

퍼지 및 신경망을 이용한 Blending Process의 최적화

황인창*, 김정남*, 주관정*

Blending Process Optimization using Fuzzy Set Theory and Neural Networks

In-Chang Hwang*, Jeong-Nam Kim*, Kuan-Jung Joo*

ABSTRACT

This paper proposes a new approach to the optimization method of a blending process with neural network. The method is based on the error backpropagation learning algorithm for neural network. Since the neural network can model an arbitrary nonlinear mapping, it is used as a system solver. A fuzzy membership function is used in parallel with the neural network to minimize the difference between measurement value and input value of neural network. As a result, we can guarantee the reliability and stability of blending process by the help of neural network and fuzzy membership function.

Key Words : Neural Network (신경 회로망), Fuzzy Membership Function (퍼지소속함수),
Error Backpropagation Learning Algorithm (오류 역전파 학습 알고리즘),
Blending Process (인액혼합공정)

1. 서 론

과거 수많은 화학제품제조공정 등 여러 관련분야에서 인액 혼합과정 (blending process) 을 포함한 생산시스템의 구축은 인액의 비선형적 특성, 제조 공정조건의 변화와 같은 이유로 많은 어려움을 가져왔다. 재래적인 방법에 의한 인액혼합비율의 계산은 숙련된 작업자에 의해 수동으로 행해지는데 이는 만족할만한 결과를 유출하기까지 가능한 다양한 혼합비를 통한 실험을 반복하여 시행착오법으로 해를 구하는 것이었다. 이 절차는 최종결과에 영향을 미치는 인자 (parameter) 의 수가 증가함에 따라 (예로서 저장채 혼합의 경우 인액의 sheet 저항값, 소성 저항값, 소성조건, 인쇄막후, blending 조건 등) 지루하고도 시간의 소모가 많을 뿐 아니라 숙련공의 개인적인 경험에 의존함으로써 일관된 최적해를

구하기가 어렵다.

더우기 실제 문제들은 많은 제한조건을 따르는데 각 문제에서 우리가 얻고자 하는바를 기술한 목적함수가 이러한 제한조건들을 다 반영하지 못할때도 있다. 이러한 요인으로 인해 지금까지의 시스템 구축방법은 제한된 시간안에 최적해가 아닌, 경험적으로 얻은 정보와 근사 알고리즘을 이용하는 최적화 기법으로 휴리스틱 (heuristic) 원리에 근거한 방법의 개발에 역점을 두어 왔었다. 그러나 이러한 방법들은 휴리스틱 정보를 적용하여 최종결과를 탐색하는데 있어서 탐색 속도를 감소시켰으며, 탐색 절차가 순차적 (sequential) 이며 결정론적 (deterministic) 인 방법들 이므로 좋은 결과를 얻기가 어려웠다. 이에 따라 최근에는 최적화 문제에 있어서 지금까지의 순차적으로 탐색해 가던 개념들로부터 병렬적 (parallel) 으로 탐색하는 기법을 구현하기 위해 신경

* 삼성전기(주) 종합연구소 CAD/CAM 연구실

회로망(neural network)을 적용하여 최적화 문제에 대한 새로운 평가와 가능성을 제시하고 있다.

최근 수년사이에 병렬처리(parallel processing)능력을 가지고 전통적인 해석적 접근법으로 해결될 수 없는 많은 문제를 처리할 수 있는 신경 회로망이 여러분야의 관심의 대상이 되고 있는데, 기존의 원형정합(template matching), 통계적(statistical) 및 구문론적(syntactic) 접근방법에 비해 신경 회로망이 적용되는 경우 기존의 방법들이 잘 해결하지 못하는 여러 분야에서 좋은 성능이 발휘되고 있는게 사실이다.

신경 회로망은 기본적으로 많은 뉴런(neuron)과 특별한 구조를 갖는 상호접속으로 구성된 것으로서, 병렬분산처리에 의거하여 입력 데이터에 대한 출력 데이터로의 비선형 사상(mapping)에 적합하고 주변 환경에 따라서 그들의 동작을 수정할 수 있는 유연성을 가지고 있다. 결국, 비선형 사상을 통해서 계통 특성을 학습(learning)하는 능력을 가지기 때문에 비선형 특성에 의해 최종 결과가 결정되는 여러 분야에 잘 적용될 수 있다.

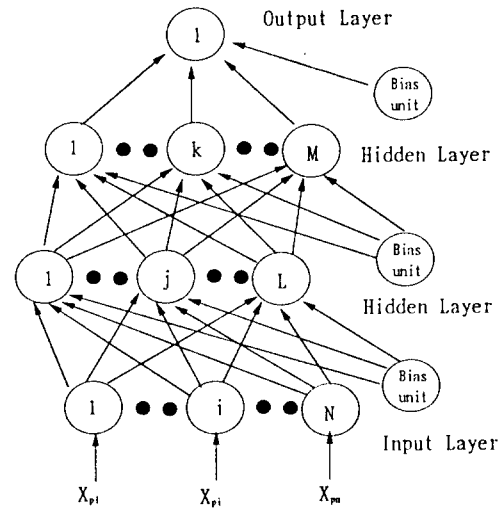
본 논문에서는 원액혼합과정을 포함하는 생산시스템 구축에 있어서, 최적화된 원액혼합비를 추론을 위해 신경 회로망과 퍼지 이론을 도입하여 시스템을 구성하고 그 결과를 확인하고자 한다.

2. 시스템 구성

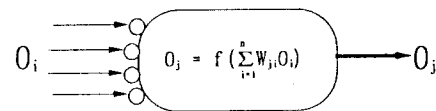
2.1 혼합비를 최적 추론을 위한 신경회로망 구조

혼합비를 최적추론을 위한 시스템 구성에 있어서 사용된 다층 신경 회로망 모델(multi-layered neural network model)은 구조의 단순함과 많은 분석, 그리고 강력한 문제해결 능력으로 가장 널리 사용 되어지고 있다. 일반적으로 다층 신경 회로망 모델은 계단형(hard limiter), 임계논리형(threshold logic), S 자형(sigmoid) 전이함수(activation function) 특성을 가지는 처리기(PE : Processing Element)들이 <Fig. 1>과 같이 사이클이 존재하지 않는 전방향 네트워크로 구성된 것이다.

다층 신경 회로망 모델은 기본 구성요소인 단층 지각소 모델(single-layered perceptron model)을 이용하여 입력층(input layer)과 은닉층(hidden layer), 그리고



(a)



(b)

Fig. 1 (a) Multilayer perceptron architecture and (b) Relationship between Neuron Input and Output in units

출력층(output layer)으로 구성된다.⁽¹⁾⁻⁽³⁾ 일반적으로 다층 신경 회로망 모델을 택한 시스템의 경우 임의의 비선형(nonlinear) 연속함수를 구성할 수 있으며 또 최근에는 S 자형 전이함수를 갖는 삼층 지각소 모델로서 임의의 비선형 연속함수를 근사시킬 수 있다는것이 증명되었다.⁽⁴⁾

혼합비를 최적추론을 위한 신경 회로망 시스템은 <Fig. 2>와 같은 S 자형 전이함수를 처리기의 활성화 함수로 사용하고 은닉층과 출력층에 속하는 처리기들의 입력(net input)은 이전 노드(node)의 가중된(weighted) 합으로 이루어진다. W_{ji} 를 처리기(PE_i)에서 처리기(PE_j)로의 연결강도라 할때 처리기(PE_j)의 출력

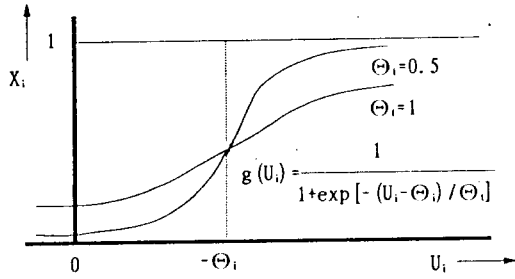


Fig. 2 The Sigmoid activation function, with bias and shape modification.

O_i 는

$$O_i = f(\text{net}_i) \quad (1)$$

$$\text{net}_i = \sum_{j=1}^n (W_{ji} O_j) \quad (2)$$

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (3)$$

로 결정된다.

결국, 다층 신경 회로망은 임의의 입력값에 대하여 원하는 출력값으로 변환되도록 할 수 있다. 이는 입력층과 출력층 사이에 있는 처리기에서의 비선형적 변환에 의한 특성으로서 임의의 입력값에 대한 사상의 경우 입력층과 출력층의 처리기는 선형 활성화 함수를 갖고, 은닉층의 처리기는 비선형 활성화 함수를 갖게 된다. 보통 비선형 활성화 함수는 단조증가하며 혼합비율 추론을 위한 신경 회로망에서는 미분가능한 함수인 sigmoid 함수를 사용하였다.

2.2 다층 신경 회로망의 학습

다층 신경 회로망의 학습¹⁰⁾은 널리 알려진 바와같이 오류 역전파 학습 알고리즘(error backpropagation learning algorithm)에 의해 실행된다. 오류 역전파 학습 알고리즘은 임의의 입력값에 대한 출력을 외부에서 가르쳐 주는 supervised 학습 방법에 의하여 원하는 출력값으로 변환되도록 다음과 같은 error measure function(cost function)

$$E = \sum_j E_j = 1/2 \sum_j \sum_i (t_{ji} - O_{ji})^2 \quad (4)$$

t_{ji} : 출력층 PE_j의 원하는 출력값

O_{ji} : 출력층 PE_j의 실제의 출력값

이 연결공간(weight space)에서 gradient descent 방법으로 감소하도록 처리기와 처리기 사이의 연결강도를 조정하는 방법이다. 그러므로, E를 최소화 하는 방향 $-\partial E / \partial W_{ji}$, 즉 gradient descent를 구해 그 반대방향으로 가중치를 변화 시킴으로서 학습을 진행한다. 이를 위해 $-\partial E / \partial W_{ji}$ 를 구해보면 오류 역전파 학습 알고리즘은 다음과 같이 요약될 수 있다.

$$\Delta_j W_{ji} = \eta (-\partial E / \partial W_{ji}) = \eta \delta_{ji} O_{ji} \quad (5)$$

η : 학습율 (learning rate)

여기에서 다층 신경회로망의 연결강도 변화는 출력층과 은닉층의 처리기에서 다음과 같이 다르게 정의된다.

출력층에서는

$$\delta_{ji} = (t_{ji} - O_{ji}) f'_{ji}(\text{net}_{ji}) \quad (6)$$

은닉층에서는

$$\delta_{ji} = \sum_k \delta_{kj} W_{ki} f'_{ji}(\text{net}_{ji}) \quad (7)$$

$$f'(\text{net}_{ji}) = O_{ji}(1 - O_{ji}) \quad (8)$$

로 계산된다.

2.3 정규화

신경 회로망 구성시 처리기나 가중치의 상태들을 나타내는 방법은 매우 중요한 부분중의 하나이다. 여러 분야에서 입력값과 출력값의 표현 방법에 따라 똑같은 시스템에서도 여러가지 다른 형태의 성능을 보여주는 것을 알 수 있다. 이를 위해 Hancock¹¹⁾, Takeda와 Goodman¹²⁾등이 여러가지 표현 방법들을 제시하기도 하였다.

값의 표현방법 이외에도 신경 회로망의 입력층에서 다른 성격의 입력값이 사용되는 경우 이 값들을 그대로 신경 회로망의 입력값으로 사용한다면 학습을 진행할 수 없다. 그 이유는 입력값의 존재범위가 너무크고, 서로의 입력값에 대한 연관성이 적어 수렴조건 성

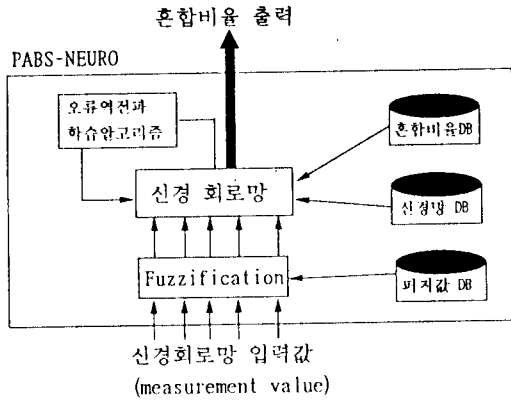


Fig. 3 Block diagram of PABS-NEURO
(PABS-NEURO : PASTE Blending System
using Neural Network)

립이 어렵기 때문이다. 오류 역전파 학습 알고리즘은 각 층에서 생긴 오차값(δ)을 입력값에 비례하여 가중치의 수정량을 결정하는데 $X_1 \gg X_2$ 인 두 입력값에 대하여 가중값의 수정량은 X_2 에 연결된 ΔW_1 이 X_1 에 연결된 ΔW_2 에 비하여 매우 적어지게 되므로 역전파되는 오차의 값은 X_1 에만 영향을 미치게 되어 X_2 는 전혀 학습에 관여를 하지 못하는 현상을 초래한다.

$$\Delta W_1 = -\eta \frac{\partial E}{\partial W_1} = \eta \delta X_1, \quad (9)$$

$$\Delta W_2 = -\eta \frac{\partial E}{\partial W_2} = \eta \delta X_2, \quad (10)$$

$$X_1 \gg X_2 \rightarrow \Delta W_1 \gg \Delta W_2 \quad (11)$$

이런 문제점들의 해결을 위해 신경회로망의 입력값을 퍼지소속함수를 통해 계산된 퍼지값으로 정규화 시켜 사용할 수 있다. 주로 사용된 퍼지소속함수는 <Fig. 2>에 나타난 sigmoid 함수이며 각각의 입력값의 특성에 따라 서로 다른 형태의 퍼지소속함수를 사용할 수 있다.

3. 실험 결과

<Fig. 3>는 신경 회로망을 이용하여 혼합비율을 추론하는 시스템의 전체구성도이다. 오류 역전파 학습방법을 적용한 다층 신경 회로망 시스템을 두가지 이상의 저항체 원액을 혼합하여 새로운 저항값의 저항체를 만

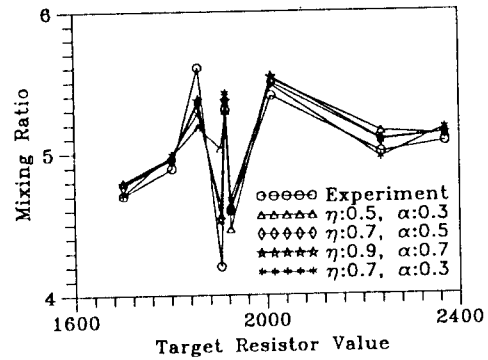


Fig. 4 Calculated Mixing Ratio with different learning rate(η) and momentum rate(α)

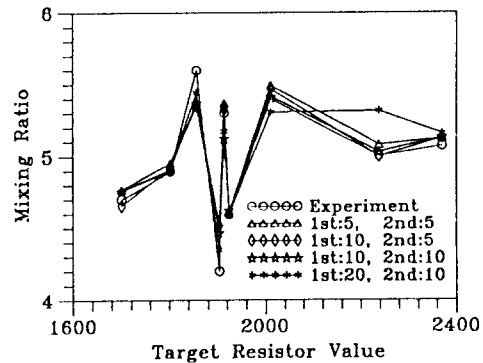


Fig. 5 Calculated Mixing ratio with different node numbers in hidden layer.

들어내는 저항제품 생산시스템 구축에 적용하였다. 두 저항체 원액의 혼합 후 새로운 저항값을 갖는 저항제품이 생산되기까지의 일련의 제조공정에서 많은 인자가 최종 저항값 산출에 영향을 미치게 되므로 이들 인자들을 신경회로망의 입력값으로 사용하였다. 특히 서로 다른 성격의 관련 인자들을 처리하고 원액 저항값과 같은 입력값은 신경회로망의 S 자형 전이함수의 초기 포화 현상에 의한 학습의 어려움을 방지하기 위해 데이터 베이스내에 존재하는 해당제품 저항값의 평균값을 이용하여 정규화 과정을 통한 퍼지값으로 변환하여 사용하였다. 입력층에서의 노드의 갯수는 관련 인자의 수와 동

일하게 처리하였고, 출력층의 노드수는 하나로 하여 최종적으로 혼합비율 추론 결과를 출력토록 하였다. 보통 신경 회로망을 사용하여 시스템을 구축할때 입력층과 출력층의 노드수는 시스템의 입력과 출력값의 갯수에 따라 결정되지만, 은닉층의 노드수를 결정하는 확정된 알고리즘이 아직 발견되고 있지 않으므로, 경험적으로 여러 경우의 은닉층 노드수를 테스트하여 수정해 나갔다.

학습에 사용된 값중 학습율(η)와 합성율(α)는 가능한 여러가지 값을 사용함으로써 최적의 추론 결과를 유출할 수 있도록 하였으며, 초기의 연결강도는 -1 과 +1 사이의 임의의 값을 사용하였다.

<Fig. 4>는 학습율(η)와 합성율(α)에 따른 혼합비율의 추론 결과를, <Fig. 5>는 2개의 은닉층으로 구성된 신경 회로망의 경우 은닉층의 노드수에 따른 혼합비율의 추론 결과중 한정된 부분만 표현한 것이다. 실험결과 혼합원액의 종류와 저항값에 따라 서로 다른 은닉층의 수와 노드수, 학습율, 합성율이 최적의 혼합비율 추론을 위해 필요함을 알게되었다. 결국, 전체 저항값(10 Ω - 100M Ω) 제품의 혼합비율을 모두 처리할 수 있는 신경 회로망은 구성할 수 없으며, 각 원액의 저항값과 혼합후 결과 저항값을 기준으로 서로 다른 신경 회로망의 구성이 필요하게 되었다.

4. 결 론

본 연구에서는 과거 숙련된 작업자에 의해 수동으로 행해지던 원액혼합과정을 신경 회로망과 퍼지값을 이용하여 최적추론 할 수 있도록 시스템을 구성하였다. 과거의 축적된 데이터나 경험적으로 얻은 정보와 휴리스틱 원리에 근거한 원액혼합비율 계산 방법과는 달리, 입력 데이터에 대한 출력 데이터로의 비선형 사상 관계를 통한 원액혼합비율 추론에 신경 회로망의 병렬 분산 처리 능력을 이용하여 관련인자들의 연관성을 처리함으로써 보다 정확한 혼합비율을 계산토록 하였다. 위와같이 신경 회로망을 이용함으로써 보다 나은 성능의 시스템을 구축할 수 있도록 하였다.

참고문헌

1. Yoh-Han Pao, "Adaptive Pattern Recognition and Neural Networks", Addison Wesley, 1989.
2. James A. Freeman and David M. Shapura, "Neural Networks - Algorithms, Applications, and Programming Techniques", Addison Wesley, 1991.
3. Patrick K. Simpson, "Artificial Neural Systems - Foundations, Paradigms, Applications, and Implementations", PERGAMON, 1989.
4. K. Fukushima, "Necognitron : A hierarchical neural network capable of visual pattern recognition," Neural Networks 1(2), pp. 119-130, 1988.
5. O. K. Ersoy and D. Hong, "Neural Network Learning Paradigms Involving Nonlinear Spectral Processing", Proceedings of ICASSP 89, Glasgow, Scotland, U. K., May, 1989.
6. Peter J. B. Hancock, "Data representation in neural nets : an empirical study", Preceeding of 1988 Connectionist Models, Summer School, Morgan Kaufmann Publishers Inc., p. 11-20, 1988.
7. M. Takeda and J. W. Goodman, "Neural Networks for Computation : Number Representations and Programming Complexity," Applied Optics, Vol. 25, no.18. September 1986.