

# 공생파트너 선택전략에 따른 공생진화 알고리즘의 성능 분석† Performance analysis of symbiotic evolutionary algorithms according to partner selection strategies

김재윤\* · 김여근\* · 박재승\* · 김동목\*\*

\*전남대학교 산업공학과, \*\*동신대학교 산업공학과

## Abstract

Symbiotic evolutionary algorithms are stochastic search algorithms that imitate the biological coevolution process through symbiotic interactions. In the algorithms, the fitness evaluation of an individual requires first selecting symbiotic partners of the individual. The symbiotic partner effects the change of individual's fitness and search direction. In this study we are to analyze how much partnering strategies can influence the performance of the algorithms. For this goal extensive experiments are carried out to compare the performance of partnering strategies. The NKC model and the binary string covering problem are used as the test-bed problems. The experimental results indicate that there does not exist statistically significant difference in their performance.

## 1. 서론

공생진화 알고리즘(symbiotic evolutionary algorithm)은 자연계에서 복수개의 종들이 공생하며 진화하는 자연현상을 모방한 확률적 탐색기법이다. 공생진화 알고리즘은 1990년 이후 연구되기 시작하여 현재까지 아직 연구 초기단계이지만, 복잡하고 동적인 문제를 해결하는 효율적인 방법으로 인식되고 있다[7]. 공생진화 알고리즘은 몇 개의 부분해(partial solution)를 통한 병렬 탐색(parallel search)이 완전해(entire solution)을 통한 탐색보다 효과적이라는 가설을 따른다[7]. 공생진화 알고리즘은 흔히 여러 부분문제들의 조합으로 구성된 문제나 하나의 큰 문제를 여러 개의 부분문제로 나눈 문제에 적용된다. 이때 각 모집단에 있는 개체들은 부분해가 되고, 완전해는 이들 부분해의 결합에 의해 표현된다. 공생진화는 부분해들의 결합(공생)에 의해 만들어진 완전해에 의해서 적응도가 평가될 수 있다. 따라서 개체와 적응도 평가시 완전해를 구성하기 위한 상대 모집단의 개체들이 - 이를 공생파트너(symbiotic partners)라 부르기로 함 - 요구된다. 단순진화알고리즘에서는 개체가 완전해로 표현되므로, 공생파트너를 고려할 필요가 없다.

공생진화 알고리즘에서는 공생파트너에 따라 개체의 적응도가 변하고, 개체의 적응도 변화는 해의 탐색방향을 변화시키며, 파트너 선택 전략에 따라 알고리즘의 계산소요시간에 차이가 발생한다. 공생파트너 선택전략에 관한 몇몇 연구[2,6]가 이루어졌으나, 기존 연구결과는 분석이 불완전하여 적절한 공생파트너 선정에 있어서 유용한 정보를 제공하는 데에는 부족하다고 판단된다. 따라서 본 연구에서는 공생파트너 선택전략이 공생진화 알고리즘의 탐색 성능에 얼마나 영향을 주는지 분석한다. 이를 위하여 NKC 모델[3]과 이진 스트링 커버링문제[8]를 대상으로 공생진화 알고리즘을 분류하고, 각 알고리즘에 다양한 공생파트너 선택 전략을 적용하여 알고리즘 성능 결과를 분석하고자 한다.

## 2. 공생파트너 선택전략

앞에서 언급했듯이, 공생진화 알고리즘에서 각 개체는 부분해를 나타내고, 완전해는 부분해의 결합으로 표현된다. 공생진화 알고리즘은 완전해를 여러 부분해로 나누어 이들을 결합함으로써 다양한 형태의 완전해를 구성할 수 있어 효과적으로 해를 탐색할 수 있다는 장점을 갖는다. 공생진화에서 개체의 적응도가 공생파트너에 의해 결정되는 것은 자연스럽다. 그리고 개체의 재생산은 적응도에 의한 선별에 종속될 것이다. 따라서 각 개체가 공생파트너를 선택하고, 이로부터 자신의 적응도를 결정하는 것은 진화 연산에서 매우 중요한 과정이다. 본 연구에서는 공생파트너 선택전략으로 다음 여섯가지를 고려하고 컴퓨터 모의실험을 통하여 이들을 분석하고자 한다.

- SPM(Same Position Mapping) : 격자구조를 갖는 모집단에서 동일한 위치의 개체들이 서로의 공생파트너가 된다.
- RMP(Random Mapping) : 각 개체마다 상대 모집단(이웃)에서 자신의 공생파트너를 임의로 선택한다.
- RND(Random) : 모집단(이웃)에서 임의 선택한 하나의 개체가 상대 모집단(이웃)내 모든 개체의 공생파트너가 된다.
- BST(Best) : 모집단(이웃)에서 가장 높은 적응도를 갖는 한 개체가 상대 모집단(이웃)내 모든 개체의 공생파트너가 된다.
- RNK(Rank) : 모집단(이웃)에서 적응도 순위가 동일한 개체들이 서로의 공생파트너가 된다.
- SLN(Selection): 모집단(이웃)내에서 확률바퀴 선별로 선택한 하나의 개체가 상대 모집단(이웃)내 모든 개체의 공생파트너가 된다.

## 3. 공생진화 알고리즘

기존 연구에서 소개된 공생진화 알고리즘은 크게 세가지로 분류할 수 있다. 첫째, 분리된

† 본 연구는 한국과학재단 특정기초연구(과제번호:98-0200-09-01-3) 지원으로 수행되었음.

모집단단위 공생진화 알고리즘(Separated and Population based symbiotic evolutionary Algorithm: SPA)은 개체들이 분리된 형태로 공생하며, 부분해들은 모집단 단위로 진화하는 것이다[2,6,8].

둘째, 분리된 이웃단위 공생진화 알고리즘(Separated and Neighborhood based symbiotic evolutionary Algorithm: SNA)은 모집단의 다양성 유지는 알고리즘의 탐색 성능을 향상시킨다는 점에서 착안하여 SPA 에 이웃진화 개념을 포함시킨 것이다[5]. SNA 는 다양한 개체들이 공생 환경이 되고, 모집단의 다양성이 유지됨으로써 SPA 보다 더 나은 성능을 보인 것으로 알려져 있다[5].

셋째, 내공생 진화알고리즘(Endosymbiotic Evolutionary Algorithm: EEA)은 부분해를 통한 부적절한 해의 분리는 오히려 알고리즘의 탐색 성능을 저하시킬수 있다[1]는 단점을 보완하기 위하여 자연계의 내공생 진화과정을 모방한 것이다[1,4]. EEA 에서는 부분해와 완전해가 공존하고 부분해는 이웃단위 진화를 통해 좋은 해를 생성하며, 높은 적응도를 갖는 부분해들은 완전해로 결합하여 스스로 진화할 수 있도록 하여 진화를 촉진한다. 지금까지 설명한 세 알고리즘의 구체적인 절차는 Kim *et al.*[4]을 참조하기 바란다.

## 4. 실험문제

### 4.1 NKC 모델

Kauffman[3]은 특성이 다른 종간의 공진화에 관한 연구를 위하여 NKC 모델을 제안하였다. NKC 모델은 문제의 복잡성, 상호 종속성, 적응도 지형(fitness landscape)의 굴곡성(ruggedness)에 대한 변화가 용이하다. 따라서 NKC 모델은 공생진화 알고리즘에 관한 기존 연구[1,2,8]에서 알고리즘의 성능을 분석하는 모델로 많이 사용되었다.

NKC 모델에서 한 개체는 N 비트의 이진수로 표현된다. 인자들은 같은 계층에 있는 K 개의 다른 인자들, 그리고 다른 종을 구성하는 C 개의 계층과 상호 관련성을 갖는다. 그러므로, NKC 모델에서 각 개체의 적응도는 동일한 개체 또는 다른 모집단에 있는 개체의 인자 구성에 의존하여 결정되고, K 와 C 에 의하여 관련성의 정도가 조절된다. 적응도 지형의 관점에서 볼 때, 평균적으로 낮은 K 값은 평탄한 적응도 지형을 생성한다. K 가 증가함에 따라 적응도 지형에 있는 부분최적의 수는 증가하며, 지형은 점점 더 굴곡성을 갖는다. 또한, 개체간의 유사성을 해밍거리로 표현할 때, K 가 증가할수록 개체의 인자형과 표현형의 관련성은 낮아지게 된다[3,8]. 한편, C 의 변화는 상호관련성을 갖는 다른 종들의 적응도 지형을 변형시키거나 뒤틀리게 한다.

NKC 모델은 중간, 그리고 종내 상호작용이 매우 복잡하여 그들이 적응도에 주는 영향은 임의값으로 할당하는 것이 적당하다고 가정한다. 이진스트링으로 표현된 개체에서 K+C 의 가능한 상호작용에 대하여,  $2^{(K+C)}$ 의 행과 N 의 열을 갖는 테이블을 생성하고, 각 셀에 [0.0, 1.0] 범위의 난수를 발생하여 채워넣는다. 각 인자의 적응도 공헌도는 적응도 테이블의 적절한 행과 열에서 선택하여 결정하고, 한 개체의 적응도는 각 인자의 적응도 공헌도를 평균하여 구한다. 이 모델에 관한 자세한 내용은 Kauffman 의 연구[3]를 참조할 수 있다.

### 4.2 이진스트링커버링문제

본 연구에서 사용한 또 다른 실험문제는 이진 스트링커버링문제이다. 이 문제는 척추동물의 면역 체계내에서 발생하는 항원-항체반응처럼 자연계의 수많은 복잡과정을 모형화 하기 위한 메커니즘으로 사용된다[8]. NKC 모델은 부분해 수의 증가에 따라 적응도 테이블 크기가 지수적으로 증가하여 다양한 실험문제의 생성이 곤란하지만, 이 문제는 모집단 수의 증가가 용이하고, 문제의 복잡성을 다양하게 변형시킬 수 있다. 또한 이 문제는 개체의 인자형과 표현형간의 차이가 없다는 특징을 갖는다[8].

이진스트링커버링문제는 주어진 목표스트링에 대하여 최대 일치하는 정해진 개수의 이진 스트링 집합을 진화시키는 것을 목표로 한다. 진화하는 스트링 집합을 일치집합, 목표스트링의 집합을 목표집합이라 부른다. 목표스트링과 일치스트링간의 일치강도는 같은 위치에서 동일한 인자값을 갖는 비트수를 더하여 계산한다. 일치집합의 적응도는 각 목표스트링에 대한 최대일치강도를 평균하여 구한다 이 문제에 대한 구체적인 내용은 Potter 의 연구[8]를 참조하기 바란다.

## 5. 실험 및 분석

NKC 모델은 2 개의 부분해, 이진스트링커버링 문제는 부분해가 3 개인 경우에 대하여 실험하였다. NKC 모델은 주어진 N, K, C 에 의한 적응도지형에서 전체 최적점을 찾는 것을 문제로 하며, 이진스트링 커버링문제는 주어진 목표스트링과 최대일치강도를 갖는 3 개의 일치스트링을 찾는 것을 문제로 한다. NKC 모델에서 두 부분해는 24 비트의 이진수로 표현하였고, 다양한 K 와 C 의 조합으로 실험문제를 만들었다. 이진스트링커버링문제에서는 목표 집합의 크기와 스트링의 비트수를 변화시킨 문제를 생성 하였다.

모집단은  $10 \times 10$  의 격자구조로 두었으며, 선별은 토너먼트 크기가 2 인 토너먼트 선별을 사용하였다. 교차는 교차율 0.5 로 이점교차를 수행했으며, 돌연변이는 0.1 의 확률로 개체를 선택하고, 선택된 개체의 각 인자에 0.1 의 확률로 대립형질 돌연변이를 수행하였다. 그리고 모든 모집단에서 생성된 개체수 합이 20,000 개일때 종료하였다. 알고리즘은 C++프로그램 언어로 구현 되었으며, 300MHz Pentium CPU 를 장착한 IBM-PC 에서 수행되었다.

### 5.1 실험결과 분석

실험은 두 실험문제에 대하여 SPA, SNA, EEA 에 각각 여섯가지 공생파트너 선택전략을 적용하여 보았다. 실험결과, 10 회 반복실험의 평균을 기준으로 가장 좋은 해를 찾는 빈도에서 각 전략은 뚜렷한 경향을 보이지 않았다. 계산시간은 RND 가 가장 적게 소요되었으며, RNK 는 가장 많이 소요되었다. 알고리즘간의 계산시간은 비슷한 수준을 유지하였다. NKC 모델에서 EEA 에 RND 전략을 적용한 경우, CPU 시간은 평균적으로 약 75~248 초, RNK 는 약 90~290 초 정도 소요되었다. 이진스트링커버링문제에서 목표스트링의 수가 6 개, 128 비트 문제에 RND 를 적용한 EEA 는 약 152 초 정도, RNK 는 약 382 초 정도 소요되었다. 구체적인 실험결과는 지면의

제약상 생략하기로 한다.

분산분석(analysis of variance)을 이용하여 각 전략이 알고리즘 성능에 미치는 영향을 분석하였다. 분산분석에서 구한 통계량으로  $F$  분포표에서 유의확률의 값(p-value)을 구하고, 유의확률이 유의수준보다 작으면 귀무가설을 기각하여 요인의 평균들이 모두 같지 않다는 결론을 내린다. <표 1>은 두 실험문제에 대한 각 문제의 유의확률이며, SPSS ver 7.5를 이용하여 구하였다. 아래 표에서 NKC 모델의 실험문제는 “K-C”, 이진스트링커버링문제는 “목표스트링수-비트수”로 표시하였다.

NKC 문제				이진스트링커버링문제			
문제	SPA	SNA	EEA	문제	SPA	SNA	EEA
00-00	0.240	1.000	1.000	3-064	0.290	0.107	0.853
00-04	0.206	0.078	0.611	3-096	0.111	0.228	0.442
00-08	0.138	0.477	0.398	3-128	0.090	0.481	0.502
00-16	0.153	0.199	0.221	4-064	0.214	0.199	0.556
04-00	0.931	0.634	0.997	4-096	0.095	0.162	0.450
04-04	0.621	0.198	0.931	4-128	0.129	0.155	0.242
04-08	0.292	0.512	0.990	5-064	0.080	0.138	0.214
08-00	0.354	0.808	0.999	5-096	0.121	0.450	0.458
08-04	0.399	0.490	0.794	5-128	0.095	0.110	0.157
08-08	0.407	0.606	0.996	6-064	0.154	0.577	0.840
12-00	0.296	0.701	0.989	6-096	0.179	0.207	0.258
16-00	0.346	0.856	0.994	6-128	0.159	0.410	0.698

<표 1> 문제별 유의확률

일반적으로 분산분석에서 사용하는 유의수준 0.05와 위의 표에 제시된 유의확률의 값을 비교하면 모든 값이 유의수준보다 높다. 이로써, NKC 모델과 이진스트링커버링문제에 대해서는 각 요인의 평균이 모두 같다는 귀무가설이 채택되어 공생파트너 선택 전략이 알고리즘에 미치는 효과는 유의한 차이를 보이지 않는다고 말할 수 있다.

## 5.2 특성분석

앞에서 제시한 공생파트너 선택전략들은 공생파트너 선택기준으로 적응도의 사용 유무에 따라 분류된다. 적응도가 선택기준으로 사용되는 전략인 BST, RNK, SLN은 해 공간의 집중 탐색이 가능하며, 모집단의 평균 적응도가 높다는 장점을 갖는다[2]. 그러나 공생파트너의 변화가 약하여 유전적 혼합의 정도가 낮고, 한번 낮은 적응도를 갖는 개체로 평가되면 진화하는 동안 공생파트너로 참여할 가능성이 낮다는 단점을 갖는다. 반면, 임의성을 갖는 전략인 SPM, RND, RMP는 해 공간의 다양한 탐색에 유리하다. 또한 임의적으로 공생 환경에 변동을 주기 때문에, 다양한 개체들이 공생파트너로 참여하여 편승현상(hitch-hiking)을 방지할 수 있을 것으로 기대된다. 그러나 이들 전략들은 공생파트너의 잦은 변동으로 해 공간을 안정적으로 탐색하지 못하고, 탐색 방향의 일관성 유지가 어려울 것으로 보인다. 이상과 같이 본 연구에서 제시한 공생파트너 선택전략들은 각각 많은 장단점을 지니고 있다. 그러나 NKC 모델과 이진스트링커버링문제에서는 이러한 장단점이 서로 보완적으로 상호작용하여 특정한 한 두가지의 공생파트너 선택전략이 다른 전략들에 비해 우월한 성능을 갖지 않은 것으로 나타났다.

<그림 1>과 <그림 2>는 목표스트링의 수가 3개이고, 128비트인 이진스트링커버링문제를 대상으로 세가지 공생진화 알고리즘에서 공생파트너

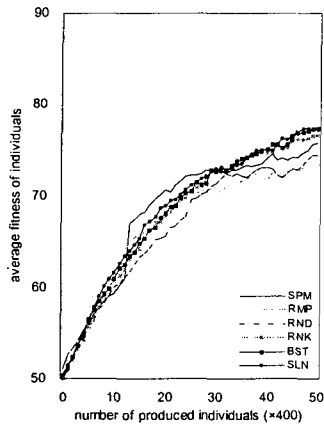
선택전략들이 보이는 모집단평균 변화와 최선해 갱신경향을 보인 것이다. <그림 1>을 통해, 적응도를 기준으로 공생파트너를 선택하는 전략들은 임의성을 갖는 전략들보다 모집단평균을 높게 함을 알 수 있다. 그리고 <그림 2>로부터 임의성을 갖는 전략들이 적응도를 기준으로 사용하는 전략에 비해 최선해 갱신회수는 적음을 알 수 있다. 그러나 진화가 안정되면 모든 전략들은 비슷한 수준의 해를 탐색한다. NKC 모델에서도 이와 유사한 경향을 보인다.

## 6. 결론

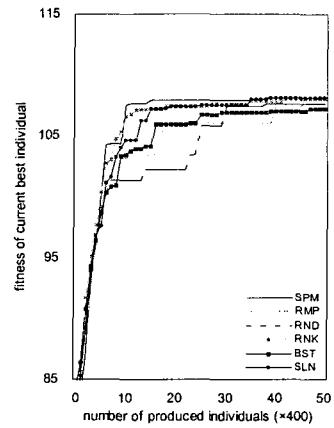
본 연구는 공생파트너 선택전략이 공생진화 알고리즘의 탐색 성능에 미치는 영향을 분석하였다. 이를 위하여 NKC 모델과 이진스트링커버링문제를 대상으로 공생진화 알고리즘을 분류하고, 다양한 공생파트너 선택전략을 적용하였다. 그리고 성능 결과를 통계적 기법에 의해 분석하였다. 분석결과, 공생진화 알고리즘에서 공생파트너 선택전략이 해의 성능에 크게 영향을 미치지 않음을 알 수 있었다. 이는 공생진화 알고리즘의 설계에 있어서 공생파트너 선택 전략을 신중하게 고려할 필요가 없음을 의미한다. 특히, 부분문제가 많아지면 공생파트너 선택에 따른 계산시간이 기하 급수적으로 증가할 수 있다. 본 연구 결과로부터 공생파트너는 임의 선택 전략(RND)과 같이 계산시간이 짧고, 구현이 용이한 전략을 선택하면 될 것으로 판단된다.

## 참고문헌

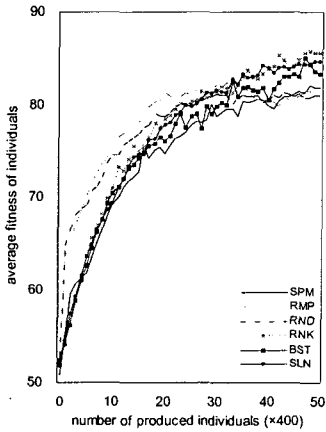
- [1] Bull, L. and Fogarty, T.C., "Artificial symbiogenesis," *Artificial Life*, 2, (1995), pp. 269-292.
- [2] Bull, L., "Evolutionary computing in multi-agent environments: partners," *Proceedings 7th International Conference Genetic Algorithms*, East Lansing, MI, pp. 370-377, 1997.
- [3] Kauffman, S.A., *The origins of order: Self-organisation and selection in evolution*, Oxford University Press: New York, Oxford, 1993.
- [4] Kim, J.Y., Kim, Y., and Kim, Y.K., "An endosymbiotic evolutionary algorithm for optimization," submitted to *Applied Intelligence*.
- [5] Kim, Y.K., Kim, J.Y., and Kim, Y., "A coevolutionary algorithm for balancing and sequencing in mixed model assembly lines," to appear in *Applied Intelligence*.
- [6] Maher, M.L. and Poon, J., "Modelling design exploration as co-evolution," *Microcomputers in Civil Engineering*, 11, (1996), pp. 195-210.
- [7] Moriarty, D.E. and Miikkulainen, R., "Forming neural networks through efficient and adaptive coevolution," *Evolutionary Computation*, 5, (1997), pp. 373-399.
- [8] Potter, M.A., "The design and analysis of a computational model of cooperative coevolution," Ph. D. dissertation, George Mason University, 1997.



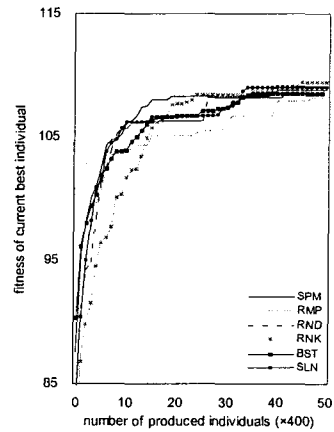
(a) SPA의 모집단평균 변화



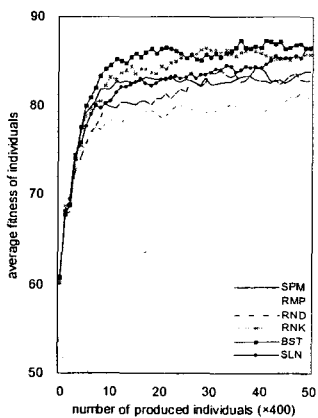
(a) SPA의 최선해 갱신경향



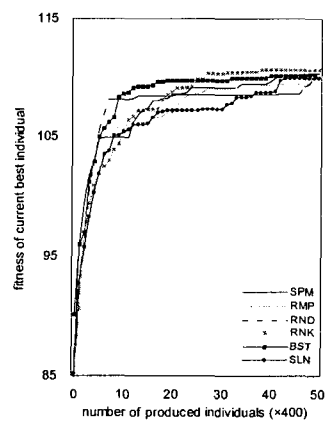
(b) SNA의 모집단평균 변화



(b) SNA의 최선해 갱신경향



(c) EEA의 모집단평균 변화



(c) EEA의 최선해 갱신경향

<그림 1> 모집단평균 변화

<그림 2> 최선해 갱신경향