

최적 통신 스페닝 트리 설계 문제를 위한 파라미터 자동조절 유전알고리즘에 대한 연구

김종율

부산광역시 사상구 주례2동 산69-1 동서대학교

Study on the Parameter Auto Tuned Genetic Algorithm for OCST Design Problems

Jong Ryul Kim

Dongseo University

E-mail : xmaskjr@gdsu.dongseo.ac.kr

요 약

최근 유무선 통신 시스템의 발전에 따른 인터넷 환경의 급격한 변화는 가상공간의 출현과 유비쿼터스 컴퓨팅 환경 구축에 대한 요구를 가속화시키고 있으며 이와 관련된 이론 및 기술의 발전을 주도해 왔다. 이와 관련한 문제들 중에 가장 근간이 되는 문제들 중 하나는 최적 통신 스페닝 트리 (OCST: Optimal Communication Spanning Tree) 문제이다. 본 논문에서는 이러한 최적 통신 스페닝 트리 문제를 해결하기 위해 파라미터를 자동 조절하는 유전 알고리즘 (Parameter Auto Tuned GA, PAT-GA)을 이용한다. 제안하는 유전 알고리즘은 교차율, 돌연변이율과 같은 파라미터를 자동조절하기 위해 퍼지 논리 제어기 (FLC: Fuzzy Logic Controller)를 이용한다. 임의로 생성된 예제에 대한 수치 실험을 통해 통신시스템의 기본 문제 중 하나인 최적 통신 스페닝 트리 문제의 해법으로서의 제안 알고리즘의 유용성과 효율성을 확인한다.

키워드

파라미터 자동 조절 유전 알고리즘 (Parameter Auto Tuned Genetic Algorithm, PAT-GA), , 최적 통신 스페닝 트리 (Optimal Communication Spanning Tree, OCST)

1. 서 론

네트워크 설계 문제들 중에 가장 근간이 되는 문제는 최적 통신 스페닝 트리 (OCST: Optimal Communication Spanning Tree) 문제이다. OCST를 적용한 대형 통신 네트워크 시스템들은 백분으로 여러 네트워크들을 함께 묶어서 구성된다. 또한 근래에 네트워크를 구축하려고 할 때, 무중계 전송 가능 거리 (최대 120Km 이상), 광대역폭 (거의 50 Tbps), 적은 신호 감쇠 (거의 0.2 dB/km) 등 [1]과 같은 잠재력을 가지고 있어 광케이블이 주로 사용되고 있으며 광케이블 접속의 비싼 비용을 고려할 때 초기의 네트워크의 구조는 주로 스페닝 트리 (Spanning Tree) 형태로 구성되는 것이 일반적이다.

한편 급격히 증가하는 인터넷 이용자에 대응하기 위해서 라우팅 프로토콜 (Routing Protocol)의 제어와 관련한 규모성 문제 (Scalability Issues)를 효과적으로 다루기 위해서 ISP (Internet Service Provider)들은 자신이 관리하는 네트워크 (자치 시스템 또는 Autonomous System (AS))를 논리적으로

분할하여 계층적으로 운용하기 시작하였으며 이에 따라 라우팅 프로토콜의 제어는 분할된 각 지역 영역 (Local Area, LA)별로 독립적으로 수행 되도록 운용하고 있다. 이러한 계층적 네트워크 시스템은 주로 스페닝 트리 형태로 구성되고 있다 [2]. 또한, 인터넷 수능 방송과 같이 다수의 사용자가 동일한 서비스를 요청할 때 서비스하는 방식인 그룹 통신을 위한 멀티캐스트 서비스에도 그 그룹을 관리하기 위해 주로 스페닝 트리 형태로 구성하여 서비스하는 것이 일반적이며 유비쿼터스 컴퓨팅을 위해 최근 그 관심이 한층 더 높아지고 있는 센서 네트워크 분야에서도 배터리 운용 등을 고려할 때 현실적으로 가능한 토폴로지는 스페닝 트리 형태를 따른다는 연구들이 많이 수행되고 있다 [3].

최적의 OCST를 찾아내는 것은 매우 복잡한 조합최적화 문제이며 NP-hard 최적화 문제로 분류된다. 전통적인 방법들을 이용하면, 네트워크의 규모가 지수적으로 확대되어지는 경우, OCST 문제를 해결해 내기가 힘들어 진다. 이러한 이유 때

문에 유전 알고리즘 (GA)을 기초로 하는 메타 휴리스틱 알고리즘이 주목을 받고 있다. GA는 많은 실세계 문제들을 풀기 위해 자주 이용되고 있으며 네트워크 최적화 문제들의 최적화 기술로서의 잠재력을 매우 주목받고 있는 기법이며 여러 문제들에서 다목적 문제에 대한 기법으로 매우 활발한 연구가 이뤄지고 있다 [4, 5].

본 논문에서는 파라미터를 자동 조절하는 유전 알고리즘 (Parameter Auto Tuned GA, PAT-GA)을 이용한다. 제안하는 유전 알고리즘은 교차율, 돌연변이율과 같은 파라미터를 자동조절하기 위해 퍼지 논리 제어기 (FLC: Fuzzy Logic Controller)를 이용한다. 제안하는 PAT-GA는 n 개의 노드에 대해 $n-2$ 개의 숫자열로 표현 가능한 최적 통신 걸침 문제에 적합하도록 하는 효율적인 유전자 표현법을 이용한다. 마지막으로 임의로 생성된 예제를 수치 실험하여 본 논문에서 제안하는 PAT-GA가 OCST문제를 효과적이고 효율적으로 풀이할 수 있는 해법임을 확인한다.

II. 문제기술

일반적으로 OCST 문제는 전형적인 MST문제와는 달리 각각 노드들의 쌍과 관련이 있는 통신요구량 (Communication requirements)과 각 링크의 연결 비용을 함께 고려한다. 이러한 시스템의 성능측도들은 주로 네트워크 토폴로지에 의해 좌우되며 클러스터링과 라우팅의 두 가지 주요한 항목들도 구해낸 토폴로지에 의해 결정된다 [7].

본 논문에서 풀이하고자 하는 OCST문제는 수학적으로 다음과 같은 전체 스패닝 트리의 통신 비용을 최소화 하는 문제로 정의할 수 있다.

$$\begin{aligned} \min z(\mathbf{x}) &= \sum_{i=1}^n \sum_{j=1}^n d_{ij} \cdot c_t(i, j) \cdot x_{ij} \\ \text{s. t. } \sum_{i=1}^{n-1} \sum_{j=i+1}^n x_{ij} &= n-1 \\ x_{ij} &= 0 \text{ or } 1, i=1, 2, \dots, n-1, \quad j=2, 3, \dots, n \end{aligned}$$

여기서 c_{ij} 는 링크 (i, j) 의 연결 비용이고 d_{ij} 는 링크 (i, j) 에서의 통신 요구량을 나타내고 x_{ij} 는 링크 (i, j) 가 스패닝 트리를 구성하기 위해 선택되는지 여부를 나타낸다. 또한 $c(i, j)$ 는 노드 i 와 노드 j 사이의 패스가 $i-k-l-m-j$ 이면 $c_{ik} + c_{kl} + c_{lm} + c_{mj}$ 와 같이 계산되는 경로패스의 비용을 나타낸다.

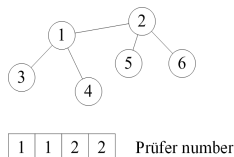


그림 1. 트리와 그에 대한 Prüfer 수 (PN)

III. 문제를 위해 제안하는 유전 알고리즘

3.1 유전자 표현법 및 초기화

본 논문에서는 PN을 기반한 유전자 표현법을

실세계의 OCST 문제에 보다 더 적합하면서도 PN의 특성을 살린 새로운 유전자 표현법을 제안한다. 즉, 본 논문에서는 중간 노드들을 PN으로 나타내고 잎 노드들을 클러스터 숫자열로 표현하는 새로운 유전자 표현법을 제안한다.

제안하는 유전자 표현법을 설명하기 위해 우선 PN에 대하여 설명하면 다음과 같다. 그래프 이론의 오래된 정리들 중에서 그래프 열거에 대한 일반적인 정리 중 하나는 k 개의 노드를 가진 완전 그래프에는 $k^{(k-2)}$ 개의 스패닝 트리가 존재한다는 Cayley의 정리이다. Prüfer는 그러한 스패닝 트리들과 k 개의 숫자들로 이뤄지는 $k-2$ 길이의 숫자열 집합사이의 일대일 대응관계를 확증함으로써 Cayley의 정리를 증명하였다. 이것은 k 개의 다른 숫자들로 구성되는 $k-2$ 개의 숫자열로 스패닝 트리를 표현할 수 있다는 것을 의미한다. 왜냐하면 임의의 스패닝 트리에 대해 적어도 2개의 잎(leaf) 노드가 존재하기 때문이다. 이러한 숫자 조합의 열을 Prüfer수 (PN)이라고 한다 [4, 5, 10].

예를 들면 그림 1에서 나타낸 것과 같이 6개의 노드로 구성되는 스패닝 트리를 PN (1 1 2 2)으로 표현할 수 있다.

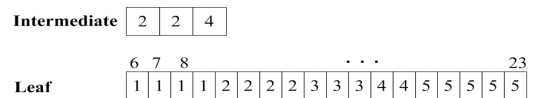


그림 2. 제안하는 유전자 표현법의 예

본 논문에서는 앞서 설명한 PN 인코딩 방법의 장점을 살리면서 보다 더 실세계의 OCST 문제에 적합한 유전자 표현을 위해 다음과 같은 두개의 유전자형을 가지는 인코딩 방법을 제안한다: 중간 노드들의 연결 상태는 PN으로 표현하고 중간 노드와 잎 노들사이의 연결 상태들은 클러스터링 문자열로 나타내는 것이다. 따라서 그림 2에 표시된 것과 같이 초기화 과정에서 서비스 센터는 영역 $[1, n]$ 에서 임의로 $n-2$ 개의 숫자 (PN)로 생성하여 나타내고 사용자는 영역 $[1, m]$ 에서 임의로 m 개 숫자 (클러스터 문자열)로 생성한다. 여기서 클러스터 문자열은 어떻게 사용자를 서비스 센터에 할당되는지에 대한 정보를 나타낸 것으로 사용자가 어떤 특정 서비스 센터에 속하는 것을 의미한다. 본 논문에서의 유전자 표현은 $n+m-2$ 길이를 가진다.

3.2 FLC를 이용한 파라미터 자동조절

본 논문에서는 유전 알고리즘의 파라미터들을 자동으로 조절하기 위해 Yun의 방법 [13]을 이용한다. 핵심 아이디어는 GA의 탐색 과정에서의 연속된 두 세대간의 평균 적합도의 변화를 고려하면서 FLC를 이용하여 교차율과 돌연변이율을 휴리스틱하게 업데이트하는 것이다.

자동 조절 알고리즘을 위해 t 세대에서의 평균 적합도 $\Delta \bar{f}(t)$ 를 다음과 같이 정의한다.

$$\Delta \bar{f}(t) = \left(\overline{f_{parSize}}(t) - \overline{f_{offSize}}(t) \right) = \left(\frac{\sum_{k=1}^{parSize} f_k(t)}{parSize} - \frac{\sum_{k=1}^{offSize} f_k(t)}{offSize} \right)$$

FLC의 입력은 두 연속적인 세대에서의 $\Delta \bar{f}(t)$ 와 $\Delta \bar{f}(t-1)$ 이고 출력은 교차율과 돌연변이율에 대한 변화량을 의미한다. 또한 퍼지 입력과 출력 변수에 대한 멤버십 함수는 각각 NR (Negative larger), NL (Negative large), NM (Negative medium), NS (Negative small), ZE (Zero), PS (Positive small), PM (Positive medium), PL (Positive large), 그리고 PR (Positive larger)과 같이 정의 된다.

따라서 본 논문에서는 다음과 같은 프로시저로 FLC를 이용한 교차율 및 돌연변이율에 대한 자동 조정이 가능하도록 한다.

Step 1: FLC의 입력값으로 두 연속적인 세대 ($t-1$ 과 t)에서의 평균 적합도의 변화량 ($\Delta \bar{f}(t)$ 와 $\Delta \bar{f}(t-1)$)으로 설정한다.

Step 2: $\Delta \bar{f}(t)$ 와 $\Delta \bar{f}(t-1)$ 을 -1에서 1사이의 값 (교차율) 또는 -0.1에서 0.1 사이의 값으로 정규화한 후 이 값에 대한 Yun [13]에서 제안한 비퍼지화 테이블에서의 인덱스값을 구해낸다.

Step 3: 교차율의 변화량 $\Delta c(t)$ 또는 돌연변이율 변화량 $\Delta m(t)$ 을 다음과 같이 계산해 낸다.

$$\Delta c(t) = 0.02 \times Z(i, j) \quad ,$$

$$\Delta m(t) = 0.002 \times Z(i, j)$$

여기서 $Z(i, j)$ 는 Step 2에서 구해 낸 인덱스 값들에 해당하는 비퍼지화 값을 의미한다.

Step 4: 교차율과 돌연변이율을 다음과 같이 수정한다.

$$p_c(t) = \Delta c(t) + p_c(t-1)$$

$$p_m(t) = \Delta m(t) + p_m(t-1)$$

여기서 $p_c(t)$ 와 $p_m(t)$ 는 각각 t 세대에서의 교차율과 돌연변이율을 의미한다.

3.3 평가 함수

네트워크의 토폴로지를 설계하는 문제를 위한 평가함수 $eval(v_k)$ 의 값을 평가하기 위해 우선 인코딩된 염색체 v_k 를 인접행렬 x_k 로 디코딩한다. 그런 후 목적함수 값을 구해 내고 평가함수 $eval(v_k)$ 을 다음과 같이 계산한다.

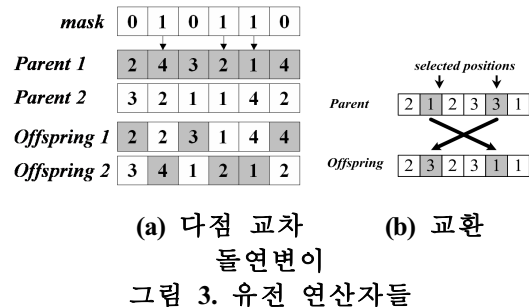
$$eval(v_k) = z(v_i) + p(v_i),$$

여기서 $p(\cdot)$ 는 페널티 함수로 만일 $R(v_k) > R_{req}$ 이면 0이고 그렇지 않으면 $z(v_i) \cdot (1 - R(v_i))$ 의 값이 된다.

3.4 선택

본 논문에서 사용된 선택 연산은 자유로이 해공간을 탐색하도록 하기 위해 토너먼트 선택법과 엘리트리스트 (elitist)방법을 혼합한 형태이다.

토너먼트 선택은 n 개의 염색체를 임의로 선택하고 이들 중에서 가장 적응도가 높은 개체를 다음 세대의 집단에 포함시킨다. 다음 세대의 염색체들을 모두 얻을 때까지 이 과정을 반복한다. 이때 선택된 염색체를 현재 집단에서 제거하지 않는다. 즉, 한 염색체가 여러 번 선택될 수 있다. 본 논문에서는 n 을 5로 설정하여 이용한다. 선택 압력은 토너먼트 크기인 n 의 값에 따라 좌우되는데 n 이 크면 적합도가 낮은 염색체들이 다음 세대에서 선택되는 기회가 적어진다. 이러한 토너먼트 선택은 최근에 주로 많이 사용되는 선택법 중에 하나이고 구현이 쉬우며 병렬 구조로 수행될 때 용이하고 선택 압력을 쉽게 조절할 수 있다는 장점을 가지고 있다.



3.4 유전자 연산자들

그림 3에서와 같이 균등교차 (uniform crossover) 이라고도 불리는 다점 교차(multi-point crossover)와 교환 돌연변이를 이용한다.

IV. 수치실험 및 결과

수치 실험을 통해 본 논문에서 두 연속적인 세대에서의 적합도 변화율을 고려한 FLC를 이용하여 자동으로 파라미터를 조정하는 제안 유전 알고리즘의 성능에 대한 우수성을 나타낸다. 수치실험은 다음과 같은 2가지 수치예제로 구성되어 수행하였다:

- 5 (최상위 및 중간노드 수: 3, 잎노드: 2)개의 노드를 가진 문제(Example 1)
- 12 (최상위 및 중간노드 수: 4, 잎노드: 8)개의 노드를 가진 문제(Example 2)

각각은 집단의 크기를 100, 최대 세대수를 1000, 초기 교차율을 0.7, 초기 돌연변이율을 0.6으로 설정하여 수행되었다.

Example 1의 통신 요구량 행렬과 비용행렬은 참고문헌 [6]에 나타내져 있는 것을 이용하였고 Example 2의 통신 요구량 행렬과 비용행렬은 각각 [0, 200]사이의 영역에서 임의로 생성한 행렬을 이용하였다. 수치 실험은 각 예제에 대해 20번씩 수행하였으며 최선, 최악의 결과와 20회 실험에 대한 평균을 내어 그 결과들을 표 1에서 요약하여 나타내었다.

표 1에서 볼 수 있듯이 제안한 유전 알고리즘

은 적절한 시간 내에 자동 조절기능을 사용하지 않은 GA보다 더 강건한(robust) 형태로 최소의 비용을 갖는 네트워크 토폴로지를 구해 낼 수 있다는 것을 확인할 수 있다.

표 1. 수치실험의 결과

Numerical Examples	GA		제안 GA		
	Objective	Comp. Time	Objective	Comp. Time	
1	Best	283	1.00	283	0.60
	Avg.	283	1.00	283	0.70
	Worst	283	1.00	283	1.10
	Std.	0	0.00	0	0.24
2	Best	31348	15.00	31348	9.00
	Avg.	31348	15.15	31348	9.95
	Worst	31348	16.00	31348	10.00
	Std.	0	0.37	0	0.22

Avg.: average, Comp. Time: computational time (second), GA's parameter settings: $pop_size=100$, $max_gen=1000$, $p_c(0)=0.7$, $p_m(0)=0.6$.

V. 결 론

본 논문은 실세계의 네트워크 문제들 중에 가장 기본이 되는 OCST 문제에 대해 네트워크 토폴로지의 설계를 가능하게 하는 유전 알고리즘을 제안하였다. 본 논문에서는 두 연속적인 세대들 간의 평균 적합도의 변화량을 고려한 FLC를 이용하는 파라미터 자동 조정 유전 알고리즘을 제안하였다. 제안 알고리즘에서 사용하는 유전자 표현법은 n 개의 노드를 가진 그래프에 대해서 $n-2$ 개의 숫자열을 이용하여 트리를 표현할 수 있으면서 보다 OCST 문제에 적합한 유전자 표현법으로 최상위 노드를 포함한 중간 노드들의 연결 상태는 PN을 이용하여 표현하고 그 외 잎 노드와 중간 노드들과의 연결 상태는 클러스터링 문자열로 표현하여 보다 더 효율적인 GA를 제안한다. 제안한 유전 알고리즘을 이용한 수치 실험을 통해 그에 대한 증명을 하였다. 앞으로의 연구는 보다 더 많은 수치예제에 적용해 보고 메시지 지연시간과 같은 제약이 있는 OCST 문제에 대한 해법을 고안하는 것에 열중할 것이다. 또한, OCST를 위한 다양하면서도 새로운 유전자 표현법들에 대한 연구 수행할 계획이며 이러한 유전자 표현법들에 대한 비교연구도 수행할 계획이다.

참고문헌

[1] Mukherjee, B., *Optical Communication Networks*, McGraw-Hill, Inc., New York, 1997.

[2] 김도훈, "계층화된 링크-상태 인터넷 라우팅에서 가상 링크 운용 최적화를 위한 다기준 유전자 알고리즘의 응용", *IE Interfaces*, Vol. 16, Special Edition, pp. 14-20, December 2003.

[3] T. Chen, H. Tsai, and C. Chu, "Gathering-Load-Balanced Tree Protocol for Wireless Sensor Networks," *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, vol.2, pp.8-13, 2006.

[4] Gen, M. and R. Cheng, *Genetic Algorithms and Engineering Design*, John Wiley & Sons, New York, 1997.

[5] Gen, M. and R. Cheng, *Genetic Algorithms and Engineering Optimization*, John Wiley & Sons, New York, 2000.

[6] S. Soak, S. Chang, S. Byun, and B. Ahn, "Evolutionary Algorithm for Solving Optimum Communication Spanning Tree Problem," *Journal of the Korean Information Science Society: SW and Applications*, vol.32, no.4, pp.268-276, 2005 (in Korean).

[7] J. R. Kim and M. Gen, "A Genetic Algorithm for Solving Bicriteria Network Topology Design Problems", *Journal of Japan Society for Fuzzy Theory and Systems*, Vol.12, No.1, pp.43-54, 2000.

[8] Elbaum, R. and M. Sidi, "Topological design of local-area networks using genetic algorithms," *IEEE/ACM Transactions on Networking*, Vol.4, No.5, pp.766-778, 1996.

[9] H. Prüfer, "Neuer beweis eines satzesüber permutationen," *Arch. Math. Phys*, vol.27, pp.742-744, 1918.

[10] 김종율, 이재욱, 현광남, "2기준 네트워크 설계를 위한 새로운 인코딩법을 기반으로 하는 유전자 알고리즘", *정보과학회 논문지: 소프트웨어 및 응용*, 제32권, 제10호, pp.963-973, 2005.10.

[11] Kershenbaum, A. and R. Van Slyke, "Recursive analysis of network reliability," *Networks*, Vol.3, pp.81-94, 1973.

[12] G. R. Raidl and B. A. Julstrom, "Edge sets: an effective evolutionary coding of spanning trees," *IEEE Transaction on Evolutionary Computation*, Vol.7, No.3, pp.225-239, June, 2003.

[13] Yun, Y. S., "Study on Adaptive Hybrid Genetic Algorithm and Its Applications to Engineering Design Problems", Ph.D dissertation, Waseda University. 2005.