

# 수색 구조 유닛 배치 개선을 위한 유전 알고리즘 파라미터 최적화

홍승열, 김용혁\*

\*광운대학교 컴퓨터과학과

ghdqkxk159@naver.com, \*yhdfly@kw.ac.kr

## Optimizing Genetic Algorithm Parameters for Improving Search and Rescue Unit Deployment

Seung-Yeol Hong, Yong-Hyuk Kim\*

\*Dept. Computer Science, Kwangwoon Univ.

### 요약

본 논문에서는 수색 구조 유닛을 배치하기 위한 유전 알고리즘의 성능을 개선하기 위해 파라미터를 최적화하는 방법에 대해 제안한다. 제안하는 파라미터 최적화 방법은 연산량을 기준으로 파라미터 최적화 순서를 결정함으로써, 적은 실험량으로도 효과적으로 파라미터 최적화를 수행한다. 또한, 여러 개의 데이터를 통해 특정 데이터에 특화되지 않은 일반화된 파라미터 최적화 결과를 도출한다. 실험 결과, 제안하는 방법을 통해 도출된 파라미터 조합을 사용한 유전 알고리즘은 기존 연구에서 제시한 파라미터 조합을 사용한 것보다 적은 연산량을 가지면서도 더 좋은 성능을 보였다.

### I. 서론

유전 알고리즘은 생물학적 진화 과정을 모방하여 개발된 최적화 알고리즘이다. 유전 알고리즘은 문제의 해를 염색체로 표현하고, 염색체의 선택, 교차, 돌연변이, 대체 연산을 통해 가장 좋은 해를 탐색한다. 이때, 유전 알고리즘의 각 연산 과정에는 유전 알고리즘의 설계자가 입력해야 하는 파라미터가 존재한다. 유전 알고리즘의 파라미터는 유전 알고리즘의 최적화 성능과 동작 시간에 영향을 미치는 중요한 요소이지만, 설계자의 지식과 경험으로 정해지고 있다는 단점이 존재한다[1]. 예를 들어, 많은 교차 연산이 일어날수록 좋은 성능을 가진다는 경험적 지식으로 교차 연산에 대한 파라미터를 결정할 수 있지만, 이를 통해 정해진 파라미터가 좋은 파라미터라고 단정할 수 없다. 따라서, 유전 알고리즘의 설계자는 좋은 파라미터를 찾기 위한 파라미터 최적화를 수행해야 한다.

하지만 파라미터 최적화를 하기 위해서는 설계자가 수많은 실험을 진행해 볼 수밖에 없다. 돌연변이 확률, 생성하는 자식 염색체의 수, 해 집단의 규모(인구수) 등 유전 알고리즘의 파라미터는 모두 숫자로 나타나며, 파라미터 최적화를 위해 모든 숫자의 조합을 실험하는 것은 매우 비효율적이다. 따라서, 본 연구에서는 두 가지를 고려하는 파라미터 최적화 방법을 제안한다. 첫째, 일반화 성능이 뛰어난 파라미터 조합을 찾기 위해 서로 다른 3개의 데이터를 사용하여 파라미터 최적화를 진행하며, 검증 데이터를 통해 파라미터 최적화 결과를 평가한다. 둘째, 유전 알고리즘의 연산량을 고려하여 파라미터 최적화 순서를 결정하여 효율적으로 파라미터 최적화를 진행한다.

### II. 본론

#### 1) 파라미터 최적화 대상

본 연구에서 파라미터 최적화를 진행하고자 하는 연구는 해양 수색 구조 유닛 배치를 위한 유전 알고리즘이다[2]. 해양 수색 구조 유닛을 배치하는 문제는 입자가 존재하는 지도에 사각형을 배치하는 문제로 표현된다. 입자는 해양 수색 구조 유닛이 수색·구조해야 하는 대상을 의미하고, 사각형은 해양 수색 구조 유닛이 수색하는 영역을 의미한다. 해당 문제에서는

$m$ 개의 사각형이 최대한 많은 입자를 포함하도록 배치하는 것이 목적이다. [그림 1]은 2개의 수색 구조 유닛 배치를 위한 유전 알고리즘의 실행 결과에 대한 예시이다.

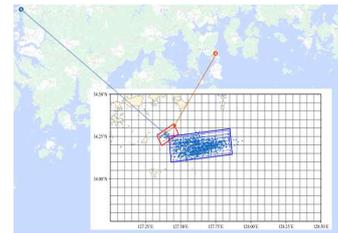


그림 1. 수색 구조 유닛 배치 예시

해당 문제를 위한 최근 연구[2]에서 설계한 수색 구조 유닛 배치를 위한 유전 알고리즘의 파라미터는 [표 1]과 같다.

표 1. 이전 연구[2]에서의 유전 알고리즘 파라미터

선택 연산자	Roulette-wheel	세대 수	1,000
교차 연산자	Uniform crossover	인구수	200
돌연변이 확률	0.05	자식 염색체 수	50
대체 연산	자식 염색체의 적합도보다 작은 적합도를 가지는 부모 염색체와 대체		
수정 연산	적절한 수색 영역의 크기를 가지도록 너비와 높이 수정		

#### 2) 파라미터 최적화

본 연구에서는 세대수, 인구수, 돌연변이 확률, 생성되는 자식 염색체 수 등 총 4가지의 유전 알고리즘 파라미터를 최적화한다. 이때, 각 파라미터 조합에 대한 올바른 비교를 위해, 유전 알고리즘의 연산량을 비슷하게 유지한다. 유전 알고리즘의 연산량은 인구수+세대수×자식 염색체 수로 계산된다. 연산량은 세대수와 자식 염색체 수에 민감하게 반응하며, 이를 고려하여 세대수는  $18,000 \div$  자식 염색체의 수로 고정하였다. 이때, 전 알고리즘의 동작 시간을 약 3분의 1로 감소시키기 위해 기존의 연산량 50,000의 36%인 18,000을 연산량의 기준으로 하였다.

연산량에 적은 영향을 미치는 파라미터 순서로 최적화를 진행하여, 먼저 진행한 실험이 이후에 진행되는 실험에 적은 영향을 미치도록 한다. 이에

따라, 돌연변이 확률, 인구수를 최적화한 이후, 세대수와 자식 염색체의 수를 동시에 최적화한다. 각 파라미터의 탐색 범위를 지정하고, 그중 가장 좋은 파라미터 조합을 선택한다. 파라미터별 탐색 범위는 다음과 같다.

- 세대수 = { 1800, 1200, 900, 720, 600, 514, 450, 400, 360 }
- 인구수 = { 50, 100, 150, 200, 250, 300, 350, 400 }
- 돌연변이 확률 = { 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 }
- 자식 염색체 수 = { 10, 15, 20, 25, 30, 35, 40, 45, 50 }

각 실험에서 좋은 결과를 보인 값을 다음 파라미터 최적화 실험에서 사용한다. 예를 들어, 돌연변이 확률에 대한 파라미터 최적화 결과 0.04를 사용했을 때 가장 좋은 결과를 보였다면, 인구수에 대한 파라미터 최적화 과정에서는 돌연변이 확률을 0.04로 고정한다. 파라미터 최적화를 진행하지 않은 부분에서는 돌연변이 확률을 제외한 나머지 파라미터는 기존의 연구에서 제시하는 파라미터를 사용하였다. 예를 들어, 인구수 최적화 실험에서, 돌연변이 확률은 이전 실험(돌연변이 확률 최적화)에서 선택된 돌연변이 확률을 사용하고, 인구수는 탐색 범위 내에서 변화하지만, 세대수, 자식 염색체 수는 이전 연구에서 사용된 값 1,000과 20을 사용한다. 이때, 파레토 집합을 사용하여 좋은 파라미터 조합을 선정한다. 파레토 집합이란, 하나의 목적 함수에 관한 결과가 좋지 않더라도, 나머지 목적 함수에 관한 결과가 가장 좋은 해의 집합을 의미하며, 여러 개의 목적 함수를 최적화할 때 중요한 역할을 하는 개념이다.

### III. 결론

한 개의 데이터를 통해 파라미터 최적화 실험(PO1)과 세 개의 데이터를 사용해 파라미터 최적화 실험(PO3)의 결과를 비교한다. PO1은 세 개의 실험 데이터 데이터1, 데이터2, 데이터3 중, 데이터1만을 사용하여 파라미터 최적화를 진행한다.

#### 1) 파라미터 최적화 결과

파라미터 최적화 실험에서는 정확한 결과 분석을 위해 각 파라미터 조합별로 30번의 반복 실험의 평균을 가지고 실험 결과를 분석했다. 또한 돌연변이 확률 최적화 실험 11회를 진행한 결과, 파레토 집합으로 3개의 돌연변이 확률이 선택되었다. 인구수 최적화 실험은 앞선 실험에서 선택된 파레토 집합 3개와 인구수 탐색 범위 8개의 값을 모두 조합하여 실험을 진행했다. 따라서 총 24회의 실험을 진행하였고, 그 결과 파레토 집합으로 총 7개의 <돌연변이 확률, 인구수> 쌍이 선택되었다. 세대수, 자식 염색체 수를 최적화하는 실험은 앞선 실험인 인구수 최적화에서 선택된 파레토 집합 7개와 세대수, 자식 염색체 수 탐색 범위 9개의 값을 모두 조합하였다. 따라서 총 63회의 실험을 진행하였으며, 그 결과 파레토 집합으로 선택된 파라미터 조합은 15개이다. 그중에서 각 데이터의 값이 골고루 높은 파라미터 조합을 최적의 파라미터 조합으로 선택했다.

데이터1만을 사용하여 진행한 파라미터 최적화 실험에는 총 11+8+9번의 실험이 진행되었으며, 각 실험에서 최솟값을 가지는 파라미터가 선택되었다. [표 2]는 최종적으로 선택된 파라미터 조합과 이전 연구[2]의 파라미터 조합(Original)을 나타낸다.

표 2. 파라미터 최적화 결과

	돌연변이 확률	인구수	세대수	자식 염색체 수
Original[2]	0.05	200	1,000	50
PO3*	0.03	350	1,200	15
PO1**	0.04	250	720	25

\*3개의 데이터를 통해 파라미터 최적화 \*\*1개의 데이터만 사용하여 파라미터 최적화

#### 2) 검증 실험 결과

검증 실험은 검증 데이터를 통해 이루어지며, Intel Core i7-13700 CPU(3.40GHz)와 64GB 메모리에서 실행되었다. 파라미터 최적화의 결과(PO3, PO1)와 이전 실험[2]의 파라미터 조합(Original)을 사용한 것을 비교한다.

[표 3]은 3.1절에서 결정된 파라미터 조합에 대한 검증 실험 결과이다. 정확한 실험을 위해 파라미터 조합별로 30회 반복 실험을 진행하였다. 실험 결과, 제안하는 방법을 통해 파라미터 최적화를 진행한 PO3가 평균적으로 가장 좋은 성능을 보이는 것을 확인할 수 있다. 이때, 이전 연구[2]의 파라미터를 사용한 Original보다 연산량을 64% 감소시켰음에도 더 좋은 적합도 평균을 가진다는 것에서, 파라미터 최적화가 잘 이루어졌음을 확인할 수 있다. 연산량 감소를 통해 평균 시간 역시 3분의 1가량으로 떨어진 것을 확인할 수 있다. 한 개의 데이터를 사용하여 파라미터 최적화를 수행한 PO1의 경우, 동작 시간의 감소는 성공적으로 이루어졌지만, 이에 따른 성능 감소는 극복하지 못한 것을 확인할 수 있다.

표 3. 파라미터 최적화 결과 별 반복 실험 30회 결과 비교

	적합도 최솟값	적합도 (평균)	표준편차	평균 시간 (초)
Original[2]	168.919	166.045	1.194	39.61
PO3*	168.799	166.140	1.342	13.98
PO1**	168.759	165.830	1.548	15.72

\*3개의 데이터를 통해 파라미터 최적화 \*\*1개의 데이터만 사용하여 파라미터 최적화

#### 3) 논의

본 연구에서는 해양 수색 구조 유닛 배치를 위한 유전 알고리즘의 파라미터 최적화를 수행하였다. 연산량을 기반으로 파라미터 최적화 순서를 결정하여 모든 파라미터 조합을 탐색(실험)하지 않으면서도 효과적으로 파라미터 최적화 실험을 진행하였다. 또한, 특정 데이터에 과적합 되어 파라미터가 조절되지 않도록 세 개의 데이터를 사용하여 일반화 성능이 좋은 파라미터 조합을 찾고자 하였다. 실험 결과, 제안하는 방법을 통해 파라미터 최적화를 수행한 결과가 기존 연구[2]보다 더 적은 연산량과 구동 시간을 가지면서도, 성능을 개선한 것을 확인할 수 있었다. 이를 통해 모든 파라미터 조합 11×8×9(=792)의 13%가량인 11+24+63(=98)회의 실험만으로도 효과적인 파라미터 최적화를 수행했음을 확인할 수 있다. PO1의 실험 결과에서, 동작 시간의 감소는 성공적으로 이루어졌지만 이에 따른 성능 감소는 극복하지 못한 것을 확인할 수 있는데, 이는 PO1이 특정 데이터에 특화된 파라미터 조합을 도출했기 때문이라고 추측할 수 있다.

### ACKNOWLEDGMENT

이 논문은 2023년도 해양경찰청 재원으로 해양수산과학기술진흥원의 지원을 받아 수행된 연구임(20220463, 지능형 해양사고 대응 플랫폼 구축).

### 참고 문헌

[1] M. Mosayebi and M. Sodhi “Tuning Genetic Algorithm Parameters using Design of Experiments”, 2020, *In Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, pp. 1937 - 1944.

[2] S.Y. Hong and Y.H. Kim “Maximizing Particle Coverage with Fixed-Area Rectangles”, 2023, *In Proceedings of the 15th International Joint Conference on Computational Intelligence - Volume 1: ECTA*, pp. 172-178