

# 이열 기계 배치 문제에서의 개체 선택 방법의 비교

김 정 집<sup>\*</sup>  
서울대학교 컴퓨터공학과  
jjkim@nova.snu.ac.kr

장 병 탁  
서울대학교 컴퓨터공학과  
btzhang@comp.snu.ac.kr

김 영 택  
서울대학교 컴퓨터공학과  
ytkim@comp.snu.ac.kr

## Comparison of Selection Schemes for Double-Row Machine Layout Problems

Jung Jib Kim  
Dept. of Computer Engineering  
Seoul National University

Byoung-Tak Zhang  
Dept. of Computer Engineering  
Seoul National University

Yung Taek Kim  
Dept. of Computer Engineering  
Seoul National University

### 요 약

이열 기계 배치 문제는 기계의 수가 많아지면 기존의 알고리즘으로는 실제로 해결이 불가능한 조합 최적화 문제이다. 본 논문에서는 최적의 기계 배치를 찾기 위하여 유전자 알고리즘을 사용하고 기존의 방법과 비교하여 선택 과정의 개선을 통해 최적화의 성능을 향상시키는 방법에 대하여 고찰한다. 실험을 통하여 제안된 방법의 성능을 분석하였다.

## 1. 서 론

시설 배치 문제(facility layout design problems)는 공장에서 쓰이는 많은 기계들의 위치를 조정함으로써 생산 시스템의 효율을 높이려는 문제이다. 이 문제의 해는 전체 시스템의 효율적인 이용에 기반이 된다. 이 효율은 기계의 수, 크기, 기계 사이의 최소 간격, 작업의 이동빈도 등의 여러 변수에 의해 영향을 받는다. 따라서 효율은 여러 변수를 가진 함수가 되는 것이고, 이런 함수의 최적화는 매우 어려운 문제이다.

이열 기계 배치 문제(double-row machine layout problem)는 시설 배치 문제에서 기계들을 두 개의 열로 배치한다는 제한을 가한 것이다. 이 문제역시 어려운 문제이므로 기존의 알고리즘으로 해결하기 어렵다.

유전자 알고리즘(genetic algorithms)은 유전학과 자연선택의 원리에 근거한 최적화 탐색방법이다[1]. 이 알고리즘은 함수 최적화에 높은 성능을 보였고, 이것을 사용한 이열 기계 배치 문제의 연구도 시도된 바 있다[2].

본 논문에서는 유전자 연산자 중에서 선택 방법을 변화시켜서 이열 기계 배치 문제에 적합한 유전자 알고리즘의 방법을 제시하고자 한다. 선택방법으로는 적합도 비례 선택(fitness-proportional selection)이 아니라 순위 선택(rank selection)과 토너먼트(tournament selection)를 사용한다[3].

순위 선택에서는 높은 적합도의 순서대로 일정 비율을 다음 세대로 선택하는 방법이고, 토너먼트 선택은 임의로 선택한 개체끼리 적합도를 비교해서 선택하는 방법이다. 이 두 방법 모두 적합도 비례 선택보다는 지역 최적값(local minima)에 쉽게 빠지지 않아서 높은 적합도 값을 얻을 수 있다[4].

논문의 구성은 다음과 같다. 2절에서는 이열 배치 문제를 수학적으로 정의하며, 3절에서는 사용된 유전자 알고리즘 방법을 기술하고 마지막절에서는 실험환경과 실험 결과를 보인다.

## 2. 이열 기계 배치 문제

이열 기계 배치 문제는 기계를 두 개의 열로 배치해서 모든 기계 사이의 이동비용(비용 $\times$ 이동빈도 $\times$ 이동거리)을 최소화하는 문제이다. 본 논문에서는 이열 기계 배치 문제를 모델링하기 위해 다음과 같은 가정을 하였다.

- 각 기계들은 직사각형이다.
- 각 기계의 방향은 결정되어 있다 즉, 모든 기계는 길이가 너비보다 작지 않다.

이열 기계 배치 문제는 다음과 같이 정형화 될 수 있다

$$z_{jk} = \begin{cases} 1, & \text{기계 } j \text{가 } k \text{열에 있을 때} \\ 0, & \text{나머지 경우} \end{cases} \quad (1)$$

$$\min \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_{ij} f_{ij} (|x_i - x_j| + |y_i - y_j|) \quad (2)$$

$$s. t. |x_i - x_j| z_{ik} z_{jk} \geq \frac{1}{2} (l_i + l_j) + d_{ij}, \quad i, j = 1, \dots, n \quad (3)$$

$$y_i = \sum_{k=1}^m l_0 (k-1) z_{ik}, \quad i = 1, \dots, n \quad (4)$$

$$\sum_{k=1}^m z_{ik} = 1, \quad i = 1, \dots, n \quad (5)$$

$$\sum_{i=1}^n z_{ik} < n, \quad k = 1, \dots, m \quad (6)$$

$$x_i, y_i \geq 0, \quad i = 1, \dots, n \quad (7)$$

$$z_{ik} \in \{0, 1\}, \quad i = 1, \dots, n, k = 1, \dots, m \quad (8)$$

위에서 사용된 변수들의 의미는 다음과 같다  $n$ 은 기계의 수,  $m$ 은 열의 수,  $f_{ij}$ 는 기계  $i, j$  사이의 이동의 빈도,  $c_{ij}$ 는 기계  $i, j$  사이의 단위 이동당 비용,  $l_i$ 는 기계  $i$ 의 길이,  $l_0$ 는 두 개의 인접한 열사이의 간격,  $d_{ij}$ 는 기계  $i, j$  사이의 최소 간격,  $x_i$ 는 수직기준선  $l_H$ 로부터 기계  $i$ 의 중심까지의 거리,  $y_i$ 는 수평기준선  $l_H$ 로부터 기계  $i$ 의 중심까지의 거리이다

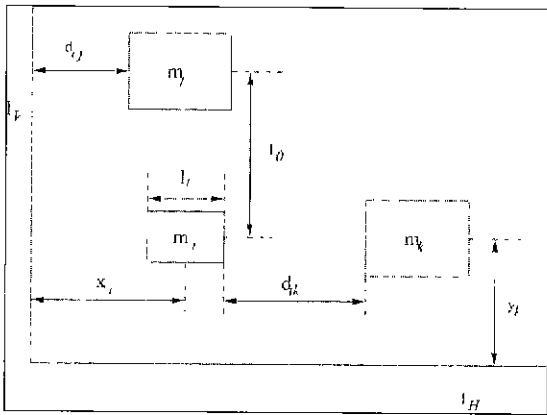


그림 1 이열 기계 배치 문제를 기술하기 위한 기본적인 인자와 매개 변수

### 3. 사용된 유전자 알고리즘

사용된 유전자 알고리즘은 기본적으로 선택 연산사를 제외하고 [5]의 것과 동일하다

본 논문에서 사용된 적합도 비례 선택은 다음과 같은 과정을 거친다. 모든 개체들에 대해 개체의 적합도 값을 총 적합도 값으로 나누어 선택 확률  $p_k$ 를 구하고 모든 개체들에 대해  $p_1$ 부터  $p_k$ 까지의 합으로 누적 확률  $q_k$ 를 구한다 그리고 난수를 발생 시켜서  $q_{k-1} < \text{random}() < q_k$ 이면  $k$ 를 선택한다

순위 선택의 방법은 모든 개체들을 적합도를 기준으로 정렬시켜서 원하는 순위 이상의 것만 선택한다

토너먼트 선택은 임의로 토너먼트 크기(tournament size)의 개체를 선택한 후, 제일 좋은 것을 선택한다

이 때, 순위 선택의 경우에는 부모로 선택되는 선택물(truncation threshold)이, 토너먼트 선택의 경우는 토너먼트의 크기(tournament size)가 새로운 인자가 된다.

### 4. 실험 결과

실험에 사용한 데이터 집합은 기계수가 20일 때에 대해 임의로 발생시킨 것이다. 이것으로 3가지 선택 방법을 사용한 유전자 알고리즘을 비교하였는데, 동일한 조건에서 비교하기 위해 모두에 임의의 돌연변이를 사용하였다. 또 빠른 최적화를 위해서 엘리트 전략(elitist strategy)을 사용하였다 종료 조건은 동일한 최고값이 50번 반복될 때로 하였다

위와 같은 과정을 50번씩 반복해서 최고값을 구하고 그 평균과 표준편차를 구했다.

| 매개 변수                    | 설정 값 |
|--------------------------|------|
| 교차율(crossover rate)      | 0.4  |
| 돌연변이율(mutation rate)     | 0.2  |
| 반복횟수(iteration)          | 50   |
| 기계수(machine size)        | 20   |
| 개체군의 크기(population size) | 500  |
| 최대 진화 횟수(max generation) | 2000 |

표 1 실험에 사용된 매개변수 값

각 선택 방법에는 여러 가지 인자가 존재하게 된다 즉 토너먼트 선택 방법에는 토너먼트의 크기, 순위 선택 방법에는 선택물 등이 있는데 이들에 의한 영향을 알기 위해 다른 값으로 변화 시켜 가면서 적합도 값과 실행 시간을 비교했다.

그림 2, 3에는 여러 가지 토너먼트 크기(tournament size)에 대한 결과를 보였다. 토너먼트 크기가 키질수록 해의 품질이 떨어지니 빠른 시간에 멈췄는데, 이것은 미숙한 수렴(premature convergence)를 의미하는 것이다

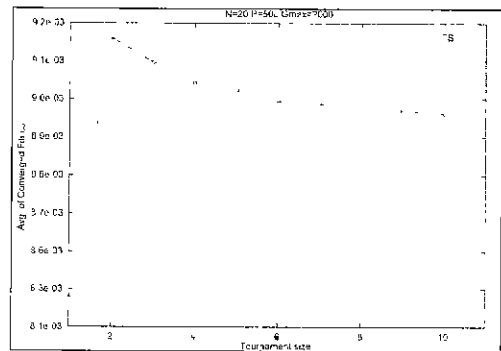


그림 2 기계수 N=20에서 토너먼트 선택 방법을 사용한 적합도 결과의 비교

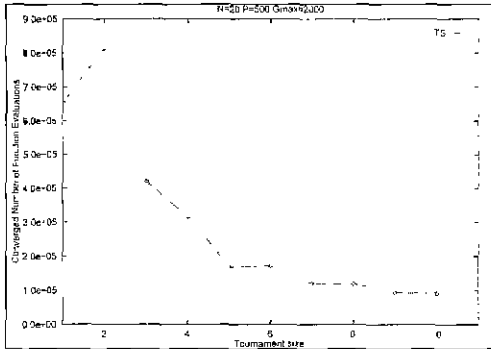


그림 3 기계수 N=20에서 토너먼트 선택 방법을 사용한 실행 시간의 비교

그림 4,5에는 여러 가지 선택률(truncation threshold)에 대한 결과를 보였다. 이 경우엔 선택률이 60%에서 최고의 해를 얻었지만 여기서 밀어질수록 해의 품질이 떨어졌다. 60%이상의 선택률에선 기존 결과의 보존이 새로운 탐색을 방해할 정도로 많아진 것이다.

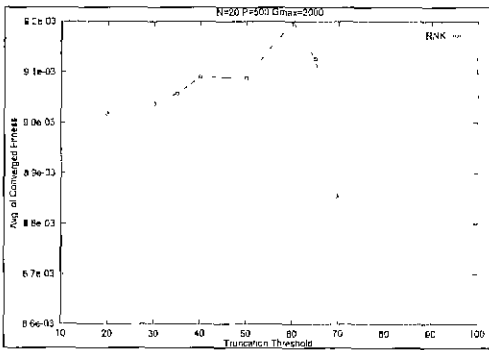


그림 4 기계수 N=20에서 순위 선택 방법을 사용한 적합도 결과의 비교

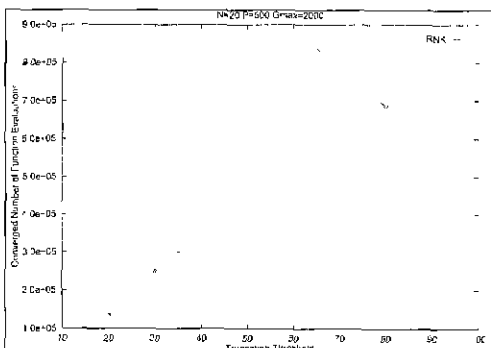


그림 5 기계수 N=20에서 순위 선택 방법을 사용한 실행 시간의 비교

각 선택 방법에서 최적의 해를 얻었을 때를 보면 적합도 비례

선택, 토너먼트 선택, 순위 선택의 순으로 좋은 해를 얻었다 표5에 의하면 토너먼트 선택과 순위 선택이 4% 정도의 시간만으로도, 2% 낮은 비용을 얻을 수 있다

여기서 토너먼트 크기가 2인 토너먼트 선택과 선택률이 60%인 순위 선택이 유사한 결과를 얻었는데, 이론적인 접수시간(takeover time)도 유사하다[7]

|        | RWS       | TS        | RNK       |
|--------|-----------|-----------|-----------|
| 적합도값   | 9.033e-03 | 9.162e-03 | 9.162e-03 |
| 평가수    | 3.331e+06 | 1.417e+05 | 1.324e+05 |
| 총비용(%) | 100       | 98.6      | 98.6      |
| 평가수(%) | 100       | 4.25      | 3.97      |

표 2 각 선택 방법의 평균 실행 결과 (단, 적합도 비례 선택이 100% 기준)

### 5. 결론

이열 기계 배치 문제의 해결을 위해 유전자 알고리즘을 사용하는 방법을 고찰하였다 유전자 알고리즘에서 다른 선택 방법을 사용함으로써 복잡한 문제에서의 유전자 알고리즘의 성능을 향상시킬 수 있음을 실험적으로 확인하였다. 또 각 선택 방법의 적합도 값과 평가수를 비교한 결과 토너먼트 선택이나 순위 선택이 더 효율이 높음을 보였다.

감사의 글: 본 연구는 한국 IBM에 의하여 일부 지원되었음.

### 참고문헌

- [1] D. E Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading, MA Addison-Wesley, 1989
- [2] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Design*, John Wiley & Sons, pp. 292-310, 1997
- [3] D. E. Goldberg, B. Korb, and K. Deb, Messy genetic algorithms Mutation, analysis, and first results, *Complex Systems*, 3(5) 493-530, 1989.
- [4] T. M. Mitchell, *Machine Learning*, McGraw-Hill, pp. 249-273, 1997
- [5] 김정집, 장병탁, 김영택, "유전자 알고리즘에 의한 이열 기계 배치 문제의 해결", *정보과학회 봄 학술 발표논문집(B)*, pp. 264-266, 1998
- [7] Thomas Back, *Evolutionary Algorithms in Theory and Practice*, Oxford University Press, pp. 163-195, 1996