

論文93-30B-8-7

초기값의 최적 설정에 의한 최적화용 신경회로망의 성능 개선

(Improving the Performances of the Neural Network for Optimization by Optimal Estimation of Initial States)

趙鏞鉉*, 崔興文**

(Yong Hyun Cho and Heung Moon Choi)

要約

본 논문에서는 전역최적해로 수렴 가능한 초기상태를 효과적으로 설정하는 방안을 제시하고 최적화용 신경회로망의 성능 개선 방법을 제안하였다. 제안된 방법에서는 확률적 근사법을 사용하여 전역최적해 근처에 초기상태를 설정한 후에, 홉필드의 상태경신 알고리즘을 사용함으로써 빠른 수렴속도로 전역최적해를 구할 수 있도록 하였다. 제안된 방법의 타당성과 성능을 확인하기 위하여 외판원 경로 문제와 최적 태스크 분할에 적용하여 시뮬레이션하였다. 그 결과 제안된 방법은 기존의 홉필드 신경회로망이나 Abe의 방법에서 랜덤하게 각 뉴우런의 초기상태 값을 설정하는 경우에 비해 전역최적해로의 수렴특성이 우수하며, 또한 볼츠만 신경회로망에 비하여 수렴속도가 빠름을 확인할 수 있었다. 또한 최적화에 있어서 문제의 규모가 클수록 초기상태 값의 설정이 어려우며 제안된 방법은 이와같이 문제 규모가 클수록 그 성능 개선의 정도가 큼을 확인하였다.

Abstract

This paper proposes a method for improving the performances of the neural network for optimization by an optimal estimation of initial states. The optimal initial state that leads to the global minimum is estimated by using the stochastic approximation. And then the update rule of Hopfield model, which is the high speed deterministic algorithm using the steepest descent rule, is applied to speed up the optimization. The proposed method has been applied to the travelling salesman problems and an optimal task partition problems to evaluate the performances. The simulation results show that the convergence speed of the proposed method is higher than conventional Hopfield model, Abe's method and Boltzmann machine with random initial neuron output setting, and the convergence rate to the global minimum is guaranteed with probability of 1. The proposed method gives better result as the problem size increases where it is more difficult for the randomized initial setting to give a good convergence.

*正會員, 嶺南專門大學 電子科
(Dept. of Elec. Yeungnam Junior College)
接受日字: 1992年 8月 24日

**正會員, 慶北大學校 電子工學科
(Dept. of Elec. Kyungpook Nat'l Univ.)

I. 서론

최근 조합형 최적화(combinatorial optimization)와 같이 NP-complete 문제들을 해결하기 위하여 대규모 병렬성을 가진 신경회로망을 이용하고자 하는 연구가 많이 진행되고 있다.^[1-9] 일반적으로 최적화에서 국소최적해가 아닌 전역최적해를 얻기 위한 보장 및 그 수렴속도는 주로 에너지 함수의 가중치 파라미터 및 뉴우런의 초기출력 설정에 매우 의존한다.

최적화 알고리즘에는 수치적인 방법과 신경회로망을 이용하는 방법이 있다. 수치적인 방법에는 기울기 경사법(steepest descent) 및 공액 기울기법(conjugate gradient method)등^[10] 과 같은 결정론적 방법과 아닐링(annealing)의 개념을 이용한 시뮬레이티드 아닐링(simulated annealing : SA)^[11], 전역해로의 수렴속도를 개선하기 위하여 빠른 온도기획 및 코오시 분포의 잡음을 이용한 빠른 시뮬레이티드 아닐링(fast SA : FSA)^[12,13], 그리고 평활화 성질을 가지는 확률적 근사법(stochastic approximation with smoothing)^[14] 등과 같은 확률적 방법이 있다. 한편 신경회로망을 이용하는 방법에도 홉필드 신경회로망(Hopfield model)^[1-4] 과 같은 결정론적 방법과 볼츠만 신경회로망(Boltzman machine)^[1-6], 그리고 가우스 신경회로망(Gaussian machine) 등^[7,8] 과 같은 확률적 방법이 있다. 일반적으로 결정론적 방법들은 수렴속도가 빠르며 초기값을 잘 설정하면 전역최적해를 얻을 수 있으나 국소최적해를 만났을 때 이를 벗어 날 수 없는 반면, 확률적 방법들은 전역최적해로의 수렴은 보장되나 수렴속도가 늦은 문제점이 있다.

최적화용 신경회로망에서 각 뉴우런의 출력값으로 표현되는 에너지함수는 여러개의 국소최소점을 가지는 다중극점(multiextremal)함수일 수 있다. 이러한 에너지함수에서 전역최적해를 찾을 수 있으면서도 수렴속도가 빠른 알고리즘의 개발이 요구된다. 이를 위한 연구들로는 각 뉴우런의 전달함수를 SA 과정으로 수정하는 불연속 시간 홉필드 신경회로망인 볼츠만 신경회로망^[5], 각 뉴우런의 입력측에 가우스 분포 잡음을 인가한 후 온도의 순차적 감소로 전역최적해를 찾는 가우스 신경회로망^[7], 그리고 FSA의 병렬화 모델인 코오시 신경회로망(Cauchy machine)^[8] 등의 확률적 방법들이 있다. 그러나 이러한 연구들은 전역최적해로의 수렴과 그 수렴속도가 온도기획 및 뉴우런 전달함수의 기울기등에 크게 의존하는 단점을 가지고 있다. 또한 Abe는 홉필드 신경회로망에서 국

소최적해로 수렴하는 영역을 해석하여 가중치 파라미터의 구간을 적절히 설정하고, 뉴우런의 초기출력을 초월입방체의 중심 근처에 설정함으로써 전역최적해로의 수렴확률을 개선하였다. 그러나 이 방법에서는 뉴우런의 초기출력 설정이 경험적이며 해의 질이 정의된 에너지함수에 크게 의존한다.^[15]

지금까지 발표된 최적화 신경회로망에 관한 연구들에서는 주로 전역최적해로의 수렴과 수렴속도의 개선을 위하여 온도기획 방법이나 에너지함수의 가중치 파라미터 조정등을 시도하고 있으며, 뉴우런의 초기출력을 전역최적해 근처에 설정함으로써 수렴속도와 전역최적해를 보장하려는 체계적인 방법은 찾아보기 힘들다.

본 논문에서는 전역최적해로 수렴하기 용이한 적정 초기상태를 설정함으로써 최적화용 신경회로망의 성능을 개선할 수 있는 방법을 제안하였다. 제안된 방법에서는 확률적 근사법(stochastic approximation method)을 사용하여 전역최적해가 보장되는 뉴우런의 초기출력을 설정한 다음, 이를 초기값으로 하여 홉필드의 상태경신 알고리즘인 기울기 경사법을 이용함으로써 빠른 수렴속도로 전역최적해를 구할 수 있도록 하였다. 제안된 방법을 외판원 경로 문제(travelling salesman problem : TSP)와 최적 태스크 분할(optimal task partition)문제에 적용 시뮬레이션하여 각 뉴우런의 초기출력을 랜덤하게 설정하는 홉필드 신경회로망, Abe의 방법 및 볼츠만 신경회로망을 이용한 결과와 각각 비교 검토 하였다.

II. 확률적 근사법과 홉필드 신경회로망에 의한 효율적 최적화

평활화(smoothing) 기능을 가지는 확률적 근사법은 n 차원의 상태 $x \in R^n$ 에 대한 다중극점함수 $f(x) \in R^1$ 에 평활화 함수 $h(\gamma)$ 를 상승적분(convolution)하여 평활된 함수 즉, 평균화된 함수 $f^-(x, \beta)$ 를 구하고, 이 평활된 함수를 이용하여 $h(\gamma)$ 에 대한 분산제어 파라미터 β 를 순차적으로 조정함으로써 전역최적해를 찾는 최적화 알고리즘이다.^[14] 여기서 랜덤교란벡터(random perturbation vector) γ 를 생성하는 평활화 함수 $h(\gamma)$ 로는 가우스, 균일(uniform), 또는 코오시 확률밀도함수등이 이용될 수 있다. 예로서 연속이고 미분 가능한 2개의 다른 최소점을 가지는 함수 $f(x)$

$$f(x) = x^4 - 16x^2 + 5x \quad (1)$$

에 확률적 근사법을 적용하면 평활된 함수 $f^-(x, \beta)$

는

$$f^{\sim}(x, \beta) = (1/2)[f(x + \beta\gamma) + f(x - \beta\gamma)] \quad (2)$$

와 같이 된다.

그림 1은 분산제어 파라미터의 순차값을 $\{\beta\} = \{5.0, 4.0, 3.0, 2.0, 1.0, 0.001\}$ 로 하고 평활화 함수 $h(\gamma)$ 로 균일 확률밀도함수, 그리고 랜덤 초기상태 $x_0 = 4.0$ 일 때에 각 β 값 대한 $f(x)$ 의 평활된 함수 $f^{\sim}(x, \beta)$ 를 나타낸 것이다.

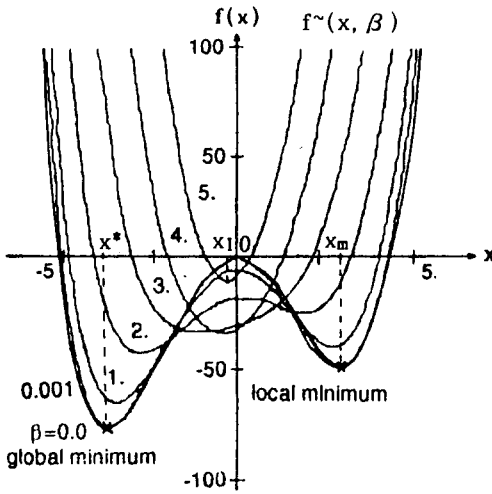


그림 1. 분산제어 파라미터 β 의 변화에 대한 평활된 함수 $f^{\sim}(x, \beta)$

Fig. 1. Smoothed function $f^{\sim}(x, \beta)$ for each of the β changes.

그림 1에서 보는 바와 같이 확률적 근사법을 이용하여 전역최적해 x^* 를 구하는 데는 β 의 순차값(β)의 결정, 각 β 값에 대한 제어 파라미터의 설정뿐만 아니라, 확률적 동작으로 느린 수렴속도, 그리고 원래 함수에 대한 정확한 해가 아닌 근사해가 구해지는 등의 문제점이 있다. 한편 β 값이 5.0, 4.0, 및 3.0일 때를 보면 이처럼 큰 β 값에 대해서는 $f^{\sim}(x, \beta)$ 가 단봉특성을 가짐을 알 수 있다. 이때 각 β 에 대한 평활된 함수 $f^{\sim}(x, \beta)$ 의 최소값은 원래의 다중극점함수 $f(x)$ 의 전역최적해의 근사값이 된다.^[14] 여기서 초기 β 는 전역최적해의 근사적인 위치를 결정하기 위해 큰 β 값으로 하여 간단한 경험적인 판단에 기초하여 결정된다.^[14] 따라서 초기에 큰 값으로 확률적 근사법을 적용하여 다중극점함수 $f(x)$ 의 전역최적해를 보장하는 상태 x_i 를 구하고, 구해진 상태 x_i 를

초기값으로 하여 기울기 경사법을 적용함으로써 빠른 속도로 원래 함수 $f(x)$ 의 전역최적해를 찾을 수 있게 된다.

따라서 상태 벡터 x 와 함수 $f(x)$ 를 신경회로망의 상태 벡터 x 와 에너지함수 $E(x)$ 로 대체하면 확률적 근사법을 신경회로망에 적용할 수 있다. 이 때 에너지함수 $E(x)$ 와 평활된 함수의 기울기 $\nabla_x E^{\sim}(x, \beta)$ 은 다음과 같다. 즉,

$$E(x) = -(1/2) \sum_i \sum_j T_{ij} x_i x_j - \sum_i I_i x_i \quad (3)$$

$$\begin{aligned} \nabla_x E^{\sim}(x, \beta) &= (1/2)(\beta/\gamma)[E(x + \beta\gamma) - E(x - \beta\gamma)] \\ &= -(1/2)\gamma^2 \sum_i \sum_j T_{ij}(x_i + x_j) - \gamma^2 \sum_i I_i \\ &= -(1/2)\gamma^2 [(\sum_i \sum_j T_{ij} x_i + \sum_i I_i) + (\sum_i \sum_j T_{ij} x_j \\ &\quad + \sum_i I_i)] \end{aligned} \quad (4)$$

이다. 여기서 i 와 j 는 뉴우런을 나타내는 번호들이다.

그러므로 확률적 근사법과 홉필드의 상태경신 알고리즘을 이용하여 최적화의 수렴특성을 개선하기 위해 제안한 최적화 알고리즘을 다음과 같이 정리할 수 있다. 즉,

- 단계 1: 최적화 문제의 목적함수와 제약조건을 포함하는 비용함수 $f(x)$ 를 정의한다.
- 단계 2: 정의된 비용함수를 신경회로망의 에너지함수 $E(x)$ 와 사상시켜 연결세기 행렬 및 입력력 바이어스를 구한다.
- 단계 3: 에너지함수 $E(x)$ 에 대한 평활된 함수의 기울기 $\nabla_x E^{\sim}(x, \beta)$ 를 계산한다.
- 단계 4: 분산제어 파라미터 β 를 크게 하고 뉴우런의 초기출력 x_0 를 랜덤하게 설정하여 평활된 함수의 기울기를 이용한 확률적 근사법을 수행한다.
- 단계 5: 단계 4에서 구해진 뉴우런의 출력을 초기값으로 하여 홉필드의 상태경신 알고리즘을 적용한다.
- 단계 6: 단계 5에서 구해진 뉴우런의 출력에 대한 에너지함수 값이 허용치 이하이면 알고리즘을 종료하고, 그렇지 않으면 단계 3으로 간다.

이와같이 제안된 방법에서는 초기에 확률적 근사법을 이용하여 뉴우런의 초기출력을 설정하고 홉필드의 상태경신 알고리즘을 적용한다. 여기서 확률적 근사법은 확률적 방법으로 최적해로의 수렴속도가 느리

다. 한편 뉴우런의 상태경신 알고리즘으로 SA를 이용하는 볼츠만 신경회로망과 같은 기존의 확률적 신경회로망에서는 온도기획이 요구되며, 이들 신경회로망에서 최적해로의 수렴속도는 온도기획에 따라 달라진다. 그러나 제안된 방법에서는 확률적 근사법이 전역최적해로의 수렴을 확실히 하는 뉴우런의 초기출력을 찾는 데에만 한 번 이용되므로 분산제어 파라미터의 순차값 $\{\beta\}$ 내의 여러개 β 를 전부 결정하여야 하는 어려움은 없다. 또한 제안된 방법에서는 뉴우런의 초기출력을 설정한 다음에 홉펠드의 상태경신 알고리즘을 적용함으로써 기존의 확률적 방법을 이용하는 신경회로망에서 온도기획에 따른 수렴속도의 문제를 없앨 수 있다.

III. 최적화 문제의 응용

잘 알려진 바와 같이 외판원 경로 및 최적 태스크 분할은 NP-complete 최적화 문제이며, 이러한 최적화 문제를 신경회로망으로 풀기 위해서는 일반적으로 먼저 문제의 정확한 정의, 목적함수, 제약조건, 그리고 목적함수와 제약조건을 포함하는 비용함수를 정의하여야 한다. 또한 정의된 비용함수를 신경회로망의 에너지함수와 비교하여 연결세기 행렬 및 입력 바이어스를 구하고, 가중치 값들을 적절히 설정한 후 해당 신경회로망을 구동시킨다.

1. 외판원 경로 문제

외판원 경로 문제는 n 개의 도시가 주어졌을 때 임의의 한 도시에서 출발하여 모든 도시를 한 번씩만 방문하고 다시 출발했던 도시로 되돌아오는 여러가지 경로중에서 총 거리의 합이 최소가 되는 경로를 구하는 최적화 문제로서, 다음과 같은 목적함수 및 제약조건들로 정의된다. 즉,

$$\min f(x) = \sum_i^{n-1} \sum_j^{n-1} \sum_m^{n-1} d_{ij} x_{im} x_{j(m+1 \bmod n)} \quad (5)$$

subject to

$$\sum_{i=0}^{n-1} x_{im} = 1, \quad m = 0, 1, \dots, n-1 \quad (6)$$

$$\sum_{m=0}^{n-1} x_{im} = 1, \quad i = 0, 1, \dots, n-1 \quad (7)$$

이다. 목적함수 (5)에서 i, j 는 도시의 좌표, m 는 방문순서를 의미하며, x_{im} 은 i 번째 도시가 m 번째에 방문되면 1 그렇지 않으면 0인 할당행렬 요소값이다. 따라서 각 방문 도시간의 거리 d_{ij} 가 최소가 되도록

모든 도시를 방문하였을 때 목적함수의 값은 최소가 된다. 제약조건 (6)과 (7)은 각각 어느 한 도시만을 반드시 한 번은 방문해야 되고, 매 방문 순서마다 반드시 한 도시를 방문하여야 한다는 조건이다.

이상의 목적함수와 제약조건들을 각각 에너지함수 $E_{obj}(x)$ 와 $E_{con}(x)$ 로 표현하면 다음과 같다. 즉,

$$E_{obj}(x) = (A/2) \sum_i^{n-1} \sum_j^{n-1} \sum_k^{n-1} d_{ij} x_{ik} (x_{j,k+1} + x_{j,k-1})$$

$$E_{con}(x) = (B/2) \sum_i^{n-1} \sum_k^{n-1} \sum_{l \neq k}^{n-1} x_{ik} x_{il} + (C/2) \sum_i^{n-1} \sum_k^{n-1} \sum_{l \neq k}^{n-1} x_{ik} x_{li} + (D/2) \left(\sum_i^{n-1} \sum_j^{n-1} x_{ij} - n \right)^2$$

이 된다. 여기서 A, B, C, 그리고 D는 가중치 파라미터들이며, n 는 전체 도시 수이다.

따라서 외판원 경로 문제의 전체 에너지함수 $E_{TSP}(x)$ 는 다음과 같다.

$$E_{TSP}(x) = E_{obj}(x) + E_{con}(x) \quad (8)$$

한편 홉펠드 신경회로망과 볼츠만 신경회로망의 표준 에너지함수 식은 각각 다음과 같다. 즉,

$$E(x) = \sum_i^{n-1} \sum_j^{n-1} \sum_k^{n-1} \sum_l^{n-1} T_{ijkl} x_{ij} x_{kl} + \sum_i^{n-1} \sum_j^{n-1} I_{ij} x_{ij} \quad (9)$$

$$E(x) = \sum_i^{n-1} \sum_j^{n-1} \sum_k^{n-1} \sum_l^{n-1} T_{ijkl} x_{ij} x_{kl} \quad (10)$$

이다. 식 (8)을 식 (9) 및 (10)과 각각 비교하여 홉펠드 신경회로망 및 볼츠만 신경회로망의 연결세기 행렬 T_{ijkl} 및 입력 바이어스 I_{ij} 를 구한 후, 가중치 파라미터 값을 설정하여 각 신경회로망을 적용한다.

2. 최적 태스크 분할

최적 태스크 분할은 n 개의 부태스크들로 구성된 태스크를 p 개의 프로세서에 분할하는 것으로 전산처리량 및 통신량의 균형있는 분할을 목적으로 하며 보통 $n \geq p$ 의 관계를 갖는 최적화 문제로서, 다음과 같은 목적함수 및 제약조건들로 정의된다. 즉,

$$\min f(x) = \sum_i^p \sum_j^p \sum_k^p W_i W_j x_{ik} x_{jk} + \sum_i^p \sum_j^p \sum_k^p e_{ij} x_{ik} x_{jk} \quad (11)$$

subject to

$$\sum_i^p \sum_j^p \sum_k^p x_{ik} x_{jk} = 0$$

$$\sum_i^p \sum_k^p x_{ik} = n \quad (12)$$

이다. 목적함수 (11)에서 i, j 는 부태스크의 번호이며 k 와 l 은 프로세서 번호이다. e_{ij} 는 i 번째 부태스크와 j 번째 부태스크 사이의 통신량이며, W_i 는 i 번째 부태스크의 전산처리량이다. x_{ik} 는 i 번째 부태스크가 k 번째 프로세서에 할당되는지 여부를 나타내는 할당행렬 요소값으로, 할당되면 1, 그렇지 않으면 0의 값을 갖는다. 식 (11)에서 첫째 항은 전산처리량의 균형을 위한 항이고, 둘째 항은 통신량의 균형을 위한 항이다. 제약조건 식 (12)는 각 프로세서에는 한개 이상의 부태스크가 할당되어야 하며, 각 부태스크는 반드시 어느 하나의 프로세서에라도 할당되어야만 한다는 제약조건이다.

이상의 목적함수와 제약조건들을 각각 에너지함수 $E_{obj}(x)$ 와 $E_{con}(x)$ 로 표현하면 다음과 같다. 즉,

$$E_{obj}(x) = (A/2) \sum_i^n \sum_j^n \sum_k^p W_i W_j x_{ik} x_{jk} + (B/2) \sum_i^n \sum_j^n \sum_k^p \sum_l^p e_{ij} (x_{ik} x_{jl} - 2x_{ik} x_{jl}) + (C/2) \left(\sum_i^n \sum_k^p x_{ik} - n \right)^2 + \sum_i^n \sum_k^p \sum_l^p x_{ik} x_{il}$$

이 된다. 여기서 A, B, C 및 D는 에너지함수의 가중치 파라미터로 상수값이다.

따라서 최적 태스크 분할의 전체 에너지함수 $E_{GTP}(x)$ 는 다음과 같다.

$$E_{GTP}(x) = E_{obj}(x) + E_{con}(x) \tag{13}$$

식 (13)을 각각 식 (9)와 (10)의 홉필드 신경회로망 및 볼츠만 신경회로망의 표준 에너지함수 식과 비교하여 연결세기 행렬 T_{ijk} 및 입력 바이어스 I_{ij} 를 구하고 가중치 파라미터 A, B, C, 그리고 D를 설정하여 홉필드 신경회로망, 볼츠만 신경회로망, 그리고 제안된 방법에 의하여 최적해를 얻을 수 있다.

VI. 시뮬레이션 및 결과 고찰

제안된 방법의 타당성과 성능을 확인하기 위하여 외판원 경로 문제에서 도시 수 n 이 7과 10인 두가지 문제와 최적 태스크 분할에서 부태스크 수 n 이 4와 6 이고 프로세서 수 p 가 각각 3인 문제들을 대상으로 시뮬레이션하였다. 그림 2는 외판원 경로 문제 각각의 도시좌표이며, 그림 3은 최적 태스크 분할 각각의 전산 및 통신량을 나타낸 것이다.

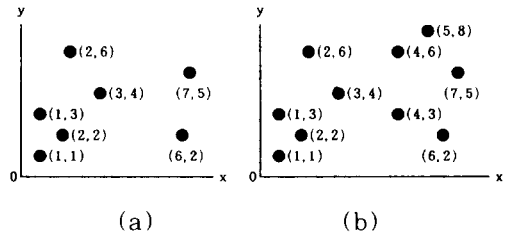


그림 2. (a) 7 도시 및 (b) 10 도시 외판원 경로 문제 예

Fig. 2. Examples of (a) 7-city and (b) 10-city travelling salesman problem.

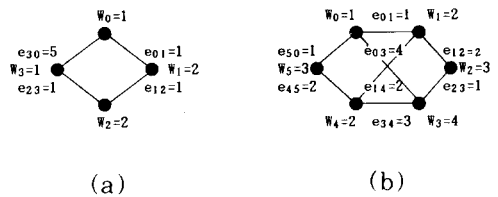


그림 3. (a) 4 개 부태스크와 3 개 프로세서 및 (b) 6 개 부태스크와 3 개 프로세서의 최적 태스크 분할 문제 예

Fig. 3. Examples of (a) 4-subtask and 3-processor, (b) 6-subtask and 3-processor optimal task partition.

그림 2에서 (x, y) 는 도시의 2차원 좌표이며, 그림 3에서 e_{ij} 는 i 와 j 번째 부태스크 사이의 통신량이고, W_i 는 부태스크 i 의 전산처리량이다. 시뮬레이션에서 각 뉴우런의 초기출력 값은 -0.5 와 $+0.5$ 사이의 랜덤값으로 random seed에 의하여 조정되도록 하였다. 이때 random seed에 따라 발생하는 상태값을 홉필드 신경회로망과 볼츠만 신경회로망에서는 초기값으로 사용하지만, 제안된 방법에서는 이 값을 확률적 근사식에 적용하여 최적 초기상태를 구한 다음 홉필드의 상태경신 알고리즘을 수행하게 된다. 여기서 이용되는 홉필드 신경회로망은 연속적인 모델이며, 감쇠정수는 0.001로 하였다. 외판원 경로 문제를 위한 가중치 파라미터의 값은 각각 $A = B = 200$, $C = 50$, 그리고 $D = 100$ 으로 하였으며, 최적 태스크 분할에서는 각각 $A = 38$, $B = 5$, $C = 20$, 그리고 $D = 30$ 으로 하였다. 볼츠만 신경회로망의 온도기획은 쌍곡선(hyperbolic) 기획으로 하여 $T = T_0 / (1 + k/\tau)$ 이며, 여기서 k , T_0 , 그리고 τ 는 각각 반복수, 초기온도, 그리고 시정수로 T_0 와 τ 는 각각 500과 5

로 하였다. 제안된 방법에서 평활화 함수 $h(\gamma)$ 는 균일 확률밀도함수로 하였고, 분산제어 파라미터 β 는 3.0으로 하였다. 알고리즘의 종료는 두 문제 모두 반복횟수(number of iterations)가 1500회 이상이거나, 외판원 경로 문제에서는 모든 뉴우런의 출력변화가 없을 때로 하였으며, 최적 태스크 분할에서는 에너지함수의 변화량이 0.0001 이하일 때를 종료시기로 하였다. 한편 외판원 경로 문제에 있어서 수정된 에너지함수^[15]를 이용하여 Abe의 방법 및 제안된 방법을 적용할 때에 가중치 파라미터의 값은 각각 $A = B = 200$ 및 $D = 100$ 으로 하였다.

그림 4는 그림 2(a)에서와 같은 7 개 도시 TSP에서 random seed를 0, 3 및 5로 하여 홉필드 신경회로망과 제안된 방법을 각각 적용한 결과로서 반복횟수 N_h 에 따른 정규화된 에너지함수 E 를 나타낸 것이다. 그림에서 보면 홉필드 신경회로망에서는 random seed가 0과 3인 경우에 반복횟수가 각각 196번째와 163번째에 최적해로 수렴하고 254번째와 286번째에 종료되었으나, 이 때의 최적해는 국소최적해이다. 또한 random seed가 5인 경우에는 반복횟수가 52번째에 최적해로 수렴되고 311번째에 종료되었다. 이것은 홉필드 신경회로망의 수렴특성이 초기값에 크게 의존함을 보여준다. 그러나 제안된 방법을 적용했을 때는 random seed에 관계없이 확률적 근사법과 홉필드의 상태경신 반복횟수를 합쳐서 69번째에 최적해로 수렴되고 313번째에 종료되었다.

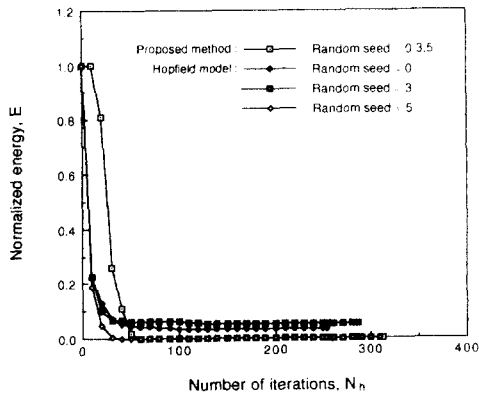


그림 4. 7 도시 외판원 경로 문제의 정규화된 에너지함수 E

Fig. 4. Normalized energy function E of 7-city travelling salesman problem.

표 1은 그림 12(a)의 7 개 도시 TSP에서 10 개의 random seed에 대한 실험결과를 나타낸 것이다. 표

에서 N_h , N_b 및 N_s 는 각각 홉필드, 볼츠만 신경회로망 및 확률적 근사법 내의 반복횟수이며, t_h 와 t_b 는 각각 N_h 와 N_b 에 소요된 CPU 시간이다. 또한 t_p 는 제안된 방법에서의 N_h 와 N_s 에 각각 소요된 CPU 시간의 합이다. E 는 종료시의 에너지함수 값이다. 표에서 *는 국소최적해로 수렴된 경우로 제약조건을 만족하지 못한 결과를 나타낸다. 표에서 나타난 바와 같이 홉필드 신경회로망에서는 random seed가 5, 15 및 30일 때에 제약조건을 만족하였으며, 그 중 5와 15일 때에만 볼츠만 신경회로망이나 제안된 방법의 결과와 동일한 에너지함수 값으로 수렴되었다. 특히 random seed가 5인 경우에는 볼츠만 신경회로망이나 제안된 방법에 비해 수렴속도가 더 빠른 데, 이는 홉필드 신경회로망에서 임의로 설정한 뉴우런의 초기출력이 우연히도 전역최적해 근처에 가깝게 설정된 경우로 보이며, 실제로 기존 홉필드 신경회로망의 경우 이와같이 적절하게 뉴우런 초기 출력의 설정을 보장하기 위한 체계적인 방안은 별로 없다. 또한 볼츠만 신경회로망에서는 최적해로의 수렴은 보장되나 그 수렴속도가 늦음을 확인할 수 있다. 제안된 방법에서는 확률적 근사법에 의하여 최적해로의 수렴을 확실히 하는 뉴우런의 출력을 결정한 다음 홉필드의 상태경신 알고리즘을 수행하며, 분산제어 파라미터 값 및 그 값에 대한 파라미터들을 고정한 결과 항상 동일한 반복횟수, 에너지함수 값 및 동일한 수렴시간으로 최적해로 수렴되었다. 또한 확률적 근사법의 1 회 반복에 소요되는 CPU 시간은 1 초로 홉필드 신경회로망의 약 0.89 초보다 오래 걸렸다. 그러나 제안된 방법에서는 확률적 근사법의 반복횟수가 적어 최적해로의 수렴속도에는 큰 영향을 미치지 않았다. 따라서 10 개의 random seed에 대한 실험결과를 평균

표 1. 7 도시 외판원 경로 문제의 실험결과
Table 1. Results of the 7-city travelling salesman problem.

Random seed	Hopfield model			Boltzmann machine			Proposed method			
	N_h	E_t	t_h	N_b	E_t	t_b	N_s	N_h	E_t	t_p
0	254	-3407*	226	197	-3481	317	1, 312	-3481	280	
3	286	-3346*	255	207	-3481	352	1, 312	-3481	280	
5	311	-3481	277	105	-3481	337	1, 312	-3481	280	
8	179	-3373*	132	318	-3481	393	1, 312	-3481	280	
10	342	-3356*	299	289	-3481	355	1, 312	-3481	280	
13	235	-3354*	208	674	-3481	504	1, 312	-3481	280	
15	351	-3481	310	118	-3481	346	1, 312	-3481	280	
30	343	-3454	304	113	-3481	334	1, 312	-3481	280	
50	190	-3398*	170	357	-3481	412	1, 312	-3481	280	
100	264	-3325*	232	406	-3481	481	1, 312	-3481	280	

t_i : CPU time in [sec]

해본 결과, 제안된 방법은 홉필드 신경회로망에 비해 최적해로의 수렴확률은 약 5 배 정도 개선되었으며, 최적해로 수렴된 경우에 대한 그 수렴속도는 약 1.2 배 정도 개선되었다. 그리고 볼츠만 신경회로망에 비해서는 같은 수렴확률에 대해 최적해로의 수렴속도가 약 1.4 배 정도 개선되었다.

표 2는 7 개 도시 TSP에서 100 개의 random seed에 대하여 최적해로 수렴된 경우에 대한 반복횟수 및 CPU 시간의 각각 평균 \bar{x} 와 표준편차 σ 를 나타낸 것이다. 표에서 나타난 바와 같이 제안된 방법은 홉필드 신경회로망에 비해 최적해로의 수렴확률 및 그 수렴속도에서 각각 약 2.5 배와 약 1.2 배 정도 개선되었음을 알 수 있다. 홉필드 신경회로망의 결과에서는 표준편차가 매우 큰 값이므로 그 수렴특성이 뉴우런의 초기값에 의존함을 확인할 수 있다.

표 2. 7 도시 외판원 경로 문제의 100 번 시도에 대한 실험결과

Table 2. Results of the 7-city travelling salesman problem for 100 trials.

	Hopfield		Proposed	
	\bar{x}	σ	\bar{x}	σ
N_h	368.9	124.9	313	0
t_h	328.2	111.1	279	0
P_r	40 %		100 %	

\bar{x} : Mean, σ : Standard deviation,
 P_r : Convergence ratio

표 3은 그림 12(b)에서와 같이 10 개 도시 TSP에서 10 개의 random seed에 대해 실험한 결과를 나타낸 것이며, 여기서도 *는 제약조건을 만족하지 못한 경우이다. 표에서 보는 바와 같이 홉필드 신경회로망에서는 random seed가 3, 5 및 30일 때에 제약조건을 만족하였다. 홉필드의 경우 random seed가 3일 때만 볼츠만 신경회로망이나 제안된 방법의 결과와 동일한 에너지함수 값을 갖는 최적해로 수렴되었다. 볼츠만 신경회로망에서는 10 개의 random seed 모두에 대해 최적해로 수렴되었으나 그 수렴속도가 늦음을 알 수 있다. 여기서도 확률적 근사법의 1 회 반복시간은 8 초로 홉필드 신경회로망의 약 3.5 초에 비하여 약 2.3 배 정도 더 오래 걸림을 알 수

있다. 따라서 10 개의 random seed에 대한 실험결과를 평균해 본 결과, 제안된 방법은 홉필드 신경회로망을 이용할 때보다 최적해로의 수렴확률은 약 9 배 정도 개선되었으며, 최적해로 수렴된 경우만을 고려할 때 수렴속도는 약 1.2 배 정도 개선되었다. 그리고 볼츠만 신경회로망에 비해서는 최적해로의 수렴속도만이 약 1.2 배 정도 개선되었다. 또한 표 3의 아랫쪽 표에서는 10 개 도시 TSP에 대하여 Abe의 방법과 제안된 방법을 각각 적용하여 실험한 결과를 비교하였다. 여기에서 그 결과는 뉴우런의 초기출력의 범위를결정하는 파라미터 α 를 1, 10^{-1} , 10^{-3} , 10^{-5} , 10^{-7} 및 10^{-9} 로 할 때, 그 각각에 대하여 100 개의 random seed에 대한 실험결과와 평균값이다. N_a 와 t_a 는 각각 Abe 방법의 반복횟수 및 CPU 시간이다. 표에서 나타난 것과 같이 Abe의 방법에서는 최적해로의 수렴정도과 그 수렴속도가 α 에 의존하며, α 의 값이 10^{-5} , 10^{-7} 및 10^{-9} 일 때에는 동일한 수렴특성을 나타내어 뉴우런의 초기출력이 0.5 가까이 설정되면 최적해로 수렴될 수 있음을 확인할 수 있다. 그러나 제안된 방법에서는 6 개의 α 및 그 각각에 대하여 100 개의 서로 다른 random seed를 이용하여도 동일한 결과를 나타내며, Abe 방법의 결과보다 우수한 수렴특성을 얻었다. 특히 Abe의 방법에서는 그 수렴특성이 random seed에 의존하나 제안된 방법에서는 그 수렴특성이 α 와 random seed에 거의 영향을 받지 않았다.

표 3. 10 도시 외판원 경로 문제의 실험결과

Table 3. Results of the 10-city travelling salesman problem.

Random seed	Hopfield model			Boltzmann machine			Proposed method			
	N_h	E_t	t_h	N_b	E_t	t_b	N_s	N_h	E_t	t_p
0	938	-7324*	3280	341	-7443	3790	2, 695	-7443	2447	
3	851	-7443	2976	329	-7443	3657	2, 695	-7443	2447	
5	1498	-7428	5238	338	-7443	3757	2, 695	-7443	2447	
8	571	-7222*	1997	382	-7443	4246	2, 695	-7443	2447	
10	1166	-7390*	4077	257	-7443	2856	2, 695	-7443	2447	
13	773	-7373*	2703	473	-7443	5257	2, 695	-7443	2447	
15	768	-7369*	2686	255	-7443	2834	2, 695	-7443	2447	
30	1498	-7412	5238	248	-7443	2756	2, 695	-7443	2447	
50	1185	-7367*	4144	257	-7443	2856	2, 695	-7443	2447	
100	540	-7389*	1888	403	-7443	4479	2, 695	-7443	2447	

α	Abe's method			Proposed method		
	N_a	E_t	t_a	N_s , N_h	E_t	t_p
1.0	199.4	-3501.0	697	4, 65	-3637	259
10^{-1}	192.6	-3503.9	674	4, 65	-3637	259
10^{-3}	232.2	-3506.7	812	4, 65	-3637	259
10^{-5}	228	-3508.1	796	4, 65	-3637	259
10^{-7}	228	-3508.1	796	4, 65	-3637	259
10^{-9}	228	-3508.1	796	4, 65	-3637	259

t_i : CPU time in [sec]

표 4에서는 10 개 도시 TSP에서 100 개의 서로 다른 random seed로 홉필드 신경회로망과 제안된 방법을 각각 적용한 결과, 최적해로 수렴된 경우에 대한 반복횟수 및 CPU 시간의 평균 \bar{x} 와 표준편차 σ 를 각각 나타낸 것이다. 표에서 보는 바와 같이 제안된 방법은 홉필드 신경회로망에 비해 최적해로의 수렴확률과 수렴속도가 각각 약 6.3 배 및 약 1.7 배 정도 개선되었다. 또한 여기서도 홉필드 신경회로망의 결과에서는 표준편차가 매우 큰 값이므로 그 수렴 특성이 뉴우런의 초기출력에 많은 영향을 받을 수 있다.

표 4. 10 도시 외판원 경로 문제의 100 번 시도에 대한 실험결과

Table 4. Results of the 10-city travelling salesman problem for 100 trials.

	Hopfield		Proposed	
	\bar{x}	σ	\bar{x}	σ
N_h	1147.6	404.9	692	0
t_h	4012.9	1415.9	2419.8	0
P_r	16%		100%	

\bar{x} : Mean, σ : Standard deviation
 P_r : Convergence ratio

관련 표 2와 4에서 보는 바와 같이 일반적으로 문제의 규모가 클수록 전역최적해로의 수렴확률 및 그 수렴속도 면에서 뉴우런의 초기출력의 설정은 더욱 어려우며, 그 성능 또한 초기값에 크게 의존함을 알 수 있다. 그러나 제안된 방법은 이런 경우에도 그 초기출력의 설정이 단순함을 확인할 수 있다.

표 5는 그림 3(a)의 최적 태스크 분할에 대하여 10 개의 random seed에 대한 실험결과를 나타낸 것이다. 여기서는 홉필드 신경회로망, 볼츠만 신경 회로망 및 제안된 방법 모두 10 개의 random seed에 대하여 최적해로 수렴되어 국소최적해가 존재하지 않는 문제임을 추측할 수 있다. 그러나 제안된 방법은 TSP에서 처럼 홉필드 신경회로망이나 볼츠만 신경회로망에 비해 각각 약 1.5 배 정도 및 약 2.7 배 정도 수렴속도가 더 빠름을 확인할 수 있다.

표 6은 그림 3(b) 문제에 대한 실험결과를 나타낸 것이며, *는 국소최적해로 수렴된 경우로 제약조건을 만족하지 못한 때이다. 표에서 보는 바와 같이 홉필드 신경회로망에서는 random seed가 10과 100일

표 5. 4 개 부태스크 및 3 개 프로세서 최적 태스크분할의 실험결과

Table 5. Results of the optimal task partition for the 4-subtask and 3-processor problem.

Random seed	Hopfield model			Boltzmann machine			Proposed method			
	N_h	E_t	t_h	N_b	E_t	t_b	N_s	N_p	E_t	t_p
0	466	-98	47	292	-98	51	5	371	-98	38
3	691	-98	68	580	-98	111	5	371	-98	38
8	629	-98	62	953	-98	209	5	371	-98	38
9	940	-98	91	861	-98	165	5	371	-98	38
10	345	-98	34	299	-98	55	5	371	-98	38
13	431	-98	43	327	-98	58	5	371	-98	38
15	458	-98	46	325	-98	57	5	371	-98	38
30	517	-98	53	422	-98	74	5	371	-98	38
50	719	-98	73	768	-98	135	5	371	-98	38
100	412	-98	42	539	-98	94	5	371	-98	38

t_i : CPU time in [sec]

때에만 최적해로 수렴되었고, 그 중에서도 10의 경우는 수렴속도가 볼츠만 신경회로망이나 제안된 방법보다 더 빠르다. 그러나 다른 8 개의 random seed에 대해서는 뉴우런의 초기출력이 국소최적해에 가까이 설정된 것으로 그 동작 특성상 국소최적해에 수렴되었다. 또한 볼츠만 신경회로망에서는 10 개의 random

표 6. 6 개 부태스크 및 3 개 프로세서 최적 태스크분할의 실험결과

Table 6. Results of the optimal task partition for the 6-subtask and 3-processor problem.

Random seed	Hopfield model			Boltzmann machine			Proposed method			
	N_h	E_t	t_h	N_b	E_t	t_b	N_s	N_p	E_t	t_p
0	464	-168*	80	445	-175	114	5	619	-175	107
3	454	-148*	78	479	-175	122	5	619	-175	107
8	583	-148*	100	487	-175	125	5	619	-175	107
9	1500	-152*	257	575	-175	147	5	619	-175	107
10	605	-175	104	437	-175	111	5	619	-175	107
13	1007	-167*	174	417	-175	102	5	619	-175	107
15	621	-152*	107	458	-175	112	5	619	-175	107
30	979	-165*	169	467	-175	120	5	619	-175	107
50	532	-149*	92	496	-175	128	5	619	-175	107
100	725	-175	125	431	-175	110	5	619	-175	107

t_i : CPU time in [sec]

seed 모두에 대해서 최적해로 수렴되었으나 그 수렴 시간이 오래 걸렸다. 10 개의 random seed에 대해 제안된 방법은 모두 최적해로 수렴되었고, 볼츠만 신

경회로망에 비해서 수렴속도는 약 1.1 배 정도 개선되었으며, 홉필드 신경회로망에 비해서 수렴확률이 약 5 배 정도 개선되었다. 또한 확률적 근사법과 홉필드 신경회로망의 1 회 반복시간에 대한 CPU 시간은 각각 약 0.2 초 정도와 약 0.17 초 정도이다.

이상의 선정된 문제의 규모에 있어서 제안된 방법은 문제의 규모나 random seed에 무관하게 최적해로 수렴되었으며, 뉴우런의 초기출력을 랜덤하게 설정하는 기존의 홉필드 신경회로망, Abe의 방법 및 볼쯔만 신경회로망에 비하여 최적화의 수렴특성이 더욱 우수함을 확인할 수 있었다.

Ⅶ. 결 론

본 논문에서는 전역최적해로 수렴하기 용이한 적정 초기상태를 설정함으로써 최적화용 신경회로망의 성능을 개선할 수 있는 방법을 제안하였다. 제안된 방법에서는 확률적 근사법을 사용하여 전역최적해가 보장되는 뉴우런의 초기출력을 설정한 다음, 이를 초기값으로 하여 기울기 경사법인 홉필드의 상태경신 알고리즘을 이용함으로써 빠른 수렴속도로 전역최적해를 구할 수 있도록 하였다.

제안된 방법의 타당성과 성능을 확인하기 위하여 도시 수가 각각 7 개 및 10 개인 외판원 경로 문제와 부태스크 수가 각각 4와 6이고 프로세서 수가 각각 3 인 최적 태스크 분할에 적용하여 시뮬레이션하였다. 10 개 도시 외판원 경로 문제의 경우, 제안된 방법은 기존의 홉필드 신경회로망을 이용할 때보다 최적해로의 수렴확률이 약 6 배 정도 개선되었고, 수렴속도는 약 1.7 배 정도 개선되었다. 그리고 볼쯔만 신경회로망에 비해서는 약 2 배, Abe의 방법에 비해서는 약 3 배 정도 수렴속도를 개선할 수 있었다. 문제의 규모가 클수록 일반적으로 초기값은 신경회로망의 성능에 더 많은 영향을 미치나 제안된 방법은 이런 경우에도 초기값의 설정이 단순하다. 또한 제안된 방법은 7 개 도시보다는 10 개 도시 외판원 경로 문제에서 그 성능 개선의 정도가 큼을 확인할 수 있었다.

향후 제안된 방법을 큰 규모의 최적화 문제나 학습용 신경회로망에도 적용하기 위한 연구^[6]가 좀 더 보완적으로 진행되어야 하겠다.

參 考 文 獻

[1] J. McClelland and D. Rumelhart, *Parallel Distributed Processing*, MIT press, Cambridge, MA, 1986.

- [2] A.J. Maren, C.T. Harston, and R.M. Pap, *Handbook of Neural Computing Application*, Academic Press Inc. London, 1990.
- [3] J.J. Hopfield and D. Tank, "Neural Computation of Decision in Optimization Problems," *Biol. Cybern.*, vol.52, pp.141-152, 1985.
- [4] J. Cervantes and R. Hilderbrant, "Comparison of Three Neuron-based Computation Schemes," *IEEE Conference on Neural Networks*, San Diego, vol.3, pp.657-671, June 1987.
- [5] P. K. Mazaika, "A Mathematical Model of the Boltzmann Machine," *IEEE Conference on Neural Networks*, San Diego, vol.3, pp.157-163, June 1987.
- [6] J. Ramanujam and P. Sadayappan, "Optimization by Neural Networks," *IEEE Conference on Neural Networks*, San Diego, vol.2, pp.325-332, June 1988.
- [7] Y. Akiyama, A. Tamashita, M. Kajimura, and H. Asio, "Combinatorial Optimization with Gaussian Machines," *International Joint Conference on Neural Networks*, Washington D.C., vol.1, pp. 533-540, June 1989.
- [8] Y. Takefuji and H. Szu, "Design of Parallel Distributed Cauchy Machine," *International Joint Conference on Neural Networks*, Washington D.C., vol.1, pp.529-532, June 1989.
- [9] J.A. Freeman and D.M. Skapura, *Neural Networks: Algorithms, Applications, and Programming Techniques*, Addison Wesley Co. Inc., 1991.
- [10] D.P. Bertsekas and J.N. Tsitsiklis, "Parallel and Distributed Computation Numerical Method," Prentice-Hall London, pp. 1-50, 1989.
- [11] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, pp. 671-680, 1983.

[12] H. Szu, "Fast Simulated Annealing." *AIP Conference Proceedings 151 : Neural Network for Computing*, Snowbird, UT, 1986.

[13] H. Szu and R. Hartley, "FastSimulated Annealing." *Physics Letters A*, vol.122 (3.4), pp.157-162, 1987.

[14] M.A. Styblinski and T.S. Tang, "Experiments in Nonconvex Optimization : Stochastic Approximation with Function Smoothing and Simulated Annealing." *IEEE Trans. on Neural Networks*, vol.3, no. 4, pp. 467-483, 1990.

[15] S. Abe, "Global Convergence and Suppression State of the Hopfield Neural Network." *International Joint Conference on Neural Networks*, Singapore, vol.2, pp. 935-940, Nov. 1991.

[16] Y. H. Cho and H. M. Choi, "Improving the training performances of the multilayer neural networks by SAS-based optimal estimation of initial weights." *JTC-CSCC*, Kyungju, pp. 475-478, July 1992.

著 者 紹 介



趙 鏞 鉉(正會員)

1957年 3月 18日生. 1979年 2月 경북대학교 전자공학과 졸업. 1981年 2月 경북대학교 대학원 전자공학과 졸업(공학석사). 1993年 2月 경북대학교 대학원 전자공학과 졸업(공학박사). 1983年 9月 ~ 1984年 2월 삼성반도체통신(주) 근무. 1984年 3月 ~ 1987年 2월 한국전자통신연구소 연구원. 1987年 3月 ~ 현재 영남전문대학 전자과 조교수. 주관심 분야는 병렬분산처리 및 최적화, 신경회로망 및 스위칭 네트워크 등임.

崔 興 文(正會員) 第 30卷 第 5號 參照

현재 경북대학교 전자공학과 교수.