

조합최적화 문제를 위한 새로운 유전연산자

석상문[○], 안병하
광주과학기술원 기전공학과
{soakbong[○], bayhay}@kjist.ac.kr

New genetic crossover operators for sequencing problem

Sangmoon Soak[○], Byungha Ahn
Dept. of Mechatronics, Kwang-ju Institute of Science and Technology

요 약

지난 10년 동안 유전 알고리즘은 어렵고 복잡한 다양한 문제들을 해결하기 위한 새로운 방법으로 인식되어왔다. 이러한 유전 알고리즘의 성능은 알고리즘 내에 구현되는 여러 연산자들에 좌우된다. 따라서 많은 연구자들이 새로운 연산자 개발에 관심을 가져 왔었다. 특히, 가장 널리 알려진 조합최적화 문제 중에 하나인 알려진 traveling salesman problem (TSP)의 경우 NP-hard 문제로 분류되어 현재까지 이를 해결하기 위한 다양한 유전 연산자들이 개발되어 왔었다. 따라서 본 논문에서는 TSP 문제를 test problem로 이용하여 이를 해결하기 위한 새로운 유전 연산자 특히 교차 (crossover operator) 연산자들을 제안하고 기존의 다양한 연산자들과 비교를 통해서 성능을 입증한다.

1. 서 론

유전 알고리즘은 처음 John Holland에 의해 제안된 이후 아주 다양한 복잡한 문제를 해결하기 위해 적용되어 왔다 [1]. 그리고 알고리즘의 성능은 이를 구성하는 연산자들에 의해 좌우되기 때문에 성능 향상을 위해서 아주 많은 연산자들이 개발되어왔다. 특히, 많은 연구자들이 교차연산자에 보다 많은 관심을 가져왔는데, 이는 일반적으로 최적해 (global optimal solution) 가 국소 최적해 (sub-optimal solution) 의 근처에 있다고 알려져 있고 [2] 또한 유전 알고리즘에서 지역탐색 (local search)의 기능을 하는 연산자가 교차연산자이기 때문이다. 따라서 본 논문에서는 유전 연산자의 속도와 정확성의 측면에서 성능을 향상시킬 수 있는 두 가지의 새로운 유전 연산자를 제안하고, 제안한 연산자의 성능을 입증하기 위해 잘 알려진 다양한 유전 연산자들과의 비교실험을 수행한다.

2. 다양한 교차 연산자들

본 논문에서, 5가지의 치환 기반 교차 연산자 (permutation-based crossover : PMX, OX, OBX, PBX, CX), 휴리스틱 교차 연산자 (heuristic crossover operator) [3] 와 휴리스틱 교차 연산자 의 변형들 [6][7] 그리고 두가지의 에지 기반 교차 연산자 (edge-based crossover operators) 즉, edge recombination 교차 연산자 (ER) [4] 그리고 enhanced edge recombination 교차 연산자 (EER) [5] 등 다양한 연산자들이 제안하는 연산자와 비교된다.

치환 기반 연산자의 경우 많은 문헌들에서 소개되고 있기 때문에 여기서는 휴리스틱 교차 연산자 및 그 변형들과 그리고 edge recombination 교차 연산자와 그 변형들에 대해 간단히 소개한다.

2.1 휴리스틱 교차 연산자와 변형들

휴리스틱 교차 연산자 (HX)는 Grefenstette [1], [3] 등에 의해 처음 소개 되었다. 이 연산자는 부모 세대의 이웃하는 에지 정보를 이용하는 최초의 연산자였다. 휴리스틱 연산자의 다음과 같이 수행된다.

우선, 두 부모를 고르고 두 부모에서 무작위로 하나의 도시를 선택한다. 만약 이것이 어떤 cycle 을 형성하지 않는다면 그 다음 선택된 도시와 이웃하는 부모세대의 두 오른쪽 도시와의 거리를 비교를 통해 더 짧은 에지가 처음에 선택된 도시의 다음에 온다. 만약 짧은 에지가 어떤 cycle을 형성한다면 이전에 선택되지 않은 도시들 중에 하나를 선택한다. 이 과정을 모든 도시가 다 선택될 때까지 반복을 한다.

휴리스틱 교차 연산자는 단지 부모의 오른쪽에 이웃하는 도시들의 정보만 사용한다. 하지만 symmetric TSP의 경우 에지 (2, 3) 과 (3, 2)는 전적으로 똑 같은 거리 값을 가지고 있다. 따라서 오른쪽에 놓여 있는 에지만을 사용할 이유가 없다. 본 논문에서는 이를 확장하여 양쪽의 에지를 모두 고려하는 휴리스틱 교차 연산자의 확장 연산자를 소개할 것이다.

그리고 휴리스틱 교차 연산자의 변형으로 greedy crossover (GX) [6] 가 있다. greedy 교차 연산자의 경우는 자손 1은 부모세대의 오른쪽 에지를 이용해서 형성되

고 자손 2는 부모세대의 왼쪽 에지를 이용해서 형성된다.

2.2 에지 재결합 교차 연산자와 변형들

에지 재결합 교차 연산자 (Edge recombination crossover : ER)는 Whitley [4] 등에 의해 TSP를 해결하기 위해 처음 개발되었다. HX가 부모세대 에지들의 거리 정보를 이용하는 대신에 ER은 단순히 부모세대의 에지 정보만을 이용한다.

즉, ER은 우선 부모세대에 있는 에지들의 표를 작성한 후 자손을 만들어내기 위해 미리 작성된 표를 이용한다.

Fig. 1은 ER의 예를 보여 준다.

먼저 무작위로 한 도시(도시 3)를 선택하고 선택된 도시와 연결 가능한 도시들 중에서 에지 표에 남아 있는 에지의 개수가 가장 작은 도시를 선택한다. 이 예의 경우 도시 4가 4개 남아 있고 도시 5가 2개 남아 있으므로 도시 5가 선택된다. 만약 도시들이 똑 같은 개수로 남아 있다면 무작위로 하나를 선택한다. 이 과정을 완전한 tour가 구성될 때까지 수행한다.

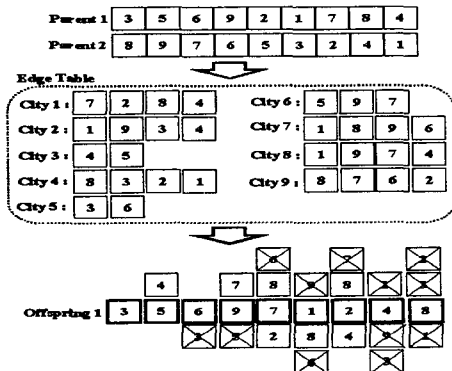


Fig. 1. Edge Recombination 교차 연산자

ER을 수정한 새로운 연산자가 Starkweather [5] 등에 의해 제안된 enhanced edge recombination crossover (EER)이다. ER과의 차이는 EER이 부모의 공통된 에지들을 보존해 나간다는 것이다. 즉 부모들의 공통된 에지를 경쟁에서 우선권을 부여한다.

그들은 부모들의 이웃하는 정보를 이용하는 ER과 EER이 다른 기존 교차 연산자들 보다 우수한 결과를 나타낸다는 것을 보여 주었다 [4],[5]. EER이후 ER을 강화하기 위한 많은 변형들이 [8],[9] 소개되었다. 하지만 본 논문에서는 ER과 EER만을 고려한다.

3. 새로운 교차 연산자들

새로운 교차연산자는 HX [3]와 EER [5]의 분석을 통해서 개발되었다. 전자는 에지의 거리정보에 초점을 맞추고 후자는 부모의 공통 에지에 초점을 맞추고 있다. 본 논문에서는 단지 에지의 거리정보에만 초점을 두는 교차 연산자와 거리 정보 뿐만 아니라 공통된 에지 정보 또한 이용하는 연산자를 개발하였다. 여기서 개발된 교차 연산자

는 더 많은 부모 세대의 정보를 자손에게 전달 시켜 주는 특징을 지니고 있다.

단순 에지 보존 교차 연산자 (Simple edge preservation crossover : SEPX)는 각 에지의 거리 정보에 초점을 둔다. HX와 달리 SEPX는 부모의 양쪽 에지의 거리정보 중 짧은 쪽을 선택하도록 한다.

두번째 교차 연산자는 에지 보존 교차 연산자 (edge preservation crossover : EPX)이다. EPX는 에지의 거리 정보와 부모의 공통된 에지 정보를 동시에 고려를 한다. 본 논문에서는 공통 에지가 진화 과정에서 누적된 중요한 정보를 더 많이 가지고 있다고 가정하고 거리 정보보다 에지 정보를 우선해서 고려한다.

SPX의 절차는 다음과 같다.

먼저, 임의의 도시를 선택하고 현재 도시의 부모의 양쪽 에지들을 확인한다. 만약 공통된 에지가 존재하면, 공통된 에지가 선택되고 그렇지 않을 경우는 현재까지 선택되지 않은 도시와 연결하는 에지 중에서 거리가 가장 짧은 에지를 선택한다. 만약 모든 에지가 cycle을 형성한다면, 이전에 선택되지 않은 도시들 중 무작위로 하나의 도시를 선택한다.

Fig. 2는 임의로 선택된 두 부모와 EPX에 의해 만들어진 자손들을 보여준다.

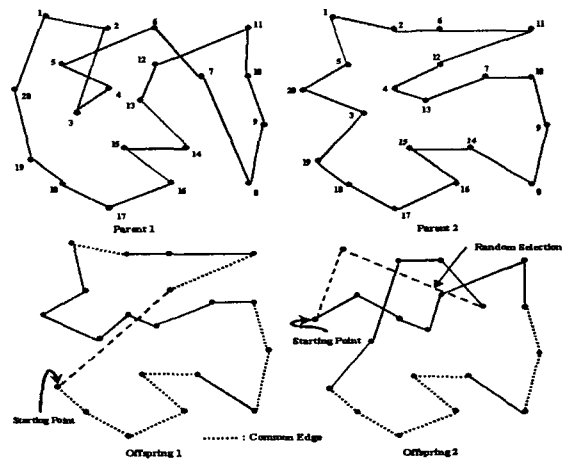


Fig. 2 EPX의 예

4. 실험 결과

본 논문에서는 기존의 교차 연산자들과의 비교를 위해 단순 유전 알고리즘을 사용하여 몇 가지 실험 데이터를 이용해 실험을 하였다.

선택전략으로는 토너먼트 선택전략 (토너먼트 크기 : 2)과 엘리트 전략 (보존 개수 : 3)을 사용하였고, 돌연변이 연산자로는 반전 돌연변이 (inversion mutation) 연산자를 사용하였다. 그리고 population의 크기 (*pop_size*), 교차 확률 (P_c) 그리고 돌연변이 확률 (P_m)은 각각 200, 0.6 and 0.4를 사용하였다. 종료 조건은 최적해를

발견하거나 N 세대 동안 해의 개선이 없을 경우로 하였으면 여기서 N 의 크기는 1000으로 설정하였다. 모든 알고리즘은 Pentium III 450Mhz 에서 Visual C++를 이용해서 구현되었고 4가지 (eil51, eil76, kroA100 and pr124) 실험 데이터를 이용하여 실험하였다. 모든 실험 데이터는 TSPLIB[10]를 참고하였다.

Table 1은 실험 결과를 보여 준다. 본 논문에서 제안하는 두 가지 교차 연산자 모두 다른 기존의 교차 연산자들 보다 거의 모든 경우에서 우수한 결과를 산출하였다. eil51의 경우, 제안하는 두 교차 연산자 모두 최적해를 발견하였지만, 다른 교차 연산자들은 최적해를 찾지 못했다. 다른 실험 데이터들에서는 최적해를 찾지는 못했지만 거의 최적에 가까운 해들을 찾아내었다.

pr124의 경우 EPX와 SEPX 각각 3.9%와 2.5% 정도 최적해를 초과하는 결과를 산출했으며 EPX와 SEPX가 다른 치환 기반 교차 연산자 (permutation-based crossover operators)들 보다 계산시간이 더 소요되었으며, 반면에 부모들의 이웃정보를 이용하는 연산자들 보다 훨씬 더 적은 계산시간을 필요로 하였다.

EPX와 SEPX 두 연산자의 비교를 통해서 SEPX가 EPX보다 더 좋은 결과를 찾아냈지만 더 많은 계산시간과 반복횟수를 필요로 하였다. 따라서 본 논문에서 제안하는 두 연산자 모두 다른 치환 기반 교차 연산자나 에지 기반 교차 연산자들 보다 훨씬 좋은 결과를 만들어 냄을 확인할 수 있었다.

5. 결론

본 논문에서 조합 최적화 문제를 위한 두 가지 새로운 교차 연산자를 제안하였다. 그리고 몇 가지 실험을 통해서 제안하는 연산자가 다른 기존 연산자들 보다 더 우수한 결과를 산출해 냄을 확인할 수 있었다. 비록 제안하는 연산자가 더 많은 계산 시간을 필요로 할지라도 성능은 훨씬 우수하였다. 또한 본 논문을 통해 부모의 정보를 이용하는 것이 좋은 해를 찾기 위한 방법이 될 수 있음을 확인할 수 있었다. 그리고 만약 적합도 함수 값만 합당하게 바뀌 준다면 제안하는 교차 연산자들은 다른 조합 최적화 알고리즘에 적용이 가능하다.

제안하는 연산자들은 단지 부모의 공통된 에지만을 고려한다. 따라서 부모의 공통된 sub-string은 고려하지 않고 있기 때문에 부모가 가지고 있는 정보를 잃어버릴 수도 있다. 이를 고려하는 것에 대한 연구는 다음으로 남겨 둔다.

참고 문헌

[1] D. Goldberg.(1989) Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley, Reading, MA
 [2] K.D. Boese, "Cost Versus Distance In the Traveling Salesman Problem", Technical Report CSD-950018, UCLA, 1995.
 [3] J. Grefenstette, R. Gopal, B. Rosmaita, and D. Gucht, "Genetic Algorithms for the Traveling Salesman Problem", Proc. the 1st Inter. Conf. on GAs and Their Applications, PP.160-168, 1985.
 [4] D. Whitley, T. Starkweather and D. Fuquay, "Scheduling problems and traveling salesman: the genetic edge recombination and operator", Proc. Third Int. Conf. G As, pp.133-140, 1989.
 [5] T. Starkweather, S. McDaniel, K. Mathias, D. Whitley and C.

Table. 1 교차 연산자 비교 결과

	eil51 (426)				
	Min	Avg.	Max	Avg. CPU Time(sec)	Avg. gen.
PMX	434	448.8	463	11.6	1614.2
OX	431	444.6	457	16.3	1916.1
OBX	431	445.3	463	17.9	1848.9
PBX	431	446.8	468	24.4	1888.9
CX	435	448.7	463	8.8	1540.5
ER	435	445.9	462	35.8	1762.9
EER	435	447.5	465	28.8	1448.8
HX	427	438.9	450	25.2	1577.2
GX	435	447	467	19.6	1308.6
EPX	426	435.4	450	17.4	1080.4
SEPX	426	434.2	450	28.6	1365.8
eil76 (538)					
PMX	568	582.8	605	18.0	1521.7
OX	545	560.6	579	26.2	2014.9
OBX	556	572.5	589	41.3	2420.8
PBX	556	575.3	596	56.3	2435.8
CX	552	583.5	608	13.9	1638.1
ER	561	578.2	609	73.7	2048.8
EER	561	580.8	604	61.7	1734.7
HX	552	566	586	52.1	1895.6
GX	551	580.5	599	37.3	1434.8
EPX	544	561.8	577	36.1	1235.9
SEPX	541	552.5	563	48.7	1553.6
kroA100 (21282)					
PMX	22165	23159	24782	29.9	1771.4
OX	21389	22712	24045	42.6	2071.6
OBX	21876	22911	24341	67.3	2581.5
PBX	21543	23280	24503	90.2	2545.4
CX	21959	23003.8	24854	21.4	1943.5
ER	21579	22882	24383	102.3	1849.2
EER	21941	23031.3	24185	92.3	1672.6
HX	21644	22504	23848	70.7	1704.5
GX	22261	23029.1	24255	62.9	1615.9
EPX	21556	22263.9	24009	59.7	1392.8
SEPX	21383	21894.4	22798	75.3	1530.8
pr124 (59030)					
PMX	60362	62654.9	66679	48.6	2230.8
OX	59777	62048.7	64401	63.2	2300.1
OBX	59576	61260.7	65125	80.6	2869.0
PBX	60163	6224.5	65693	137	2780.9
CX	59652	62061.3	64837	29.6	2389.6
ER	60049	62265.5	70275	163.2	2102.4
EER	59903	62104.5	66618	155.7	2023.6
HX	59413	61147.8	63700	116.5	2007.7
GX	60560	62362.2	67275	104.4	1910.9
EPX	59087	61359.5	63906	90.6	1521.1
SEPX	59323	60561.5	63297	104.3	1538.7

(): Optimal solution

Whitley, "A Comparison of Genetic Sequencing Operators", Proc. the Fourth Inter. Conf. on GA, 1991.
 [6] R. Yang, "Solving Large Traveling Salesman Problems with Small Populations" Genetic Algorithms in Engineering Systems: Innovations and Applications, pp157-162, 1997.
 [7] L. Yang and D.A. Stacey, "Solving the Traveling Salesman Problem Using the Enhanced Genetic Algorithm", Lecture Notes in Computer Science, Vol. 2056, pp307-316
 [8] K. Mathias and D. Whitley, "Genetic operators, the fitness landscape and the traveling salesman problem", Parallel Problem Solving from Nature 2, pp.219-228, 1992.
 [9] H. D. Nguyen, I. Yoshihara and M. Yasunaga, "Modified Edge Recombination Operators of Genetic Algorithms for the Traveling Salesman Problem", IECON 2000. 20 th Annual Conference of the IEEE, Volume:4, pp 2815-2820, 2000.
 [10] TSPLIB. Web Site, <http://www.iwr.uni-heidelberg.de/iwr/comopt/soft/TSPLIB95/TSPLIB.html>.