2001년 한국정보처리학회 추계 학술발표논문집 제8권 제2호

# 평균장 어닐링 알고리즘의 개선 및 응용 연구

권오준*
*동의대학교 전산통계학과
e-mail:ojkwon@dongeui.ac.kr

# Improvement of Mean Field Annealing Algorithm and its Application

Oh-Jun Kwon*
*Dept of Computer Science and Statistics, Dong-Eui University

요약

현실 세계의 많은 조합 최적화 문제들은 변수들이 강하게 상호 작용함에 따라 만족해야 하는 목표 함수가 매우 복잡하게 주어진다. 복잡한 목표 함수에는 많은 지역 최적해들이 존재하기 때문에 전역 최적해를 얻는 것은 엄청난 시간을 필요로 한다. 이러한 문제에 대해 해를 찾는 방법의 하나로 평균장 어닐링 알고리즘(MFA)이 있다. 본 논문에서는 기존의 이진 상태 공간에만 적용할 수 있었던 평균장 어닐링 알고리즘을 연속 상태 공간을 갖는 조합 최적화 문제에 적용할 수 있도록 알고리즘을 수정, 제안한다. 그리고 제안된 알고리즘을 제한된 연속 상태 공간을 가지는 단순 회귀 모델의 D-최적 설계에 적용하였다. 실험결과 제안된 알고리즘이 시뮬레이티드 어닐링 알고리즘(SSA)과 비교하여 비슷한 수준의 결과를 내면서도 계산 속도면에서는 수 배 정도의 빠른 좋은 결과를 얻었다.

## 1. 서론

패턴 매칭, 음성 처리, 전자회로 설계, 높은 통신 트래픽하에서의 경로 제어 등에서 많은 조합 최적화 문제들이 존재한다. 이러한 문제들에서는 목표함수가 복잡하게 주어짐에 따라 전역 최적해를 얻는 것은 엄청난 시간을 필요로 하는 경우가 많다. 일부 문제들은 NP-complete 부류의 문제에 속하기도 한다. 따라서 이러한 조합 최적화 문제에 대해서는 항상 최적에 가까운 근사해를 가능한 한 빠른 시간에 발견할 수 알고리즘을 개발하는 것이 중요하다.

최근 시뮬레이티드 어닐링(simulated annealing, SSA) 알고리즘과 평균장 어닐링(mean field annealing, MFA) 알고리즘은 그래프 분할[1], 전자회로 설계[2], 위성 방송 스케쥴링[3], 이미지 재생[4] 등 다양한 조합 최적화 문제에 대해 성공적인 것으로 보고된 바 있다. 더우기 이들 어닐링 알고리즘들은 병렬로 수행되는 것이 가능하므로 기존의 일반 휴리스틱 알고리즘에 비해 비교할 수 없을 정도로 계산 속도를 개선할 수 있는 장점이 있다[1, 5].

시뮬레이티드 어닐링 알고리즘은 경사 강하 방법과 지역 최소점을 빠져나올 수 있는 접근 방법으로 확률적 랜덤 과정을 통해 목표 함수의 최적해를 찾는다. 평균장 어닐링 알고리즘은 시뮬레이티드 어닐링 알고리즘과 홉필드 신경망의 결합이라 할 수 있다. 이 알고리즘은 시스템이 평형 상태에 있다고 가정함으로써 특정 노드가 다른 모든 노드들로 부터 받는 평균장을 사용하여 그 노드의 값을 갱신한다. 그리고 주어진 온도에서 모든 노드들이 더 이상의 움직임이 없는 평형점에 도달하게 되면 온도 변수 $T$를 감소시키면서 이러한 상태 천이 과정을 계속하여 최적해에 도달한다.

일반적으로 평균장 어닐링 알고리즘은 시뮬레이티드 어닐링 알고리즘에 비해 동일한 수준의 해를 내면서도 상당한 빠른 계산 속도를 보여주는 것으로 실험적으로 보고되고 있다[1, 6]. 그러나 불행히도 기존의 평균장 어닐링 알고리즘은 단지 탐색 공간이 이진 상태를 가지는 문제들에만 적용할 수 있는 단점이 있다[1, 3, 8]. 하지만 현실 세계의 많은 조합

최적화 문제들에서 각 변수들이 연속 상태 공간을 갖는 경우가 많이 있다. 본 논문에서는 기존의 평균장 어닐링 알고리즘을 연속적인 상태 공간을 갖는 조합 최적화 문제에도 적용할 수 있도록 알고리즘을 수정, 제안한다.

2. 평균장 어닐링 알고리즘의 확장

일반적으로 조합 최적화 문제에서 목표 함수는 상호 작용하는 일련의 변수들을 포함하는 복잡한 비선형 식으로 주어진다. 이러한 문제의 해는 목표 함수값을 최소로 하는 변수들의 특정한 조합으로 얻어진다. 이러한 문제에 대해 물리학자들은 종종 평균장 근사 기법을 사용하여 문제를 해결한다. 평균장 근사 기법은 열 평형 상태에서 입자나 스핀들로 구성된 시스템의 움직임에 대한 단순 해석적인 근사 방법이다.

$N$개의 스핀들이 상호 작용하면서 결합된 시스템의 에너지 함수는 식 (1)과 같이 표현된다.

$$H(s)=\sum_{i=1}^{N}\sum_{j\neq i}\beta_{ij}s_is_j+\sum_{i=1}^{N}h_is_i \qquad (1)$$

위 식에서 스핀들의 상호 작용 $\beta_{ij}$는 대칭적(즉, $\beta_{ij}=\beta_{ji}$)이고 어떠한 자기 상호 작용도 없다(즉, $\beta_{ii}=0$)고 가정한다. 그리고 $h_i$는 각 스핀과 관련된 지역장(local field), 즉 상대적인 중요도를 나타낸다.

이 시스템은 볼쯔만 분포를 따르므로 모든 스핀들이 평형 상태에 있다고 가정하면 $i$번째 스핀의 평균, $\langle s_i\rangle$와 다른 모든 스핀들의 작용으로 인해 $i$번째 스핀에 의해 느껴지는 평균장, $\Phi_i$를 결정할 수 있다. 이 때 하나의 스핀이 다른 스핀에 미치는 영향이 전체 장에 비해서 매우 작을 정도로 스핀들의 수가 매우 많다고 가정하면 특정 스핀 $s_i$와 다른 스핀 $s_j$사이의 상호 작용를 무시함으로써 평균장 근사를 사용할 수 있다. 이러한 평균장 시스템은 self-consistent하므로 각 온도에서의 해 또는 평형점은 다음과 같은 상태 천이 과정을 반복함으로써 얻을 수 있다.

$$\langle s\rangle^{(0)} \Rightarrow \Phi^{(0)} \Rightarrow \langle s\rangle^{(1)} \Rightarrow \Phi^{(1)} \Rightarrow \cdots \Rightarrow \langle s\rangle^{(L-1)} \Rightarrow \Phi^{(L-1)} \Rightarrow \langle s\rangle^{(L)}$$

각 변수들이 제한된(bounded) 연속 상태 공간을 가지는 조합 최적화 문제를 고려하자.

$$v_i\in[x_1, x_2] \quad \text{for} \quad 1\leq i\leq N \qquad (2)$$

그러면 $i$번째 노드 $v_i$가 $\alpha$ 값을 가질 확률은

$$\Pr\{v_i=\alpha\}=Z^{-1}\times e^{-H_\alpha/T} \qquad (3)$$

위 식에서 $H_\alpha=\langle H(v)\rangle|_{\langle v_i\rangle=\alpha}$이다. 그리고 $Z$는 분할 함수이다.

$$Z=\int_{x_1}^{x_2}e^{-H_\alpha/T}d\alpha \qquad (4)$$

그리고 모든 노드들이 주어진 온도에서 평형 상태에 있다고 가정하면, $I$번째 노드의 평균 $\langle v_i\rangle$를 계산하는 식은 식 (5)와 같이 얻을 수 있다.

$$\langle v_i\rangle=\frac{\int_{x_1}^{x_2}\alpha e^{-H_\alpha/T}d\alpha}{\int_{x_1}^{x_2}e^{-H_\alpha/T}d\alpha} \qquad (5)$$

위 식에는 분모와 분자에 각각 적분항을 가지고 있어 계산이 복잡하다. 이를 간단히 수행하기 위한 방법으로 동일한 확률 분포하에서 표본을 추출하여 이들의 평균값으로 노드의 값을 갱신하는 방법을 사용한다. 시스템이 평형 상태에 있다고 가정하고 모든 노드들의 값을 고정하면 한 노드의 움직임에 대한 시스템의 평균장의 영향은 볼쯔만 분포를 따르게 된다. 이 확률 분포을 이용하여 적절한 수의 노드 변화에 대한 표본을 추출하여 이 표본들의 평균값을 새로운 노드의 값으로 결정하면 시스템의 평균장을 간접적으로 이용할 수 있다.

수정 알고리즘에서 $\lambda$는 어닐링 과정의 온도 변수 감소율을 나타낸다. 그리고 평형 상태라 함은 시스템의 모든 변수 노드들의 값이 더 이상 움직임이 없는 상태를 의미한다. $N_s$가 증가함에 따라 더 높은 신뢰도(confidence)을 가지는 표본 평균이 얻어질 수 있다.

1. 온도 변수 $T$를 초기 온도 $T_s$로 초기화한다. 그리고 모든 노드들을 작은 랜덤 값으로 초기화한다.
2. 임의의 한 노드 $v_i$를 선택한다. 노드 $v_i$를 제외한 모든 노드들의 값을 고정시킨다.

```
k=0;
WHILE(k < N_s)
    노드 v_i의 다음 상태 표본값을 임의로 선택;
    목표 함수의 변화량 ΔE을 계산;

    if ( ΔE≤0)
      P(A)=1의 확률로 수락;
    else
      P(A)=e^(-ΔE/T)의 확률로 수락;

    k = k + 1;

END_OF_WHILE

모든 수락된 상태값들의 평균을 계산
```

3. 단계 2에서 얻어진 상태값들의 평균을 새로운 노드 평균 $\langle v_i \rangle$로 갱신한다. 만약 시스템이 평형 상태에 도달했다면 다음 단계로 간다. 그렇지 않으면 단계 2로 간다.
4. 만약 온도 변수 $T$가 최종 온도 설정값, $T_f$에 도달했다면 알고리즘을 종료한다. 그렇지 않으면 $T$를 아래 식으로 감소하고 단계 2로 간다.

$$T = \lambda T$$

수정된 평균장 어닐링 알고리즘(MMFA)

## 3. 실험 및 토론

제안된 알고리즘의 응용문제로 단순 회귀 모델의 D-최적 설계 문제를 선택하였다. 차수가 $k$인 일반적인 단순 회귀 모델은 다음과 같다.

$$y_i = f'_k(x_i)\beta + \varepsilon_i, \qquad i=1,\cdots,n. \tag{6}$$

위식에서 $y_i$는 독립 변수 $X$의 한 값 $x_i$에 대해 얻어지는 종속 변수 $Y$의 확률 변수을 나타낸다. 벡터 $f_k(x_i)$는 $(k+1)\times 1$ 차원의 벡터로 공간 $X$상의 연속 함수이다. $k$는 단순 회귀 모델의 복잡도를 의미하는 정도를 나타낸다. $k$가 1인 경우는 선형(linear) 모델이고, $k\geq 2$인 경우는 다항(polynomial) 모델이다. ' ''은 벡터의 전치(transposition)를 의미한다. $\omega$를 공간 $X$상에서의 임의의 $n$개의 점을 나타내는 벡터라 하면, D-최적 설계 기준은 식 (7)과 같이 주어지는 목표 함수 $E(\omega)$를 최대로 하는 $\omega$를 찾는 것이다.

$$\begin{aligned} E(\omega) &= |M(\omega)| \\ &= |\sum_{i=1}^{n} f(x_i)\, f'(x_i)| \end{aligned} \tag{7}$$

한편 Welch는 유한 설계 공간에 대하여 단순 회귀 모델의 D-최적 설계 문제는 NP-hard하다는 것을 보였으며 Haines는 시뮬레이티드 어닐링 알고리즘을 사용하여 이 문제를 해결한 바 있다[7].

제안된 알고리즘을 이 단순 회귀 모델의 D-최적 설계에 적용하였다. 본 실험에서는 식 (6)의 모델에서 $f'_k(x_i)$와 $x_i$의 상태 공간이 식 (8)과 같은 경우를 고려하였다.

$$f'_k(x_i) = (1, x_i, x_i^2, \cdots, x_i^k), \quad x_i \in [-1, 1] \tag{8}$$

이제 평균장 어닐링 알고리즘을 적용하기 위해서는 변수 노드 $x_j$을 다른 노드와 분리된 목표 함수를 얻어야 한다. 식 (7)로 부터 최소화해야 하는 목표 함수 $E(\omega)$의 일반식을 다음과 같이 얻을 수 있다.

$$\begin{aligned} E(\omega) &= -|M(\omega)| \\ &= -|\sum_{i \neq j} f_k(x_i)\, f'_k(x_i) + f_k(x_j)\, f'_k(x_j)| \end{aligned}$$

| $n$ | 알고리즘 | 결과 | $-\lvert M(\omega)\rvert^{-1}$ |
|---|---|---|---|
| 6 | SSA | ±1.0, ±0.765, ±0.285 | -241.545 |
| | MMFA | ±1.0, ±0.765, ±0.285 | -241.545 |
| 7 | SSA | ±1.0, 0.768, -0.774, -0.392, 0.309, -0.154 | -118.358 |
| | MMFA | ±1.0, 0.768, -0.774, -0.392, 0.309, -0.152 | -118.358 |
| 8 | SSA | 2×1.0, -1.0, 0.768, -0.774, -0.392, 0.309, -0.154 | -59.236 |
| | MMFA | ±1.0, 2×0.766, -0.773, -0.386, 0.300, -0.175 | -59.615 |
| 9 | SSA | ±1.0, 2×0.765, -0.765, 2×±0.285 | -30.193 |
| | MMFA | 2×1.0, 2×±0.765, 2×0.285, -0.285 | -30.193 |
| 10 | SSA | 1.0, 2×-1.0, 0.765, 2×-0.765, 2×±0.285 | -15.097 |
| | MMFA | 2×1.0, -1.0, 2×0.765, -0.765, 2×±0.285 | -15.097 |
| 11 | SSA | 2×1.0, -1.0, 2×±0.765, 2×±0.285 | -7.548 |
| | MMFA | 1.0, 2×-1.0, 2×0.765, -0.765, 3×0.285, 2×-0.285 | -7.548 |

표 1 제안된 알고리즘(MMFA)의 결과와 시뮬레이티드 어닐링(SSA)에 의한 결과

실험은 $k$=5로 하고 $n$을 다양하게 변화하면서 수행하였다. 제안된 알고리즘에 의한 실험 결과들은 Haines[7]에 의해 수행된 시뮬레이티드 어닐링 알고리즘에 의한 결과와 함께 표 1에 요약하였다. 표 1

로 부터 우리는 제안된 평균장 어닐링 알고리즘 2에 의한 설계 결과가 시뮬레이티드 어닐링 알고리즘에 의해 얻어진 결과와 비교하여 동일한 수준의 결과임을 알 수 있다. 이 실험에서는 표본 추출 시도 횟수 $N_s = 100$을 사용하였다.

| $n$ | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|
| SSA | 15.10 | 15.72 | 16.03 | 16.57 | 17.12 | 17.86 |
| MMFA | 2.04 | 2.16 | 2.31 | 2.41 | 2.52 | 2.59 |

표 2 알고리즘 수행 시간 비교(CPU 시간, 초)

표 2는 233MHz PC를 사용하여 서로 다른 초기 조건하에서 각 5번씩 알고리즘들을 수행하여 수렴하는데 걸리는 평균 계산 시간이다. 제안된 알고리즘이 시뮬레이티드 어닐링 알고리즘에 비해 약 7배 정도 빠르게 최적해에 수렴한 것을 알 수 있다.

## 4. 결론

본 논문에서는 기존의 이진 상태 공간을 가지는 조합 최적화 문제에만 적용할 수 있었던 평균장 어닐링 알고리즘을 연속 상태 공간을 가지는 조합 최적화 문제에도 적용할 수 있도록 확장하였다. 그리고 제안된 알고리즘을 D-최적 기준에 따른 단순 회귀 모델의 최적 설계에 적용하여 실험하였다. 실험결과에 따르면 시뮬레이티드 어닐링 알고리즘의 결과와 비교하여 동일한 수준의 결과를 내면서도 최적해로 수렴하는데 걸리는 계산 시간은 약 7배 정도 빠른 결과를 보여 주었다.

결론적으로 본 논문에서 제안된 새로운 평균장 어닐링 알고리즘은 연속 상태 공간을 가진 조합 최적화 문제에도 쉽게 적용할 수 있을 뿐만 아니라 기존의 평균장 어닐링 알고리즘의 장점도 그대로 가지고 있음을 알 수 있다.

## 참고문헌

[1] D. E. Van den Bout, and T. Miller, "Graph partitioning using annealed neural networks," *IEEE Trans. on Neural Networks*, Vol. 1, 1990, pp. 192-203.

[2] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* Vol. 220, 1983, pp. 671-680.

[3] Nirwan Ansari, Edwin S. H. Hou and Youyi Yu, "A New Method to Optimize the Satellite Broadcasting Schedules Using the Mean Field Annealing of a Hopfield Neural Network," *IEEE Trans. on Neural Networks*, Vol. 6, 1995, pp. 470-483.

[4] Geman Stuart and Geman Donald, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, 1984, pp. 721-741.

[5] Emile Aarts and Jan Korst, *Simulated Annealing and Boltzmann Machines*, John Wiley & Sons, 1989.

[6] Ikuo Matsuba and Hironari Masui, "Asymptotic Behaviors of Simulated Annealing and Mean Field Approximate Annealing," *Proc. of International Joint Conference on Neural Networks*, Vol. 2, Singapore, 1991, pp. 923-928.

[7] Linda M. Haines, "The Application of the Annealing Algorithm to the Construction of Exact Optimal Designs for Linear-Regression Models," *Technometrics*, Vol. 29, 1987, pp. 439-447.

[8] J. Anderson, "A mean field computational model for PDP," *Connectionist Models*, 1983, pp. 217-223.