

퍼지 추론 기반의 유전알고리즘 선택 연산자

Fuzzy Reasoning based Selection Operator for Genetic Algorithm

서기성, 현수환

서울시 성북구 서경대학교 전자공학과

E-mail: ksseo@skuniv.ac.kr, xjavalov@shhyun.com

요 약

본 논문은 퍼지추론을 통해 개체의 유사성과 적합도의 종합적 평가를 이용한 유전알고리즘의 선택연산자를 제안한다. 단일 집단을 가상적으로 임의의 n 개의 개체군을 나누고, 개체의 적합도와 유사도에 기반한 퍼지추론을 통해, 효율적인 계층화를 구성하고자 한다. 동시에 점진형(steady-state) 진화방식과 결합시켜 계층화된 군집내에서 개체들이 조기에 수렴하는 현상을 방지해 줄 수 있도록 하고, 적은 개체를 이용하여 효율적인 진화가 가능하도록 구현하였다. 2가지 기만적 문제에 대해서 다른 선택 연산자들의 결과와 비교하였으며, 만족할만한 성능을 얻었다.

Key Words : Genetic Algorithm, Fuzzy Reasoning based Selection Operator, Deceptive Problem

1. 서 론

유전알고리즘[1]의 수행에 있어서 조기 수렴 문제는 매우 중요한 문제로 다루어지고 있다. 조기 수렴(premature convergence)은 전역 최적점에 도달하지 못하고, 대신 지역 최적 값으로 수렴되는 현상을 의미한다.

조기 수렴문제의 해결책으로 제시된 여러 기법 중의 하나는 복수개 집단(multi-pop) 모델을 사용하여 개체의 다양성을 유지하도록 하는 것이다. 그러나 단순한 복수개 집단(multi-pop) 모델은 조기 수렴 개선에 기대보다 큰 효과가 없으며, 이를 개선하기 위하여, 복수개 집단을 기반으로 하여 발전된 모델인 HFC (Hierarchical Fair Competition)[2], ALPS (Age-Layered Population Structure)[3] 등이 조기 수렴을 방지하는 데 매우 높은 효율성을 가지고 있음이 보고되었다.

이들 모델의 공통점은 개체들을 계층적으로 나누어 각 계층별로 공정한 경쟁을 하는 것이다. HFC 모델은 적합도를 기반으로 했고, ALPS 모델에서는 각각의 개체들의 발전되어 온 역사(나이)를 기반으로 했다.

하지만 이들 모델은 섬 모델을 기반으로 한 다중군집화에 기초하고 있기 때문에, 많은 수

의 개체를 확보해야 좀 더 효율적인 연산이 가능하다는 단점이 있다. 즉, 임베디드 장비나 로봇 등의 소규모 메모리의 이용만이 가능한 환경에서 온라인으로 유전알고리즘을 적용하기에는 적합하지 않다.

따라서 단일집단에 대해서 기존의 조기수렴 문제의 해결을 위한 계층적 구분의 모델을 기본 원리를 이용하면서, 더 적은 개체수로서 효율적 연산이 가능한 퍼지추론 기반의 선택 연산자를 제안한다. 단일 집단을 가상적으로 임의의 n 개의 개체군을 나누고, 개체의 적합도와 유사도에 기반한 퍼지추론을 통해, 효율적인 계층화를 구성하고자 한다.

비교대상인 적합도는 개체가 갖는 적합도를 스케일링을 통해 0부터 1까지의 범위의 수치로 변환했으며, 유사도는 해밍 거리를 이용해서 해당 군집, 혹은 전체 개체에 대한 평균적 유사도를 측정하여 이용했다.

부가적으로 제한된 자원의 사용과 연산량 측면에서 세대형(generational) 진화방식보다 유리한 점진형(steady-state) 진화방식을 적용하였다.

2. 계층적 고립 모델의 문제점

HFC나 ALPS 모두 집단적 계층화 모델로 조기수렴 문제를 크게 개선하였으나, 아직도 다음과 같은 문제점들이 남아 있다.

첫째, 적합도나 나이등의 하나의 평가 기준으로 계층화를 하므로, 같은 계층내에서 우수 개체들에 의해서 지배되는 지역수렴이 발생 될 수 있다. 즉, 적합도는 낮지만 가능성 있는 형질을 가진 개체들이 상위 군집으로 이동되기 전에 선택에서 탈락될 확률이 매우 높기 때문에 다양성을 잃을 수가 있다.

둘째, 특히, 세대형(generational) 진화방식의 경우에 매 세대마다 너무 많은 개체들이 유전 연산에 참여함으로 이러한 현상을 가속화 시킬 수 있다.

따라서, 본 연구에서는 새로운 퍼지 추론 기반의 조합 선택 연산자를 제안한다. HFC(그림 1)의 다수의 군집을 통한 계층적 구조를 하나의 군집에서 퍼지를 통한 여러 가지 정보의 조합으로 간접적 계층화를 이루어 내고, 추가적으로 점진형(steady-state) 진화방식을 적용해서 기존의 대규모의 군집을 필요로 하던 연산을 적은 수의 연산으로도 가능하게 한다.

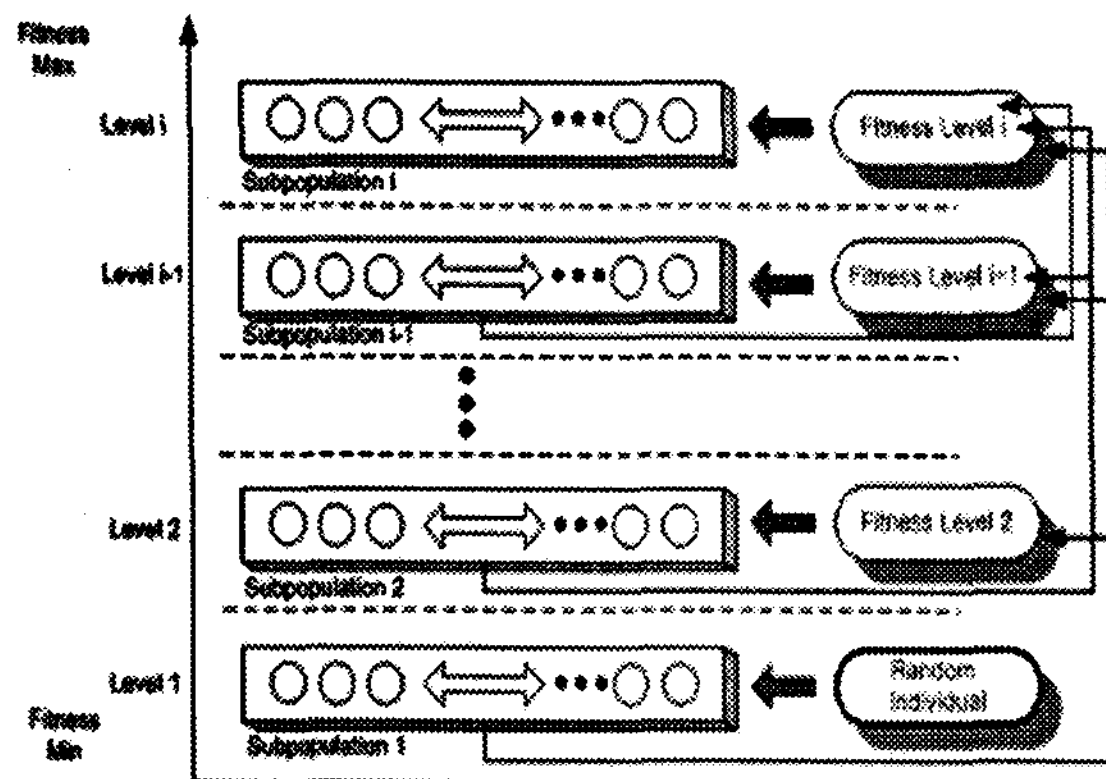


그림1. 적합도에 의한 HFC 모델의 계층화 방식

3. 퍼지 추론 기반의 조합 선택 연산자

제안하는 퍼지 추론 기반의 조합 선택 연산자의 모델은 아래의 그림2와 같은 흐름으로 이루어진다. N 개의 개체를 통한 임의의 선택을 한 후에 유사성과 적합도에 대해 각각 재조정 작업이 이루어진 후에 퍼지추론 모델을 통해 퍼지수를 결과로 나타내고, 비퍼지화를 통해 실제 이용 가능한 개체의 소속도를 평가한다.

그리고 이것을 이용한 경쟁을 통해 적합도 뿐 아니라 유사도에 대한 비교를 조합시킨 형태로 선택연산이 이루어진다.

선택은 토너먼트(크기 = 9) 방식을 사용하였

고, 토너먼트에 속한 개체들은 퍼지추론 모델을 통해 계산된 퍼지값을 기준으로 경쟁한다.

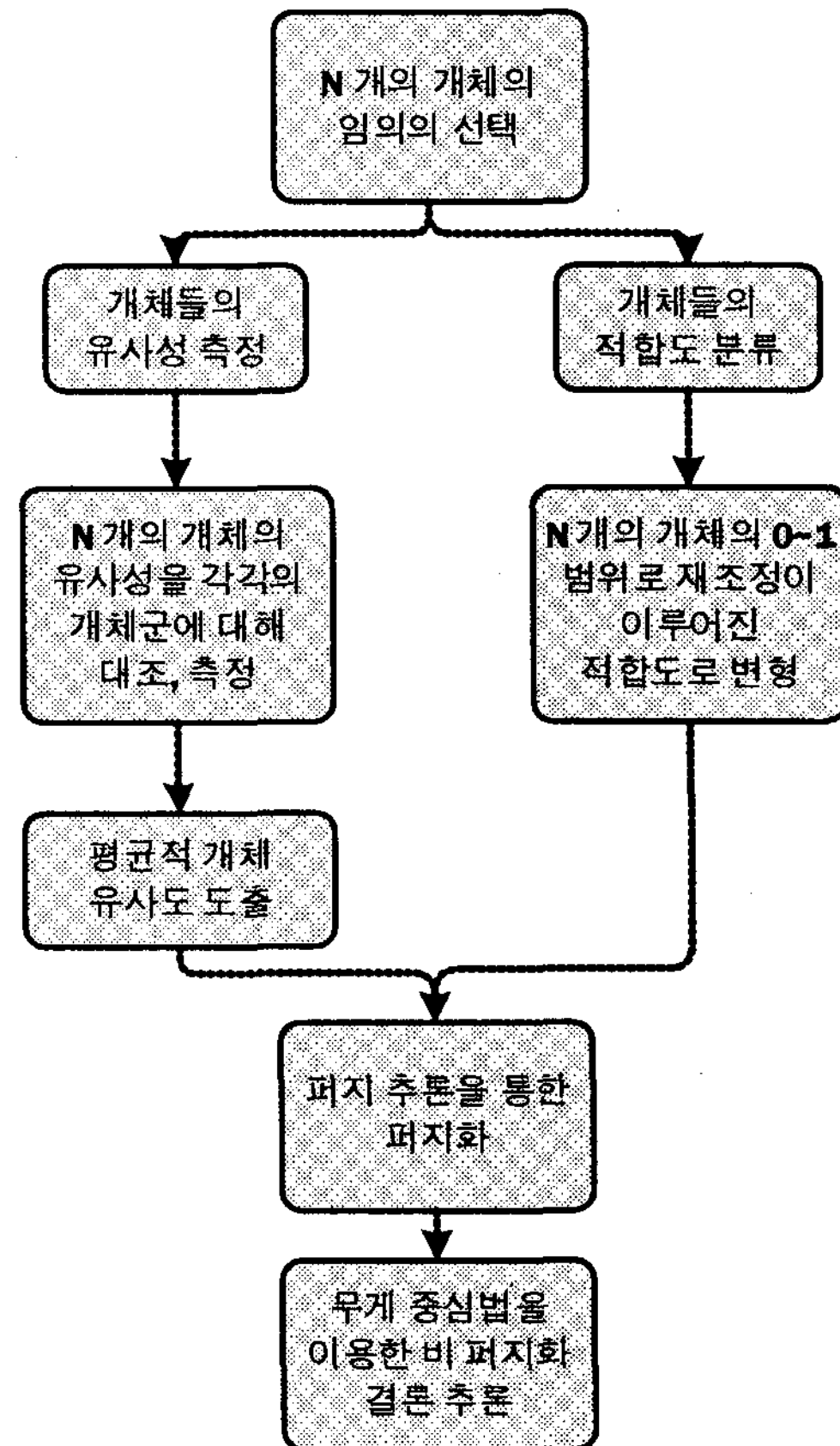


그림2. 퍼지를 이용한 조합방식

4. 퍼지 모델

퍼지 모델은 전반부 변수 2개와 후반부 변수 1개로 구성했다. 전반부 변수 $X1, X2$ 는 각각 적합도와 유사도를 나타내고, 후반부 변수는 G_i 은 소속계층을 의미한다. 전반부 변수의 멤버쉽 함수는 각각 Big, Middle, Small 로 구성했으며, 가우시안 분포를 따른다. 후반부 변수는 다섯 개의 상수 값으로 구성된다. 각 변수들의 범위는 다음과 같다.

$$0 \leq X1, X2 \leq 1, \\ Y \in \{G1 = 1, G2 = 2, G3 = 3, G4 = 4, G5 = 5\}$$

추론 방식은 간략 추론방식을 이용했으며, 각 규칙은 다음과 같이 구성하였다.[4]

- R1 : IF $X1$ is Big and $X2$ is Big,
Then Y is $G5$
- R2 : IF $X1$ is Big and $X2$ is Middle,
Then Y is $G3$
- R3 : IF $X1$ is Big and $X2$ is Small,
Then Y is $G1$

- R4 : IF X1 is Middle and X2 is Big,
Then Y is G4
- R5 : IF X1 is Middle and X2 is Middle,
Then Y is G3
- R6 : IF X1 is Middle and X2 is Small,
Then Y is G2
- R7 : IF X1 is Small and X2 is Big,
Then Y is G3
- R8 : IF X1 is Small and X2 is Middle,
Then Y is G2
- R9 : IF X1 is Small and X2 is Small,
Then Y is G1

이는 각각의 적합도와 유사도에 따라 계층화 시킨 방식으로써, G1 은 개선될 확률이 가장 적은 개체가 속한 계층을 의미하고, G5 는 개선될 확률이 가장 높은 개체가 속한 계층을 의미한다.

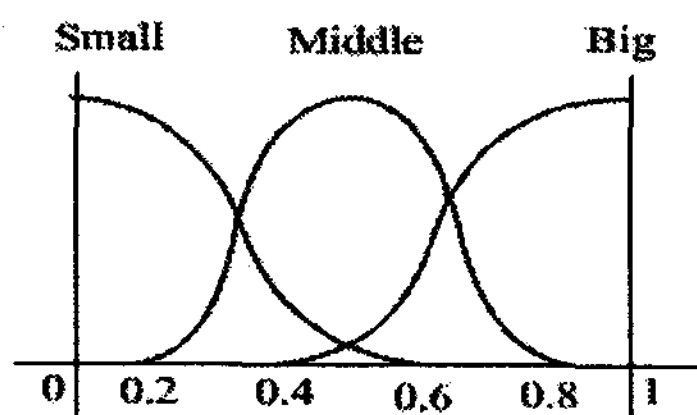


그림3. 퍼지함수의 구성

4. 시뮬레이션 결과 및 고찰

4.1 테스트 함수

실험에서 이용된 문제는 f3deceptive 와 f5trap 함수이다[5-6]. 다음의 수식(1), (2)와 같이 정의된 함수로서 각각의 3비트, 5비트가 하나의 부분문제로 정의된다. f3deceptive 는 연결되어 있는 3개의 비트로 구성된 스트링에 포함되어있는 1의 개수로 적합도를 평가하고, f5trap 은 연결되어 있는 5개의 비트에 속해있는 1의 개수로 적합도를 평가 한다. 연속된 단위문제를 묶어서 문제의 크기가 커질수록 해결이 매우 힘들어지는 문제로 알려져 있다.

$$f3deceptive(u) = \begin{cases} 0.9 & u = 0 \\ 0.8 & u = 1 \\ 0 & u = 2 \\ 1 & otherwise \end{cases} \quad (1)$$

$$f5trap(u) = \begin{cases} 4-u & \text{if } u < 5 \\ 5 & otherwise \end{cases} \quad (2)$$

여기서, u값은 한 단위문제당 포함되어 있는 1의 개수이다. 즉, f3deceptive 는 3비트를 한

단위로 볼 때, 1의 개수가 1,2 개로 증가될 때 점점 낮아지기 때문에 모든 비트가 서서히 0으로 수렴해 나가는 현상을 볼 수 있다. 그리고 전역 최적점은 반대방향에 존재하기 때문에 알고리즘이 지역 최적점을 회피해 전역 최적점의 탐색 능력을 실험해보기 좋은 함수로 널리 이용되고 있다.[5-6] GA 코드는 Gallops[7]를 수정하여 사용하였다.

4.2 실험결과

f3deceptive 문제에 대한 첫 번째 실험환경은 popsize=30, pcross=0.7, pmutation=0.1를 이용했고, 사용된 연산자는 교배연산자는 2점 교배 연산자, 변이연산자는 다중비트 변이 연산자를 사용했다. 각각의 실험방법은 FC, T+DC, T, R 로 구분하였는데, FC는 제안된 선택연산자이고, T는 토너먼트 선택연산자이다. R은 룰렛휠 선택 연산자를 나타내고, DC는 DeJong Crowding[8] 기법을 의미한다. 각 실험은 10회씩 반복 실험 되었으며, 평균값을 기록하였다.

표 1. 테스트 결과(평균값, f3deceptive)

	R	T	T+DC	FC
20	548632	44254	24819	15017
40	1476106	103766	67562	52170
80	3842609	320227	222893	185723
160	5842967	783645	534458	471572
200	8986015	1243961	975250	732389

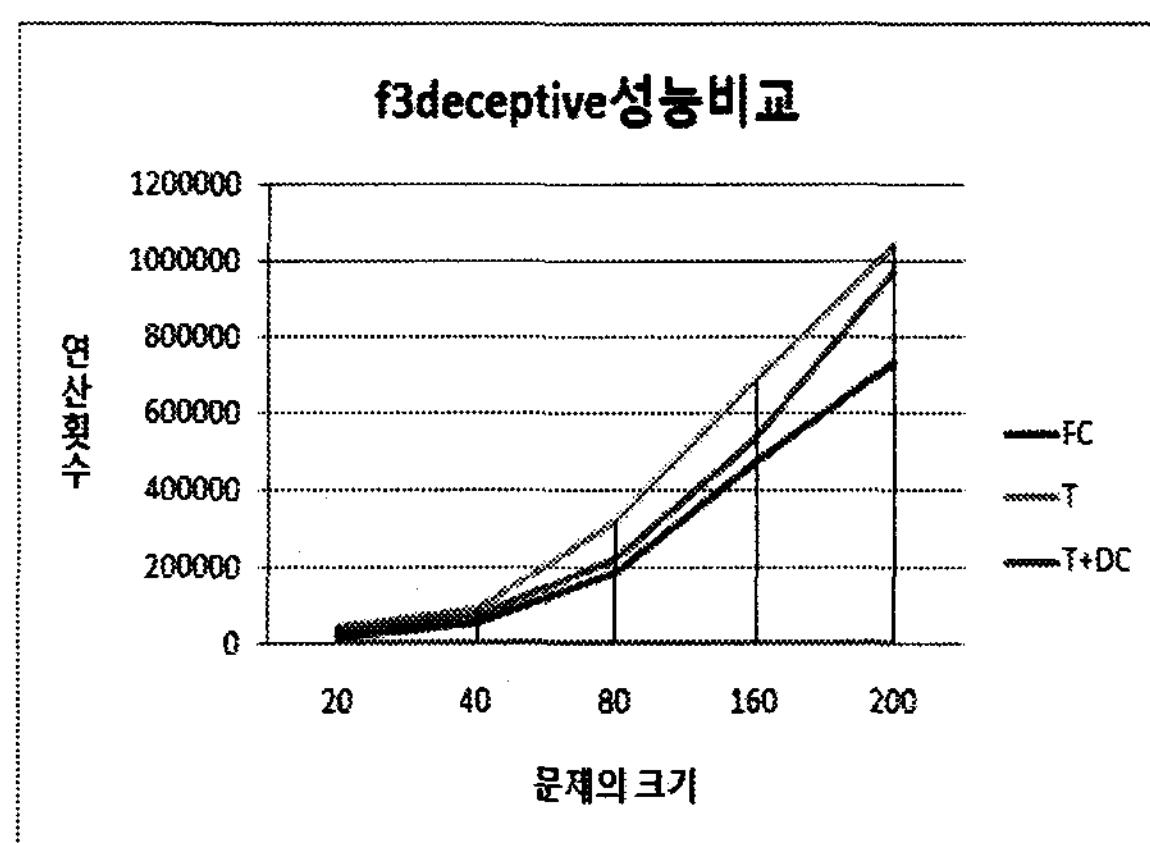


그림4. f3deceptive 함수의 성능 비교

제안된 퍼지기반 선택연산(FC)이 다른 3가지 방법보다 성능이 우수함을 알 수 있었고, DC 가 두 번째로 성능이 우수한 것은 개체의 유사도를 이용하므로 인해 얻어진다고 사료된다. 룰렛휠(R)이나 토너먼트(T)에 비하여는 월

등한 성능 향상을 보여주고 있음을 확인할 수 있다.

그림4의 그래프에 연산량의 증가 추이가 나와 있다. 제안된 방법이 문제의 크기에 따라 증가되는 연산량의 기울기가 다른 알고리즘에 비해서 낮은 것을 확인할 수 있다. 플랫폼의 경우 성능의 차이가 너무 커서 그래프에는 표시하지 않았다.

f5trap 테스트도 f3deceptive 와 마찬가지로 popsize=30, pcross=0.7, pmutation=0.1 의 환경에서 수행되었다.

표 2. 테스트 결과(평균값, f5trap)

	R	T	T+DC	FC
20	6846248	533344	371889	282461
40	X	1678021	963200	678476
80	X	3662225	2285588	1245239
160	X	8080191	6087087	3554323

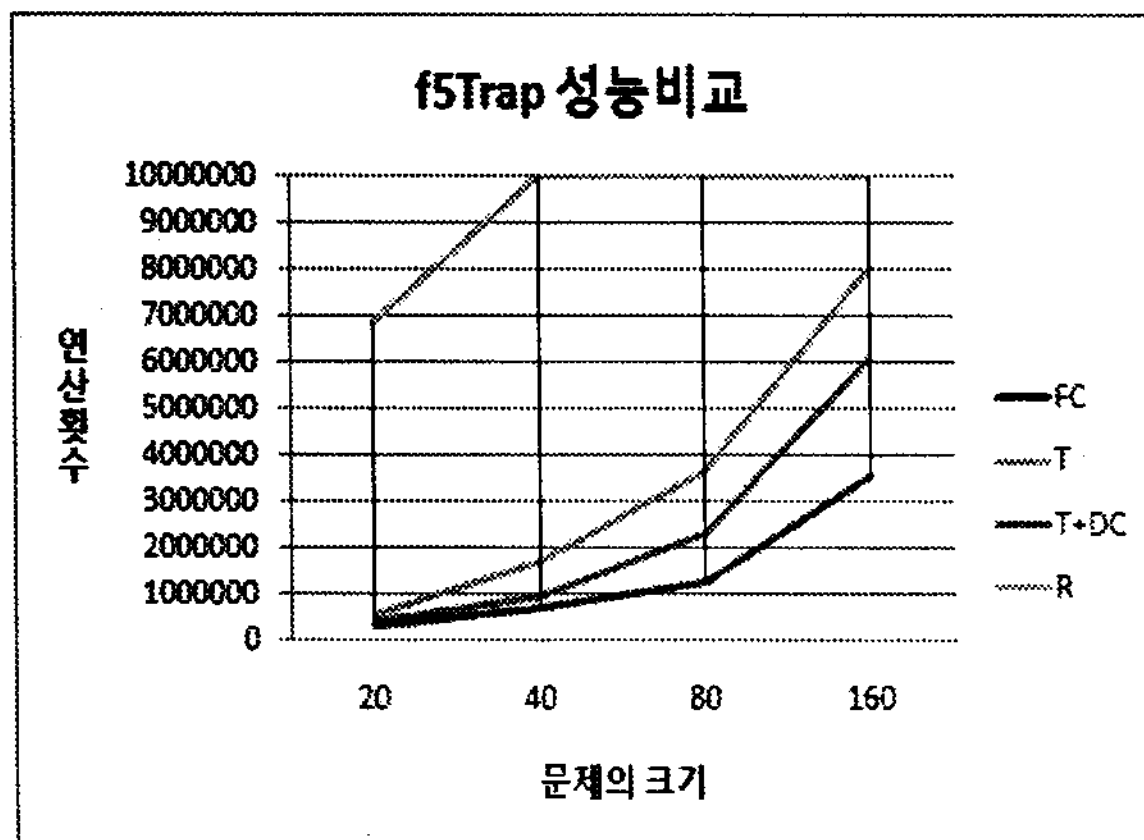


그림5. f5trap 함수를 통한 성능비교

표 2의 결과에 표시된 X 는 천만번 이상의 연산을 거치고도 전역최적점을 찾아내지 못한 경우다. f5trap 문제의 경우에는 5개나 되는 비트를 하나의 단위 묶음으로서 보기 때문에, 난이도가 더 높은 탐색 문제이다. 표2와 그림 5에서 보는 바와 같이, 제안된 방법이 다른 알고리즘과 성능차이가 f3deceptive에 비해서 더 커진 것을 알 수 있다. 더 어려운 문제에 대해 더 좋은 성능을 발휘하고 있는 것을 확인할 수 있었다.

6. 결론

본 연구에서는 퍼지추론을 기반으로 하여 유전 알고리즘의 개체에서 중요한 두 가지 속성인 적합도와 개체간의 유사성을 통한 분류 및

선택 알고리즘을 구현하였고 이를 기존의 대표적인 선택 방식들과 비교하였다. 대규모의 군집 및 개체수를 사용한 방식과는 달리, 일반적으로 이용되는 개체보다도 훨씬 적은 30개의 개체만으로 기만적인 어려운 문제를 탐색할 수 있음을 보였다. 따라서 고용량의 메모리를 지원하지 않는 로봇이나 임베디드 시스템등에 온라인으로 효율적인 진화연산을 할 수 있는 가능성을 기대할수 있다.

제안된 방법에서 사용한 적합도와 유사도의 비교외에 추가적으로 개체들의 나이, 개체의 개선속도등의 다른 특성 요소를 포함하는 방법과 이에 적합한 퍼지 추론법의 구현등이 필요하다고 본다.

참 고 문 헌

- [1] J. H. Holland, "Genetic Algorithms in Search, Optimization and Machine Learning," Addison- Wesley, Reading, MA, 1989.
- [2] J. Hu, E. Goodman, K. Seo, Z. Fan, R. Rosenberg, " The Hierarchical Fair Competition(HFC) Framework for Continuing Evolutionary Algorithms,". Evolutionary Computation, The MIT Press, Vol. 13, Issue 2, pp 241-277, 2005.
- [3] G. S. Hornby. "ALPS: The Age-Layered Population Structure for Reducing the Problem of Premature Convergence", GECCO' 06, July 8 - 12, 2006, Seattle, Washington, USA. ACM 1595931864/06/0007.
- [4] S.-K. Oh, "Fuzzy Model & Control System by C-Programming", Naeha Publishing Co., 2002.
- [5] M. Pelikan, D. E. Goldberg, and E. Cant' u-Paz, "BOA: The Bayesian Optimization Algorithm", Proceedings of GECCO'99, pp. 525 - 32, 1999.
- [6] J. Hu and E. Goodman, "Robust and Efficient Genetic Algorithms with Hierarchical Niching and a Sustainable Evolutionary Computation Model", GECCO 2004,LNCS 3102, pp. 1220-1232, 2004.
- [7] E. D. Goodman, *Gallops User's Manual*. Michigan State University, July 1996.
- [8] DeJong, K. A.: "An analysis of the behavior of a class of genetic adaptive systems." (Doctoral dissertation, University of Michigan). Dissertation Abstracts International, 36(10),514B. (University Microfilms No. 76-9381). (1975).