

타부 리스트가 결합된 유전자 알고리즘을 이용한 트리형 네트워크의 경제적 설계

이성환* · 이한진** · 염창선**†

*수산자원사업단 경영기획실

**부경대학교 경영학부

Economic Design of Tree Network Using Tabu List Coupled Genetic Algorithms

Seong-Hwan Lee* · Han-Jin Lee** · Chang-Sun Yum**†

*Management Planning Division, Korea Fisheries Resources Agency

**Division of Business Administration, Pukyong National University

This paper considers an economic design problem of a tree-based network which is a kind of computer network. This problem can be modeling to be an objective function to minimize installation costs, on the constraints of spanning tree and maximum traffic capacity of sub tree. This problem is known to be NP-hard.

To efficiently solve the problem, a tabu list coupled genetic algorithm approach is proposed. Two illustrative examples are used to explain and test the proposed approach. Experimental results show evidence that the proposed approach performs more efficiently for finding a good solution or near optimal solution in comparison with a genetic algorithm approach.

Keywords : Economic Design, Tree-Based Network, Tabu List, Genetic Algorithm

1. 서 론

트리형 네트워크는 루트(root) 노드와 여러 노드가 계층적으로 구성 되어 있는 네트워크를 의미한다. 트리형 네트워크의 설계 문제는 루트 노드와 여러 노드의 위치 등이 주어진 상태에서 신장트리(spanning tree)와 최대 트래픽 용량을 제약조건으로 하면서 설계비용을 최소화하는 노드들의 연결 형태를 결정하는 것이다. 이러한 문제는 노드의 수가 증가할수록 계산 시간과 해의 탐색 공간이 지수적으로 증가하는 NP-hard 문제에 속한다[4].

그 동안 트리형 네트워크의 설계를 위해 메타 휴리스틱

기법을 이용한 다음과 같은 연구들이 수행되어 왔다. Ersoy et al.[2]은 시뮬레이티드 어닐링(SA : Simulated Annealing)을 이용하여 신장트리 제약조건 하에서 평균 네트워크 지연(average network delay)을 최소화하는 네트워크 설계 방법을 제시하였다.

Torres-Jiménez et al.[7]은 SA을 이용하여 신장트리와 노드 간 링크의 최대 트래픽 용량 제약조건 하에 설계비용을 최소화하는 네트워크 설계 방법을 제시하였고, Sharaiha et al.[5]은 타부 서치(TS : Tabu Search)를 이용하여 신장트리와 노드 간 링크의 최대 트래픽 용량 제약조건 하에서 설계비용을 최소화하는 네트워크 설계 방법을 제시하

였다. Zhou et al.[8]은 유전자 알고리즘(GA : Genetic Algorithm)을 이용하여 신장트리와 서브트리의 최대 트래픽 용량 제약조건 하에서 설계비용을 최소화하는 네트워크 설계 방법을 제시하였다. Gamvros et al.[3]은 GA를 이용하여 신장트리와 노드 간 링크의 최대 트래픽 용량 제약조건 하에서 네트워크의 설계비용을 최소화하는 설계 방법을 제시하였다. 트리형 네트워크의 설계 문제는 최근에 메타휴리스틱 알고리즘 중에서는 GA가 가장 많이 이용되고 있다.

GA의 해 탐색과정은 확률에 의한 연산과정을 포함하고 있어 지속적인 진화 연산과정 중 개체집단 내에 여러 개체가 동일한 해를 가질 수 있다. 개체집단 내의 동일한 해들의 존재는 GA의 효율적인 최적해 탐색을 방해하는 요인이 될 수 있다. 이러한 GA의 비효율적인 탐색과정을 개선하기 위한 한 방법으로 TS의 타부 리스트 기능이 이용될 수 있다. 타부 리스트 기능은 개체집단 내의 개체들이 동일한 해를 갖는 것을 배제하는 역할을 수행한다.

본 연구에서는 트리형 네트워크 설계를 위해, 유전자 알고리즘에 타부 리스트 기능을 결합시킨 형태인 타부 리스트가 결합된 유전자 알고리즘을 제안한다. 그리고 새로 제안한 알고리즘을 이용하여 신장트리와 서브트리의 최대 트래픽 용량을 제약조건으로 하고 네트워크 총 설계비용을 최소화하는 네트워크의 설계 방법을 제시한다. 그리고 실험적 문제들을 가지고 본 연구에서 제시한 설계 방법과 GA를 이용하여 설계하는 방법을 비교 평가함으로써 본 연구에서 제시한 설계 방법의 우수성을 입증하고자 한다.

2. 트리형 네트워크 설계 문제의 표현

2.1 기본 가정 및 문제 모형

네트워크 설계를 위한 기본 가정은 다음과 같다.

- 각 노드(허브 또는 터미널)의 위치는 주어져 있다.
- 노드는 완전히 신뢰할 수 있다.
- 각 링크는 무방향이다.
- 노드의 메모리는 무한하다.
- 네트워크에 있어 중복된 링크는 존재하지 않는다.

네트워크 설계와 관련된 표기는 다음과 같다.

- n : 전체 노드의 수
- s_i : i 번째 노드, $i = 1, 2, 3, \dots, n$
- s_{ij} : s_i 와 s_j 간 링크
- d_{ij} : s_i 와 s_j 간 거리 또는 s_i 와 s_j 간 비용
- T_j : j 번째 서브트리($s_i \in T_j$)

- $C(x)$: 네트워크 x 의 총 설치비용(\$)
- q_i : i 번째 노드의 발생 트래픽 용량(KB/sec)
- K : 하나의 서브트리에서 전송 가능한 최대 트래픽 용량

타부 리스트가 결합된 유전자 알고리즘을 위한 표기는 다음과 같다.

- g : 세대 수
- g_{max} : 최대 세대수
- s : 개체집단 크기
- $m_{\%}$: 각 개체집단에 발생하는 돌연변이 비율
- r_m : 염색체내의 돌연변이 비율
- T_s : 타부 리스트의 크기

본 연구에서 사용되는 트리형 네트워크 설계 문제는 아래와 같이 모형화 될 수 있다.

목적함수 :

$$\text{Minimize } C(x) = \sum_{i=0}^{n-1} \sum_{j=1}^n d_{ij} s_{ij} \quad (1)$$

제약조건 :

$$s_{ij} = \begin{cases} 1, & \text{if } (i, j) \text{ is selected} \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$\sum_{i=0}^{n-1} \sum_{j=1}^n s_{ij} = n - 1 \quad (3)$$

$$\sum_{i=0}^n s_{ij} \geq 1, \quad j = 1, 2, \dots, n \quad (4)$$

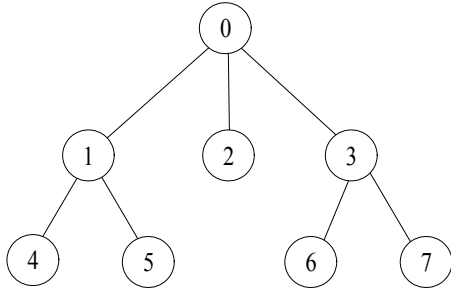
$$\sum_{s_i \in T_j} q_i \leq K \quad (5)$$

식 (1)은 네트워크 총 설치비용 $C(x)$ 인 네트워크 전체에 대한 링크연결 비용의 최소화를 의미한다. 식 (2)는 노드 s_i 와 s_j 간의 링크가 있으면 1, 없으면 0을 나타낸다. 식 (3)은 네트워크가 신장트리 조건을 만족하기 위해 $n-1$ 개의 링크를 가져야 한다는 것을 표현한 것이다. 식 (4)는 각 노드는 반드시 하나 이상의 다른 노드와 연결되어야 한다는 것을 의미한다. 식 (5)는 어떤 서브트리(T_j)에 속하는 노드들의 발생 트래픽 용량의 합이 최대 트래픽 용량(K) 이하가 되어야 한다는 것을 의미한다.

2.2 트리형 네트워크의 염색체 표현 및 서브트리 트래픽 용량

<그림 1>은 1개의 루트 노드와 7개의 노드로 구성된 트리형 네트워크의 한 후보해이다. 이 후보해는 {1, 4, 5},

{2}, {3, 6, 7}의 3개 서브트리로 구성되어 있다. <표 1>은 네트워크의 노드 간 링크 유무를 보여 주고 있다.



<그림 1> 트리형 네트워크의 후보해

<표 1> 네트워크의 노드 간 링크 유무

	0	1	2	3	4	5	6	7
0	-	1	1	1	0	0	0	0
1		-	0	0	1	1	0	0
2			-	0	0	0	0	0
3				-	0	0	1	1
4		-			-	0	0	0
5						-	0	0
6							-	0
7								-

이 후보해는 다음과 같은 염색체(chromosome)로 표현될 수 있다.

해 : {1 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0}

이는 <표 1>의 우 상향 삼각형 모양의 각 행을 1줄로 나열한 것이다. 각 형질(allele)은 링크 유무를 나타낸다.

루트 노드를 제외한 각 노드들의 발생 트래픽 용량이 순서대로 2, 3, 1, 1, 2, 2, 2일 경우, {1, 4, 5} 서브트리의 트래픽 용량은 노드 1, 4, 5의 발생 트래픽 용량 합인 5이고, {2} 서브트리의 트래픽 용량은 노드 2의 발생 트래픽 용량인 3이고, {3, 6, 7} 서브트리의 트래픽 용량은 노드 3, 6, 7의 발생 트래픽 용량 합인 5이다.

3. 타부 리스트가 결합된 유전자 알고리즘을 이용한 트리형 네트워크 설계

3.1 트리형 네트워크 설계를 위한 절차

타부 리스트가 결합된 유전자 알고리즘을 이용한 트리형 네트워크 설계의 주요 절차는 다음과 같다.

- (1) 초기 세대의 개체집단 생성($g = 1$)
 - (a) 초기 세대의 개체집단을 무작위로 생성한다.
 - (b) 초기 세대의 개체집단을 비용 계산함수로 보낸다.
 - (c) 타부 리스트가 비어 있다면, 초기해를 타부 리스트에 등록한다. 타부 리스트가 비어 있지 않다면, 타부 리스트의 해와 중복되는 초기해는 버리고 중복되지 않는 초기해는 타부 리스트의 해보다 우수할 경우 타부 리스트에 등록한다.
 - (d) 타부 리스트에 등록된 초기해는 초기 개체집단에 넣는다.
 - (e) 초기 세대의 가장 우수한(best) 해를 검사한다.

(2) 반복적인 교배 및 돌연변이 연산

- (a) 교배
 - 현재 세대의 개체집단으로부터 서로 다른 두 부모를 선택한다.
 - 균등교배를 통해 자손을 생성한다.
 - 생성된 자손을 비용 계산함수로 보낸다.
 - 타부 리스트의 해와 중복되는 자손은 버리고 중복되지 않는 자손은 타부 리스트의 해보다 우수할 경우 타부 리스트에 등록한다.
 - 타부 리스트에 등록된 자손은 개체집단에 넣는다.
- (b) 돌연변이
 - 개체집단으로부터 돌연변이 대상 해를 무작위로 선택한다.
 - 돌연변이 비율에 따라 돌연변이 연산을 한다.
 - 생성된 돌연변이 해를 비용 계산함수로 보낸다.
 - 타부 리스트의 해와 중복되는 돌연변이 해는 버리고 중복되지 않는 돌연변이 해는 타부 리스트의 해보다 우수할 경우 타부 리스트에 등록한다.
 - 타부 리스트에 등록된 돌연변이 해는 개체집단에 넣는다.
- (c) 현재 세대의 가장 우수한(best) 해를 검사한다.
- (d) 최대 세대 수($g = g_{max}$)까지 반복한다.

3.2 교배 및 돌연변이

교배 연산을 위해 서로 다른 두 개의 부모 염색체를 순위 기반 선택(rank based quadratic procedure)을 적용하여 선택한다[6]. 다음으로 선택된 두 부모 염색체를 대상으로 균등 교배를 실시한다. 예를 들어, 아래와 같이 부모 1과 부모 2가 선택되었고, 각 형질의 난수가 발생되었다고 가정하자.

자손의 염색체는 난수가 임계확률(0.5)보다 같거나 큰 2, 4, 5, 7, 9, 11, 12, 13, 18, 19, 20, 21, 23, 26, 28번째

부모 1	{	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	}
부모 2	{	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	}
난수	{	.3	.9	.4	.5	.6	.2	.5	.1	.5	.3	.8	.6	.7	.2	.4	.1	.3	.6	.7	.8	.8	.3	.8	.2	.1	.6	.2	.5	}
자손	{	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	}

형질은 부모 1의 형질을 물려받고 난수가 임계확률보다 작으면 나머지 형질에 대해서는 부모 2의 형질을 물려받아 생성된다.

다음으로 행해지는 돌연변이 연산은 먼저 개체집단으로부터 돌연변이 염색체의 비율($m\%$)에 따라 돌연변이 대상 염색체를 무작위로 선택한다. 그리고 선택된 돌연변이 대상 염색체에 대해 돌연변이 형질의 비율(r_m)에 따라 형질을 무작위로 선택한다. 예를 들어, 돌연변이 대상 염색체가 {1110000001000001000011000000} 이고, 돌연변이 형질의 비율($r_m = 0.15$)일 때 11, 16, 17, 21번째 형질이 무작위로 선택되면, 돌연변이 된 염색체는 {111000000110000010000100000}이다.

3.3 타부 리스트

타부 리스트는 개체집단 내 동일한 해의 생성을 방지하는 기능이다. 교배 및 돌연변이 연산을 통해 생성된 해는 일단 타부 리스트의 해와 비교된다. 생성된 해가 타부 리스트의 해와 중복되지 않고 가장 열악한 해 보다 우수할 경우, 가장 열악한 해는 생성된 해로 대체된다[1, 5].

4. 수치적 예제 및 성능 실험

이 장에서는 두 개의 수치적 예제를 통해 본 연구에서 제시한 타부 리스트가 결합된 GA를 이용한 설계 방법의 성능을 실험한다. 실험은 Intel Core2 Duo(2.2GHz)와 1GB RAM으로 구성된 PC 환경에서 이루어졌다.

첫 번째 문제는 13개의 노드를 연결하는 경우이고, 두 번째 문제는 17개의 노드를 연결하는 경우이다. 두 문제를 위해 사용된 각 노드의 좌표는 <표 2>와 같다.

<표 2> 각 노드별 좌표

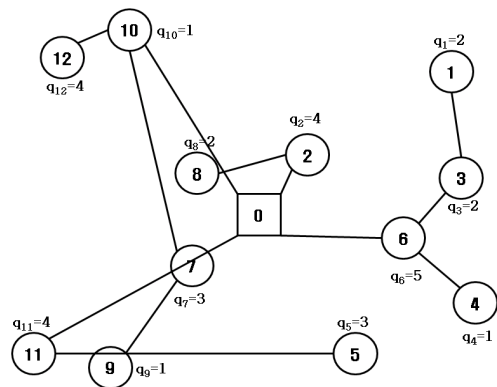
노드	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
가로 좌표	50	90	59	92	94	70	80	35	37	19	23	3	9	75	96	76	14
세로 좌표	50	3	63	8	31	20	45	9	59	17	90	20	85	76	36	13	33

주) 노드 0은 루트 노드이다. 문제 1은 노드 0~12를 사용한다.

4.1 문제 1 : 13개 노드 연결

13개 노드를 갖는 네트워크 설계 문제 1을 위한 서브트리 네트워크의 최대 트래픽 용량 $K = 10$ (KB/sec)이다. 그리고 루트 노드를 제외한 각 노드의 발생 트래픽 용량은 순서대로 2, 4, 2, 1, 3, 5, 3, 2, 1, 1, 4, 4이다.

초기 탐색 실험을 통해 초기화 값($g_{max} = 100, s = 100, m\% = 30, r_m = 0.2, Ts = 100$)을 얻은 후, $K = 10$ (KB/sec)에 대해 실험을 수행하였다. 초기 세대($g = 1$)에서는 <그림 2>와 같은 네트워크 비용이 396.64\$, 최대 서브트리 네트워크 트래픽이 10(KB/sec)인 초기 세대의 가장 우수한 해{010001000110010000000000000001000000100000001000000000001000000001100000000010}을 찾았다. <표 3>은 초기 세대의 가장 우수한 해에 대한 서브트리 트래픽 용량을 나타내고 있다. 그리고 최대 세대 수($g = g_{max}$)까지 실험을 수행한 결과, <그림 3>과 같은 설계구조를 갖고, 네트워크 비용이 277.66\$, 최대 서브트리 네트워크 트래픽이 9(KB/sec)인 가장 우수한 최종해 {01000111000001000000001000000000000000000001100000000000000000000000010000100010010}을 찾았다. <표 4>는 가장 우수한 최종해의 서브트리 트래픽 용량을 나타내고 있다.



<그림 2> 가장 우수한 초기해(문제 1 : 396.64\$)

<표 3> 가장 우수한 초기해의 서브트리 트래픽 용량(문제 1)

서브트리	{2, 8}	{6, 3, 1, 4}	{11, 5}	{10, 12, 7, 9}
트래픽용량 (KB/sec)	6	10	7	9

