유전자 알고리즘을 이용한 최단경로탐색

Development of the shortest path algorithm by genetic algorithm

김 형 진 (연세대학교 도시공학과 교수)
유 완 (연세대학교 도시공학과 교수)
양 인 철 (연세대학교 도시공학과 석사과정)

목 차

Ⅰ. 서 론
Ⅱ. 유전자 알고리즘
1. 유전자 알고리즘의 개요
2. 유전자 알고리즘의 기본 구조
Ⅲ. 유전자 알고리즘을 이용한 최단경로탐색 알고리즘
1. 네트워크의 표현
2. 임의 경로 발생 알고리즘
3. 선택, 돌연변이, 교배 알고리즘
Ⅳ. 알고리즘의 적용
1. 가상 네트워크
2. 실제 네트워크
Ⅴ. 결론 및 향후 연구과제

Ⅰ. 서론

네트워크의 링크별 통행시간, 통행거리 등의 저항요소(impedance)를 이용하여 존간 최단경로를 탐색하는 문제는 통행배정에 있어서 가장 먼저 이루어져야 할 과정이며 동시에 매우 중요한 문제이다. 또한 최근 교통분야에서 연구가 활발하게 이루어지고 있는 ITS 분야의 하나인 첨단 여행자 정보체계(ATIS)에 있어서 효율적인 경로안내 및 유도를 위한 신속·정확한 최단경로탐색 알고리즘의 필요성이 매우 높아지고 있다.

최단경로탐색 알고리즘은 1950년대 Ford(1956)와 Bellman(1957)의 알고리즘을 기반으로 하여 Moore(1957)와 Dijkstra(1957)에 의해 그 기틀이 다져졌으며, 그 이후 많은 휴리스틱한 알고리즘들이 제안되어 왔다. Bellman(1958)의 최적원리에 기초한 노드 표지기법(node labeling)의 대표적인 수형망 알고리즘(tree building algorithm)인 Dijkstra algorithm은 네트워크의 좌회전, U턴, P턴 등 회전에 따른 지체를 고려하지 못하는 단점이 있다. 수형망 알고리즘의 단점을 극복하기 위해 개발된 덩굴망 알고리즘(vine building algorithm)은 U턴, P턴을 고려하지 못하고 또한 교차로 좌회전금지가 2개 이상 연속하는 경우에 비현실적인 최적해를 도출하는 한계를 들어내고 있다. 이러한 기존 최단경로탐색 알고리즘의 문제점을 극복하기 위해 가로망 표현방식의 변화, 알고리즘 내 적용 자료의 구조적인 변형, 표지 방식의 변화 등의 연구가 이루어지고 있으나 대부분 각각의 노드를 기준으로 최단경로를 탐색하는 분할 탐색기법에 기반하고 있는 실정이다.

본 연구의 목적은 전역해 탐색기법의 하나인 유전자 알고리즘(genetic algorithm)을 이용하여 덩굴망 알고리즘의 한계로 지적된 연속 좌회전 금지 교차로 문제를 해결하고 적절한 U턴, P턴을 포함한 경로를 제시할 수 있는 최단경로탐색 알고리즘을 개발하는데 있다. 이는 분할 탐색기법 위주의 최단경로탐색 알고리즘 개발에 있어서 새로운 방향을 제시할 뿐만 아니라, 장래 첨단정보체계(ATIS)에서 보다 정확한 최단경로를 사용자에게 안내할 수 있다.

Ⅱ. 유전자 알고리즘

1. 유전자 알고리즘의 개요

인공 유전 시스템(artificial genetic system)의 하나인 유전자 알고리즘(genetic algorithm)은 자연세계의 진화 현상에 기초한 계산 모델로서 John Holland에 의해서 1975년에 개발된 전역적(global)인 최적화 알고리즘이다. 유전자 알고리즘은 자연 선택의 원리와 자연계의 생물 유전학에 기본 이론을 두며 병렬적이고 전역적인 탐색 알고리즘으로서, 모든 생물은 주어진 다양한 환경 속에 적응함으로써 살아남는다는 다윈의 적자생존(survival of the fittest)의 이론을

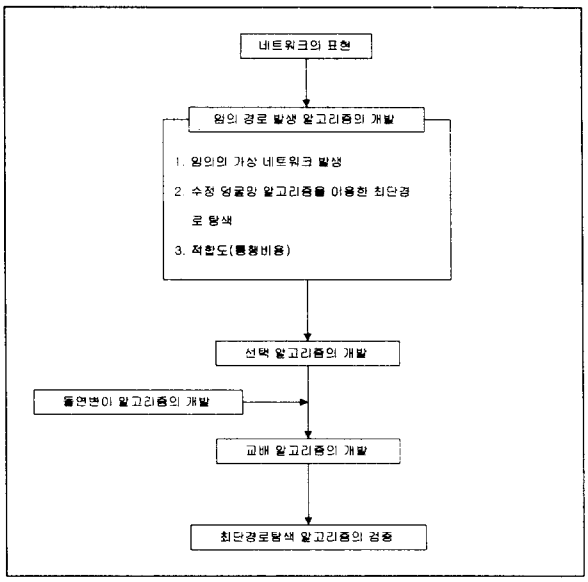
기본개념으로 한다.
유전자 알고리즘은 탐색공간에서 임의의 N개의 점들을 선택하여 개체집단(population)을 형성한다. 이 개체 집단에서 주어진 문제에 적합한 개체를 확률적으로 선택(reproduction)하여, 교배(crossover)와 돌연변이(mutation)를 통해 다음 세대(offsprings)를 만들고 그 개체 집단에 우리가 찾고자 하는 점이 있거나 또는 여러 번의 세대반복을 통해 미리 정해 놓은 세대수를 넘으면 탐색이 끝난다.

2. 유전자 알고리즘의 기본 구조

유전자 알고리즘은 자연계의 선택과 유전의 메카니즘을 기반으로 하는 확률적 탐색기법이다. 이는 기존의 탐색기법과는 다르게 개체집단이라는 임의의 해(solution) 집단에서부터 탐색과정이 시작된다. 개체집단을 구성하는 각 개체를 염색체(chromosome)라 하는데 이는 문제의 해를 나타낸다. 이러한 염색체는 계속적인 반복과정을 거쳐 진화를 하게 되는데 각 반복과정을 세대(generations)라고 한다. 각 세대마다 염색체의 적합도(fitness)를 평가(evaluate)하는데 적합도가 높은 염색체일수록 다음 세대의 개체집단에 선택될 확률이 높다. 다음세대는 현 세대의 두 염색체(parents)를 교배하거나 임의의 염색체를 돌연변이화함으로써 만들어진다. 새로운 세대(new generation)는 현 세대의 염색체(parents)와 교배, 돌연변이에 의해 새롭게 만들어진 염색체(offsprings)를 적합도에 따라 선택 또는 도태시킴으로써 형성되어 진다.

III. 유전자 알고리즘을 이용한 최단경로탐색 알고리즘

유전자 알고리즘을 이용한 최단경로탐색 알고리즘은 <그림 1>과 같은 과정에 의해 개발하였다. 우선 최단경로탐색 문제를 유전자 알고리즘을 이용하여 해결하기 위해 교통 네트워크를 유전자 형식으로 표현하였다. 링크를 표현하는 유전자 값의 임의 생성을 통해 가상 통행비용과 실제 통행비용으로 구성된 가상 네트워크를 발생시킨 후 수정 덩굴망 알고리즘으로 최단경로를 탐색하여 임의 경로를 발생시켰다. 각 임의 경로의 적합도는 수정 덩굴망 알고리즘을 통해 얻은 최단경로 통행비용을 이용하였다. 선택, 돌연변이, 교배 알고리즘 등의 경로 개선 알고리즘을 이용하여 최종적으로 최단경로 탐색 알고리즘을 개발하였다. 알고리즘의 효율성 및 적용가능성을 검토하기 위해 가상 네트워크와 실제 네트워크(서울특별시 강남구 일부지역)에 적용해 보았다.



<그림 1> 알고리즘 개발 과정

1. 네트워크의 표현

유전자 알고리즘은 개체의 특성을 결정하는 유전자(gene)와 이러한 유전자를 포함하는 염색체(chromosome)를 이용하여 선택(selection), 교배(crossover), 그리고 돌연변이(mutation) 과정을 여러 세대(generation)를 통해 반복하면서 최적의 해(solution)를 탐색하는 방법이다. 최적화 문제를 해결하기 위해 유전자 알고리즘을 이용할 때 가장 기본적으로 결정되어야 할 점은 문제(problem)를 유전자 알고리즘에 적절한 형식으로 표현하는 것이다.
본 연구에서는 최단경로탐색을 위해 네트워크를 유전자 알고리즘에서 이용하기 위한 형식으로 재구성하기 위해 노드와 노드간을 연결하는 링크(link)를 유전자로, 노드와 링크로 구성되는 네트워크(network)를 염색체로 표현하였다. 그리고 알고리즘의 계산에 있어서 축차계산(iteration)의 횟수를 세대로 표현하였다.

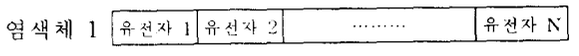
1) 유전자 - 링크
링크를 표현하는 유전자의 기본형은 다음과 같이 표현할 수 있다.

$X_{(세대)}[염색체번호][링크번호]$

유전자 '$X_3[12][15]$' 는 제 3세대 12번째 염색체의 15번 유전자 값을 나타낸다. 다시 말하자면 3번째 계산시 12번째 가상 네트워크의 15번 링크의 유전자 값를 의미한다. 일반적인 2 진 표현 유전자 알고리즘에서 유전자는 0 또는 1 의 값을 갖는다. 이러한 유전자 값을 임의로 생성하기 위해 초기에 1~99 사이의 임의의 수를 발생시킨 후 50 보다 작은 수의 값을 갖는 유전자는 0 의 값을 부여하고 50 보다 큰 수인 경우 1 의 값을 부여하였다. 어떤 링크의 유전자 값이 0 일 경우 그 링크의 통행비용은 고유의 통행비용을 가상의 통행비용(50)으로 대체하였고, 반대로 유전자 값이 1 일 경우 그 링크는 고유의 통행비용을 갖게 하였다. 임의의 링크 통행비용을 가상의 통행비용으로 대체하는 이유는 다음 절에 설명하게 될 임의 경로 발생 알고리즘에서 수정 덩굴망 알고리즘을 이용하여 최단경로를 탐색할 때 임의 링크로의 통행을 어렵게 하여 다양한 경로를 발생시킬 수 있기 때문이다.

2) 염색체 - 네트워크

N개의 링크로 구성된 네트워크를 표현하는 염색체의 기본 형식은 다음과 같다.

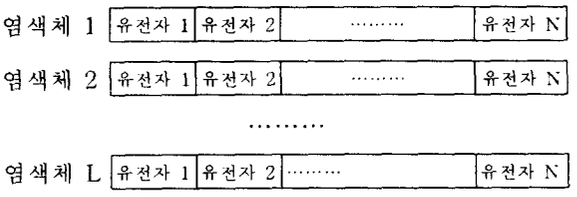


<그림 2> 염색체

염색체는 다양한 형태의 유전자로 구성되어 일정한 형태의 특성을 갖는 개체를 표현한다. 본 연구에서 임의로 발생된 N개의 유전자 값으로 구성된 하나의 염색체는 하나의 가상 네트워크를 표현한다.

3) 개체집단

개체집단은 유전자 알고리즘에서 다양한 형태의 해(solution)를 탐색할 수 있는 공간을 제공하는 것으로 다수의 염색체로 구성된다. 하나의 염색체가 하나의 가상 네트워크를 표현하기 때문에 개체집단은 다수의 가상 네트워크를 표현하게 된다. 유전자 알고리즘은 이러한 개체집단 내의 염색체를 이용하여 선택, 교배, 돌연변이를 시행함으로써 문제의 최적해를 탐색하게 된다. L 개의 염색체로 구성된 개체집단의 기본형식은 다음과 같다.



<그림 3> 개체집단

2. 임의 경로 발생 알고리즘

유전자 알고리즘을 이용한 최단경로탐색 알고리즘에서 가장 중요하게 다루어져야 할 부분이 임의 경로(random path) 발생 알고리즘이다. 기존에 유전자 알고리즘에서 많이 다루어지던 최적경로 문제인 TSP(Travelling Salesman Problem)는 주어진 네트워크 상의 모든 노드를 연결하는 경로를 찾는 문제이기 때문에 노드집합의 순열을 정하는 문제였다. 하지만 최단경로 탐색 문제는 네트워크 상의 임의 개수만큼의 노드를 거쳐 출발노드와 도착노드를 연결하는 하나의 연결된 경로를 찾는 문제이다. 따라서 임의 경로는 반드시 출발노드와 도착노드를 연결하는 경로가 되어야만 하고, U턴과 P턴을 표현할 수 있어야 하며, U턴과 P턴을 제외한 순환경로가 발생하지 않아야 한다. 또한 유전자 알고리즘 상에서 지역해(local solution)의 도출을 방지하기 위해 넓은 탐색공간을 제공할 수 있도록 다양한 형태의 경로를 표현할 수 있어야 한다.

1) 가상 네트워크 발생

가상 네트워크는 임의 발생된 유전자 값에 따라 링크의 통행비용을 변화시켜 구축하였다. 염색체 내의 각 유전자 별로 1~99 사이의 값을 발생시킨 후 50이상인 경우 유전자 값을 1 로, 50이하인 경우 0 으로 하였다. 유전자 값이 1 인 링크의 통행비용은 고유의 통행비용을 그대로 이용하고, 유전자 값이 0 인 링크의 통행비용은 가상의 통행비용(50)을 이용하였다.

2) 임의 경로 발생

새롭게 구축된 가상 네트워크 상에서 최단경로 탐색을 실시하여 임의 경로를 발생시켰다. 각각의 가상 네트워크는 링크별로 고유의 비용 또

는 가상비용을 다르게 가지기 때문에 같은 방법의 최단경로탐색 알고리즘을 이용하여 최단경로를 탐색하면 가상 네트워크별로 다양한 임의 경로를 발생시킬 수 있다. 본 연구에서는 덩굴망 알고리즘을 일부 수정하여 U턴과 P턴을 효과적으로 표현할 수 있도록 만든 수정 덩굴망 알고리즘을 이용하였다.

기존 덩굴망 알고리즘을 이용하여 임의 경로를 발생시킬 경우 U턴과 P턴을 적절히 표현할 수 없기 때문에 일부 과정을 추가 또는 수정하여 문제를 해결하였다. 덩굴망 알고리즘의 수행과정와 추가·수정된 과정은 다음과 같다. (추가·수정된 과정은 굵은 이탤릭체임.)

h : 출발노드
d : 도착노드
m_j : 출발노드 h 에서 노드 j 까지의 통행비용
m_j' : 출발노드 h 에서 노드 j 까지의 임시 통행비용
c_{ij} : 노드 i 에서 노드 j 까지의 통행비용
$NEXT$: 탐색의 기준이 되는 노드
j : 노드 $NEXT$ 와 하나의 링크로 연결된 노드
k : 노드 j 와 하나의 링크로 연결된 노드
L : 탐색 기준노드 $NEXT$ 가 될 노드의 집합
R : 모든 노드에 대한 정보(비용, 전노드, 전전노드)의 집합
p_k : 노드 k 의 전(前)노드
pp_k : 노드 k 의 전전(前前)노드
P_{NEXTjk} : 시작노드 $NEXT$ 에서 회전노드 j 를 거쳐 끝 노드 k 에 이르는 차량회전의 회전 벌점(turn penalty)

[단계 1] (Initialization)

(a) 집합 R 에서 모든 노드 i 의 비용, 전노드 및 전전노드의 노드표지를 초기화한다.
$m_i = \infty$, $p_i = 0$, $pp_i = 0$

(b) 집합 R에서 출발 노드 h 와 하나의 링크로 연결된 모든 노드 j 에 대해 다음과 같이 초기화한다.
$m_j = c_{hj}$, $p_j = h$, $pp_j = 0$

(c) 출발노드 h와 노드 j 를 집합 L 에 포함시킨다.

(d) $NEXT$ = h 로 놓는다.

[단계 2]

(a) 노드 $NEXT$ 를 집합 L 의 목록에서 제거한다.

(b) 링크 ($NEXT$, j)에 대해 다음의 과정을 시행한다. 만약 링크 ($NEXT$, j)에 해당하는 링크가 없을 경우 [단계 3]으로 간다.

① $m_j' = m_{NEXT} + c_{NEXTj} + P_{rNEXTj}$ 를 계산한다. (여기서 r = p_{NEXT})

(i) 만약 $m_j' \le m_j$ 이면 집합 R 에서 pp_j = -1, p_j = $NEXT$, m_j = m_j' 로 노드 j의 정보를 바꾼 후 노드 j 를 집합 L 에 추가한다. $m_j' \ge m_j$ 이면 아무 것도 하지 않는다.

② 링크 (j, k)에 대해 다음의 과정을 시행한다. 만약 링크 (j, k)에 해당하는 링크가 없을 경우 (a) 로 간다.

(i) $m_k'' = m_j' + c_{jk} + P_{NEXTjk}$ 를 계산한다.

(ii) 만약 $m_k'' \le m_k$ 이면 집합 R 에서 pp_k, = $NEXT$, p_k = j, m_k = m_k'' 로 노드 k 의 정보를 바꾼 후 k 를 집합 L 에 추가한다. $m_k'' \ge m_k$ 이면 아무 것도 하지 않는다.

(c) [단계 3] 으로 간다.

[단계 3]

(a) 집합 L 로부터 다음 탐색 기준노드 $NEXT$ 를 선택한 후 [단계 2]로 간다. 선택할 다음 탐색 기준노드 $NEXT$ 가 없으면 [단계 4] 로 간다.

[단계 4] (최단경로와 통행비용의 출력)

(a) 집합 R 에서 r = d 로 하고 d 를 집합 P_d 에 저장한다.

(b) 만약 pp_r = 0 or -1 이면 p_r 을 P_d 에 저장하고 r = p_r 로 하고 (c) 로 간다. 만약 pp_r ≠ 0 or -1 이면 p_r 과 pp_r 을 P_d 에 저장한 후 (c) 로 간다.

(c) r = h 이면 P_d 를 출력한 후 [단계 5]로 간다. r ≠ h 이면 (a) 로 간다.

[단계 5] (종료)

기존 덩굴망 알고리즘에 추가된 과정은 [단계 2] 부분의 굵은 이탤릭체 부분이고 수정된 과정은 [단계 4] 부분으로, 수정된 덩굴망 알고리즘을 이용하여 최단경로를 탐색한 결과 국부적으

로 U턴 및 P턴을 적절히 표현할 수 있었다.

3) 적합도(통행비용)

개체집단 내 L 개의 염색체가 표현하는 각각의 가상 네트워크 상에서 수정 덩굴망 알고리즘을 이용하여 최단경로를 탐색한 결과 발생된 임의 경로의 적합도(통행비용)는 수정 덩굴망 알고리즘에서 산출된 최단경로 통행비용을 이용하였다.

3. 선택, 돌연변이, 교배 알고리즘

유전자 알고리즘을 이용한 최단경로의 탐색은 세대가 넘어감에 따라 많은 변화가 이루어지지 않으면 자칫 지역해(local solution)에 도달할 확률이 매우 높다. 단순한 교배와 돌연변이, 선택 연산자를 이용할 경우 그에 따라 발생하는 임의 경로가 초기의 임의 경로에서 크게 벗어나지 못하기 때문에 다양한 경로(넓은 탐색공간)에 대한 검토가 이루어지지 못해 지역해가 도출될 위험이 높다.

본 연구에서는 보다 넓은 범위의 탐색을 가능케 하여 지역해 도출의 위험율을 최대한 낮추는 것에 초점을 맞추었다. 따라서 최대한 다양한 경로의 검토를 위해 유전자 알고리즘에서 일반적으로 사용되어 오던 교배 후 돌연변이 발생의 순을 바꾸어 돌연변이를 실시한 후 돌연변이가 발생한 염색체를 포함한 개체집단 내에서 교배를 실시하였다.

개체집단 내 염색체를 임의 경로의 적합도(통행비용)에 따라 랭킹(ranking)한 후 하위 50%를 도태시키고 상위 30%는 보존, 나머지 20%는 돌연변이를 실시하였다. 보존된 염색체와 돌연변이된 염색체를 교배하여 50%의 염색체를 생성시켜 다음 세대를 구성하였다.(돌연변이율 0.2, 교배율 0.5)

1) 돌연변이 알고리즘

① 돌연변이율에 따라 임의의 염색체를 선택한다.
② [1, 유전자수(링크수)] 사이의 임의의 정수를 전체 유전자수의 20%만큼 발생시킨다.
③ 임의의 정수에 해당하는 유전자의 값을 반전한다. 즉, 유전자의 값이 1 일 경우 0 으로, 0 일 경우 1 로 바꾼다.

2) 교배 알고리즘

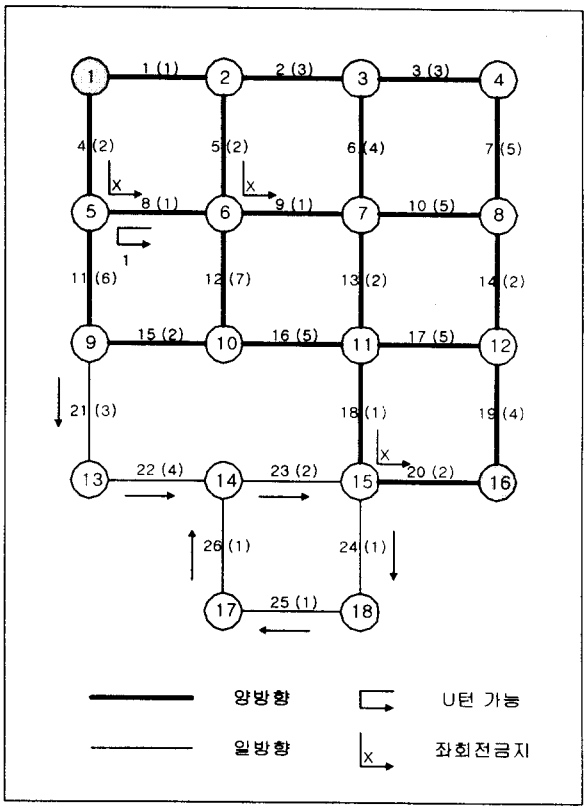
① 교배율에 따라 임의의 염색체 2개를 선택한다.
② [1, 유전자수(링크수)] 사이의 임의의 정수를 발생시킨다.
③ 임의의 정수에 해당하는 유전자부터 마지막 유전자를 서로 교환한다.

IV. 알고리즘의 적용

본 연구에서 개발된 알고리즘의 실제 적용성 및 효율성을 검토하기 위해 가상 네트워크와 실제 네트워크(서울특별시 강남구 일부지역)에 적용하여 보았다.

1. 가상 네트워크

가상 네트워크는 <그림 4>와 같이 연속 좌회전 금지 교차로와 U턴, P턴을 모두 고려할 수 있도록 구성하였다.



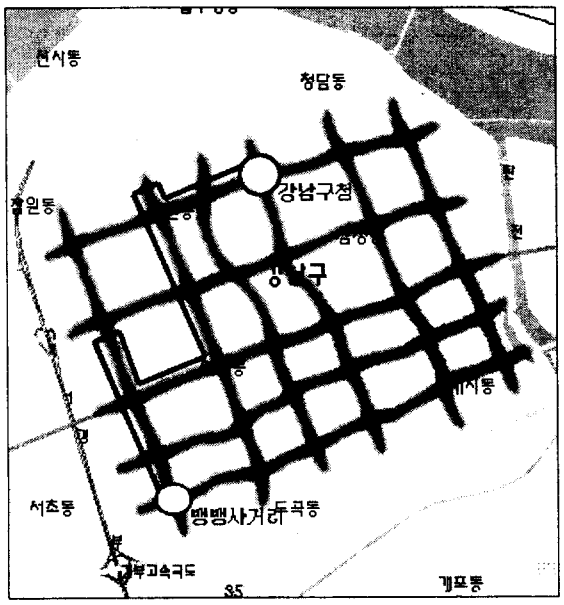
<그림 4> 가상 네트워크

개체집단의 크기는 60개, 최대 세대수는 50세대

로 하여 출발노드 1번부터 도착노드 16번까지의 최단경로를 탐색한 결과 최단경로는 1 - 2 - 6 - 5 - 6 - 7 - 11 - 15 - 18 - 17 - 14 - 15 - 16 이고 최단경로 통행시간은 17 로 나타났다. 이는 U턴, P턴을 전혀 고려하지 못한 덩굴망 알고리즘을 이용한 최단경로(1 - 2 - 3 - 4 - 8 - 12 - 16, 최단경로 통행시간 18)에 비해 매우 효율적인 최단경로인 것으로 분석되었다.

2. 실제 네트워크

알고리즘의 현실 적용성을 검토하기 위해 서울특별시 강남구 일부지역을 대상으로 알고리즘을 적용하여 보았다. 대상 네트워크의 노드수는 52개, 링크수는 71개이고, 개체집단의 크기와 최대 세대수는 가상 네트워크와 동일하게 설정하였다. 강남구청에서 뱅뱅사거리까지의 최단경로를 탐색한 결과는 다음 <그림 5>와 같다.



<그림 5> 실제 네트워크의 최단경로탐색 결과

<그림 5>에서 볼 수 있듯이 유전자 알고리즘을 이용한 최단경로탐색 알고리즘은 강남 일부지역의 실제 네트워크 상에서 U턴을 효과적으로 이용하여 정확한 최단경로를 탐색할 수 있었다.

V. 결론 및 향후 연구과제

본 연구에서는 전역적 탐색기법의 하나인 유전자 알고리즘을 이용하여 덩굴망 알고리즘의 한계로 지적된 연속 좌회전 금지 교차로 문제를 해결하고 적절한 U턴, P턴을 포함한 경로를 제시할 수 있는 최단경로탐색 알고리즘을 제시하였다. 임의의 유전자 값에 의해 가상 네트워크를 구축한 후 일부 과정을 추가·수정한 수정 덩굴망 알고리즘을 이용하여 가상 네트워크 상의 최단경로를 탐색하여 임의 경로를 발생시키는 알고리즘을 제시하였다. 효율적인 최단경로 탐색을 위해 다양한 임의 경로를 발생시켜 최단 경로를 개선할 수 있는 선택 알고리즘, 돌연변이 알고리즘, 교배 알고리즘을 제시하였다. 제시된 최단경로탐색 알고리즘을 가상 네트워크와 서울특별시 강남구 일부지역에 적용한 결과 연속 좌회전 금지 교차로가 존재하는 네트워크, 또는 U턴, P턴이 필요한 네트워크 상에서 정확한 최단경로를 탐색할 수 있었다.

향후 연구과제로 덩굴망 알고리즘에 비해 탐색시간이 상대적으로 긴 단점을 보완할 수 있는 연구가 수행되어야 할 것으로 판단된다. 또한 본 연구에서 제시한 임의 경로 발생 알고리즘은 기존의 최단경로탐색 알고리즘인 덩굴망 알고리즘을 이용하였기 때문에 임의경로 발생에 한계를 나타내고 있다. 따라서 다양한 형태의 임의 경로를 발생할 수 있는 알고리즘의 개발이 이루어져야 할 것으로 판단된다.

참고 문헌

1. 김현명, 임용택(1999), 유전 알고리즘을 이용한 전역탐색 최단경로 알고리즘 개발, 대한교통학회지, 제17권 제2호, pp. 163~178
2. 김익기(1998), ATIS를 위한 수정형 덩굴망 최단경로 탐색 알고리즘의 개발, 대한교통학회지, 제16권 제2호, pp. 157~167
3. Gen M. and Cheng R.(1997), Genetic Algorithms and Engineering Design, John Wiley & Sons
4. Thomas. R.(1991), Traffic Assignment Techniques, Avebury Technical
5. Zbigniew Michalewicz(1994), Genetic Algorithm + Data Structures = Evolution Programs, Springer-Verlag
6. Soojung Jung, Ali Haghani(1998), A Genetic Algorithm for the Vehicle Routing Problem with Time-Dependent Travel Times, Proceeding of 5th World congress on ITS