최적의 TSP문제 해결을 위한 유전자 알고리즘의 새로운 집단 초기화 및 순차변환 기법

강래구* · 임희경* · 정채영**

New Population initialization and sequential transformation methods of Genetic Algorithms for solving optimal TSP problem

Rae-Goo Kang* · Hee-Kyoung Lim* · Chai-Yeoung Jung**

이 논문은 2005년도 조선대학교 학술 연구비의 지원을 받아 연구되었음

요 약

TSP(Traveling Salesman Problem)는 N개의 도시가 주어질 때 어떠한 임의의 도시에서 출발하여 모든 도시를 단 한번만 방문하여 다시 출발지로 되돌아오는 여러 경로들 중 가장 짧은 거리를 구하는 문제이다. 방문 도시수가 증가함에 따라 계산량이 기하급수적으로 증가하게 되는 문제로 인해 NP-Hard문제로 분류되며 유전자 알고리즘이 대표적으로 이용된다. TSP문제에 있어서 보다 우수한 결과를 얻기 위해 현재까지 다양한 연산자들이 개발되고 연구되어왔다. 본 논문에서는 새로운 집단 초기화 방법과 순차변환 방법을 제안하여 기존의 방법들과 비교를 통해 성능 향상을 입증하였다.

ABSTRACT

TSP(Traveling Salesman Problem) is a problem finding out the shortest distance out of many courses where given cities of the number of N, one starts a certain city and turns back to a starting city, visiting every city only once. As the number of cities having visited increases, the calculation rate increases geometrically. This problem makes TSP classified in NP-Hard Problem and genetic algorithm is used representatively. To obtain a better result in TSP, various operators have been developed and studied. This paper suggests new method of population initialization and of sequential transformation, and then proves the improvement of capability by comparing them with existing methods.

키워드

유전자 알고리즘, 집단 초기화, TSP, Traveling Salesman Problem

* 조선대학교 전산통계학과

** 교신저자

접수일자 : 2005. 10. 11

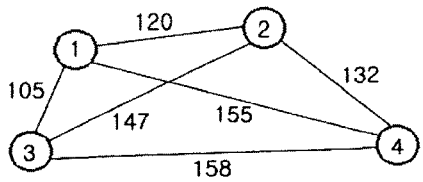
Ⅰ. 서 론

TSP(Traveling Salesman Problem : 순회판매원 문제)는 N개의 도시와 도시 사이의 거리가 주어질 때, 어떤 도시에서 시작하여 모든 도시를 단 한번만 방문하고 원래의 출발지로 되돌아오는 최단거리를 찾는 문제이다. TSP의 탐색공간은 모든 여행의 집합 $\{T_1, T_2, \ldots\ldots, T_{n!}\}$이 되고 그 크기는 총 $N!$이며 이 중 순회 거리가 가장 짧은 것이 해가 된다.

TSP는 네트워크 최적화, 공장에서 처리 순서 결정의 문제, 택배운송경로 문제 등 광범위한 분야에 적용 될 수 있다.[1] 하지만 도시의 수가 증가하게 되면 최적해의 계산량이 기하급수적으로 증가하기 때문에 NP(Nondeterministic Polynomial)-hard 문제로 분류된다.

이러한 문제로 인해 TSP문제에 있어 가장 최적의 해를 계산하기 위하여 John Holland에 의해 제안된 유전자 알고리즘(Genetic Algorithms : GA)을 대표적으로 이용한다.[2][3]

본 논문에서는 각 종 유전자 알고리즘을 적용하기 위해 선행되는 초기화방법에 있어 무작위 초기화법과 유도된 초기화법을 동시에 이용하는 새로운 혼합 초기화법을 유추해 냈고, 선택연산자를 적용할 때 미리 순차변환 방법을 통해 국소 최적해에 가장 근접한 값을 구할 수 있도록 실험하였다.



T_1 = 1-2-3-4 Dis_1 = 580
T_2 = 1-2-4-3 Dis_2 = 515
T_3 = 1-3-2-4 Dis_4 = 539
⋮
T_{22} = 4-2-3-1 Dis_{22} = 539
T_{23} = 4-3-1-2 Dis_{23} = 515
T_{24} = 4-3-2-1 Dis_{24} = 580

그림 1. 4개 도시들 간의 거리
Fig 1. Distance between 4 cities

그림 1은 4개의 도시들 간의 구할 수 있는 총 경로를 나타낸 계산식이다.($4! = T_1, \ldots, T_{24}$)

Ⅱ. 실험에 사용된 유전자 알고리즘

표 1은 본 논문에서 제안한 방법의 성능을 증명하기 위해 사용한 연산자이다.

표 1. 실험에 사용한 연산자
Table 1. Operator used in experiment

선택(Selection)	Roulette wheel
	Stochastic universal
교배(Crossover)	PMX
	Edge Recombination
	CX
	OX
돌연변이(Mutation)	Inversion
	Swapping mutation

2.1 선택연산자

지금까지 다양한 선택 연산자가 발표되었는데, 공통된 원칙은 우수한 해가 선택될 확률이 높아야 한다는 것이다. 본 논문에서는 Roulette wheel selection과 Stochastic universal sampling 연산자를 이용하였다.

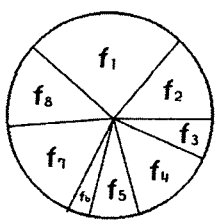


그림 2. Roulette wheel selection
Fig 2. Roulette wheel selection

Roulette wheel selection은 가장 대표적인 선택 연산자로 각 개체 i의 적합도(적합도는 0이 아니어야한다.) $f(i), i = 1, 2, 3, \ldots, N$의 총합을 구해 각 개체 i의 선택확률을 구했다. 이때 적합도가 큰 개체 일수록 선택되는 확률은 높아진다.

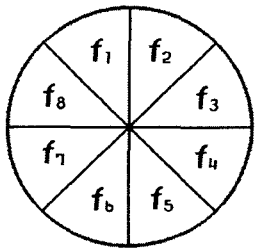


그림 3. Stochastic universal sampling
Fig 3. Stochastic universal sampling

Stochastic universal sampling은 모든 개체를 동일한 확률로 선택 가능하도록 하고 있다.[3]

2.2 교배연산자

교배연산자는 GA 중 가장 다양하고 가장 대표적인 연산자로써 본 논문에서는 PMX, OX, CX, Edge Recombination(ER)을 사용했다.[4] PMX, OX, CX 연산자와는 달리 ER 연산자는 Grefensetette[5]에 의해 처음 소개된 휴리스틱 교차 연산자의 일종으로 부모세대의 유전자만으로 유전자를 결정하던 기존의 연산 방식과 달리 부모세대 이웃에 있는 에지정보를 이용하는 기법으로 Starkweather[6]에 의해 소개되었다.

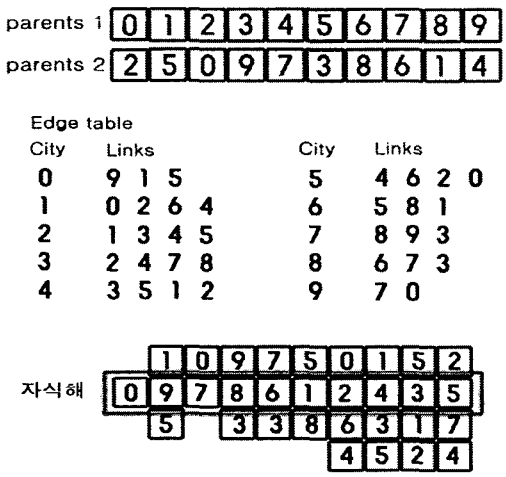


그림 4. Edge Recombination
Fig 4. Edge Recombination

그림 4는 ER 연산자의 예를 나타낸 것이다. 먼저 무작위로 하나의 도시를 선택한 다음(그림 4에서는 0번 도시를 선택하였다.) 선택한 도시 중 연결 가능한 에지수가 제일 적은 도시를 고른다. 만약 에지수가 같은 경우에는 무작위로 하나를 선택한다. 이 과정을 마지막까지 수행하여 새로운 자식해를 구한다.

2.3 돌연변이 연산자

재생산과 교배를 통해 해당 집단을 더욱 강하고 서로 닮아가게 만든다. 그러나 세대가 거듭해 갈수록 유전인자의 다양성은 없어진다. 이러한 단점을 보완하기 위해 돌연변이 연산자를 이용한다. 즉, 돌연변이 연산을 적용할 때 초기 세대부터 특정비트가 고정되는 것을 방지함으로써 새로운 집단을 만들 수 있다. 본 논문에서는 돌연변이 연산자로 Swapping mutation과 Inversion을 이용하였다.[1]

Ⅲ. 제안한 방법

본 논문에서는 우수한 TSP결과 값을 얻기 위하여 혼합 초기화법과 순차변환 방법을 제안하였다.

3.1 혼합 초기화법

각종 연산자를 통해 국소 최적해에 가장 근접한 결과 값을 얻기 위해서는 무엇보다도 최초에 집단을 초기화 시키는 방법이 중요하다.

집단 초기화 방법에는 크게 두 가지 방법이 있다. 하나는 집단 생성에 있어서 특별한 규칙 없이 무작위로 추출해서 집단을 생성하는 무작위 초기화법이 있고 다른 하나는 주어진 값을 통해 어떠한 사전 지식이나 정보를 바탕으로 일관성을 유지하여 집단을 생성하는 유도된 초기화법이 있다.

지금까지 TSP문제에서 사용된 집단 초기화 방법으로는 무작위 초기화법이 많이 사용되어 왔다. 본 논문에서는 난수발생기를 이용한 무작위 초기화법과 사전 정보나 경험을 바탕으로 한 유도된 초기화법을 동시에 사용하는 혼합 초기화법을 제안한다.

그림 5와 같이 먼저 난수발생기를 이용하여 시작 도시를 선택한 다음 각 도시간의 사전에 이미 알고 있는 거리 정보를 가지고 선택된 도시를 기준으로 거리가 가장 짧은 도시부터 순서대로 나열한다. 이러한 순서로 N개의 모든 도시를 나열하면 총 $N \times N$ 개의 행렬이 나오게 된다. 이렇게 생성된 행렬을 최초 집단으로 사용하였다.

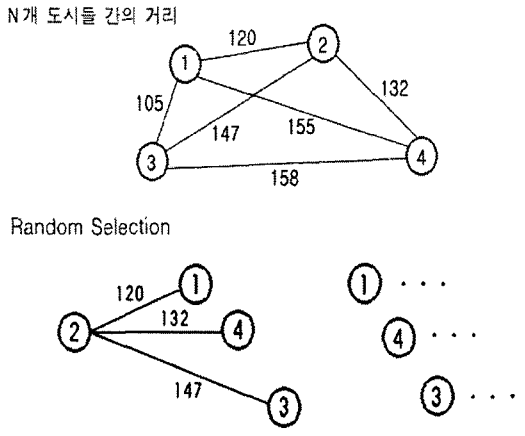


그림 5. 혼합 초기화법
Fig 5. Mixture initialization method

3.2 순차변환 방법

각 세대별 선택 연산이 진행되기 전에 전 세대에서 생성된 집단의 해를 순차적으로 변환 시킨 후 선택 연산을 적용 시키는 순차변환 방법을 제안한다.

< 1 세대 진행 결과 >

T_1 = 2-1-4-3 Dis_1 = 580
T_2 = 1-3-2-4 Dis_2 = 539
T_3 = 3-1-2-4 Dis_3 = 515
T_4 = 4-2-1-3 Dis_4 = 515

⇩

< 순차변환 적용 >

T_3 = 3-1-2-4 Dis_3 = 515
T_4 = 4-2-1-3 Dis_4 = 515
T_2 = 1-3-2-4 Dis_2 = 539
T_1 = 2-1-4-3 Dis_1 = 580

⇩

< 2 세대 진행 >

그림 6. 순차변환 진행 과정
Fig 6. Sequential transformation progress process

그림 6은 순차변환 과정을 설명한 것으로 선택 연산이 진행 될 때 전 세대의 해를 가지고 가장 우수한 해를 만들어낸 개체부터 순차적으로 재배열함으로써 다음 세대에서 사용될 개체의 집단을 생성 시 우성인자끼리 선택 될 확률을 높여준다.

Ⅳ. 실험 결과

본 논문에서 제안한 혼합 초기화법과 순차변환 방법의 결과 분석을 위해 2개의 선택연산자, 4개의 교배연산자, 2개의 돌연변이연산자를 사용하였고, 엘리트전략을 사용하여 2개의 우성인자를 보존하였다.

총 도시 수는 100개, 집단의 크기는 1000으로 한정 하였으며 총 세대 수는 제한을 두지 않고 100세대 동안 해의 값에 변화가 없을 경우 자동으로 종료되도록 설계 하였다. 그리고 교배 확률(P_c)과 돌연변이 확률(P_m)은 0.7과 0.1을 사용하였다.

실험은 P-4 2.4GHz에서 PowerBuilder8.0을 이용하여 구현하였으며 Oracle 9i를 이용하여 데이터 저장 및 분석을 하였다.

그림 7. 실험에 이용한 어플리케이션
Fig 7. Application used in experiment

그림 7은 본 논문에서 제안한 두 가지 방법의 성능을 입증하기 위해 PowerBuilder8.0으로 개발한 TSP 프로그램이며 가장 우수한 결과 값을 얻었을 때의 화면이다.

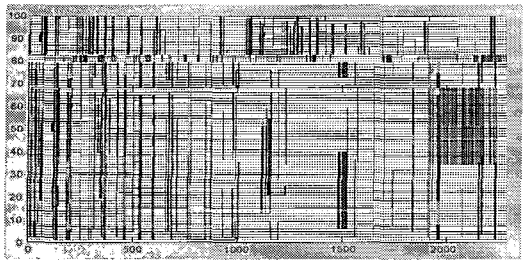


그림 8. 각 도시 방문 순서의 변화 과정
Fig 8. Change process of visit order of each city

그림 8은 세대별 도시 방문 경로를 나타내는 것으로써 1세대부터 마지막 세대까지 해당 세대에서 가장 우수한 값을 나타낸 개체의 경로가 변화 하는 모습을 보여주고 있다. 즉, 본 논문에서 제안한 방법을 바탕으로 유전자 알고리즘을 통해 세대가 거듭해 갈수록 최단거리를 찾아가는 과정을 나타내고 있다.

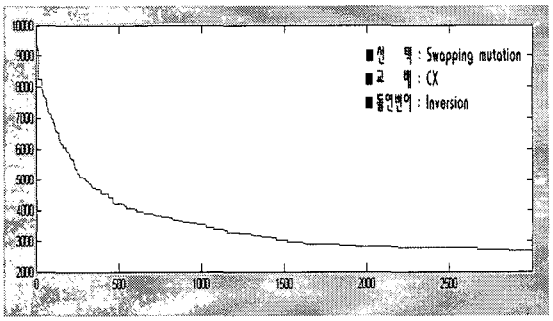


그림 9. 무작위 초기화법을 이용한 최적해 그래프
Fig 9. Optimization value graph of using random initialization method

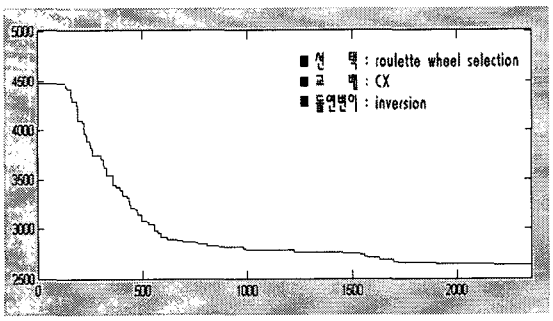


그림 10. 제안한 방법에 의한 최적해 그래프
Fig 10 Optimization value graph of using newly-proposal method

표 2. 실험 결과 비교
Table 2. experiment result comparison

선택	교배	돌연변이				제안 방식		
			Min	Avg.	gen	Min	Avg.	gen
RW	PMX	Inv	2746	3813	2696	**2660**	2827	2517
		Swa	3728	3958.2	2829	**3414**	3585.3	2734
	ER	Inv	2989	4913.2	2870	**2879**	4013.5	2633
		Swa	3326	3544	2369	**3282**	3512.1	2178
	CX	Inv	2701	2817.2	2890	**2637**	2754.4	2370
		Swa	3739	3939	2996	**3520**	3669	2466
	OX	Inv	3210	5892.7	2764	**3129**	6387.9	2997
		Swa	3808	6315.6	2938	**3599**	6673.2	2984
SU	PMX	Inv	3056	4589.7	2670	**2888**	4896.2	2964
		Swa	3051	3246.9	2444	**3040**	3240.4	2349
	ER	Inv	2988	4282.5	2871	**2677**	4009.7	2572
		Swa	3728	3998.2	2829	**3550**	3759.7	2876
	CX	Inv	2678	2831.3	2246	**2647**	2997.6	2932
		Swa	3461	3672.3	2587	**3414**	3593.2	2852
	OX	Inv	3461	6370	2901	**3179**	6111.1	2701
		Swa	4075	6971.2	2923	**3713**	6662.2	2958

* 단위 : Km * RW : Roulette wheel selection
* SU : Stochastic universal sampling
* Inv : Inversion * Swa : Swapping mutation

표 2에서 보는 바와 같이 기존에 주로 사용되던 무작위 초기화법을 적용했을 때 최단 거리는 2678Km 이였으나 본 논문에서 제안한 혼합 초기화법과 순차변환 방법을 적용 했을 때의 최단거리는 2637Km 라는 더 좋은 결과를 얻었으며 모든 결과에서 제안한 방법에 의한 거리가 항상 짧게 나타나는걸 알 수 있다. 또한 결과 값 중 최고 개선률은 10.4%였으며 최저 개선률은 0.3%로써 평균 4.6%의 개선률을 나타냈다. 그림 9과 그림 10은 무작위 초기화법과 본 논문에서 제안한 방법을 적용한 실험 중 각각 가장 우수한 결과를 나타내는 그래프이다. 그래프에서 보는 바와 같이 제안한 방법에 의한 결과 값은 최종값 뿐만 아니라 제 1세대부터 우수한 값을 생산해 내고 있음을 확인 할 수 있다.

실험 결과를 통해 알 수 있듯이 최초에 집단을 초기화 할 때 주어진 거리에 대한 정보를 무시한 무작위 초기화 법 보단 해당 거리를 활용하여 어떠한 특징을 찾아 이용하는 유도된 초기화법이 보다 효과적임을 알 수 있고 또한, 순차변환을 통해 우성인자들을 일괄적으로 재배열함으로써 선택연산자를 적용 시 우성인자끼리 선택될 확률을 높일 수 있다.

V. 결 론

본 논문에서는 보다 효과적으로 TSP문제를 해결하기 위해 두 가지 방법을 제안하였다. 하나는 무작위 초기화 법과 유도된 초기화법을 동시에 사용하는 혼합 초기화법이고 다른 하나는 각 세대에서 개체 집단을 재배열하여 우성인자끼리 선택 될 확률을 높여주는 순차변환 방법이다. 이 두 가지 방법을 적용하여 실험한 결과 제안한 방법이 평균 4.6% 정도의 개선된 결과를 나타내었고 또한, 기존의 방법 보다 제 1세대부터 우수한 값을 생산해내고 있음을 확인하였다.

즉, 사전에 미리 알고 있는 각 도시간의 거리정보를 활용하는 방법과 개체를 재배열함으로써 우성인자끼리 선택될 확률을 높여주는 방법이 성능 개선에 많은 도움이 된다는 걸 알 수 있다.

본 논문에서 제안한 방법은 기존에 발표된 GA만을 적용시켜 실험했기 때문에 차후 새로운 알고리즘에 대한 적용과 대규모 도시 또는 좀 더 복잡한 구조를 가진 경우 최적해를 구하기 위해서는 또 다른 연구가 지속적으로 진행되어야 할 것이다.

참고문헌

[1] 진강규. 유전알고리즘과 그 응용. 교우사. 2000

[2] D. Goldberg. Genetic Algorithms in search, Optimization, and Machin Learning. Addison Wesley, Reading, MA.1989

[3] K.D.Boese, "Cost Versus Distance In the Traveling Salesman Problem", Technical Report CSD-950018, UCLA, 1995.

[4] Z.Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer - Verlag, 1992.

[5] J.Grefenstette, R.Gopal, B.Rosmaita, and D. Gucht, "Genetic Algorithms for the Traveling Salesman Problem", Proc. the 1st Inter. Conf. on GAs and Their Applications, pp.160-168, 1985.

[6] D.Whitley, T.Starkweather and D.Fuquay, "Scheduling problems and traveling salesman: the genetic edge recombination and operator", Proc. Third Int. Conf. G As, pp.133-140,1989.

저자소개

강 래 구 (Rae-Goo Kang)

2005년 조선대학교 전자공학과 공학사
2005년 - 조선대학교 전산통계학과 석사과정

※관심분야 : 신경망,인공지능,정보보호,Bioinformatics

임 희 경 (Hee-Kyoung Lim)

2003년 조선대학교 전산통계학과 박사과정수료
2006년 - 서강정보대학 컴퓨터정보과겸임교수

※관심분야 : 신경망,인공지능,정보보호,Bioinformatics

정 채 영 (Chai-Yeoung Jung)

1989년 조선대학교 컴퓨터공학과 공학박사
1986년 - 조선대학교 컴퓨터통계학과 교수

※관심분야 : 신경망,인공지능,정보보호,Bioinformatics
※ e-mail : cyjung@mail.chosun.ac.kr ※062)230-6625