

## 2-단계 병렬 유전자 알고리즘

길원배<sup>o</sup>, 이승구  
포항공과대학교 전자전기공학과  
{wbkil<sup>o</sup>, slee}@postech.ac.kr

### A Two-Phase Parallel Genetic Algorithm

Wonbae Kil<sup>o</sup>, Sunggu Lee  
Dept. of Electronic and Electrical Engineering, POSTECH

#### 요 약

본 논문에서는 유전자 알고리즘(Genetic Algorithm: GA)의 새로운 병렬화 방법을 제안 하고 있다. 기존의 병렬 유전자 알고리즘(Parallel Genetic Algorithm: PGA)은 전체 개체집단을 부개체집단(Subpopulation)으로 나누어 해의 가능 영역을 동시에 탐색하는 것이 일반적인 방법인데 반해, 본 논문에서 제안하는 병렬화 방법은 전체 해의 영역을 나누어 각각의 영역에서 독립된 개체집단들이 서로 다른 영역을 탐색하게 하는 방법이다. 이 방법은 두 가지 단계의 병렬 유전자 알고리즘으로 구성된다. 먼저 적응 교배 연산자(Adaptive Crossover Operator: ACO)를 이용한 PGA를 통해 지역해에 인접한 범위들로 해의 영역을 나누고, 이렇게 나누어진 각각의 영역들에서 다시 병렬로 GA를 적용시켜 자세하게 탐색하는 방법이다. 첫 번째 수행되는 PGA 단계에서는 탐색 시간을 줄이고 두 번째 PGA 단계에서는 보다 자세한 탐색을 하기 위해 정밀도(Precision)의 조정을 유전자 알고리즘의 병렬화에 적용하였으며, 이를 통해 빠르고 자세한 탐색이 가능한 유전자 알고리즘의 병렬화 방법을 제안하고 있다.

#### 1. 서 론

유전자 알고리즘(Genetic Algorithm: GA)은 유전적 계승과 다윈의 생존경쟁이라는 자연의 현상을 모델링한 확률적인 탐색 방법이다. 다시 말해 주어진 환경에 잘 적응하는 유전자만을 선택(Selection)하여 교배(Crossover)하고 교배 후 발생할 수 있는 경우로 돌연변이(Mutation)도 생성하여 다음 세대에 우수한 유전 형질이 전달(Reproduction)되게 된다. 따라서 진화(Evolution)가 거듭될수록 주어진 환경에 더 적합한 유전자들만이 남아 있게 되는 것이다[1]. 이러한 GA는 근본적으로 병렬성을 가지고 있으며, 방대한 해의 영역을 가지는 문제에서 조금이라도 더 빠른 탐색을 하기 위해 병렬 유전자 알고리즘(Parallel Genetic Algorithm: PGA)에 관한 연구가 많이 행해져 왔다.[2,3] 본 논문은 분산 환경에서 기존의 병렬 유전자 알고리즘에서 가지는 통신 시간에 의한 부담을 줄이고, 병렬 유전자 알고리즘보다 좀 더 빠르고 정밀한 탐색을 하기 위해 유전자 알고리즘의 새로운 병렬화 방법을 제안하고 있다. 이를 위해 [4]에서 제안된 적응 교배 연산자(Adaptive Crossover Operator)를 본 논문의 목적에 맞게 변형한 적응 교배 연산자를 유전자 알고리즘에 이용하고 있다. 또한 정밀도(Precision)의 조정을 유전자 알고리즘의 병렬화에 적용하여 빠르고 자세한 탐색이라는 궁극적 목적에 부합되는 병렬화 방법을 제안하고 있다.

#### 2. 유전자 알고리즘의 병렬화

기존의 일반적인 병렬 유전자 알고리즘에서 개체집단을 부개체집단(Subpopulation)으로 나누어 해의 가능 영역

을 동시에 탐색하는 것과 다르게, 본 논문에서 제안하는 병렬화 방법은 해의 영역을 나누고 각각의 영역에서 독립된 개체집단들이 서로 다른 영역을 탐색하게 하는 2-단계 병렬 유전자 알고리즘이다. 즉, 하나의 문제에 두 가지 단계의 유전자 알고리즘을 적용하여 최적해를 찾아 내는 방법이다. 첫 번째 단계의 유전자 알고리즘은 전역 최적해를 포함한 지역 최적해들에 인접한 범위들로 해의 영역을 나누기 위해 수행되는 것이며, PGA에 의해 수행된다. 이 과정에서 소요되는 탐색 시간을 줄이기 위해 계산의 복잡도가 낮고 수행속도가 빠른 적응 교배 연산자를 이용한 PGA를 사용하며, 보다 빠른 탐색을 위해 원하는 정밀도보다 낮은 정밀도를 사용하여 탐색하는 방법을 사용한다. 이 부분에서 정밀도를 낮춘다는 말은 하나의 개체를 이진수로 표현 할 때 사용하는 비트수를 줄인다는 말이며, 이렇게 할 수 있는 이유는 첫 번째 유전자 알고리즘(PGA)의 탐색 목적이 두 번째로 수행될 유전자 알고리즘(GA)을 위한 지역해들의 대략적인 범위를 먼저 찾는 것이므로 자세한 탐색을 필요로 하지 않기 때문이다.

두 번째 단계의 유전자 알고리즘에서는 이전 과정에서 얻은 지역해의 인접 영역들에서 병렬로 여러 개의 GA들이 수행되게 되며, 실제 원하는 만큼의 높은 정밀도를 사용하여 각각의 분할된 해의 영역에서 GA를 수행함으로써 자세한 탐색을 하게 된다. 그림1은 제안된 병렬화 방법에 대한 개요도를 보여주고 있다. 그림1을 보면, 정밀도를 낮춤으로 축소된 해의 영역에서 PGA를 먼저 수행하게 되며, 적응 교배 연산자를 이용하여 빠른 탐색 결과를 얻게 된다. PGA의 결과로 얻어진 지역해들을 통해 자세하게 탐색할 범위를 설정하게 되고, 다시 각각의 탐색영역에서 GA를 수행하게 된다. GA가 수행될 줄어드

탐색영역을 보다 많은 비트로 표현함으로써 보다 자세한 탐색을 할 수 있게 되며, 분산 환경에서 각각의 GA를 수행하는 노드들은 통신을 필요로 하지 않기 때문에 일반적인 병렬 유전자 알고리즘에 비해 통신량을 현저하게 줄일 수 있게 된다.

은 lsb(least significant bit)를 나타낸다.

(3)과 같은 각 변수에 대한 이진수 열을 하나의 개체로 표현하기 위해 아래의 (4)와 같이 이진수 열로 표현 할 수 있다.

$$\underbrace{b_n^m b_{n-1}^m \dots b_2^m b_1^m}_{msb\ bits} b_n^{m-1} b_{n-1}^{m-1} \dots b_2^{m-1} b_1^{m-1} \dots \underbrace{b_n^1 b_{n-1}^1 \dots b_2^1 b_1^1}_{lsb\ bits} \quad (4)$$

각 변수들의 이진수 열을 (4)와 같이 msb bits부터 lsb bits까지 나열함으로써 교차점을 msb bits에 인접한 지역에서 lsb bits 인접 지역으로 쉽게 이동 시킬 수 있으며, 이는 유전자 알고리즘이 진화되어 가는 상태에 따라 탐색 구간을 조절 할 수 있는 적응 교배 연산자의 구현을 쉽게 하기 위한 개체의 표현방법이다.

### 2.2 변형된 적응 교배 연산자를 이용한 PGA

[4]에서 제안된 기존의 적응 교배 연산자는 일반적인 유전자 알고리즘(GA)에서 전역탐색과 지역탐색의 균형을 목적으로 하여 제안된 방법이다. 즉, 유전자 알고리즘의 초기 단계에서는 탐색 공간에서 더 넓은 구간으로 탐색 하고, 최적해에 가까이 근접할수록 더 좁은 구간으로 탐색하려는 의도에서 제안된 방법이다. 본 논문에서는 보다 빨리 지역해들의 범위를 찾기 위해 본 논문의 의도에 맞게 변형된 적응 교배 연산자를 PGA에 적용하여 사용하고 있다. 또한, PGA의 기본적인 동작 방법은 [3]에 제시된 방법과 같으며, 적응 교배점에 관련된 PGA의 동작만을 부각시켜 설명한다.

먼저 교배점을 유전 알고리즘의 진화에 따라 변화시키기 위해 교배점 변화의 기준으로써  $T_1$ 부터  $T_j$ 까지 경계 값을 정의하고, 현재 개체 집단의 평균 평가함수 값,  $F_{avg}$ 가  $T_1$ 보다 작으면 교배점의 범위를 lsb bits 방향으로 한 칸 이동 시키고, 다시 차례대로 평균 평가함수 값이 다음 경계값보다 작으면 다음 교배 범위로 이동하게 된다. 이러한 교배점 변화 시기 결정은 병렬 유전자 알고리즘에 존재하는 주 노드(Master Node)에서 수행되며, 부개체집단들에게 전달되게 된다. 실제 경계값의 크기와 개수는 병렬 유전자 알고리즘의 수행시간을 고려해서 설정 되어야 하며, 적용되는 문제에 따라 다르게 설정 될 수 있다. 본 논문에서는 다변수 함수의 최소값을 구하는 문제이므로 평균 평가 함수값이 경계값보다 작을 때 교배 범위를 줄여 주게 된다. 주 노드에서 평균 평가함수 값은 부개체집단에서 계산된 평균 평가함수 값들의 평균 값이며, 부개체집단에서의 평균 평가함수 값은 아래 식 (5)과 같다.

$$f_{avg} = \frac{1}{N} \sum_{k=1}^N f_k(x_1, x_2, x_3, \dots, x_n) \quad (5)$$

여기서  $N$ 은 부개체집단의 크기이며,  $f_k$ 는 현재 부개체집단에 속해 있는 임의의 개체  $k$ 의 평가함수 값을 말한다. 아래의 그림(1)은 이진수로 표현된 하나의 개체에서

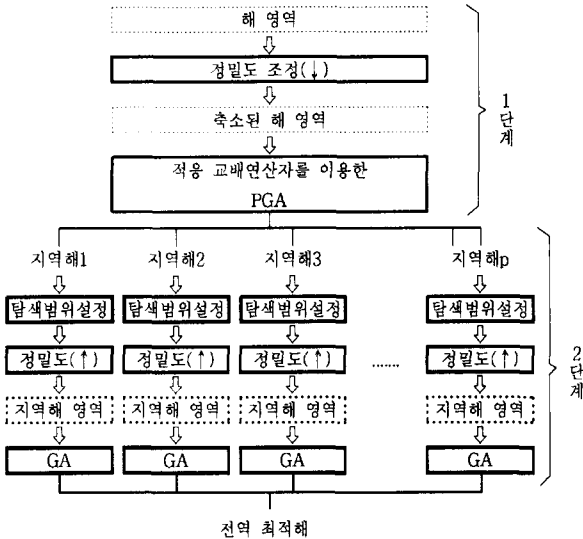


그림 1. 2-단계 PGA의 개요도

### 2.1 수치 최적화와 개체의 표현

유전자 알고리즘에 기초한 수치 최적화를 위해 본 논문에서는 아래의 식(1)과 같은  $n$ 개의 변수를 가지는 다변수함수의 최소값을 구하는 문제를 사용하고 있으며, 식 (2)의 Rastrygin 함수 [5]를 예를 들어 설명하고 있다.

$$\min_{x \in R^n} f(x_1, x_2, x_3, \dots, x_n), \quad f: R^n \rightarrow R \quad (1)$$

위에서 함수  $f$ 는 반드시 연속적일 필요는 없지만, 특정 범위 내에서 제한된 값을 가져야 한다.

$$f(x_1, x_2, x_3, \dots, x_n) = \sum_{i=1}^n \{x_i^2 - 10 \cos(2\pi x_i)\} \quad (2)$$

위와 같은 다변수 함수의 해를 유전자 알고리즘에서 하나의 개체로 표현하기 위해 이진수 표현기법을 사용하고 있으며, 각각의 변수는  $m$ 개의 비트로 다음과 같이 표현 가능하다고 가정한다. 다음은 하나의 변수  $x_k$ 에 대한 이진수 표현이다.

$$b_k^m b_k^{m-1} b_k^{m-2} \dots b_k^2 b_k^1 \quad b \in \{0, 1\} \quad (3)$$

여기서  $b_k^m$ 은 msb(most significant bit)를 나타내며,  $b_k^1$

교배점의 변화를 보여 주고 있으며, 유전알고리즘의 초기 단계에서는 msb bits 인접 지역의 초기 교배 범위에서 교배점을 임의로 선택하여 교배하게 되며, 알고리즘이 진화해 나감에 따라 lsb bits 방향으로 교배 범위가 이동 하게 된다. 그럼에서는 두개의 경계값이 있음을 보여 주고 있다.

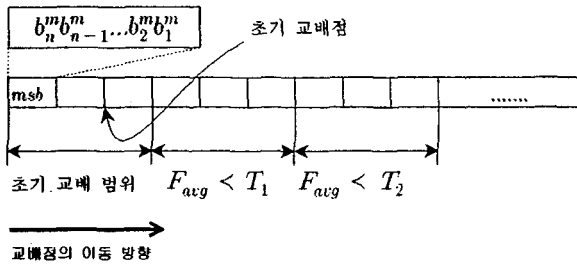


그림 2. 교배점 조정 방법

### 2.3 정밀도 조정

위와 같이 변형된 적응 교배 연산자 알고리즘을 사용하여 찾아진 지역해들의 평가함수 값과 분포를 고려하여 탐색 범위를 결정하게 되면, GA를 수행하기 위해 각각의 지역해 인접 지역에 대해 탐색범위와 정밀도를 결정할 후 GA를 수행하여 진역해를 얻게 된다.

정밀도의 조정은 위에서 언급한 바와 같이 먼저 수행되는 PGA에서는 보다 빠른 탐색을 위해 사용되며, 병렬로 수행되는 GA에서는 보다 자세한 탐색을 위해 사용된다. 정밀도의 조정은 위에서 대략적으로 설명했듯이 하나의 개체를 이진수로 표현 할 때 탐색 영역을 표현 할 수 있는 비트수를 제한하는 것이다. 즉,  $n$ 개의 변수를 가지는 다변수 함수의 최소값을 구하는 문제에서 모든 변수들의 탐색영역 크기가 같고, 하나의 변수를 표현 할 수 있는 비트수가  $m$ 이라고 가정 한다면, 하나의 개체를 표현하기 위해  $(n \times m)$ 개의 비트를 필요로 하게 된다. PGA에서는 하나의 변수를  $m$ 보다 작은 비트를 사용하여 표현함으로써 하나의 개체에 대한 표현 비트수가 줄어들게 되므로 탐색 영역이 이전보다 줄어들게 되는 것이다. 또한, PGA에 의해 결정된 일정 범위에서 GA를 다시 실행할 때를 살펴보면, PGA에서와 같이  $m$ 개의 비트로 하나의 변수를 다시 표현 할 경우, 이미 현저히 줄어든 탐색영역을 같은 수의 비트 수  $m$ 으로 표현하게 되므로, 좁은 영역에서 굉장히 정밀하게 탐색을 할 수 있게 되는 것이다. 예를 들어, Rastrygin함수에서 특정 변수의 탐색 범위가  $[-5.11, 5.12]$ 일 때, 필요한 비트수는 10이며 0.01의 간격으로 표본추출(Sampling)된 탐색영역을 가지게 된다. 여기서 표현 비트수를 9로 낮추게 되면, 똑같은 범위에서 0.02간격으로 표본 추출된 반으로 줄어든 탐색영역을 가지게 되는 것이다. 그리고  $[-0.511, 0.512]$ 로 줄어든 탐색 범위를 똑같은 비트수 10으로 표현하게 되면, 0.001 간격으로 표본추출을 할 수 있으며

로 보다 자세한 탐색을 할 수 있는 것과 같은 원리이다.

이와 같이 교배점 조정과 정밀도 조정을 이용하여 유전자 알고리즘의 병렬화를 수행하기 위해서는 PGA를 통해 얻은 지역해들로부터 GA들이 실행될 탐색범위들을 중복되지 않게 선택하는 것이 중요하며, 분산 환경에서 노드들의 수를 고려하여 지역해의 수를 결정하는 것이 제안된 병렬화 방법의 효율성을 증대시키기 위해 필요한 부분이다.

### 3. 결론

본 논문에서 제안하는 2-단계 병렬 유전자 알고리즘은 기존의 병렬 유전자알고리즘과는 다른 관점에서 제안된 방법이다. 즉, 기존의 병렬 유전자알고리즘이 개체집단의 분할을 통해 병렬성을 개발한 알고리즘이라고 한다면, 제안된 방법은 탐색영역의 분할을 통해 병렬성을 개발한 방법이라고 할 수 있다. 이는 기존의 병렬 유전자알고리즘이 부개체집단들의 중복된 영역 탐색 가능성으로 인해 병렬성을 감소시킬 수 있는 반면, 제안된 병렬화 방법은 탐색 영역의 중복 가능성이 전혀 없으므로 유전자알고리즘의 병렬화에 새로운 가능성을 보여준 방법이라고 할 수 있다. 또한 이러한 특성은 분산 환경에서 각 노드들 간의 통신량에서 잘 나타나게 된다. 즉, 기존 병렬 유전자 알고리즘에서 노드들 사이의 통신량이 알고리즘 수행 시간의 큰 비중을 차지하는 것에 비해 제안된 알고리즘에서는 독립된 탐색영역들에서 각각 유전자 알고리즘이 수행되기 때문에 통신량에 의한 부담이 거의 없다는 장점을 가지게 된다. 그리고 본 논문에서 제안된 병렬화 방법은 변형된 적응교배연산자와 정밀도 조정을 이용해 빠른 범위탐색과 정밀한 부분탐색을 의도하고 있다는 측면에서 유전자알고리즘의 전역탐색과 지역탐색의 균형을 다른 차원에서 구현하고 있다는 것에 큰 의미를 둘 수 있다.

### 참고문헌

- [1] Y. Svirzhev and V. Passekov, "Fundamentals of Mathematical Evolutionary Genetics," Kluwer Academic Publishers, Mathematics and Its Application, Vol.22, 1989.
- [2] T. Starkweather, D. Whitley, and K. Mathias, "Optimization using distributed genetic algorithm," Parallel Problem Solving from Nature, Springer Verlag, 1991.
- [3] L. Tan, D. Taniar and K. A. Smith, "A New Parallel Genetic Algorithm," Proc. of Int. Symposium on Parallel Architectures, Algorithms and Networks, pp.284-289, 2002.
- [4] M. S. Ko, T. W. Kang and C. S. Hwang, "Adaptive Crossover Operator based on Locality and Convergence," IEEE Int. Joint Symposia on Intelligence and Systems, pp.18-22, 1996.
- [5] G. Lin and X. Yao, "Analysing Crossover Operators by Search Step Size," IEEE Int. Conference on Evolutionary Computation, pp.107-110, 1997.