}^k R_j \cdot f_j}{\sum_{j=1}^k R_j} = \sum_{j=1}^k (A_j f_j) \quad (9)$$

여기서, A_j 는 j 번째 규칙의 정규화된 적합도, 즉 $A_j = R_j / \sum_{j=1}^k R_j$ 이며, f_j 는 식 (6)-(8)로 표현되는 j 번째 퍼지규칙의 로컬모델이다.

3. 다항식 RBF 뉴럴네트워크(pRBFNN)의 설계

3.1 pRBFNN 조건부의 커널함수

제안된 pRBFNN의 조건부는 입력공간의 데이터를 커널함수인 RBF에 의하여 고차원 공간으로 비선형 변환하는 부분으로 커널함수 구조는 공간의 분할 수(커널함수의 수 또는 퍼지규칙의 수), 각 분할공간의 중심점(커널함수의 중심점), 그리고 분할공간의 활성화 영역(커널함수의 활성화영역)을 결정하는 분포상수에 의해 이루어진다. 본 논문에서는 RBF의 형태로써 가장 많이 사용되는 가우시안 형태 이외의 함수들을 비교 분석하고, 원뿔형태의 선형 RBF를 제안한다. 또한 비선형 변환을 위한 분할공간의 수와 중심, 그리고 분포상수를 클러스터링 알고리즘을 통하여 얻는다.

3.1.1 커널함수 (RBF)

본 연구에서는 RBF 뉴럴네트워크의 퍼지추론 메커니즘 적용을 위하여 종형 모양의 정규 분포 특징을 갖는 가우시안 타입과 중심점에서만 매우 큰 활성화 값을 갖는 Inverse Multiquadratic, 마지막으로 삼각형 형태의 특징을 갖는 제안된 원뿔형 RBF를 은닉층의 활성화 함수, 즉 퍼지추론의 멤버십함수로 사용한다. Weighted Euclidean Distance를 적용한 3가지 타입의 RBF는 아래와 같다.

- 1) Gaussian type : $R_j = e^{-\frac{\|\mathbf{x} - \mathbf{v}_j\|^2}{2}}$
- 2) Inverse Multiquadratic type : $R_j = \frac{1}{\sqrt{w\|\mathbf{X} - \mathbf{v}\|^2 + c^2}}$
- 3) Cone type : $R_j = c - w\|\mathbf{X} - \mathbf{v}\|$ (단, $R_j < 0$ 이면 $R_j = 0$)

3.1.2 커널함수(RBF)의 수 및 중심점 결정

본 논문에서는 RBF의 수와 중심을 결정하기 위해 데이터들 간의 거리를 기준으로 근접한 정도를 측정하고, 이를 바탕으로 데이터를 특성별로 분류하는 HCM 클러스터링을 이용한다. 설계자에 의해 결정된 초기 HCM의 클러스터 수를 RBF 수로 사용하고, 각 클러스터의 중심은 RBF의 중심이 된다. HCM 클러스터링에 의한 학습 데이터로부터의 정보 추출은 다음 절차에 의한다.

[단계 1] 클러스터의 개수 ($2 \leq c \leq n$)를 결정하고, 소속행렬 U 를 $U^{(0)} \in M_c$ 으로 초기화한다.

$$M_c = \left\{ U \mid u_{ij} \in (0, 1), \sum_{k=1}^c u_{ik} = 1, 0 < \sum_{k=1}^n u_{ik} < n \right\} \quad (10)$$

여기서, u_{ij} 는 소속행렬의 파라미터이다.

[단계 2] 각각의 클러스터에 대한 중심벡터 V_i 를 구한다.

$$V_i = \{v_{i1}, v_{i2}, \dots, v_{ij}, \dots, v_{im}\} \quad (11)$$

$$v_{ij} = \frac{\sum_{k=1}^n u_{ik} \cdot x_{kj}}{\sum_{k=1}^n u_{ik}}$$

여기서, i 는 i 번째 클러스터, j 는 j 번째 입력 변수($j=1, \dots, m$), n 은 데이터 수이다.

[단계 3] 각각의 클러스터 중심과 데이터와의 거리를 계산하여 새로운 소속행렬 $U^{(r)}$ 을 생성한다.

$$d_{ik} = d(x_k - V_i) = \|x_k - V_i\| = \left[\sum_{j=1}^m (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (12)$$

$$u_{ik}^{(r+1)} = \begin{cases} 1 & d_{ik}^{(r)} = \min \{d_{jk}^{(r)}\} \text{ for all } j \in c \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

여기서, d_{ik} 는 k 번째 데이터 sample x_k 와 i 번째 클러스터 중심 V_i 의 기하학적 거리이다.

[단계 4] 만일 식 (23)을 만족한다면 종료하고, 그렇지 않으면 $r = r+1$ 로 놓고 [단계 2]로 간다.

$$\|U^{(r+1)} - U^{(r)}\| \leq \epsilon (\text{tolerance level}) \quad (14)$$

본 논문에서는 [단계 1]에서 결정한 클러스터의 개수를 은닉층 노드의 수로 사용한다.

$$[x_1, Y] \rightarrow [v_{1i}, m_{1i}], [x_2, Y] \rightarrow [v_{2i}, m_{2i}] \quad (15)$$

여기서 i 는 은닉층 노드의 수(클러스터의 수) 이고, v_{1i} 는 x_1 입력에 대한 중심, v_{2i} 는 x_2 입력에 대한 중심이 된다. 클러스터의 수는 설계자에 의해 선택되며, 각 클러스터에 속한 데이터와 중심점과의 표준편차에 의해 분포상수를 결정한다.

3.1.3 커널함수(RBF)의 분포상수 결정

RBF의 중심이 결정되면 식 (1)의 가우시안 함수에 적용한다. 여기서 σ 는 분포상수(spread constant)로서 입력에 대해 어떻게 응답하는지를 결정하는 파라미터이다.

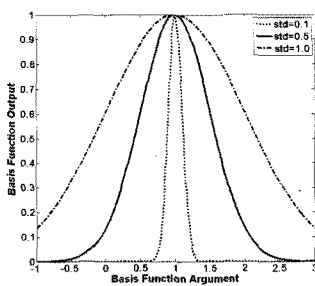


그림 3 가우시안 형태의 RBF 활성화 영역
Fig. 3 Activation field of gaussian type of radial basis function

그림 3은 가우시안 형태의 RBF 활성화 영역을 나타낸다. 여기서 보듯이 σ 가 클수록 RBF 활성화영역은 증가한다.

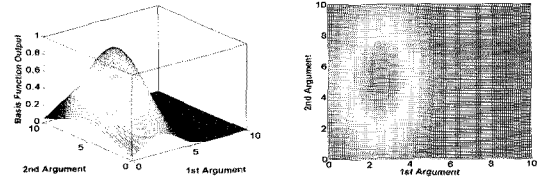
본 논문에서는 모든 입력변수에 대하여 동일하게 적용된 활성화 영역의 개선을 위해 식 (16)과 같이 각 클러스터에 속한 입력벡터 각각의 표준편차를 분포상수 σ 로써 사용한다.

$$\sigma_{ji} = \Delta\sigma_j \cdot std_{ji}(x_i) \quad (16)$$

여기서 j 는 j 번째 클러스터를 나타내며, i 는 i 번째 입력변

수이다. $\Delta\sigma_j$ 는 분포상수 가중치로 초기값은 1이다. 그림 4는 식 (16)에 의해 표현된 활성화 영역으로 각 입력에 대하여 독립된 분포를 갖게 됨을 보여준다.

또한 분포상수 가중치 $\Delta\sigma$ 을 조절함으로써 각각의 입력변수에 적용된 활성화영역을 조절하고 모델의 출력을 향상시킬 수 있다. 본 논문에서는 모델 출력 향상을 위한 최적 활성화 영역의 형성은 PSO 알고리즘을 이용하여 분포상수 가중치를 동조함으로써 수행한다.



(a) 3차원 공간상의 활성화 영역 (b) Z축에서 내려다 본 활성화 영역

그림 4 식 (28)에 의한 가우시안 RBF 활성화 영역 ($\Delta\sigma_j=1$ 인 경우)

Fig. 4 Activation field of gaussian RBF using Eq. (28) ($\Delta\sigma_j=1$)

3.2 pRBFNN 결론부의 다항식 연결가중치 학습

퍼지추론 기반 pRBFNN의 결론부에 존재한 연결가중치의 학습은 경사하강법과 모멘텀을 통하여 이루어진다. 네트워크의 연결가중치, 즉 다항식 계수 동정을 위한 오차식 정의는 식 (17)로 표현되며, 전체 학습데이터에 대한 모델의 최소 오차를 얻기위해 학습이 이루어진다.

$$E_p = (y_p - \hat{y}_p)^2, \quad (p=1,2,\dots,N) \quad (17)$$

여기서, E_p 는 p 번째 데이터에 대한 모델의 오차이며, N 은 입력 데이터 수 이다. y_p 는 p 번째 입력에 대한 원 출력이며, \hat{y}_p 는 p 번째 입력에 대해 추론된 모델의 결과로 식 (9)에 의한다.

경사하강법에 의한 다항식 계수의 학습은 식 (17)로 표현되는 오차가 최소가 되도록 식 (18)과 같이 계수의 변화분 Δa_{ji} 값을 이용하여 현재의 계수값을 갱신한다.

$$a_{ji}(\text{new}) = a_{ji}(\text{present}) + \Delta a_{ji} + \alpha \cdot \Delta a_{ji}(\text{old}) \quad (18)$$

여기서, η 는 학습율이고 α 는 모멘텀 계수로 연결가중치 변화에 관성을 부여함으로써, 연결가중치가 아주 얇은 지역 극소에 빠지는 것을 방지해 주고 네트워크의 학습 속도를 개선해 주는 효과가 있다[8].

pRBFNN의 가중치 학습은 2.2.2절에서 언급한 결론부의 다항식 형태에 따라 다음 3가지로 나뉘어 진행된다.

1) $f_j = a_{j0}$: E_p 와 a_{j0} 사이의 관계를 Chain rule로 나타내면 식 (19)와 같다. 식 (19)의 우변항은 식 (20)-(22)와 같이 정리된다.

$$\Delta a_{j0} = -\eta_w \cdot \frac{\partial E_p}{\partial a_{j0}} = -\eta_w \left(\frac{\partial E_p}{\partial \hat{y}_p} \cdot \frac{\partial \hat{y}_p}{\partial f_j} \cdot \frac{\partial f_j}{\partial a_{j0}} \right) \quad (19)$$

$$\frac{\partial E_p}{\partial \hat{y}_p} = -2(y_p - \hat{y}_p) \quad (20)$$

$$\frac{\partial \hat{y}_p}{\partial f_j} = A_j \quad (21)$$

$$\frac{\partial f_j}{\partial a_{j0}} = 1 \quad (22)$$

연결 가중치가 상수인 경우 변화분 Δa_{j0} 는 식 (19)-(22)로부터 식 (23)과 같이 표현되며, 모멘텀을 고려한 새로운 연결가중치의 최종식은 식 (24)와 같다.

$$\Delta a_{j0} = 2 \cdot \eta_w \cdot (y_p - \hat{y}_p) \cdot A_j \quad (23)$$

$$a_{j0}(new) = a_{j0}(present) + \Delta a_{j0} + \alpha \cdot \Delta a_{j0}(old) \quad (24)$$

2) $f_j = a_{j0} + \sum_{i=1}^n a_{ji} \cdot x_i$: pRBFNN 결론부의 다항식 형태가 선형식의 경우 상수항인 a_{j0} 의 학습은 식 (24)에 의한다. a_{ji} 의 변화분은 Chain rule에 의하여 식 (25)와 같이 계산되며, 모멘텀을 고려한 새로운 a_{ji} 는 식 (26)으로부터 결정된다.

$$\begin{aligned} \Delta a_{ji} &= -\eta_w \cdot \frac{\partial E_p}{\partial a_{ji}} \\ &= -\eta_w \left(\frac{\partial E_p}{\partial y_p} \cdot \frac{\partial y_p}{\partial f_j} \cdot \frac{\partial f_j}{\partial a_{ji}} \right) \quad (25) \\ &= -\eta_w \cdot (-2(y_p - \hat{y}_p) \cdot A_j \cdot x_i) \\ &= 2 \cdot \eta_w \cdot (y_p - \hat{y}_p) \cdot A_j \cdot x_j \end{aligned}$$

$$a_{ji}(new) = a_{ji}(present) + \Delta a_{ji} + \alpha \cdot \Delta a_{ji}(old) \quad (26)$$

3) $f_j = a_{j0} + \sum_{i=1}^n a_{ji} \cdot x_i + \sum_{i=1}^n \left(\sum_{k=1}^m a_{jik} \cdot x_i \cdot x_k \right)$: 모델 오차를 최소화 하기 위한 a_{j0} 와 a_{ji} 의 학습은 식 (24)와 식 (26)에 의해 각각 이루어진다. 뿐만 아니라, a_{jik} 의 변화분은 구하기 위해 위에서 적용된 동일한 방법을 이용하여 식 (27)을 유도한다. 모멘텀을 고려한 새로운 a_{jik} 의 최종식은 식 (28)과 같이 정리된다.

$$\Delta a_{jik} = 2 \cdot \eta_w \cdot (y_p - \hat{y}_p) \cdot A_j \cdot (x_i \cdot x_k) \quad (27)$$

$$a_{jik}(new) = a_{jik}(present) + \Delta a_{jik} + \alpha \cdot \Delta a_{jik}(old) \quad (28)$$

3.3 Particle Swarm Optimization을 이용한 pRBFNN의 최적화

PSO(Particle Swarm Optimization) 알고리즘은 1995년 Eberhart와 Kennedy에 의해서 제안되었으며, 개체의 군집이론을 근본으로 사회적인 행동양식에 기반을 두고 있다[9].

PSO 알고리즘은 다음의 과정으로 수행된다[12].

[단계 1] D-차원의 초기입자 $x[[d]$ 와 속도 $v[[d]$ 를 초기화 한다.

[단계 2] 목적함수(최적화 하려는 함수)로 $x[i][d]$ 를 평가한다.

[단계 3] 현재와 과거의 $pbest[[]]$ 를 비교, 평가한다.

$present[] < pbest[]$;

$$pbest[] = \text{현재 } pbest[[d]$$

[단계 4] 각 입자에 대한 이웃에서 현재 $present[gbest]$ 를 찾는다.

$present[] < present[best]$;

$$pbest[] = \text{particle's array index}$$

[단계 5] 새로운 속도 $v[i][d]$ 를 계산한다.

$$v[[d] = w \times v[[d] + c_1 \times rand() \times (pbest[[d] - present[[d]) +$$

$$c_2 \times rand() \times (present[best] - present[[d])$$

[단계 6] [단계 5]에 의해 새로운 위치 $x[i][d]$ 를 계산하고 조건에 맞을 때까지 [단계 2]부터 반복한다.

본 논문에서 제안된 pRBFNN의 파라미터로는 가중치 학습에 관련된 학습률 η 와 모멘텀 계수 α 가 있으며, 은닉층 뉴런의 구조를 결정하는 분포상수의 가중치가 있다. 표 2는 식 (16)에 의하여 분포상수 가중치를 동정 하였을 때 Particle의 구조를 나타내고 있다.

표 2 Particle 구조

Table 2 Structure of particle

구분	Particle의 구성	Particle 크기
식 (28)	$[\eta, \alpha, \Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_j, \dots, \Delta\sigma_k]$	2+퍼지규칙 수

표 2에서, k는 은닉층 뉴런의 수(클러스터 수 또는 퍼지규칙 수)를 나타낸다.

4. 실험 결과 및 고찰

본 논문에서는 제안된 pRBFNN(Polynomial RBF Neural Networks)의 성능평가를 위해 2입력 1출력의 비선형 함수와 함수처리장 활성화오니 공정의 수치리 데이터, Box와 Jenkins가 제안한 가스로 시계열 데이터를 이용한다. 근사화 성능과 일반화 성능을 평가하기 위한 성능지수는 식 (29)와 같다.

$$PI \text{ or } EPI = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (29)$$

여기서 PI(Performance Index)는 학습데이터에 대한 성능지수를, EPI(Evaluated Performance Index)는 테스트 데이터에 대한 성능지수를 나타낸다.

제안된 pRBFNN의 최적화를 위해 사용된 PSO 알고리즘의 파라미터는 표 3과 같다.

표 3 PSO 파라미터

Table 3 Parameters of PSO

파라미터	값	
세대 수	60	
particle 수	50	
v_{max}	탐색공간의 20%	
$[w_{min} \ w_{max}]$	[0.4 0.9]	
c_1, c_2	2.0	
탐색 범위	학습률	[0 1.0]
	모멘텀 계수	[0 1.0]
	분포상수 가중치	[1.0 9.0]

4.1 안정도 측도

다양한 RBFNN 모델 중에서 적합한 구조를 선택하는 것에 대한 의문이 증가 할 것이다. 분명히, 학습 데이터와 테스트 데이터에 대한 모델의 성능들(PI, EPI)은 우리가 고려해야 할 두 가지의 중요한 관점들이다. PI와 EPI 값들은 지능형 모델의 근사화와 일반화 능력을 나타내는 좋은 척도가 된다.

이러한 PI와 EPI를 이용하여 안정도 측도를 위한 다음 식 (30)을 도입한다.

$$k = \frac{EPI}{PI} \quad (30)$$

이 지수는 테스트 데이터에 의하여 모델의 성능이 얼마만큼 저하되는지를 나타낸 모델의 안정도 측도이다.

4.2 2입력 1출력의 비선형 함수

pRBFNN의 평가에 이용될 2입력 1출력의 비선형 함수는 식 (31)로 표현되며, 이 식에 의한 데이터 집합은 그림 5에 보인다.

$$z = 1.9 \{1.35 + e^{x_2} e^{-x_1} \sin(13(x_2 - 0.6)^2) \sin(7x_1)\} \quad (31)$$

$$0 \leq x_1, x_2 \leq 1$$

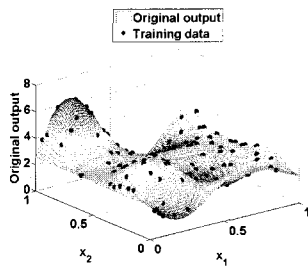


그림 5 2입력 1출력의 비선형 함수
Fig. 5 Nonlinear function with 2 inputs and 1 output

그림 5에 보인 데이터 집합으로부터 200개의 데이터를 랜덤하게 선택하고 그 중 140개를 학습데이터로, 나머지 60개의 데이터를 테스트 데이터로 사용하여 제안된 pRBFNN을 평가한다. 뿐만 아니라, 표 3의 파라미터에 의한 PSO를 사용하여 제안된 모델을 최적화 한다.

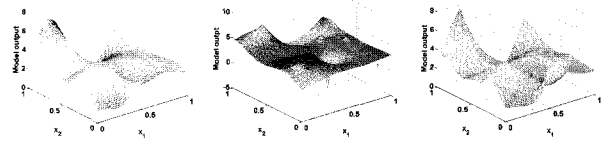
표 4 pRBFNN 모델의 결과
Table 4 Results of the pRBFNN models

RBF type	퍼지규칙	다항식 형태	PI	EPI	k(EPI/PI)
Gaussian	11	Quadratic	0.0231	0.0272	1.1775
Inverse Multiquadratic	9	Linear	0.0873	0.0920	1.0538
Cone	11	Quadratic	0.0227	0.0222	0.9780

표 4는 최적 모델로써 선정된 pRBFNN 모델의 결과를 보여준다. 이 결과는 근사화와 일반화를 고려하여 선정되었다. EPI/PI는 학습 결과와 테스트 결과를 이용한 일반화와 근사화 사이의 관계를 보여준다. 즉 EPI/PI이 큰 값을 가지는 경우 모델이 학습데이터에 의해 오버피팅(Overfitting)됨으로써 제안된 모델이 새로운 데이터에 대하여 적합하지 않음을 나타낸다. 1에 근접한 값은 학습 데이터로의 오버피팅 현상이 적게 나타나고 모델 형성에 참여하지 않은 데이터 집합에 대하여서도 모델이 적합함을 의미한다.

표 4에서 보인 결과로부터, Gaussian RBF를 이용한 pRBFNN 모델은 다른 형태의 RBF 모델에 비하여 EPI/PI 값이 약간 크게 나타나며, Inverse MultiQuadratic RBF를 이용한 모델은 다른 모델보다 성능이 떨어짐을 알 수 있다. 즉 본 논문에서 제안한 원뿔형태의 RBF를 이용한 pRBFNN 모델이 우수한 근사화 및 일반화 특징을 가진다. 표 4의

RBF 형태에 따른 제안된 pRBFNN의 출력은 그림 6과 같다. 그림에서 알 수 있듯이, 가우시안을 이용한 모델은 그림 5의 원 출력과 가장 유사한 결과를 보이지만 일반화 성능이 낮게 나타나며, 원뿔 RBF를 이용한 pRBFNN의 결과는 원 출력과 차이를 보이지만 근사화와 일반화에 있어서 균형을 이룬다. 즉, 제안된 모델은 복잡한 비선형 특성을 갖는 플랜트를 잘 묘사할 수 있다.



(a) Gaussian (b) Inverse Multiquadratic (c) Cone

그림 6 RBF 형태에 따른 제안된 pRBFNN의 출력
Fig. 6 Output of the proposed pRBFNN for RBF types

4.3 하수처리장의 활성오니 공정

하수처리장의 활성오니 공정에 대한 데이터 집합은 하수처리장의 1년간 수질 데이터를 토대로 한 데이터이다. 활성오니 공정은 측정된 데이터의 혼합액 부유물(Mixed Liquid Suspended Solid; MLSS) 농도, 잉여오니 흐름율(Waste Sludge Ratio; WSR), 반송율 설정치(Recycle sludge Ratio Set-Point; RRSP), 용존 산소 설정치(Dissolved Oxygen Set-Point; DOSP)를 입력으로 하고, 부유물(Effluent Suspended Solid; ESS) 농도를 출력으로 하는 4입력 1출력 구조를 가진다. 데이터의 수는 52개로 전반부 26개는 학습 데이터, 후반부 26개는 테스트 데이터로 이용한다[20].

근사화 성능과 일반화 성능을 고려한 최적모델의 결과는 표 5에 보인다. 가우시안의 경우 후반부 다항식은 1차식(Linear)이, Inverse Multiquadratic의 경우 후반부 다항식은 2차식(Quadratic) 그리고 Cone type은 1차식(Linaer)이 선정되었으며 모든 RBF 형태에 대하여 pRBFNN은 4개의 퍼지규칙을 선택하였다. 표 5에서는 또한 제안된 pRBFNN 모델과 기존에 연구된 모델과의 성능비교를 보여준다.

표 5 pRBFNN 모델의 결과
Table 5 Result of the pRBFNN models

입력변수	모델	규칙 수	다항식 형태	PI	EPI	k(EPI/PI)	
WSR, DOSP	gHCM_FIS[20]	4	Linear	6.108	26.913	4.4061	
MLSS, RRSP	Hybrid Fuzzy model[11]	4	Linear	6.892	35.967	5.2186	
MLSS, WSR	Oh and Pedrycz model[18]	4	Constant	13.726	16.206	1.1806	
			Linear	6.396	54.233	8.4792	
	(HCM+GA) Fuzzy model[19]	6	Constant	12.352	11.179	0.9050	
MLSS, WSR, RRSP, DOSP	pRBFNN	Gaussian	4	Linear	8.7780	7.2530	0.8263
		Inverse Multiquadratic	4	Quadratic	6.8150	9.8910	1.4514
		Cone	4	Linear	8.3810	7.4050	0.8835

표 5에서 보이듯이 본 논문에서 제안한 pRBFNN은 다른 모델과 비교하여 근사화 결과인 PI는 다소 성능이 떨어지나

평가 데이터에 대한 모델의 오차, 즉 일반화 성능은 매우 우수함을 알 수 있다. 더욱이 오버피팅 측면에서도 월등함을 보인다.

4.4 가스로 시계열 데이터

이 절에서는 Box와 Jenkins가 사용한 가스로 시계열 데이터로부터 입출력 데이터인 가스 흐름율과 연소된 이산화탄소 농도의 가스로 공정을 제안된 pRBFNN을 이용하여 모델링 한다. 입력 $u(t)$ 는 가스 흐름율이고 출력 $y(t)$ 는 이산화탄소 농도인 1입력 1출력의 데이터 296쌍으로 구성되며, 가스로 공정 모델링을 위해 $u(t-3)$, $u(t-2)$, $u(t-1)$, $y(t-3)$, $y(t-2)$, $y(t-1)$ 의 6개 입력변수와 한 개의 출력변수 $y(t)$ 를 설정하고 학습데이터와 테스트 데이터로써 각각 145개의 데이터 쌍을 사용한다.

데이터 집합 $[u(t-3), y(t-1); y(t)]$ 이외의 데이터 집합은 분포가 균일하지 못하고 흩어진 분포 구조를 보여줌으로 가스로 공정의 모델링에 적절하지 못하다.

최적 pRBFNN 모델 결과는 앞 예제에서 보인 것과 같이 동일한 방법에 의해 결과를 정리하였다. 표 6에서는 기존 모델과의 성능 비교를 보여준다.

본 논문에서 제안한 pRBFNN은 기존 모델들 보다 낮은 근사화 성능을 보이지만 테스트 성능 및 근사화와 일반화 사이의 관계는 기존 모델들보다 우수하다. 뿐만 아니라, 새로운 형태의 RBF인 제안된 원뿔형태 역시 기존의 RBF 형태보다 우수한 결과를 보여준다.

표 6 기존 모델과 pRBFNN 모델과의 성능비교

Table 6 Comparison of original output and output of pRBFNN for RBF types

모델	퍼지규칙 수	PI	EPI	k(EPI/PI)	
SVR(SVM)	x	0.023	1.028	44.695	
Standard Neural Network	15(node)	0.018	0.411	22.833	
Fuzzy model[14]	8	0.020	0.333	16.650	
George's model[17]	7	0.022	0.266	12.090	
Our Model (pRBFNN)	Gaussian	8	0.031	0.234	7.548
	Inverse Multiquadratic	10	0.033	0.258	7.818
	Cone	8	0.041	0.226	6.487

5. 결 론

본 논문에서는 퍼지 추론 기반 다항식 RBF 뉴럴 네트워크(Polynomial Radial Basis Function Neural Network; pRBFNN) 구조를 제안하였다. 이 구조는 일반적인 RBF 뉴럴네트워크를 기반으로한 확장된 구조로써 커널함수인 RBF는 HCM을 사용하여 구조를 결정하고 상수항의 연결가중치를 다항식으로 확장하여 비선형적인 결합을 제안하였다.

제안된 퍼지 추론 기반 pRBFNN은 RBF를 사용하여 퍼지공간 분할과 멤버쉽 함수를 결정한다. 여기서 기존에 주로 사용하는 가우시안 타입의 RBF 이외에 원뿔형태의 RBF를 제안하고, 이를 적용함으로써 플랜트의 특성을 알 수 없는 데이터에 대하여 일반화 성능을 향상 시켰다. 비선형 함수와 활성오니 공정, 가스로 공정 데이터의 적용을 통하여 얻어진 실험 결과로부터 제안된 pRBFNN 모델이 기존 모델

과 비슷한 근사화 성능을 가지며, 일반화 성능에는 보다 우수함을 증명하였다. 더욱이 기존 모델이 학습 데이터에 대하여 오버피팅 되는 특성이 강했던 반면 제안된 pRBFNN 모델은 근사화와 일반화 관계 측면에서 좀 더 유연한 특성을 보여 줌으로서 모델의 우수성과 타당성을 입증하였다.

감사의 글

본 연구는 경기도의 경기도지역협력연구센터사업의 일환으로 수행하였음 [GGA0801-45700, U-city 보안 감사 기술협력센터] 그리고 이 논문은 2008년도 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2008-314-D00376).

참 고 문 헌

- [1] M. L. Kothari, S. Madnani, and R. Segal, "Orthogonal Least Square Learning Algorithm Based Radial Basis Function Network Adaptive Power System Stabilizer", Proceeding of IEEE SMC, Vol. 1, pp. 542-547, 1997.
- [2] J.S Roger Jang, C.T. Sun, "Functional equivalence between radial basis function networks and fuzzy inference systems", IEEE Trans. Neural Networks, Vol. 4, No. 1, pp. 156-158, 1993.
- [3] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York, 1981.
- [4] K. Z. Mao, "RBF Neural Network Center Selection Based on Fisher Ratio Class Separability Measure", IEEE Trans. Neural Network, Vol. 13, No. 5, pp. 1211-1217, 2002.
- [5] A. C. Micchelli, "Interpolation of scattered data: Distance matrices and conditionally positive definite functions", Construct. Approx, Vol. 2, pp. 11-22, 1986.
- [6] M. J. D. Powell, "Radial basis functions for multivariable interpolation : A review", in Proc. IMA Conf. Algorithms for the Approximation of Functions and Data, Shrivensham, U.K., 1985.
- [7] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks", Complex Syst., Vol. 2, pp. 321-355, 1988.
- [8] Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern Classification ; Second Edition", John Wiley&Sons, INC., 2000.
- [9] R. C. Eberhard, P. Simpson, R. Dobbins, "Computational Intelligence PC Tools", Boston: Academic Press Professional, 1996.
- [10] 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.

[11] S.-K. Oh, W. Pedryz, B. -J. Park, "Hybrid Identification of Fuzzy Rule-Based Models", International Journal of Intelligent System, Vol. 17, pp. 77-103, 2002.

[12] Y. Shi, R. C. Eberhart, "Extracting Rules from Fuzzy Neural Network by Particle Swarm Optimization", Proceedings of the IEEE International Conference on Evolutionary Computation, 1998.

[13] F. Behloul, B.P.F. Lelieveldt, A. Boudraa, J.H.C. Reiber, "Optimal design of radial basis function neural networks for fuzzy-rule extraction in high dimensional data", The Journal of the Pattern Recognition Society, Vol. 35, pp. 659-675, 2002.

[14] 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.

[15] 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.

[16] W. Pderycy and G. Vukovich, "Granular neural networks", Neurocomputing, Vol. 36, pp. 205-224, 2001.

[17] George E. Tsekouras, "On the use of the weighted fuzzy c-means in fuzzy modeling", Advances in Engineering Software, Vol. 36, pp. 287-300, 2005.

[18] S.-K. Oh and W. Pedryz, "Identification of Fuzzy Systems by means of an Auto-Tuning Algorithms and Its Application to Nonlinear Systems", Fuzzy Sets and Syst., Vol. 115, No. 2, pp. 205-230, 2000.

[19] B. -J. Park, W. Pedrycz and S. -K. Oh, "Identification of fuzzy models with the aid of evolutionary data granulation", Vol. 148, No. 5, pp. 406-418, 2001.

[20] 오성권, 프로그램에 의한 컴퓨터지능(퍼지, 신경회로망 및 진화 알고리즘을 중심으로), 내하출판사, 2002.



박 병 준 (朴 炳 俊)

1998년 : 원광대 제어계측학과 졸업.
 2000년 : 동 대학원 제어계측학과 졸업 (공학석사). 2003년 : 동 대학원 제어계측학과 졸업(공학박사). 2005~2007년 : 캐나다 Alberta대학 전기 및 컴퓨터공학과 Post-Doc. 2008년~현재 : 한국전자통신연구원(ETRI) 선임연구원. 관심분야 : 신경회로망 응용 및 제어, 컴퓨터 지능, 패턴인식

Tel : 042-860-1503
 Fax : 042-860-1085
 E-mail : bj_park@etri.re.kr



오 성 권 (吳 聖 權)

1981년 : 연세대 전기공학과 졸업.
 1983년 : 동 대학원 전기공학과 졸업(공학석사). 1983~1989년 : 금성산전연구소(선임연구원). 1993년 : 연세대 대학원 전기공학과 졸업(공학박사). 1996~1997년 : 캐나다 Manitoba대학 전기 및 컴퓨터공학과 Post-Doc. 1993~2005년 : 원광대 전기전자 및 정보공학부 교수. 2005년~현재 : 수원대 전기공학과 교수. 2002년~현재 : 대한전기학회, 퍼지및지능시스템학회 및 제어자동화시스템 공학회 편집위원. 관심분야 : 시스템 자동화, 퍼지이론, 신경회로망 응용 및 제어, 컴퓨터 지능 등.

Tel : (031) 229-8162
 Fax : (031) 220-2667
 E-mail : ohsk@suwon.ac.kr

저 자 소 개



백 진 열 (白 珍 烈)

2007년 : 수원대 전기공학과 졸업.
 2007년~현재 : 동 대학원 전기공학과 석사과정. 관심분야 : 퍼지추론시스템, 신경회로망, 유전자 알고리즘 및 최적화이론, 지능시스템 및 제어, 마이크로프로세서 등

Tel : (031) 222-6544
 Fax : (031) 220-2667
 E-mail : qorwlsduf@suwon.ac.kr