

균일분포의 파레토 최적해 생성을 위한 다목적 최적화 진화 알고리즘

장 수 현* · 윤 병 주**

요 약

진화 알고리즘은 여러 개의 상충하는 목적을 갖는 다목적 최적화 문제를 해결하기에 적합한 방법이다. 특히, 파레토 지배관계에 기초하여 개체의 적합도를 평가하는 파레토 기반 진화알고리즘들은 그 성능에 있어서 비교적 우수한 평가를 받고 있다. 그러나 일반화된 다목적 최적화 진화알고리즘은 복잡한 문제들에서 찾아진 해들의 분포가 전체 파레토 경계면에 대하여 균일하지 못하고 특정 지역에서 집중적으로 해를 생성하는 문제점을 가지고 있다. 본 논문에서 우리는 이러한 문제점을 보완하기 위한 다목적 최적화 진화알고리즘을 제안한다. 제안한 알고리즘은 현재까지 찾아진 최적해들 중 특정 지역에 편중되지 않은 해를 우수 종자로 복제 연산에 참여시킨다. 따라서 특별한 지역탐색 기법을 사용하지 않아도 종자가 되는 개체 주위에 새로운 개체를 생성할 확률이 높기 때문에 지역탐색의 효과를 가질 수 있고, 비교적 고른 분포의 파레토 최적해를 생성할 수 있다. 5개의 테스트 함수에 대한 실험 결과, 제안한 알고리즘은 모든 문제에서 전체 파레토 경계면에 균일한 분포의 해들을 생성할 수 있었으며, 많은 지역해를 가지는 문제를 제외한 모든 문제에서 NSGA-II보다 우수한 수렴 결과를 보였다.

Evolutionary Multi-Objective Optimization Algorithms for Uniform Distributed Pareto Optimal Solutions

Su-Hyun Jang* · Byungjoo Yoon**

ABSTRACT

Evolutionary algorithms are well-suited for multi-objective optimization problems involving several, often conflicting objectives. Pareto-based evolutionary algorithms, in particular, have shown better performance than other multi-objective evolutionary algorithms in comparison. However, generalized evolutionary multi-objective optimization algorithms have a weak point, in which the distribution of solutions are not uniformly distributed onto pareto optimal front. In this paper, we propose an evolutionary algorithm for multi-objective optimization which uses seed individuals in order to overcome weakness of algorithms published. Seed individual means a solution which is not located in the crowded region on pareto front. And the idea of our algorithm uses seed individuals for reproducing individuals for next generation. Thus, proposed algorithm takes advantage of local searching effect because new individuals are produced near the seed individual with high probability, and is able to produce comparatively uniform distributed pareto optimal solutions. Simulation results on five testbed problems show that the proposed algorithm could produce uniform distributed solutions onto pareto optimal front, and is able to show better convergence compared to NSGA-II on all testbed problems except multi-modal problem.

키워드 : 다목적 최적화(Multi-objective Optimization), 진화알고리즘(Evolutionary Algorithms), 균일 분포 파레토 경계면(Uniform Distributed Pareto Front), 종자 개체(Seed Individual)

1. 서 론

대부분의 실제 문제들은 여러 개의 비교 불가능하고 중 상충하는 목적들을 동시에 만족시켜야 하는 최적화 문제들이다. 하나의 목적을 최적화 하는 문제의 경우 최적해가 일반적으로 명확하게 정의 되지만 다목적 최적화(multi-ob-

jective optimization)의 경우는 그렇지 못하다. 다목적 최적화 문제는 하나의 최적해가 아니라 일반적으로 파레토 최적해(pareto optimal solutions)라 불리는 여러 개의 상호 보완적인 해들의 집합을 가진다. 넓은 의미에서 최적해들은 모든 목적들을 고려할 때, 그 해보다 좋은 해가 탐색공간에 존재하지 않음을 의미한다.

진화 알고리즘은 병행 탐색 능력으로 인하여 한번의 실행으로 여러 개의 파레토 최적해들을 찾을 수 있는 가능성을 내재하고 있다. 그러나 많은 복잡한 문제들에서 전체 파레

* 정 회 원 : 명지대학교 대학원 컴퓨터공학과
 ** 종신회원 : 명지대학교 컴퓨터공학과 교수
 논문접수 : 2004년 7월 23일, 심사완료 : 2004년 10월 2일

토 경계면(parato front)의 해들에 비해 상대적으로 충분하지 못한 수의 해들을 생성할 가능성을 가지고 있고, 특정 지역의 해들만을 찾음으로서 문제에 대한 다양한 해를 생성하지 못할 수도 있다. 따라서 진화 알고리즘을 이용한 다목적 최적화는 다음과 같은 세 가지 목적을 가지는 일반적인 형태로 재구성 될 수 있다.

- 알고리즘의 결과로 얻어진 파레토 경계면과 최적 파레토 경계면의 거리는 최소화 되어야 한다. 즉, 최적 파레토 경계면에 잘 수렴하여야 한다.
- 찾아진 해들의 분포는 균일 분포를 가져야 한다.
- 찾아진 해들은 전체 파레토 경계면을 모두 포함하여야 한다. 즉, 특정 지역의 해들이 모두 소멸되지 않아야 한다.

다목적 최적화 진화알고리즘들은 위에서 제시한 3가지 목적을 만족시키기 위한 여러 가지 방법들을 제시하고 있다[1-8]. 이는 어떤 적합도 할당 전략과 선택 전략을 사용할 것인가의 문제, 어떤 방법으로 개체집단의 다양성을 유지하도록 할 것인가의 문제, 그리고 어떤 최적해 보존 전략을 사용할 것인가의 문제로 귀착된다. 그리고 지금까지 고안되었던 대부분의 다목적 최적화 진화알고리즘들은 위의 문제들에 대한 서로 다른 접근 방법들을 제시하고 있다.

본 논문에서 우리는 종자 개체를 이용한 다목적 최적화 진화알고리즘(SIEA : Seed Individual based Evolutionary Algorithms for Multi-objective Optimization Problems)을 제안한다. SIEA의 기본적인 아이디어는 농생물학의 우수종자 개량 과정을 다목적 최적화 진화알고리즘에 적용한 것으로, 보다 우수한 개체를 만들기 위해 현재까지 찾아진 최우수 개체를 우수종자로 복제 연산에 참여시키는 것이다. 따라서 특별한 지역탐색의 기법을 사용하지 않아도 최우수 개체 주위에 비교적 우수한 성능을 가지는 다음 세대 개체를 생성할 확률이 높기 때문에 지역탐색의 효과를 가질 수 있다. 또한 최우수 개체의 선택 기준이 비지배(nondominated) 개체들 중 이웃하는 개체와의 거리가 가장 큰 개체를 선택하고, 그 개체를 중심으로 복제연산을 수행하기 때문에 비교적 고른 분포의 파레토 최적해를 구할 수 있다.

2. 다목적 최적화문제

다목적 최적화(multi-objective optimization)문제는 실세계에 많이 존재한다. 휴대폰, 자동차 등과 같은 복잡한 하드웨어/소프트웨어 시스템의 디자인문제를 예로 들 수 있다. 그러한 시스템은 비용은 최소화해야 하는 반면 최대의 성능을 요구한다. 응용에 따라서는 신뢰도와 같은 더 많은 목적이 중요한 요소가 되기도 한다. 다목적 최적화문제는 서로 다른 최적화 기준과 제약들로 정형화 될 수 있다. 다목적 최

적화문제는 다음과 같이 정의될 수 있다.

정의 1 : (다목적 최적화 문제) 일반적으로 다목적 최적화 문제는 n 개의 결정변수(decision variables)들과 k 개의 목적 함수(object function)들, 그리고 m 개의 제약(constraint)들을 포함한다. 그리고 목적함수들과 제약들은 결정변수들의 함수이다. 즉,

$$\begin{aligned} & \text{maximize/minimize } y = f(x) = (f_1(x), f_2(x), \dots, f_k(x)) \\ & \text{subject to } e(x) = (e_1(x), e_2(x), \dots, e_m(x)) \\ & \text{where } x = (x_1, x_2, \dots, x_n) \in X \\ & \quad y = (y_1, y_2, \dots, y_k) \in Y \end{aligned} \quad (1)$$

식 (1)에서 x 는 결정벡터(decision vector)이고, y 는 목적 벡터(object vector), X 는 결정 공간(decision space), Y 는 목적 공간(object space)이다. 그리고 제약 $e(x)$ 는 가능한 해들의 집합(feasible solution set)을 결정한다.

정의 2 : (가능한 해집합) 가능한 해집합 X_f 는 제약 $e(x)$ 를 만족하는 결정벡터 x 의 집합으로 식 (2)와 같이 정의된다.

$$X_f = \{x \in X \mid e(x)\} \quad (2)$$

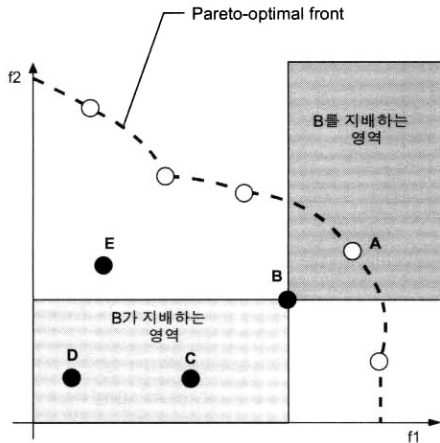
따라서 목적 공간에서 가능한 영역은 X_f 와 사상된 영역으로 $Y_f = f(X_f) = \bigcup_{x \in X_f} f(x)$ 이다.

최대화 문제를 가정하였을 때, 단목적 최적화문제에서 가능한 해집합내의 모든 원소는 목적함수 f 에 따라 완전하게 순서화 된다. 예를 들어 두개의 결정벡터 $a, b \in X_f$ 에 의해 사상되는 목적 벡터 $f(a), f(b) \in Y_f$ 는 $f(a) \geq f(b)$ 이거나 $f(b) \geq f(a)$ 의 관계로 완전히 순서화 된다. 그리고 단목적 최적화 문제의 목적은 함수 f 의 값을 최대로 하는 해(결정 벡터)를 찾는 것이다. 그러나 몇 개의 목적을 갖는 다목적 최적화 문제에서 X_f 내의 해들은 일반적으로 완전히 순서화 되지 못하고 부분 순서를 갖게 된다. (그림 1)에서 B로 표현되는 해는 C로 표현되는 해보다 우수하다. 그러나 B는 E와 비교했을때 f_1 에 대하여 우수하나 f_2 에 대하여 우수하지 못하다. 즉 B와 E는 해의 순서화가 불가능하다. 해들의 순서관계에 기초하여 다음의 파레토 지배관계를 정의할 수 있다.

정의 3 : (파레토 지배관계) 두개의 결정벡터 a 와 b 에 대한 파레토 지배관계는 다음과 같이 정의된다.

$$\begin{aligned} a > b & \text{ (} a \text{ dominates } b\text{)} \\ & \text{iff } f_i(a) > f_i(b) \quad \forall i = 1, \dots, k \\ a \geq b & \text{ (} a \text{ weakly dominates } b\text{)} \\ & \text{iff } f_i(a) \geq f_i(b) \quad \forall i = 1, \dots, k \\ a \sim b & \text{ (} a \text{ is indifferent to } b\text{)} \\ & \text{iff } [f_i(a) \neq f_i(b)] \wedge [f_i(b) \neq f_i(a)] \quad \forall i, j = 1, \dots, k \end{aligned}$$

(그림 1)에는 B로 표현된 결정 벡터에 의해 지배되는 목적공간의 영역과 B를 지배하는 목적공간의 영역이 표현되어 있다. 그리고 그 이외의 영역은 명확하게 지배관계를 정의할 수 없는 영역이다. 결정벡터 A는 B를 지배하고, C와 D는 B에 의해 지배되는 결정벡터이다. 그리고 E와 B사이의 지배관계는 정의되지 않는다.



(그림 1) 목적 공간에서 파레토 최적해와 해들 사이의 관계

파레토 지배관계의 개념에 기초하여 다목적 최적화 문제의 최적화 개념을 유도할 수 있다. (그림 1)에서 결정벡터 A는 다른 어떤 결정벡터에 의해서도 지배되지 않는다. 그리고 이러한 해를 파레토 최적해라 한다.

정의 4 : (파레토 최적해) 결정벡터 $x \in X_f$ 는 $\forall a \in A : a \succ x$ 인 결정벡터들의 집합 $A \subseteq X_f$ 에 대하여 비지배 (nondominated) 된다고 한다. 특히, x 가 X_f 에 대하여 비지배 된다면, x 를 파레토 최적해(parato optimal solution)라 한다.

(그림 1)에서 흰 점으로 표현된 해들은 파레토 최적해들이다. 그들은 각 목적에 대한 서로 다른 가중치가 존재하지 않는다면 더 좋은 해를 찾을 수 없는 서로 상호 보완적인 해들의 집합이다. 이러한 파레토 최적해들의 집합을 파레토 최적 해집합(pareto optimal set)이라하고, 각 해에 대응되는 목적 공간의 목적벡터들은 파레토 최적 경계면(pareto optimal front)을 구성한다.

3. NSGA-II (Nondominated Sorting GA-II)

이 장에서는 우리가 제안하는 SIEA와 실험적 비교 대상이 되는 NSGA-II[6] 알고리즘을 설명한다. NSGA-II는 NSGA[5]를 개선한 알고리즘으로 비지배 개체 정렬(non-dominated sorting)의 높은 연산 복잡도를 개선하고, 엘리티즘(elitism)을 보다 강화하였다. 그리고 밀도측정(density estimation) 연산

자를 추가 함으로서 적합도 평가에서 NSGA와 같이 공유 매개변수(sharing parameter) σ_{share} 를 정의할 필요성을 제거한 알고리즘이다.

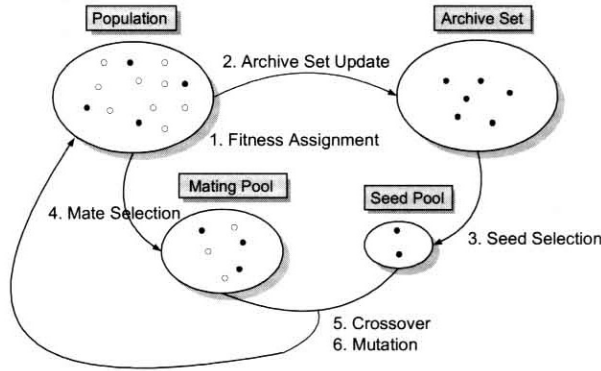
NSGA-II의 다목적 최적화 과정은 다음과 같다. 먼저 개체집단내의 모든 개체들이 파레토 지배관계에 따라 서로 다른 경계면들로 분리된다. 그리고 각 경계면에 속하는 개체들에 순위(rank)를 할당한다. 이때 첫 번째 비지배 경계면에 속하는 개체는 가장 높은 순위가 할당되고, 두 번째 비지배 경계면에 속하는 개체들은 두 번째 높은 순위가 할당되는 방식으로 모든 경계면에 순위가 할당된다. 같은 순위에 속하는 개체들에 대하여 각 목적에 따라 인접한 두 개체의 거리의 합으로 밀도를 측정한다. 각 개체의 밀도는 같은 순위의 개체들에 대한 부분 순서로서 사용된다. 즉, 같은 순위를 가지고 있는 개체일 지라도 높은 밀도를 가지는 개체는 낮은 밀도를 가지고 있는 개체와의 토너먼트 선택연산에서 탈락된다. NSGA-II는 엘리티즘을 강화하는 방법으로 부모 개체집단과 자식 개체집단을 결합한 후, 결합된 개체집단에 대한 순위 할당과 밀도측정을 한다. 그리고 성능이 낮은 50%를 탈락시킨 후, 나머지 50%를 다음 세대의 자식 개체집단을 만들기 위한 이진 토너먼트 선택(binary tournament selection)에 사용한다.

발표된 논문들에 따르면[6-8], NSGA-II의 수렴 성능과 만들어진 해의 균일 분포 상태는 대부분의 진화적 최적화 방법들 보다 우수한 결과를 보이고 있다.

4. 종자 개체를 이용한 다목적 최적화 진화알고리즘 (SIEA : Seed Individual based Evolutionary Algorithms for Multi-objective Optimization Problems)

SIEA는 다른 진화적 최적화 알고리즘들과 다른 진화 방식을 사용한다. 일반적인 진화알고리즘에서는 다음세대의 개체를 생성하기 위해 하나의 교배개체집단(mating pool)을 구성하고 교배와 돌연변이연산을 수행한다. 그러나 SIEA는 두개의 교배 개체집단을 사용한다. 하나는 현재까지 찾아진 가장 우수한 개체들을 위한 작은 수의 우수종자 보존소(seed pool)이고 다른 하나는 일반적인 교배 개체집단이다. 그리고 교배 방법은 우수종자 보존소에 있는 각각의 개체가 교배 개체집단에 있는 모든 개체와 교배를 한다. 종자 개체는 archive set에 있는 개체들 중에서 선택된다. archive set은 현 세대까지 찾아진 모든 비지배 개체들의 집합이다. 따라서 현 세대의 archive set에 속하는 개체들의 분포를 고려하여 종자개체를 선택함으로써 알고리즘의 탐색 방향을 조절 할 수 있다. 이를 위한 종자 개체의 선택 방법은 archive set에 있는 개체들 사이의 거리를 계산하여 이웃과의 거리가 가장 큰, 즉 특정지역에 편중되지 않은 개체를 선택한다. 따라서 현 세대 까지 찾아진 최적해의 분포가 어느 특정지

역에 편중되어 있다면, 다음 세대 개체집단의 탐색점이 비교적 편중되지 않은 지역으로 이동하게 됨으로 비교적 고른 분포의 파레토 최적해들을 생성할 수 있다. SIEA의 전체 진화과정은 (그림 2)와 같다.



(그림 2) SIEA 진화과정

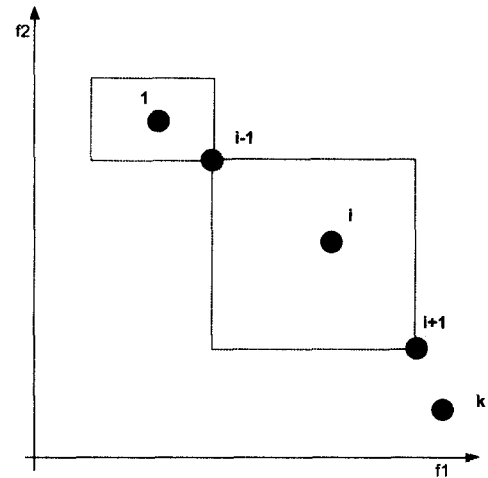
4.1 개체 적합도 계산

개체집단 P 내의 개체들에 대한 적합도는 개체의 순위와 같은 순위의 개체들 사이의 거리를 사용하여 계산한다. 개체의 순위는 NSGA-II에서 사용한 Nondominated sorting 방법을 사용하여 첫 번째 비지배 경계면의 개체들에는 순위 1을, 다음 비지배 경계면의 개체들에는 순위 2를 할당하는 방법을 사용하였다. 그리고 같은 순위의 개체들 사이의 부분 순서관계를 부여하기 위해 이웃개체와의 거리 정보를 이용한다. 개체의 거리는 같은 순위를 갖는 개체 중 이웃하는 개체들과의 거리의 합으로 계산한다.

(그림 3)에서 개체 i 의 거리는 같은 순위의 개체 중 이웃하는 $i-1$ 과 $i+1$ 을 연결한 입방체의 각 차원의 거리 합이 된다. 그리고 양 끝에 있는 1과 k 의 경우는 이웃하는 개체가 1개밖에 없기 때문에, 이웃한 개체와의 거리를 2배로 확장한 입방체의 각 차원의 거리 합으로 계산한다. 이와 같이 각 개체의 순위와 거리를 계산한 후, 개체의 적합도는 식 (4)와 같이 계산한다.

$$Fit(i) = R(i) + (1 - (D(i)/D_{max})) \quad (4)$$

식 (4)에서 $R(i)$, $D(i)$ 는 개체 i 의 순위와 거리이고, D_{max} 는 개체 i 와 같은 순위를 가지는 개체들 중 이웃과의 거리가 가장 큰 개체의 거리이다. 식 (4)에 의해 거리가 가장 큰 개체의 적합도는 그 개체의 순위와 같고, 거리가 작을수록 큰 적합도 값을 가지게 된다. 여기에서 적합도는 전체 개체집단 내에서 첫 번째 비지배 되는 개체들의 순위가 1이고 다음 비지배 되는 개체들의 순위가 1씩 커지기 때문에 일반적인 적합도 개념의 역이된다. 따라서 큰 적합도 값을 가진다는 의미는 작은 적합도 값을 가지는 개체와의 토너먼트 선택연산에서 탈락하게 됨을 뜻한다.



(그림 3) 개체의 거리 계산

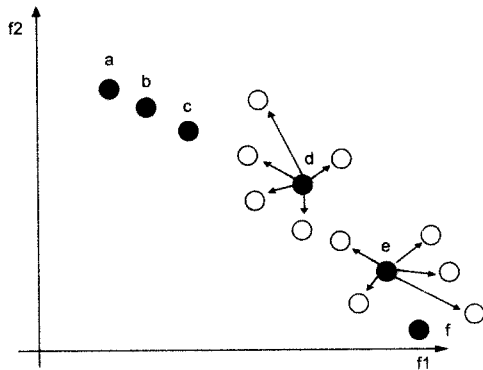
4.2 비지배 개체집합 구성

비지배 개체집합 즉, archive set의 구성은 현재 세대까지 찾은 파레토 경계면의 개체들을 다음 세대들에서 잃어버리지 않도록 사용하는 부가적 메모리이다. 또한 다목적 최적화 진화 알고리즘의 실행 결과가 된다. SIEA에서 비지배 개체집합은 매 세대마다 갱신한다. 비지배 개체집합의 갱신은 다음의 과정으로 수행한다.

- ① 현 세대의 개체집단 내에서 첫 번째 비지배 개체들을 비지배 개체 집합에 추가한다.
- ② 추가된 비지배 개체집합 내의 모든 개체들의 지배관계를 고려하여 어떤 개체에 의해 지배되는 개체들을 제거한다.
- ③ 비지배 개체집합의 최대 크기보다 많은 비지배 개체들이 만들어지면 비지배 개체집합 내에서 4.1에서 설명한 이웃개체와의 거리 평가 방법을 사용하여 거리가 작은 개체부터 비지배 개체집합으로부터 제거한다.

4.3 복제 연산을 위한 개체 선택

다음 세대 개체들을 생성하기 위한 복제 연산에 참여할 개체의 선택은 두 가지 방법으로 수행된다. 첫째 교배집단 (mating pool)의 구성은 개체집합(population)에서 개체의 적합도를 기반으로 이진토너먼트 선택에 의해 교배집단을 구성한다. 둘째 우수 종자보존소(seed pool)의 구성은 현 세대의 비지배 개체집합(archive set)에서 지역편중이 되지 않은 개체들을 선택한다. 지역 편중의 정도는 4.1에서 설명한 개체의 이웃 거리 계산 방법을 사용한다. (그림 4)의 비지배 개체집합의 경우 가장 지역편중이 되지 않은 개체는 d 이고, 다음은 e 이다. 따라서 d 와 e 가 우수종자 보존소의 개체로 선택되어진 후, 교배 집단의 개체들과 교배와 돌연변이 연산을 수행한 자식 개체들의 분포가 d 와 e 부근에 집중적으로 생성될 확률이 높다.

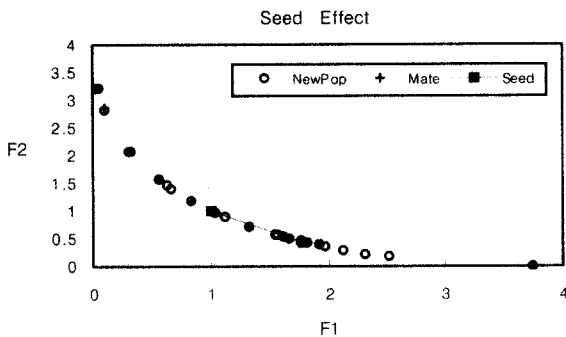


(그림 4) 비지배 개체집합의 분포와 Seed 선택 결과의 예

종자 개체 선택의 효과를 입증하기 위한 사전 실험으로 우리는 Schaffer[9]에 의해 사용되었던 비교적 간단한 다목적 최적화 문제를 적용하였다. Schaffer 문제는 식 (5)와 같이 정의된다.

$$\begin{aligned}
 & \text{minimize } f(x) = (f_1(x), f_2(x)) \\
 & \text{subject to } f_1(x) = x^2 \\
 & \quad \quad f_2(x) = (x - 2)^2
 \end{aligned} \tag{5}$$

(그림 5)는 Schaffer 문제에서 2개의 우수종자 개체와 교배집단에 있는 10개의 개체를 교배하여 생성된 40개의 새로운 개체의 분포이다. 새로 생성된 개체들의 분포는 교배집단내의 개체들 보다는 우수종자 개체의 주위에 많이 생성되었다. 따라서 우수종자 개체를 중심으로 하는 지역탐색의 효과를 가질 수 있음을 보여주고 있다.



(그림 5) 우수 종자 개체와 교배집단 개체의 교배를 통해 생성된 개체들의 분포

5. 실험 및 토의

5.1 평가 방법

본 논문에서 우리는 최적화 알고리즘의 성능을 평가하기 위한 척도로서 C metric(Coverage Metric)[10]과 S metric(Spacing Metric)[11]을 사용한다. C metric은 한 알고리즘의 결과가 다른 알고리즘의 결과를 얼마나 지배하는가를 보

여줄 수 있는 평가방법이다. C metric은 다음과 같이 정의된다. 두 결정벡터 집합 X', X'' 에 대해, 함수 C 는 순서쌍 (X', X'') 을 $[0, 1]$ 사이의 값으로 사상시키는 함수이다.

$$C(X', X'') := \frac{|\{a'' \in X''; \exists a' \in X': a' \succ a''\}|}{|X''|} \tag{6}$$

$C(X', X'') = 1$ 은 X'' 의 모든 결정벡터가 X' 의 결정벡터에 의해 지배되거나 같음을 의미한다. 반대로, $C(X', X'') = 0$ 은 X'' 의 어떤 결정벡터도 X' 의 결정벡터에 의해 지배되지 않음을 의미한다. C metric 평가는 $C(X', X'')$ 과 $C(X'', X')$ 모두를 고려해야 한다. 즉 $C(X'', X')$ 의 평가 값이 $1 - C(X', X'')$ 과 같지 않기 때문이다.

S metric은 파레토 경계면에 해들이 얼마나 균일하게 분포 하는가를 평가하기 위한 평가방법이다. 즉 S metric은 파레토 경계면에서 가장 이웃한 해 사이의 거리에 대한 분산을 측정한다. S metric은 식 (7)과 같이 정의된다.

$$\begin{aligned}
 S &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\bar{d} - d_i)^2}, \\
 d_i &= \min_j (|f_1^i(x) - f_1^j(x)| + |f_2^i(x) - f_2^j(x)| \\
 &\quad + \dots + |f_k^i(x) - f_k^j(x)|)
 \end{aligned} \tag{7}$$

식 (7)에서 $i, j = 1, 2, \dots, n$ 이고, n 은 파레토 경계면에 있는 해의 수, k 는 목적 함수의 개수, \bar{d} 는 전체 d_i 의 평균이다. S의 값이 0이면 파레토 경계면의 모든 해들이 같은 거리를 유지하면서 분포하고 있음을 의미한다.

본 논문에서 사용한 C metric과 S metric은 서론에서 제시한 다목적 최적화의 일반화된 세 가지 목적 중 첫 번째와 두 번째 목적에 대한 평가방법이다. 즉, C metric은 첫 번째 목적의 수렴 성능에 대한 상대적 평가방법이다. 최적 파레토 경계면을 이미 알고 있다면 절대적 평가 방법을 사용할 수 있지만, 대부분의 다목적 최적화 문제의 경우 최적 파레토 경계면을 미리 알고 있지 못하기 때문에 알고리즘의 수렴성능 평가로 C metric과 같은 상대 평가 방법을 사용하였다.

5.2 테스트 함수

제한한 SIEA의 성능을 평가하기 위해 Deb[12, 13]이 제시한 테스트 함수들을 실험대상으로 하였다. 테스트 함수들은 다목적 최적화 진화알고리즘들이 최적 파레토 경계면에 수렴하고 균일분포의 파레토 해들을 만드는데 있어 어려움의 원인이 될 수 있는 몇 가지 특징을 가지고 있다. 각 실험 함수들과 그 특징은 다음과 같다.

각 테스트 함수들은 식 (8)과 같이 3개의 함수 f_1, g, h 로 정의된다.

$$\begin{aligned}
 & \text{Minimize } T(X) = (f_1(x_1), f_2(X)) \\
 & \text{subject to } f_2(X) = g(x_2, \dots, x_m) \times h(f_1(x_1), g(x_2, \dots, x_m)) \\
 & \text{where } X = (x_1, \dots, x_m)
 \end{aligned} \tag{8}$$

함수 f_1 은 첫 번째 결정변수 x_1 만을 파라메타로 가지는 함수이고, g 는 나머지 $m-1$ 개의 변수로 이루어진 함수이다. 그리고 h 는 f_1 과 g 의 함수 값을 파라메타로 가진다.

<표 1>은 식 (8)의 형태로 표현된 5개의 함수들이다.

<표 1> 테스트 함수

함수명	함수 정의	비고
T_1	$f_1(x_1) = x_1$ $g(x_2, \dots, x_m) = 1 + 9 \cdot \sum_{i=2}^m x_i / (m-1)$ $h(f_1, g) = 1 - \sqrt{f_1/g}$	$m = 30$ $x_i \in [0, 1]$
T_2	$f_1(x_1) = x_1$ $g(x_2, \dots, x_m) = 1 + 9 \cdot \sum_{i=2}^m x_i / (m-1)$ $h(f_1, g) = 1 - (f_1/g)^2$	$m = 30$ $x_i \in [0, 1]$
T_3	$f_1(x_1) = x_1$ $g(x_2, \dots, x_m) = 1 + 9 \cdot \sum_{i=2}^m x_i / (m-1)$ $h(f_1, g) = 1 - \sqrt{f_1/g} - (f_1/g) \sin(10\pi f_1)$	$m = 30$ $x_i \in [0, 1]$
T_4	$f_1(x_1) = x_1$ $g(x_2, \dots, x_m) = 1 + 10(m-1) + \sum_{i=2}^m (x_i^2 - 10 \cos(4\pi x_i))$ $h(f_1, g) = 1 - \sqrt{f_1/g}$	$m = 10$ $x_1 \in [0, 1]$, $x_2, \dots, x_m \in [-5, 5]$
T_5	$f_1(x_1) = 1 - \exp(-4x_1) \sin^6(6\pi x_1)$ $g(x_2, \dots, x_m) = 1 + 9 \cdot ((\sum_{i=2}^m x_i) / (m-1))^{0.25}$ $h(f_1, g) = 1 - (f_1/g)^2$	$m = 30$ $x_i \in [0, 1]$

테스트 함수 T_1 은 볼록한(convex) 최적 파레토 경계면을 가지는 문제이다. 볼록한 파레토 경계면을 가지는 문제는 중간해가 지배하는 영역이 양쪽 끝에 있는 해가 지배하는 영역보다 커지는 경향이 있기 때문에 순수 파레토 지배관계만을 가지고 적합도를 평가하는 진화알고리즘들은 보다 많은 중간해를 찾으려는 경향을 갖게 된다.

테스트 함수 T_2 은 T_1 과 반대로 오목한 최적 파레토 경계면을 가지는 문제이다. 오목한 파레토 경계면을 가지는 문제는 볼록한 파레토 경계면을 가지는 문제와는 반대로 양쪽 끝에 존재하는 해를 찾으려는 경향을 가지게 된다.

테스트 함수 T_3 은 파라메타 공간에서는 연속적이지만 목적 공간에서는 몇 개의 비연속적인 볼록한 부분으로 이루어진 최적 파레토 프론트를 가지는 이산적인 특징을 갖는다. 이산적인 문제는 비연속적인 서로 다른 부분에 속한 해들 사이의 경쟁으로 인하여 어떤 부분의 해들이 멸종될 수도 있다.

테스트 함수 T_4 는 21^9 개의 지역적인(local) 최적 파레토 경계면들을 가지는 문제이다. 그리고 모든 지역적인 최적 파레토 해들은 목적 공간에서 구분가능하지 않다. 이 문제는 많은 지역적인 파레토 경계면을 가지게 되기 때문에 다목적 최적화 진화 알고리즘이 전역해에 수렴하기 어려운 문제이다.

테스트 함수 T_5 는 탐색공간의 비균일 분포에 의해 만들어 지는 두 가지 어려움을 가지고 있다. 첫 번째는 최적 파레토 해들이 전역적인 파레토 경계면에서 균일하지 못하게 분포 한다는 것이고, 두 번째는 해들의 밀도가 전역 파레토 경계면 근처에서 가장 작고 경계면으로부터 멀어질수록 커진다.

5.3 실험 결과

실험은 각 문제에 대해 각각 30회 수행 하였으며, 각 알고리즘에서 사용한 공통 파라메타는 다음과 같다.

세대수 : 300세대

개체집합 크기(population size) : 100

교배연산 확률(crossover rate) : 0.8

돌연변이연산 확률(mutation rate) : 1/개체 비트 스트링 수

각 결정변수들의 표현 길이 : 30비트

그리고 SIEA에서 우수종자 집단의 크기는 2, 교배 집단의 크기는 25로 설정하였다.

<표 2>는 각 문제에 대한 C metric 결과이다. C metric 결과는 T_1, T_2, T_3 문제에서 SIEA의 결과가 NSGA-II의 결과보다 상대적으로 최적 파레토 경계면에 잘 수렴함을 보이고 있다. 특히 T_5 문제에서는 SIEA의 결과가 월등히 최적 파레토 경계면에 잘 수렴하는 결과를 보였다. 그러나 T_4 문제에서는 NSGA-II의 결과가 보다 우수한 결과를 보였다. 이는 T_4 문제와 같은 많은 지역해를 가지는 문제의 경우 전체 개체집단의 다양성을 유지하는 것이 지역해를 벗어날 수 있는 보다 많은 기회를 가질 수 있기 때문이라고 사료된다. SIEA는 우수종자 개체를 중심으로 복제 연산이 이루어지기 때문에 NSGA-II에 비해 전체 개체집단의 다양성 유지 측면에서는 약점을 가지고 있다. 그러나 T_4 와 같이 많은 지역해를 가지는 문제를 제외한 대부분의 다목적 최적화 문제에서 SIEA의 수렴 성능은 우수종자 개체에 의한 지역탐색 효과에 의해 NSGA-II보다 우수하다.

<표 2> C metric 결과

문 제		C (NSGA-II, SIEA)	C (SIEA, NSGA-II)
T_1	평균	0.0307	0.571
	최대	0.18	0.84
	최소	0	0.22
T_2	평균	0.0207	0.7342
	최대	0.15	0.9588
	최소	0	0.4
T_3	평균	0.0511	0.3189
	최대	0.12	0.53
	최소	0	0.15
T_4	평균	0.7578	0.2193
	최대	1	1
	최소	0	0
T_5	평균	0.0006	0.9845
	최대	0.01	1
	최소	0	0.924

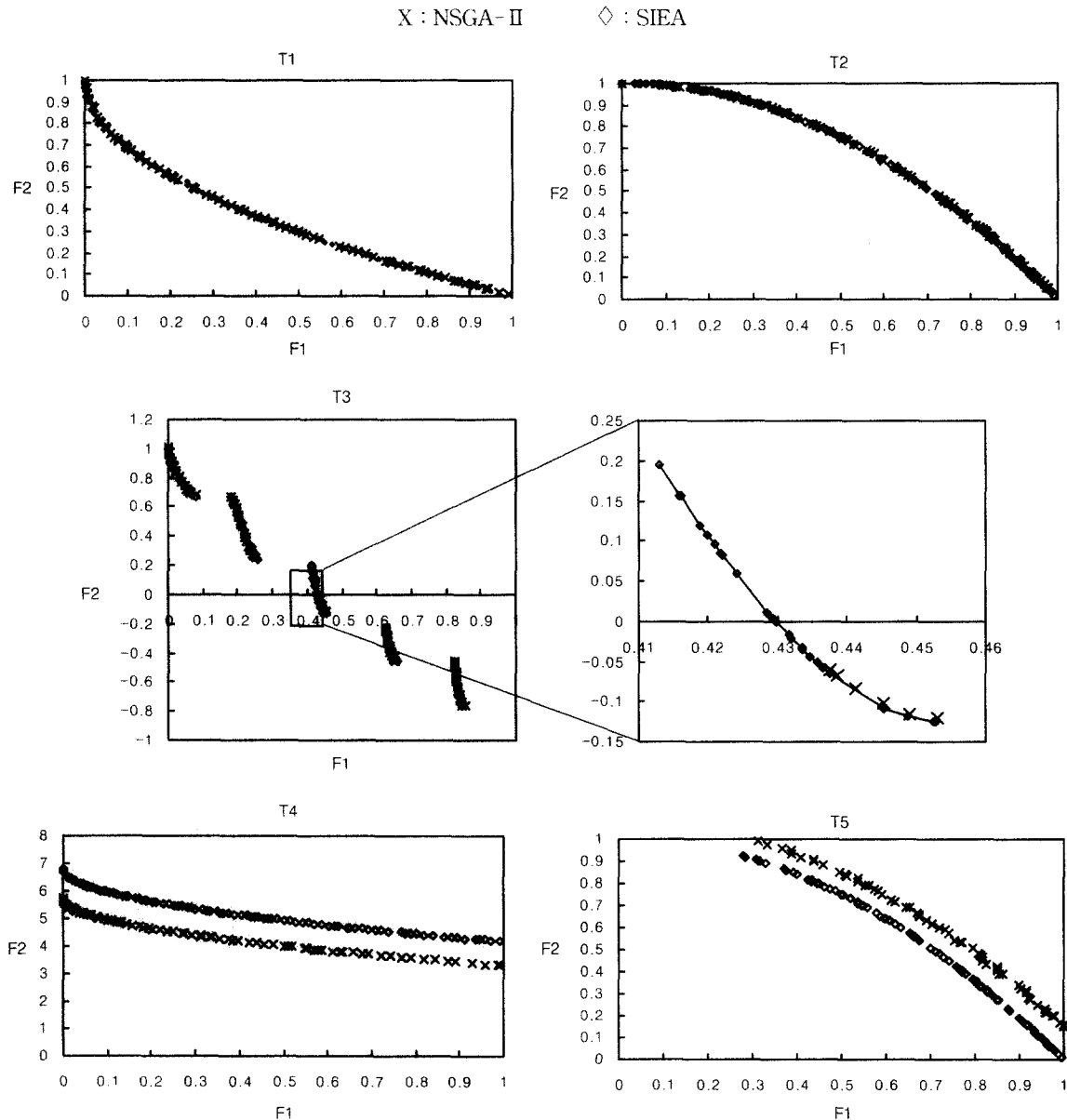
<표 3>은 SIEA와 NSGA-II 알고리즘의 S metric 결과이다. 두 알고리즘의 S metric 결과는 T_3 문제를 제외한 모든 문제에서 SIEA의 해들의 분포가 비교적 고른 분포를 보이고 있음을 보여준다. T_3 문제는 이산적인 파레토 경계면을 가지는 문제이기 때문에 SIEA에서는 우수종자의 선택이 이산적으로 끊

어지는 각 부분의 끝에 위치한 개체가 선택되는 경향을 가진다. 따라서 새로 생성되는 개체도 각 부분의 끝에서 중복적으로 생성되기 때문에 NSGA-II 알고리즘보다 전체 해들의 균일분포를 갖도록 하는 측면에서는 어려움을 보일 수도 있다.

그러나 SIEA는 비지배 개체 구성 알고리즘에서 이미 찾은 각 부분의 중간 해를 잃어버리지 않기 때문에 특정 부분의 해가 소멸되는 경우는 드물다. 반면에 NSGA-II는 전체적인 분포에 있어서는 비교적 고른 분포를 가지고 있지만, (그림 6)의 T_3 의 확대된 영역에서 보이는 것처럼 특정 부분의 해가 소멸되는 결과를 가끔 보이기도 한다.

(그림 6)은 각 문제에 대하여 가장 잘 수렴된 결과들이다. T_1, T_2, T_3 의 경우 SIEA가 NSGA-II보다 조금 더 잘 수렴

하였다. T_4 의 경우는 앞의 C metric에서 설명하였듯이 전체 개체집단의 다양성 유지 측면에서 SIEA가 약점을 갖고 있기 때문에 NSGA-II가 보다 잘 수렴하는 결과를 보였다. 이는 다목적 최적화 문제에서도 지역 최소치를 극복하기 위해 균일 돌연변이 연산자에 의한 개체집단의 다양성 유지가 우수한 성능을 보이는 것과 같은 맥락이다[14]. 반면에 T_5 문제는 전역 파레토 경계면 근처로 갈수록 해의 밀도가 작아지는 문제로, 전역해 부분이 평원을 이루는 문제이다. 이러한 문제에서는 특정 탐색 점을 중심으로 좁은 범위의 탐색이 필요한 문제이다. SIEA는 우수 종자 개체를 중심으로 좁은 범위에서 탐색을 수행하는 특징을 갖기 때문에 다른 최적화 알고리즘들보다 잘 수렴하는 결과를 보였다.



(그림 6) NSGA-II와 SIEA의 수렴 결과

〈표 3〉 S metric 결과

문 제	NSGA-II	SIEA
T_1	0.010058	0.006841
T_2	0.009371	0.00759
T_3	0.00826	0.012556
T_4	0.031623	0.027327
T_5	0.009912	0.008122

6. 결 론

본 논문에서 우리는 우수종자 개체를 중심으로 진화를 수행하는 다목적 최적화 진화알고리즘을 제안하였다. 제안한 알고리즘은 우수종자의 선택 기준을 현재까지 찾아진 최적해들 중 특정 지역에 편중되지 않은 해로 하였고, 선택된 우수종자를 중심으로 복제연산을 수행하였다. 따라서 특별한 지역탐색 기법을 사용하지 않더라도 종자가 되는 개체 주위에 새로운 개체를 생성할 확률이 높기 때문에 부분적인 지역탐색의 효과를 가질 수 있음을 실험적으로 보였다. 또한 부분적인 지역 탐색의 효과로 대부분의 문제에서 전체 해집합의 분포가 비교적 고른 분포를 가질 수 있고, 최적해로의 수렴이 우수함을 실험적으로 보였다.

그러나 제안한 알고리즘은 전체 개체집단의 다양성 유지에 약점을 가지고 있기 때문에 많은 지역해를 포함하는 문제에서는 전역 최적해로의 수렴에 어려움이 있었다. 따라서 제안된 알고리즘의 장점을 유지하면서 전체 개체집단의 다양성을 보다 잘 유지할 수 있도록 하는 방안에 대한 연구가 추가적으로 이루어져야 할 것이다.

참 고 문 헌

[1] Carlos A. Coello Coello, "An Updated Survey of GA-Based Multiobjective Optimization Techniques," ACM Computing Surveys, Vol.32, No.2, pp.109-143, June, 2000.

[2] Frank Kursawe, "A Variant of evolution strategies for vector optimization," In Parallel Problem Solving from Nature. 1st Workshop, PPSN I, Vol.496 of Lecture Notes in Computer Science, pp.193-197, 1991.

[3] Carlos M. Fonseca and Peter J. Fleming, "Genetic Algorithms for Multiobjective Optimization : Formulation, Discussion and Generalization," In Proceedings of the Fifth International Conference on Genetic Algorithms, pp.416-423, 1993.

[4] Jeffrey Horn and Nicholas Nafpliotis, "Multiobjective Optimization using the Niche Pareto Genetic Algorithm," Technical Report IlliGA1 Report 93005, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1993.

[5] N. Srinivas and Kalyanmoy Deb, "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms," Evolutionary Computation, Vol.2, No.3 pp.221-248, 1994.

[6] Kalyanmoy Deb, Samir Agrawal, Anrit Pratab, and T. Meyarivan, "A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization : NSGA- II,"

Proceedings of the Parallel Problem Solving from Nature VI Conference, pp.849-858, Springer, 2000.

[7] Eckart Zitzler, Marco Laumanns and Lothar Thiele, "SPEA 2 : Improving the Strength Pareto Evolutionary Algorithm," EUROGEN 2001, Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, pp.12-21, 2001.

[8] Carlos A. Coello Coello and Nareli Cruz Cortes, "Solving Multiobjective Optimization Problems using an Artificial Immune System," Technical Report EVOCINV-05-2002, Evolutionary Computation Group at CINVESTAV, Kluwer Academic, 2002.

[9] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," In Genetic Algorithms and their Applications : Proceedings of the First International Conference on Genetic Algorithms, pp.93-100, 1985.

[10] Eckart Zitzler and Lothar Thiele, "Multiobjective optimization using evolutionary algorithms - a Comparative study," In Parallel Problem Solving from Nature V, pp.292-301, 1998.

[11] Jason R. Schott, Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization. Master's thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1995.

[12] Kalyanmoy Deb, "Multi-Objective Genetic Algorithms : Problem Difficulties and Construction of Test Problems," Evolutionary Computation, Vol.7, No.3, pp.205-230, 1999.

[13] Eckart Zitzler, Kalyanmoy Deb, and Lothar Thiele, "Comparison of Multiobjective Evolutionary Algorithms : Empirical Results," Evolutionary Computation, Vol.8, No.2, pp. 173-195, 2000.

[14] 장수현, 윤병주, "유전자알고리즘에서의 실수처리 방법 비교", 정보처리학회논문지, Vol.5, No.2, pp.361-371, 1998.



장 수 현

e-mail : tngus68@naver.com
 1993년 명지대학교 전자계산학과(공학사)
 1996년 명지대학교 대학원 컴퓨터공학과 (공학석사)
 1999년 명지대학교 대학원 컴퓨터공학과 박사과정 수료
 2001년~2003년 세경대학 인터넷정보과 전임 강사

관심분야 : 인공지능, 지능형에이전트, 진화알고리즘 등



윤 병 주

e-mail : yoonbj@mju.ac.kr
 1975년 경북대학교 수학과(학사)
 1982년 한국과학기술원 전산학과(석사)
 1994년 Florida State University 전산학과 (박사)
 1982년~현재 명지대학교 컴퓨터공학과 교수
 관심분야 : Machine Learning, Knowledge-Based System, Hybrid Intelligent Systems