

Nash 유전 알고리즘을 통한 네트워크 설계

김종율^o 윤태수 이동훈

^o동서대학교 첨단기술연구센터

동서대학교 디지털콘텐츠학부

{xmaskjr^o tsyun dhl}@gdsu.dongseo.ac.kr

Network Design through Nash Genetic Algorithm

Jong Ryul Kim^o, Tae Soo Yun, Dong Hoon Lee

Research Center for Advanced Science and Technology, Dongseo University

Div. of Digital Contents, Dongseo University

요 약

본 논문에서 대상으로 하는 네트워크 구조는 서비스 센터, 터미널(사용자), 그리고 연결 케이블로 이뤄져 있고 서비스 센터의 구성을 위한 의사 결정자와 서비스 센터와 사용자의 터미널을 연결하는 의사 결정자들이 존재하고 각자의 목적함수를 최적화 하기위해 비타협적으로 의사 결정과정에서 참여한다고 가정한다. 이러한 문제는 Nash 게임으로 정식화될 수 있다. 본 논문에서는 연결비용, 평균 메시지 지연, 네트워크 신뢰도를 고려하여, Nash 게임으로 정식화되는 광대역 통신 네트워크의 네트워크 토폴로지 설계 문제를 풀기 위해 Nash 유전 알고리즘을 이용한다. 수치 실험을 통해 본 논문에서 이용한 Nash 유전 알고리즘이 효율적이며 효과적인 방법이라는 것을 살펴본다.

1. 서 론

통신 네트워크의 시장 규모가 점차 확대됨에 따라 광대역 통신 네트워크를 위한 설계 문제는 네트워크 설계자, 네트워크 분석가, 네트워크 관리자 와 같은 많은 관련 연구자들에 의해 많은 주목을 받고 있다. 고비용의 하드웨어/소프트웨어를 공유하고 원거리로부터의 정보에 대한 접근을 제공하기 위해 통신 네트워크의 이용은 급속히 늘어가고 있다. 텔레커뮤니케이션, 컴퓨터 네트워킹, 하수도관 시스템, 기름 및 가스관 시스템 등과 같은 많은 실세계의 네트워크 설계 문제에서, 어떤 제약조건들을 만족하면서 목적함수들을 최적화하는 네트워크 시스템을 설계하는 것은 매우 중대한 문제이다.

일반적으로 대형 통신 네트워크 시스템들은 백본으로 여러 네트워크들을 함께 묶어서 구성된다. 또한 근래에 네트워크를 구축하려고 할 때, 광대역폭 (거의 50 Tbps), 적은 신호 감쇠 (거의 0.2 dB/km), 적은 신호 왜곡, 낮은 전력 요구 등 [5]과 같은 잠재력을 가지고 있어 광케이블이 주로 사용되고 있으며 광케이블의 비싼 비용을 고려할 때 초기의 네트워크의 구조는 스페닝 트리 형태로 구성되는 것이 일반적이다. 따라서 전반적인 네트워크의 구조로 적당한 모델은 고용량의 케이블로 이뤄진 백본 네트워크로 서비스 센터들을 연결하고 분산 네트워크로 그 사용자들을 연결하는 형태라고 할 수 있다. 이러한 통신 네트워크 시스템을 설계할 때 중요한 것 중에 하나는 구축비용, 평균 메시지 지연시간, 통신량, 네트워크 신뢰도 등과 같은 성능 측도를 최적화하는 최상의 레이아웃을 탐색해 내는 것이다. 이러한 시스템의 성능측도들은 매우 중요하며 네트워크 토폴로지에 의해 좌우되는 경우가 많다. 그리고 클러스터링과 라우팅의 두 가지 주요한 항목들도 네트워크 토폴로지 설계 문제에 포함되어 있다. 클러스터링 문제는 얼마나 많은 서비스 센터들 (clusters)로 네트워크가 나뉠 수 있는가

와 어떻게 사용자들 (terminals)이 클러스터에 연결되어 질 것인가와 같은 두 가지 항목들로 이뤄져 있으며 라우팅 문제는 스페닝 트리 토폴로지를 어떻게 상호 연결할 것인가를 결정하는 것이다.

최적의 네트워크 토폴로지를 찾아내는 것은 매우 복잡한 조합최적화 문제이며 NP-hard 최적화 문제로 분류된다. 전통적인 방법들을 이용하면, 네트워크의 규모가 지수적으로 확대되어지는 경우, 네트워크 토폴로지 설계 문제를 풀이하기가 힘들어진다. 이 때문에 유전 알고리즘 (GA)을 기초로 하는 휴리스틱 알고리즘이 주목을 받고 있다 [1].

본 논문에서 대상으로 하는 네트워크 구조는 백본 네트워크의 구성을 위한 의사 결정자와 분산 네트워크의 구성을 위한 의사 결정자들이 존재하고 각자의 목적함수를 최적화 하기위해 비타협적으로 의사 결정과정에서 참여한다고 가정한다. 이러한 네트워크 설계 문제는 Nash 게임으로 정식화될 수 있다. 본 논문에서는 연결비용, 평균 메시지 지연, 네트워크 신뢰도를 고려하여, Nash 게임으로 정식화되는 광대역 통신 네트워크의 네트워크 토폴로지 설계 문제들을 풀기 위해 Nash 유전 알고리즘을 이용한다. 본 논문에서 제안한 Nash 유전 알고리즘은 n 개의 서비스 센터들과 m 개의 사용자들로 구성되어지는 문제를 위해 필요한 메모리는 단지 $n+m-2$ 뿐인 Prüfer수 (PN)와 클러스터 문자열을 이용한 새로운 인코딩 방법을 이용하여 유전자를 표현한다. 수치 실험을 통해 본 논문에서 이용한 Nash 유전 알고리즘이 효율적이며 효과적인 방법이라는 것을 살펴본다.

2. 네트워크 설계 문제

n 개의 서비스 센터들과 m 개의 사용자들을 구성되는 네트워크를 생각해 보자. 예를 들어 그림 1과 같이 5개의 서비스 센터와 18개의 사용자들로 구성되는 네트워크

를 생각해 보자. 사용자들간의 통신 통화량 수요는 사용자 통화량 행렬로 명명되는 $m \times m$ 행렬 U 로 정의되며 행렬 U 의 u_{ij} 는 사용자에서 사용자로의 데이터 통화량을 의미한다. 통화량 특성이 미리 정해져 있어서 사용자 통화량 행렬 U 에 요약된다. 서비스 센터들의 연결 형태를 나타내는 $n \times n$ 서비스 센터 토폴로지 행렬 X_1 를 정의한다. 행렬 X_1 의 x_{1ij} 는 다음과 같이 정의된다.

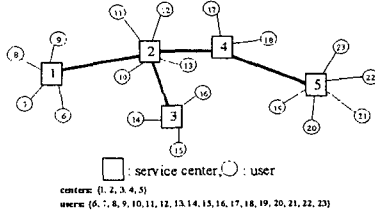


그림 1. 네트워크 구조의 예

$$x_{1ij} = \begin{cases} 1, & \text{센터 } i \text{와 센터 } j \text{가 서로 연결되어 있으면 } (i \neq j) \\ 0, & \text{그렇지 않으면} \end{cases}$$

그리고 네트워크가 n 개의 세그먼트(서비스 센터 또는 클러스터)로 나뉘질 수 있다고 가정한다. 사용자들은 n 개의 서비스 센터들로 분산된다. 사용자가 어떤 센터에 속하는 지를 나타내는 $n \times m$ 클러스터링 행렬 X_2 를 정의한다. 따라서

$$x_{2ij} = \begin{cases} 1, & \text{사용자 } j \text{가 센터 } i \text{에 연결되어 있으면} \\ 0, & \text{그렇지 않으면} \end{cases} \quad \text{이다.}$$

한 개의 터미널은 반드시 하나의 센터에만 연결된다. $n \times (n+m)$ 행렬 X 를 스페닝 트리 행렬($[X_1 X_2]$)이라고 한다. $n \times n$ 행렬 T 는 서비스 센터 통화량 행렬이라고 정의한다. 행렬 T 의 요소인 t_{ij} 는 센터 i 에 접속되어 있는 사용자들로부터 센터 j 에 접속되어 있는 사용자들에게로 전송되는 통화량을 나타낸다. 즉, $T = X_2^T U X_2$ 이다.

본 논문에서는 M/M/1 모델 [2]을 하나의 클러스터(세그먼트)의 동작을 묘사하기 위해 이용하고 0-1 비선형 Nash 게임으로 정식화되는 네트워크 설계문제는 다음과 같이 정식화할 수 있다.

$$\begin{aligned} \min_{X_2} & \sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{1ij} \cdot x_{1ij} + \sum_{i=1}^n \sum_{j=1}^m w_{2ij} \cdot x_{2ij} \\ \min_{X_1} & \frac{1}{\Gamma} \left[\sum_{i=1}^n \frac{F_i(X)}{C_i - F_i(X)} + \sum_{i=1}^n \sum_{j=1}^m \beta_{ij} \cdot f_{ij}(X) \right] \\ \text{s.t. } & R(X) > R_{\min} \\ & \sum_{i=1}^n x_{1ij} = 1, \quad j = 1, 2, \dots, n \\ & \sum_{i=1}^n x_{2ij} = 1, \quad j = 1, 2, \dots, m \\ & F_i(X) < C_i, \quad i = 1, 2, \dots, n \end{aligned}$$

$$a_{ij}^k(X) = \begin{cases} 1, & \text{센터 } k \text{를 통해 센터 } i \text{에서 센터 } j \text{로} \\ & \text{트래픽이 발생하면} \\ 0, & \text{그렇지 않으면} \end{cases}$$

$$b_{ij}^{(k,l)}(X) = \begin{cases} 1, & \text{센터 } k \text{와 센터 } l \text{을 연결하는 링크들을 통해} \\ & \text{센터 } i \text{에서 센터 } j \text{로의 트래픽이 존재한다면} \\ 0, & \text{그렇지 않으면} \end{cases}$$

$$R_{\min} \quad \text{네트워크 신뢰도의 최소 요구량}$$

$\beta_{ij} = p_{ij}/l$ 센터 i 와 센터 j 간의 링크에서의 비트당 지연 p_{ij} 패킷당 록업 지연, l 패킷의 길이 여기서 $R(X)$ 는 이후의 절에서 소개할 네트워크 신뢰도이며, w_{1ij} 는 센터 i 와 센터 j 간의 링크 가중치이고, w_{2ij} 는 센터 i 와 사용자 j 간의 링크 가중치이다. 또한, g_i 는 센터 i 에 연결할 수 있는 최대수이고 제공되는 총 통화량 Γ 는 다음과 같다.

$$\Gamma = \sum_{i=1}^n \sum_{j=1}^n t_{ij}$$

센터 k 의 총 통화량 $F_k(X)$ 는 다음과 같이 정의된다.

$$F_k(X) = \sum_{i=1}^n \sum_{j=1}^n t_{ij} \cdot a_{ij}^k(X), \quad k = 1, 2, \dots, n$$

링크 (k, l) 의 총 통화량 $f_{kl}(X)$ 는 또한 다음과 같이 정의된다.

$$f_{kl}(X) = \sum_{i=1}^n \sum_{j=1}^n t_{ij} \cdot b_{ij}^{(k,l)}(X), \quad k = 1, 2, \dots, n \quad l = 1, 2, \dots, n$$

3. Nash 유전 알고리즘

게임이론에 기반하여 경합하는 목적 함수들로 구성되는 다목적 최적화 문제를 위한 해법을 구하고자 하는 경우 게임을 진행하는 동안 게임에 참가하는 의사결정자들은 각자 자신들의 목적함수를 가지며 주어진 제약조건을 고려하여 더 이상 각자의 목적함수를 최적화할 수 없는 평형 상태 (Nash 균형)에 도달 할 때까지 의사결정 과정 (게임)을 진행하는 것을 Nash 게임이라고 한다. 이러한 Nash 게임을 풀기 위해 유전 알고리즘 (GA)를 적용한 것이 Nash GA라고 하고 이는 Nash 균형을 탐색하기 위해 제안된 내쉬 전략과 GA를 결합한 형태이다 [6,7].

3.1 유전자 표현 및 초기화

유전자 표현은 코딩 공간에서 문제의 후보해를 표현하기 위한 자료구조를 의미한다. 보통 다른 종류의 문제들은 다른 자료구조 또는 유전자 표현을 사용하게 된다. 그림 2에 표시된 것과 같이 본 논문에서는 n 개의 노드에 대하여 Prüfer 수(PN)를 기반하는 표현 방법은 $n-2$ 개의 메모리안으로도 트리를 표현 할 수 있기 때문에 서비스 센터들의 연결을 결정하는 의사 결정자를 위한 유전자 표현으로 이용하고 센터와 터미널(사용자)과의 연결을 결정하는 의사 결정자를 위해서는 클러스터 문자열로 유전자 표현법을 이용한다.

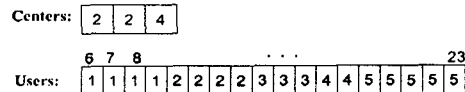


그림 2. 그림 1에 대한 유전자 표현

초기화 과정에서 서비스 센터는 영역 $[1, n]$ 에서 임의로 $n-2$ 개의 숫자(PN)로 생성하여 나타내고 사용자는 영역 $[1, n]$ 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

다. 본 논문에서의 유전자 표현은 $n+m-2$ 길이를 가진다.

3.2 평가 함수

본 논문에서는 다음과 같이 각각의 의사결정자를 위한 목적함수치로 염색체들의 적합도를 계산한다.

$$f_1(X^*) = \min f_1(X_1, X_2^*)$$

$$f_2(X^*) = \min f_2(X_1^*, X_2)$$

3.3 트리에 기초한 신뢰도 계산

네트워크의 신뢰도의 평가는 전체 트리를 부분 트리 및 하나의 노드로부터 가능하다[3]. 트리의 재귀관계를 기반으로 하는 신뢰도 계산을 위한 프로시저 기호들을 다음과 같이 정의하자: p_i^f 는 노드 i 의 노드 고장 확률이고 $p_i^o (= 1 - p_i^f)$ 는 노드의 동작확률이다. 그리고 링크 (i, j) 에도 유사하게 l_{ij}^f 와 l_{ij}^o 는 각각 링크가 고장날 확률과 동작할 확률을 의미한다. 또한 각 부분트리들을 위한 상태벡터들을 다음과 같이 정의한다: e_i 는 노드 i 의 부분 트리의 고장확률이며, o_i 는 루트노드인 노드 i 를 포함하여 부분트리의 동작 노드들 집합들이 서로 연결될 확률이고, r_i 는 부분트리의 루트노드는 고장나고 부분트리의 동작노드들 집합이 서로 연결될 확률이다.

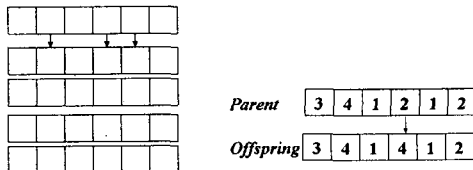
루트노드를 노드 1으로 하고 n 개의 노드를 가진 트리에 대해 모든 동작 노드들이 통신할 수 있는 네트워크 신뢰도 $R(X)$ 는 참고 문헌 [8]을 통해 살펴 볼 수 있다.

3.4 선택

본 논문에서는 토너먼트 선택을 이용한다. 토너먼트 선택은 n 개의 염색체를 임의로 선택하고 이들 중에서 가장 적응도가 높은 개체를 다음세대의 집단에 포함시킨다. 다음 세대의 염색체를 모두 얻을 때까지 이 과정을 반복한다. 이때 선택된 염색체를 현재 집단에서 제거하지 않는다. 즉, 한 염색체가 여러번 선택될 수 있다. 본 논문에서는 n 을 5로 설정하여 이용한다.

3.5 유전자 연산자들

균등교차 (uniform crossover)로도 불리는 다점 교차 (multi-point crossover)를 이용한다. 이러한 종류의 교차는 그림 3에서 보는 바와 같이 두개의 부모 염색체를 임의로 선택하여 하나의 부모염색체에서 임의로 유전자 표현형들을 선택하여 자손 염색체를 구성한다.



(a) 다점 교차 연산 (b) 단순 돌연변이 연산자
그림 3. 유전자 연산

또한, 본 논문에서의 돌연변이 연산은 하나의 염색체에서 임의로 선택된 한 개의 위치에 있는 유전자 표현형들을 변환하는 단순 돌연변이를 사용한다.

5. 실험 및 결론

4개의 서비스 센터($n=4$), 8개의 사용자($m=8$)를 가지고, g_i 는 3으로 설정되며, 센터 i 와 센터 j 간의 링크 가중치 (w_{1ij}) 및 센터 i 와 사용자 j 의 링크 가중치 (w_{2ij}), 그리고 사용자 통화량 행렬 U 는 참고 문헌 [8]에 의해 설정되고 센터 i 의 통화 용량(C_i)은 50으로 설정되는 예제이다. 또한, 센터들의 동작 확률은 0.95로 설정되며, 사용자의 동작 확률은 0.9로, 센터간의 링크는 0.9로, 센터와 사용자간의 링크는 0.85로 설정되고 R_{min} 는 0.9로 설정된다.

GA를 위한 파라미터들은 다음과 같이 설정하였다: 센터를 위한 의사결정자의 집단 크기는 20, 최대 세대수는 500, 센터 및 터미널 간을 위한 의사 결정자의 집단 크기는 20, 최대 세대 수는 500, 교차율은 0.2, 돌연변이율은 0.6으로 각각 설정하여 총 10번 실험하였다.

대부분의 경우 만족할 만한 결과를 구할 수 있었고 최선의 준 최적해, Nash 균형해를 구한 결과는 다음과 같다.

Cost: 528 Delay: 0.061042 Reliability: 0.9387
Centers: 0 2
Users: 1 3 0 2 0 1 2 3

좀더 다양한 문제에 대한 적용으로 보다 더 정밀한 분석을 통해 제안한 Nash 유전 알고리즘의 성능에 대한 제고가 필요할 것이다.

감사의 글

본 연구는 산업자원부의 지역혁신 인력양성사업의 연구결과로 수행되었음

참고문헌

- [1]Gen, M. and R. Cheng, Genetic Algorithms and Engineering Optimization, John Wiley & Sons, New York, 2000.
- [2]Bertsekas, D. and R. Gallager, Data Networks, 2nd ed., Prentice-Hall, New Jersey, 1992.
- [3]Kerhsenbaum, A., Telecommunication Network Design Algorithms, McGraw-Hill, Inc., New York, 1993.
- [4]Kerhsenbaum, A. and R. Van Slyke, "Recursive analysis of network reliability," Networks, Vol.3, pp.81-94, 1973.
- [5]Mukherjee, B., Optical Communication Networks, McGraw-Hill, Inc., New York, 1997.
- [6]J. F. Wang and J. Periaux, "Multi-Point Optimization using GAs and Nash/Stackelberg Games for High Lift Multi-airfoil Design in Aerodynamics," Proc. of the 2001 Congress On Evolutionary Computation, Vol.1, pp.552-559, 2001.
- [7]K. B. Sim, J. Y. Ki, and D. W. Lee, "Optimization of Multi-objective Function based on the Game Theory and Co-Evolutionary Algorithm," Journal of Fuzzy Logic and Intelligent Systems, Vol. 12, No. 6, pp.491-496, 2002 (in Korean).
- [8]J. R. Kim, J. U. Lee, and M. Gen, "An Algorithm based on Evolutionary Computation for Highly Reliable Network Design," Journal of Korea Information System Society, Vol.32, No.4, 2005.4 (in Korean).