차수 제약 걸침 나무 문제를 해결하기 위한 트리 표현법

석상문[O] 안병하
광주 북구 오룡동 1번지 광주과학기술원 기전공학과
{soakbong, bayhay@kjist.ac.kr}

Tree Representation for solving Degree Constraint Minimum Spanning Tree Problem

Sangmoon Soak[0] , Byungha Ahn
Dept. of Mechatronics, Kwang-ju Institute of Science and Technology

요 약

최소 걸침 나무는 널리 알려진 순회 판매원 문제와 같이 전통적인 최적화 문제 중에 하나이다. 특히나 최소 걸침 나무와는 달리 차수 제약 최소 걸침 나무의 경우는 일반적으로 NP-hard 문제로 알려져 있다. 이러한 NP-hard 문제를 해결하기 위한 다양한 접근법들이 소개되었는데 유전 알고리즘은 효율적인 방법 중에 하나로 알려져 있다. 유전 알고리즘과 같이 진화에 기반을 둔 알고리즘을 어떤 문제에 적용하기 위해서 가장 우선적으로 고려되어야 하는 것은 해를 어떻게 표현할 것인가 인데 본 논문에서는 차수 제약 최소 걸침 나무를 해결하기 위한 새로운 트리 표현법을 제안한다.

1. 서 론

최소 걸침 나무(Minimum Spanning Tree : MST)는 통신 네트워크 설계 및 교통망 네트워크 설계와 같이 다양한 네트워크 설계 문제에 많이 적용되어왔다. 특히 차수 제약을 가진 차수 제약 최소 걸침 나무(Degree Constraint MST : DCMST)는 NP-hard 문제로 알려져 있으며 이러한 문제를 해결하기 위한 많은 시도들이 있어왔다. 그 중 최근 들어 가장 많이 사용되는 방법이 진화의 개념에서 탄생한 다양한 진화 알고리즘들이다. 이러한 진화 알고리즘을 이용해서 어떤 문제를 해결하기 위해서는 가장 먼저 해의 표현에 대한 고려가 있어야 한다. 따라서 DCMST의 경우도 다양한 해 표현법들이 제안되어 왔다 [2], [4], [5], [8], [9], [10]. 하지만 대부분의 해 표현법들이 항상 트리 (tree)를 형성해 내는 것은 아니기 때문에 트리를 만들기 위한 repair 과정을 필요로 하였다. 이에 반해서 Pr□fer 수를 이용한 경우 [3], [6], [7]나 NetKey를 이용한 방법들 [10], [11]은 항상 트리를 형성한다는 장점을 지니고 있다.

본 논문에서는 DCMST를 해결하기 위한 새로운 해 표현 방법을 제안한다. 제안하는 표현법은 기존의 TSP를 해결하기 위한 해표현법을 확장한 개념으로써 Pr□fer 수와 NetKey를 이용한 방법처럼 항상 트리를 형성하기 때문에 어떤 repair 과정이 필요 없다는 특징을 지니고 있다. 또한 제안하는 표현법에 맞는 새로운 연산자도 제안한다. 따라서 본 논문에서는 기존의 항상 트리를 만들어 내는 두 방법과 제안하는 방법의 비교를 통해 제안하는 방법의 성능을 입증하도록 한다.

2. 새로운 트리 표현법

본 논문에서는 TSP 를 EAs 를 이용해서 해결할 경우 사용하는 표현법을 확장한 새로운 트리 표현 방법을 제안한다. 일반적으로, TSP 에서는 각각의 도시가 해 표현에 단지 한번씩만 나타나는데 이를 시작 도시와 종료 도시를 연결하지 않고 남겨두면 degree 가 2 인 하나의 트리가 형성된다.

이 개념을 확장하여 $N\times 2$ 의 길이를 가진 해 표현을 고려해 보자. 단 여기서 각각의 도시는 해 표현에 두 번씩 나타나야만 한다. 예를 들면, (8, 7, 2, 5, 4, 1, 4, 3, 5, 2, 6, 1, 8, 7, 6, 3)가 형성된 하나의 해이다. 이는 8 개의 노드가 있는 경우의 예이다. 다음은 이 해 표현에서 어떻게 트리를 형성할 것인가 이다. 본 논문에서는 다음과 같은 트리 구성 방법을 제안한다.

우선, $SelNode\{\phi\}$ 와 $EdgeSet\{\phi\}$ 를 설정한다. 여기서 $SelNode\{\cdot\}$ 와 $EdgeSet\{\cdot\}$ 는 각각 이전 단계에서 선택된 노드의 집합과 에지의 집합을 나타낸다. 첫 번째 노드는 8 이고 두 번째 노드는 7 이다. 이 두 노드는 이전에 선택된 적이 없기 때문에 연결하고 노드 집합과 에지 집합에 각각 포함시킨다. 따라서 선택된 노드 집합은 *SelNode*{8, 7}이 되고 에지 집합은 *EdgeSet*{(8, 7)}이 된다. 다음은 한 칸 옆으로 이동해서 7 과 2 을 비교한다. 2 는 이전에 선택된 노드가 아니므로 다시 7 과 2 를 연결하고 각각 선택된 노드 집합과 에지 집합에 추가한다. 따라서 각각 *SelNode*{8, 7, 2}과 *EdgeSet*{(8, 7), (7, 2)}가 된다. 다음은 2 와 5 이다. 5 도 이전에 선택된 적이 없으므로 연결한다. 따라서 *SelNode*{8, 7,

2, 5}과 *EdgeSet*{(8, 7), (7, 2), (2, 5)}가 된다. 만약에 두 노드 모두 선택된 경우는 그냥 넘어가고 이러한 과정을 N-1개의 에지가 에지 집합에 추가될 때까지 반복 수행한다. 그 다음 에지 집합에 있는 에지들을 이용해서 트리를 나타내면 항상 트리를 만들어 낼 수 가 있다.
그림. 1 은 제안하는 방법을 통한 트리 구성 예를 보여준다.

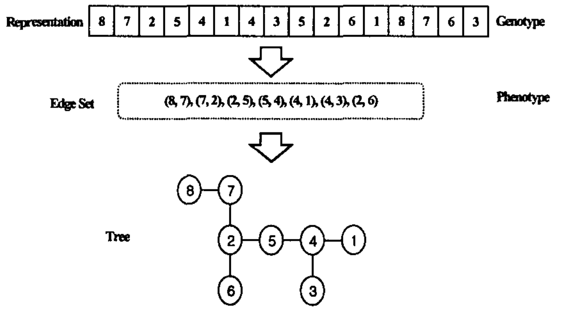

그림.1. 새로운 트리 표현 예

위에서 제시한 예는 degree 가 3 인 경우의 해 표현법이다. 즉, 해 표현에 같은 노드가 나타나는 횟수가 d-1 번 이면 degree d 를 만족하는 트리를 구성할 수 있다. 이를 확장하면 degree 가 d 일 경우는 길이가 $N\times(d-1)$ 이고, 각각의 노드가 d-1 번씩 나타나게 형성하면 된다. 이를 이용하면 모든 경우의 DCMST 문제에 간단하게 적용이 가능하며 또한 트리를 대상으로 하는 모든 문제에 적용이 가능하다. 그리고, 본 논문에서 제안하는 방법은 앞에서 설명한 두 표현법 (Pr□fer 수와 NetKey 를 이용한 방법)과 같이 항상 트리를 형성하기 때문에 다른 repair 과정이 필요가 없으며, 각각의 노드가 d-1 씩 나타나야 한다는 제약조건만 만족시키면 기존의 어떤 연산자들을 이용해서도 구현이 가능하다. 그리고 encoding 과정과 decoding 과정이 하나의 해 표현으로 가능하기 때문에 부모가 지닌 정보를 이용한 연산자 개발이 용이하다.

3. 새로운 교차 연산자

본 논문에서는 제안하는 해 표현법에 맞는 에지-거리 교차연산자 (Edge-Distance Crossover : EDX)라는 새로운 교차 연산자를 제안한다.
이 교차 연산자는 각각의 부모의 똑 같은 위치에 있는 에지들 사이의 거리를 비교하고 더 짧은 에지에 있는 노드를 자손에 넘겨준다. 만약 두 에지 모두 어떤 Cycle을 형성한다면 임의로 하나의 노드를 선택한다.
예를 들면, 부모 1의 해가 {1, 2, 3, 4} 이고 부모 2의 해가 {2, 3, 1, 4}라면, 자손 1의 첫번째 노드는 부모1의 첫번째 노드를 가져와서 {1}이 되고 다음 두번째 노드는 부모 1의 두번째 노드 2와 부모 2의 두번째 노드 3과의 거리 비교 즉, 에지 (1, 2)와 에지 (1, 3) 중 거리가 가까운 노드를 자손 1의 다음 노드로 선택한다.

4. 실 험

본 논문에서는 제안하는 방법처럼 항상 트리를 만들어내기 때문에 repair 과정이 필요 없는 Pr□er 수와 NetKey 를 이용한 표현법과 비교 실험을 수행하였다. 실험에 사용된 데이터는 500 X 500 크기의 grid 상에서 임으로 노드를 생성하였고, 각 노드간의 거리는 Euclidean distance 를 이용하였다. 그리고 각 노드의 degree 제약은 똑같이 3 으로 제한을 하였다. 또한 각 알고리즘은 Visual C++를 이용해서 작성하였고 Pentium IV 1.6 Ghz 컴퓨터상에서 실험을 하였다.
유전 알고리즘에 사용된 파라메타로 population 크기는 100, 교차 확률은 0.6, 돌연 변이 확률은 0.1 로 설정하였고 종료조건은 generation 이 1000 이 될 때까지로 설정하였다. 그리고 선택전략으로는 토너먼트 크기가 2 인 토너먼트 선택전략을 사용하였고, 3 가지의 교차연산자 (1-point, 2-point, uniform crossover)와 제안하는 EDX 연산자 그리고 3 가지의 돌연변이 연산자 (inversion, reciprocal, random perturbation)을 이용해서 실험을 하였다. 각 실험은 10 번씩 수행하였고 실험 결과는 이를 평균한 값이다.
표. 1 은 다른 표현 방법들과의 실험 결과를 보여준다.
실험결과를 통해 알 수 있듯이 모든 경우에서 제안하는 방법이 기존의 Pr□er 수를 이용하는 방법이나 NetKey 를 이용하는 방법 보다 훨씬 우수한 결과 값을 산출하였고 문제의 크기가 커질수록 해의 차이가 훨씬 커짐을 확인할 수 있다. 그리고 연산자들 간의 비교를 보면 제안하는 EDX 교차 연산자와 Inversion 돌연변이 연산자를 이용한 경우가 가장 좋은 실험결과를 산출하였다.
Pr□er 수 표현의 경우는 다른 문헌들에서 제안하는 Uniform 교차 연산자와 Random Perturbation 돌연변이 연산자를 이용한 경우가 본 실험에서도 가장 우수한 것으로 나타났다. NetKey 를 이용한 표현의 경우 pr□er 수를 이용한 표현보다 결과 값은 좋게 나왔지만 제안하는 방법에 비해서는 상당한 격차를 보임을 확인 할 수 있다. 또한 이 방법은 노드 수가 적은 경우에 아주 좋은 결과를 찾아내는 특징을 보였다.

5. 결 론

본 논문에서는 진화 알고리즘을 위한 새로운 트리표현 방법과 이를 위한 새로운 교차 연산자를 제안하였다. 새로운 방법은 기존의 TSP 를 위한 해 표현법의 확장으로 표현이 아주 간단하고 항상 트리를 만들어 낸다는 장점을 지니고 있으며 다양한 DCMST 문제로의 확장 또한 간단하다. 그리고 제안하는 교차 연산자의 경우 부모가 가지고 있는 유전정보를 이용한다는 점에서 기존의 알고리즘들이 사용하는 일반적인 연산자들과는 차이를 보이며, 기존의 표현법과의 비교 실험을 통해서도 훨씬 더 좋은 결과를 산출해냄을 확인할 수 있었다.
유전 알고리즘의 경우 각각의 연산자들을 어떻게 구성하느냐가 알고리즘의 성능을 좌우한다. 따라서 다양한 연산자들의 실험이 필요하다.

참고문헌

1. B. Dengiz, F. Altiparmak, and A.E. Smith : Local Search Genetic Algorithm for Optimal Design of Reliable Networks, IEEE Transactions on Evolutionary Computation, Vol. 1, No. 3, pp.179-188, (1997).
2. M. Gen and R. Cheng : Genetic Algorithms and Engineering Design, JOHN WILEY & SONS, (1997).
3. G.R. Raidl and B.A. Julstrom : Edge-Sets: An Effective Evolutionary Coding of Spanning Trees, TR-186-1-01-01, Vienna University of Technology, (2002).
4. B.A. Julstrom and G.R. Raidl : Initialization is Robust in Evolutionary Algorithms that Encode Spanning Trees as Sets of Edges, ACM Symposium of Applied Computing 2002, (2002).
5. G. Zhou and M. Gen : An Effective Genetic Algorithm Approach to The Quadratic Minimum Spanning Tree Problem, Computers Operations Researches, Vol. 25, NO. 3, PP. 229-237, (1998).
6. M. Krishnamoorthy, A.T. Ernst and Y.M. Sharaiha : Comparision of Algorithms for the Degree Constrained Minimum Spanning Tree, Journal of Heuristics, Vol. 7, pp. 587-611, (2001).
7. M.R. Garey and D.S. Johnson, Computers and Intractability, A Guide to the Theory of NP-Completeness, San Francisco, Freeman, 1979.
8. F.N. Abuali, R.L. Wainwright and D.A. Schoenefeld, Determinant Factorization: A New Encoding Scheme for Spanning Trees Applied to the Probabilistic Minimum Spanning Tree Problem, pp.470-477.
9. F. Rothlauf, D.E. Goldberg and A. Heinzl, Network Random Keys □ A Tree Network Representation Scheme for Genetic and Evolutionary Algorithms, IlliGAL Report No. 2000031, July 2000.
10. B. Schindler, F. Rothlauf and H.J. Pesch, Evolution Strategies, Network Random Keys, and the One-Max Tree Problem, EvoWorkshops 2002, LNCS Vol.2279, pp. 143-152, 2002.

		New method				Prufer			NetKey		
	M₩C	1-P	2-P	Uniform	Edge	1-P	2-P	Uniform	1-P	2-P	Uniform
N=10 [1147]	Inver.	1155.8	1147.7	1155.9	**1147**	1195.4	1190.9	1191.9	1157.	**1147**	**1147**
	Reci.	**1147**	1147.7	1148	**1147**	**1147**	1179.2	1148.2	**1147**	**1147**	**1147**
	Rand.	1149.7	1150.3	1153.9	1147.7	**1147**	1163.4	1167.1	**1147**	**1147**	**1147**
N=20 [1613]	Inver.	1862.5	1751.3	1994.3	**1631.1**	2117.2	2164.3	2105.6	1876	1920.2	1763.1
	Reci.	1861.5	1819.7	1899.6	1678.2	2333.9	2377.1	2434.7	1729	1712.5	1682.2
	Rand.	1919.8	1901.8	1893.3	1709.3	2429.1	2505.8	2485.5	1644	1680.1	1667.8
N=30 [1619]	Inver.	2445.3	2449.1	2639.8	**1811.2**	2877.2	2517.7	3198.7	2559.5	2651.1	2354.6
	Reci.	2348.7	2374.3	2432	1922.5	3164.1	3079.3	3235.2	2180.7	2152.4	2152.3
	Rand.	2502.6	2433.3	2441	2033.7	2770.4	2672	2607.6	2105.8	2008.2	1950.6
N=40 [2236]	Inver.	3835.4	3690.1	4524.8	**2544.9**	4145.8	4089.5	5368.7	4307.2	4203.7	3722.7
	Reci.	3837.3	3770.1	3741.5	2884	5032.6	4964.1	5127.1	3609.3	3635.3	3608.2
	Rand.	3980.7	3722.4	3679.5	3178.2	5047.9	5122.6	4785.8	3184.3	3211	3213
N=50 [2244]	Inver.	4907.2	4819.2	5683.9	**2911.4**	5097.7	5504.4	7602.4	5630.7	5582.5	4404.9
	Reci.	4673.9	4453.9	4405.6	3427.1	6393.1	6283.1	6622.8	4598.4	4683.3	4502.7
	Rand.	4773	4599.3	4570.7	3700.8	6090.6	6180	6397.5	4494.1	3933.9	4151.7

[]: Prim 알고리즘을 적용한 결과값, M : Mutation, C : Crossover.

표 1. 실험 결과 비교