

# Simulated Annealing법의 적용시 Starting Temperature 결정에 관한 연구

이 영 진\*, 이 권 순  
동아대학교 전기공학과

## A Study on Determination of Starting Temperature for the Method of Simulated Annealing

Young Jin Lee\*, Kwon Soon Lee  
Dept. of Electrical Eng.  
Dong-A University, Pusan, Korea

### Abstract

The method of simulated annealing is a technique that has recently attracted significant attention as suitable for optimization problem of very large scale. If the temperature is too high, then some of the structure created by the heuristic will be destroyed and unnecessary extra work will be done. If it is too low then solution is lost, similar to the case of a quenching cooling schedule in the Simulated Annealing (SA) phase. Therefore, a crucial issue in this study is the determination of the starting temperature and cooling schedule for SA phase.

### 1. 서론

TSP(Traveling Salesman Problem), 비선형 파라미터 추정(nonlinear parameter estimation), LSI의 설계에서 부품의 최적화 배치등과 같이 시스템의 해를 쉽게 구할 수 없는 문제들에 최적화 알고리즘들이 지금까지 많이 적용되어 왔다. 그러나, 이 알고리즘에 의한 해가 국부적으로 수렴하기 때문에 보다 나은 개선책으로 simulated Annealing(SA)법이 개발되었고, 최근 들어 이러한 NP(nondeterministic polynomial)-complete 문제의 최적 상태인 global minimum energy를 찾기 위하여 SA법이 많이 적용되고 있다[1]. 그러나, 계산시간이 너무 막대하게 걸려서 실용적이지 못하다는 단점이 있다. 그러므로, 실행 속도를 빠르게 하기 위한 가장 적절한 초기온도(starting temperature) 설정과 초기 온도로 부터 보다 빠른 속도로 수렴하도록 온도를 낮추어 나가는 방법을 선택하는 것은 보다 정확하고 빠른 최적화 관점에서 매우 중요한 선택중의 하나이다. 실제로 매우 높은 온도에서 부터 시작하게 되면, 불필요한 수행을 계속 행하게 되고 매우 낮은 온도에서는 global minimum 값을 얻지 못하는 상태를 가질 수도 있다[2]. 따라서, 본 논문에서는 SA 알고리즘이 가지는 많은 문제점들 중 계산 시간을 줄이기 위한 방법을 제시하고자 한다.

### 2. Simulated annealing법

Simulated annealing법의 개념 도입은 야금학 담금질(metallurgical annealing)의 유사성으로 부터 시작되었다 [3]. 야금학 기술에 의하면 담금질법은 금속을 거의 녹는 점까지 가열한 후 서서히 실온 상태까지 냉각시키는 것을 말한다. 즉, 담금질의 개념은 전위(轉位)나 다른 크리스탈 격자 분열(crystal lattice disruption)이 높은 온도에서 온도 격동(temperature agitation)으로 제거되어질 수 있고, 금속을 천천히 식힘으로서 새로운 전위(轉位) 형성을 막을 수 있지만, 온도 강하(dropping)에 따라 발생하는 어떤 전위(轉位)를 회복시키는데는 시간이 필요하다는 것이다. 다시 말해서, 이 과정의 본질은 금속의 전(全)에너지 함수(global energy function)가 결과적으로 어떤 절대적 최소값(absolute minimum value)에 도달하는 것이다. 만약, 그 온도가 급격히 떨어지면 격자 전위는 금속 내부에서 동결되기 때문에 금속 격자 에너지는 전(全)에너지 함수의 최소값 보다 더 높아지게 된다.

Simulated annealing법의 기본 개념은 Nicholas Metropolis, Arianna Rosenbluth등에 의해서 1953년에 제안된 이후 그들은 각 분자들의 상호작용으로 이루어진 물질의 상태 방정식과 같은 특성을 조사할 수 있는 Monte Carlo법에 관심을 가졌다[3]. 시스템 에너지(system energy)를 최소화시키는 전체적 과정은 에너지 변화의 국부적 과정(localized process)으로 요약될 수 있는데, 그 개념은 임의로 선택된 시스템의 일부분을 한번에 일시적으로 변화시킬 수 있다는 것이다. 이 변화가 전(全)시스템 에너지 상태를 낮추어 간다면 그 변화가 받아들여지고, 이 변화가 전(全)시스템 에너지 상태를 높게 되면 이 변화는 불츠만 분포로 주어질 확률 P로 받아들여진다. 온도를 서서히 내려가면서 확률적인 다이내믹스(dynamics)를 실행하여 최소점을 구해 나가는 simulated annealing법은 이진 벡터(binary vector) X의 실수값 함수(real-valued function) f(x)를 최소화 하기 위한 조합문제(combinatorial problem) 최적화의 한 방법으로 제안되었다. 그 알고리즘은 다음 7가지 순서로 즉

약될 수 있다:

- 1) 난수적으로 vector  $X$ 와 초기온도  $T$ 중 큰값을 선택한다(모든 에너지 변화  $\Delta$ 에 대하여  $\exp(-\Delta/T) \geq 0.999$ 인 초기온도를 선택한다).
- 2) 균등하게 발생시킨 난수(uniform random numbers)에서  $X$ 의 성분  $X_i$ 를 임의로 선택한다.
- 3) 0에서 1사이 혹은 1에서 0사이중 적절하게 변화 되어진 성분  $X_i$ 값을 제외하고는  $X = X'$ 로 두면,  $\Delta = f(X') - f(X)$  이다.
- 4) 만약  $\Delta < 0$  이면,  $X = X'$ 라 놓고 6)으로 간다.
- 5) 만약  $\Delta \geq 0$  이면, 확률  $P = \exp(-\Delta/T)$ 를 가지며  $X = X'$ 라 놓는다. 즉, 균일한 확률 분포를 통하여 0과 1사이의 난수  $\xi$ 를 선택하여  $\xi < P = \exp(-\Delta/T)$ 이면,  $X = X'$ 로 놓고, 그렇지 않으면  $X$ 를 그대로 둔다.
- 6) 만약 현재의 온도값에서  $X$ 의 성공적인 변화값이나 최대 허용 변화량이 일어나면,  $T$ 를  $\alpha$ 의 크기로 감소시킨다. 여기서,  $\alpha$ 는 일반적으로 0.8 - 0.9999 사이의 온도 cooling schedule 상수이다.
- 7) 만약  $f$ 의 최소값이 주어진  $\epsilon$ 보다 더 감소하지 않는다면 멈추고, 그렇지 않으면 2)로 되돌아 간다.

예를들어, TSP에 대하여 살펴보면,  $N$ 개의 도시가 주어지고 우리는 한번에 모든 도시를 최단거리로 여행하여 다시 시작지점으로 돌아오기를 원한다. 이때 각 도시간의 거리를 비용함수(cost function)로 두고 cooling schedule에 따라 비용함수를 최소화 시켜나가는 것이며, 생각할 수 있는 모든 가능한 여행경로의 수가  $N!/2N$ 로 되는 이 문제는  $N$ 이 증가함에 따라 계산량은 거의 지수적으로 증가하므로 좋은 해법이 요구되어지며, SA법에 의하여 근사적으로 해결되어질 수 있었고 이 근사해는 거의 최적의 범위 내에 있다고 할 수 있다.

### 3. Starting temperature 결정과 cooling schedule

다음 식에서와 같이 볼츠만 분포함수(Boltzmann distribution function)를 이용한 simulated annealing법에서 온도는 대단히 중요한 파라미터이다. 만약, 어떤 변화가 시스템 전체 에너지의 증가로 귀착한다면 이 변화는 다음의 식

$$P = \exp(-\Delta E/T) \quad (1)$$

에 의한 확률  $P$ 에 의하여 이루어지며  $\Delta E = E_{\text{new}} - E_{\text{old}}$  이고,  $T$ 는 상태 천이 규칙에서 이용한 온도 파라미터이다.

식 (1)로 부터 에너지가 최소가 될때 상태가 최대 확률로 나타난다는 것을 알 수 있으며 온도  $T$ 가 0의 극한에서 최소 에너지 상태를 가질 확률이 1로 된다. 어쨌든, 온도가 낮을수록 최소값을 가질 확률이 매우 크게 될 것이다. 만약 온도  $T$ 를 처음부터 0의 상태에서 출발시킨다면 상태는 에너지 함수 극소점으로 수렴하기 때문에, 반드시 최소점을 취할 수가 없으므로 초기온도의 선택과 이 온도를 줄여나가는 방법은 계산시간적 측면에서 보면 대단히 중요한 요소중의 하나이다. 온도를 줄여 나가는 방법으로는 지금까지

많은 사람들에 의하여 제안되어 왔다. 특히, 그 중에서 응용 수학자 Geman형제[4]가 화상 해석에 볼츠만 기계를 응용한 논문에서 화상을 탠덤 마르코프 과정으로 해석하여 볼츠만 기계의 최적 상태로 수렴하도록 다음의 조건을 만족하게 낮추어가면 반드시 최소점에 수렴할 수 있다고 증명하였다.

$$T(t) = T_0 / \log(1+t) \quad (2)$$

여기서  $T_0$ 는 초기온도이며,  $t$ 는 시간을 나타낸다.

물론 simulated annealing이나 볼츠만 기계(boltzmann machine)에서는 온도를 내리면 내릴수록 에너지 함수의 어떤 극소에서 다른 극소로 천이하는데 시간이 많이 걸린다. 때문에 현재로서는 시뮬레이션 시간이 너무 걸려서 실용적이지 못한 단점도 있지만, 이 방법은 현재로서는 극소에 빠지지 않고 최소를 구하면서 해의 수렴이 보장되고 있는 유일한 방법이라 할 수 있다. 실제로 식 (2)의 온도 변화는 대수적으로 매우 완만하므로 시뮬레이션 시간이 너무 막대하게 될 문제점이 있다. 때문에 SA법 적용시에 다소간의 발견법적(heuristic)인 고속화를 행하고 있다.

이외에도 코시 분포(Cauchy distribution)라고 불리는 분포 함수를 선택함으로써 SA법의 속도를 대수적인 식 (2) 대신 역수적인 속도로 고속화할 수 있는 것도 보고되고 있다[5]. 식 (1)에서 보는 바와같이 실현 확률은 초기 온도가 높으면 에너지의 작은 차이에서는 둔감하게 되며, 초기 온도가 매우 낮으면 매우 민감하게 되므로 시스템에 대한 가장 적절한 초기온도(실행 시간을 최소화 하면서 해의 최소값을 가지도록 하기위한 온도)를 설정하는 문제는 대단히 중요하다고 생각된다.

### 참고 문헌

- [1] H. William, et al., Numerical Recipes: The Art of Scientific Computing, pp. 326-334, Cambridge Univ. Press, New York, 1986.
- [2] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi, "Optimization by Simulated Annealing," Science, Vol. 220, pp. 671-680, 1983.
- [3] Robert Hecht-Nielsen, Neurocomputing, pp. 192-197, Addison-Wesley Publishing Company, New York, 1990.
- [4] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions and Bayesian Restoration of Images," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 6, pp. 721-741, 1984.
- [5] Philip D. Wasserman, Neural Computing: Theory and Practice, pp. 81-92, Van Norstrand Reinhold, New York, 1989.