

# UNDX연산자를 이용한 계층적 공정 경쟁 유전자 알고리즘을 이용한 퍼지집합 퍼지 모델의 최적화

## Optimization of Fuzzy Set Fuzzy Model by Means of Hierarchical Fair Competition-based Genetic Algorithm using UNDX operator

김길성\*, 최정내\*\*, 오성권\*\*\*

Gil-Sung Kim, Jeoung-Nae Choi, Sung-Kwun Oh

**Abstract** - In this study, we introduce the optimization method of fuzzy inference systems that is based on Hierarchical Fair Competition-based Parallel Genetic Algorithms (HFCGA) and information data granulation. The granulation is realized with the aid of the Hard C-means clustering and HFCGA is a kind of multi-populations of Parallel Genetic Algorithms (PGA), and it is used for structure optimization and parameter identification of fuzzy model. It concerns the fuzzy model-related parameters such as the number of input variables to be used, a collection of specific subset of input variables, the number of membership functions, the order of polynomial, and the apexes of the membership function. In the optimization process, two general optimization mechanisms are explored. The structural optimization is realized via HFCGA and HCM method whereas in case of the parametric optimization we proceed with a standard least square method as well as HFCGA method as well. A comparative analysis demonstrates that the proposed algorithm is superior to the conventional methods. Particularly, in parameter identification, we use the UNDX operator which uses multiple parents and generate offsprings around the geographic center off mass of these parents.

**Key Words** : Fuzzy Set-based Fuzzy Model, Information Granulation(IG), Hard C-Means(HCM), Genetic Algorithms(GAs), Hierarchical Fair Competition Genetic Algorithms (HFCGA)

### 1. 서론

1965년 Zadeh[1]에 의해 창안된 "퍼지집합" 이론을 계기로 하여 비선형적이고 다변수인 시스템을 대상으로 한 퍼지 모델링 기법의 유용성은 이미 잘 알려져 있으며 이들은 퍼지추론 시스템에 기초하고 있다. 퍼지 모델의 성능은 퍼지규칙의 구성 방법에 의존하며 보다 좋은 성능을 위해서는 퍼지규칙의 동정이 필연적이다. Pedrycz 와 Oh는 진보된 퍼지모델들을 제안하고, 유전자 알고리즘을 이용하여 구조와 전반부 파라미터를 최적화 하는 방법을 제시하였다[2-4].

본 논문은 퍼지집합 퍼지모델의 최적화에 중점을 두었으며, 최적화를 위해서 계층적 공정 경쟁 기반 병렬유전자 알고리즘(HFCGA)을 개발하고 적용하였다. 일반적으로 유전자 알고리즘은 전역해를 찾을 수 있는 최적화 알고리즘으로 잘 알려져 있으며, 많은 응용에 성공적으로 사용되고 있다. 그렇지만 탐색공간이 크거나, 파라미터 수가 많은 경우 세대가 증가해도 더 좋은 최적해를 찾지 못하고 지역해에 빠지는 조기수렴 문제를 내재하고 있다. 조기 수렴을 억제하고 전역해를 찾기 위한 해결방안으로써 병렬유전자 알고리즘이 개발되었으며[5], 특히 병렬유전자 알고리즘의 한 구조인 계층적 정당 경쟁 구조는 큰 탐색 공간을 갖는 문제에 효과적이다[6].

퍼지모델의 설계는 구조와 파라미터의 최적화 두 부분으로 수행된다. 먼저 HFCGA에 기반하여 퍼지집합 기반 모델

의 입출력변수, 입출력 변수에 대한 멤버쉽함수의 수, 후반부의 구조 등을 선택 및 멤버쉽함수의 초기정점이 구해지며, 파라미터 최적화에서는 최소자승법에 의하여 후반부 파라미터와 HFCGA에 의하여 전반부 파라미터가 최적화된다. 본 연구는 파라미터 동정시 보다 최적화된 멤버쉽함수의 정점을 구하기위해 실수코딩 유전자 알고리즘에 좋은 성능을 보이는 UNDX 연산자를 사용함으로써 보다 좋은 비선형 시스템의 모델링 방법을 제안한다.

### 2. 퍼지 모델

#### 2.1 전반부 동정

퍼지 모델링에서 전반부 동정, 즉 구조 동정 및 파라미터 동정은 비선형 시스템을 표현하는데 있어서 매우 중요하다. 본 논문에서는 멤버쉽 함수를 입력 변수의 최소값과 최대값 사이에서 임의의 개수로 등분하여 일률적으로 정의하는 Min-Max 방법을 사용하여 멤버쉽함수의 초기 정점을 동정한다.

#### 2.2 후반부 동정

후반부 다항식 함수의 형태에 따라 네 가지 구조의 모델이 있다.

##### i) Type 1(간략 퍼지추론)

후반부가 단일 상수항만을 가지는 것으로, 이와 같은 추론법을 간략 퍼지추론법이라 한다. 퍼지모델은 식 (1)과 같은 형태를 가지는 구형 규칙들로 구성되며, 퍼지추론에 의해 추론된 값  $y_j^*$ 는 식 (2)과 같다.

$$R_j : I^{A_1} \text{ is } A_1 \text{ c } \text{and} \text{ not } I^{A_2} \text{ is } A_2 \text{ c } \text{then } y_j = y_j^* \quad (1)$$

\* 正會員 : 水原大學校 電氣工學科

\*\* 正會員 : 水原大學校 電氣工學科

\*\*\* 正會員 : 水原大學校 電氣工學科 正教授

$$y^* = \frac{\sum_{j=1}^n w_j y_j}{\sum_{j=1}^n w_j} = \sum_{j=1}^n \hat{w}_j a_{j0} \quad (2)$$

후반부 파라미터  $a_{j0}$ 은 최소자승법에 의해 결정된다.

$$\hat{\lambda} = (X^T X)^{-1} X^T Y \quad (3)$$

### ii) Type 2(선형 퍼지추론)

후반부가 일차 선형식을 가지며, 이와 같은 추론법을 선형 퍼지추론법이라 한다. 퍼지모델은 식 (4)의 형태를 가지는 구형 규칙들로 구성되고, 추론된 값  $y^*$ 는 식(6)과 같다.

$$R^j \text{ IF } x_1 \text{ is } A_{1c} \text{ and } \dots \text{ and } x_k \text{ is } A_{kc} \text{ then } y_j = f_j(x_1, \dots, x_k) \quad (4)$$

$$f_j(x_1, \dots, x_k) = a_{j0} + a_{j1}x_1 + \dots + a_{jk}x_k \quad (5)$$

$$y^* = \frac{\sum_{j=1}^n w_j y_j}{\sum_{j=1}^n w_j} = \sum_{j=1}^n \hat{w}_j f_j(x_1, \dots, x_k) \quad (6)$$

### iii) Type 3(2차식 퍼지추론)

후반부가 2차식 함수의 다항식 형태를 가지며, 퍼지모델은 식 (7)과 같이 구성된다.

$$R^j \text{ IF } x_1 \text{ is } A_{1c} \text{ and } \dots \text{ and } x_k \text{ is } A_{kc} \text{ then } y_j = f_j(x_1, \dots, x_k) \quad (7)$$

$$f_j(x_1, \dots, x_k) = a_{j0} + a_{j1}x_1 + \dots + a_{jk}x_k + a_{j(k+1)}x_1^2 + \dots + a_{j(2k)}x_k^2 + a_{j(2k+1)}x_1x_2 + \dots + a_{j((k+2)(k+1)/2)}x_{k-1}x_k \quad (8)$$

모델의 추론된 값  $y^*$ 는 식 (6)과 같다.

### iii) Type 4(변형된 2차식 퍼지추론)

입력 변수의 2차 항이 생략된 구조 3의 변형된 형태로서, 다음과 같은 구형 규칙으로 구성된다.

$$R^j \text{ IF } x_1 \text{ is } A_{1c} \text{ and } \dots \text{ and } x_k \text{ is } A_{kc} \text{ then } y_j = f_j(x_1, \dots, x_k) \quad (9)$$

$$f_j(x_1, \dots, x_k) = a_{j0} + a_{j1}x_1 + \dots + a_{jk}x_k + a_{j(k+1)}x_1^2 + \dots + a_{j(2k)}x_k^2 \quad (10)$$

모델의 추론된 값  $y^*$ 는 식 (6)과 같다.

## 3. HFCGA 기반 퍼지모델의 최적화

GAs는 많은 분야에 효과적으로 사용되고 있지만 복잡한 최적해 탐색 문제에 대해서는 세대가 증가해도 최적해를 찾지 못하고 지역해에 수렴하는 조기수렴 문제가 발생하게 되며, 이것은 진화 알고리즘에서 해결해야 할 문제이다. 조기수렴문제를 해결하기 위한 방법으로 여러 개의 집단을 사용하여 집단간의 개체가 이주하는 구조를 갖는 병렬유전자 알고리즘이 있으며, 그중에서도 계층적 경쟁 방법(Hierarchical Fair Competition ; HFC)을 도입한 HFCGA가 있다. 이 방법은 사회적 교육 시스템에서 동등 수준을 가진 그룹으로 나누어 학습에 대한 경쟁을 하며, 우수한 인재를 상위그룹으로 진급하는 체계를 병렬 유전자 알고리즘에 적용한 것으로서, 유전자 알고리즘의 수행과정에서 각 개체들을 적합도별로 나누어 경쟁을 시킴으로써 공평한 경쟁을 통해 처음 생성된 낮은 수준의 개체가 개선될 기회도 없이 탈락하는 경우를 방지한다. 왜냐하면 돌연변이 연산 등에 의해 처음 임의로 생성된 개체가 상당세대를 통해 진화된 높은 수준의 개체와 유사한 성능을 갖기는 매우 희박하기 때문이다. 낮은 수준의 개체도 다음 세대에서 높은 성능을 갖는 개체로 진화하면, 그 수준에

맞는 집단으로 바로 이동할 수 있다. 이와 같은 원리를 통해서 항상 아래 수준의 집단으로부터 높은 수준의 개체가 올라올 수 있는 통로를 만들어 놓음으로써, 조기 수렴을 최대한 방지할 수 있다.

HFCGA는 여러 개의 집단이 계층적으로 구성되어있으며, 각 집단은 특정 범위의 적합도를 갖는 개체들로 구성되어 있으며, 하위수준의 집단으로부터 개체의 진입을 위한 문턱적합도를 갖고 있다. 각 집단을 독립적으로 진화하게 되며 진화과정 중 우수한 개체는 자기 수준에 맞는 상위 집단으로 이주하게 된다. 가장 열등한 집단(Access deme)의 개체는 모든 상위집단으로 이동할 수 있으며, 가장 우수한 집단(Elite deme)은 단지 다른 집단으로부터 개체를 받을 수 만 있다. 본 논문에서 사용된 HFCGA는 실수형 GAs를 기반으로 했으며, 5개의 집단을 사용하였다. 각각의 집단은 30, 30, 30, 30, 30개의 개체를 사용하였으며, 교배 연산자는 모두 UNDX 연산자를 사용하였다.

퍼지모델의 최적화를 위해서 구조 동정에서는 입력변수의 수, 사용될 입력의 선택, 선택된 입력변수에 대한 멤버쉽함수의 수 그리고 후반부 추론방식의 탐색에 사용되었으며, 전반부 파라미터 동정에서 멤버쉽함수의 정점의 최적화에 사용되었다. 그림 1에 염색체의 구성을 보인다.

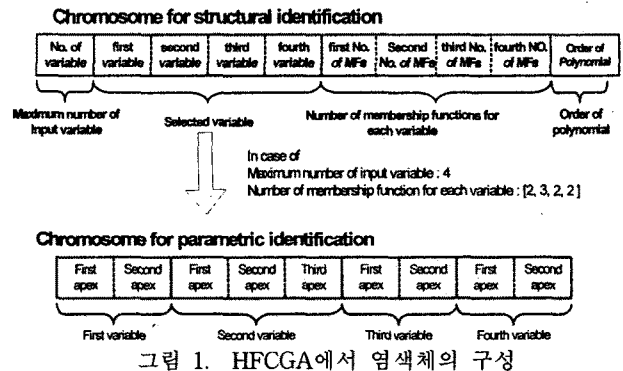


그림 1. HFCGA에서 염색체의 구성

## 4. Unimodal Normal Distribution Crossover (UNDX)

기존의 다른 교배 연산자는 두개의 부모로부터 자손이 생성되는 반면에 UNDX는 다수의 부모(multiple parents)를 이용하여 부모들의 기하학적 중심에 근접하게 정규분포를 이루며 생성된다.[7] 정규분포를 따르므로 작은 확률로 부모들의 중심으로부터 먼 자손들도 생성된다. 그림 2에서 정규분포의 중심은 parent1과 parent2의 중심 점이 되고 그 중심으로부터 표준편차를 따르는 랜덤변수(Gaussian random variable)로서 거리가 구해진다. parent3으로부터 parent1과 parent2를 잇는 선분에 수직인 벡터를 구하고 이 벡터와 또 다른 가우시안 랜덤변수와 비례하는 점에 자손이 위치하게 된다.

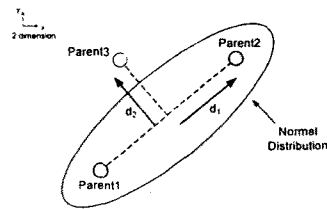


그림 2. UNDX의 기하학적 표현

## 5. 시뮬레이션

본 연구의 Sugeno가 사용한 다음의 비선형 시스템을 이용하여 시뮬레이션하였다.

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2, 1 \leq x_1, x_2 \leq 5 \quad (11)$$

본 연구에서는 50개의 데이터 중 30개를 모델을 동조하기 위한 Training 데이터로 사용하였고, 나머지 20개를 모델을 평가하기 위한 Testing 데이터로 사용하였다. 또한 HFCGA의 성능비교 대상으로 한 개의 집단을 갖는 실수코딩 유전자 알고리즘(RCGA)을 사용하였다. 표1은 RCGA와 HFCGA를 사용하여 파라미터 동정시 설정된 진화알고리즘의 파라미터이다.

표 1. HFCGA와 RCGA의 파라미터

	RCGA	HFCGA
세대수	300	300
집단수	1	5
집단크기	150	30×5
교배율	0.75	0.75
돌연변이율	0.1	0.1

구조 동정결과 사용될 변수는  $x_1, x_2$  모두 사용하고 각 변수에 대한 멤버쉽함수의 수는 3, 4개로 선택되었고, 후반부 다항식은 선형식(Type 2)이 선택되었다. 표 2에서는 파라미터 최적화후 모델의 성능지수를 보인다. 성능지수로 Mean Square Error (MSE)를 사용하였다. 그림 3에서는 파라미터 동조과정에서 RCGA와 HFCGA의 세대에 따른 목적함수의 수렴과정을 보인다. HFCGA+UNDX를 이용하여 파라미터를 동정시 RCGA보다 이용한 동정보다 좋은 성능을 보임을 알 수 있다.

표 2. HFCGA+UNDX와 RCGA와 성능지수비교

진화알고리즘	PI	EPI
RCGA	1.0101E-4	0.000248
HFCGA+UNDX	9.1683E-5	0.000169

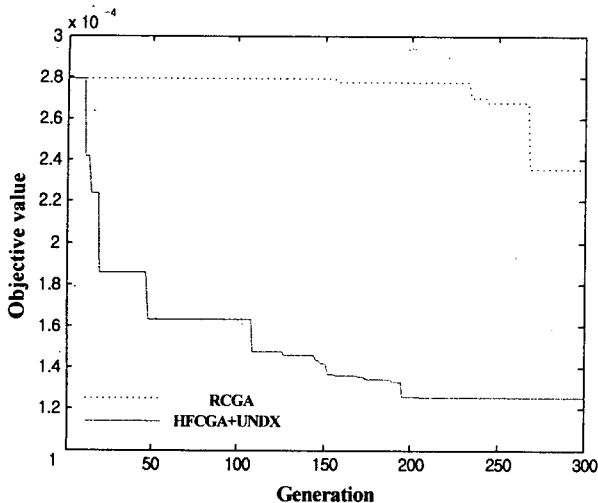
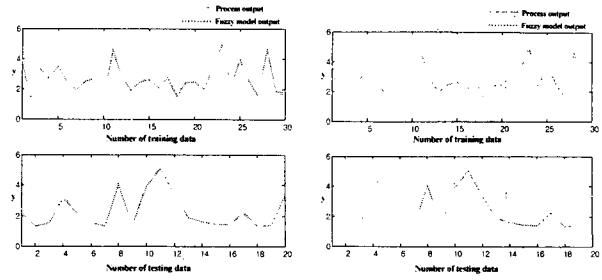


그림 3. 진화알고리즘에서 수렴곡선

그림4에서는 구조동정 후에 실제데이터와 모델출력데이터와 파라미터 동정 후 출력을 비교한다. 구조동정만 했을 경우에는 Testing 데이터에 대하여 오차가 비교적 많이 발생하지만 파라미터 최적화를 통하여 오차를 더욱더 줄일 수가 있다.



(a) 구조동정 (b) 파라미터동정  
그림 4. 실제데이터와 퍼지모델의 비교

## 6. 결론

과라본 논문에서는 비선형 공정에 대하여 효율적인 모델링을 위하여 계층적 공정 경쟁 유전자 알고리즘(HFCGA)과 실수코딩 유전자 알고리즘에서 좋은 성능을 보이는 교배 연산자인 UNDX 연산자를 이용하여 보다 최적화된 비선형 공정을 모델링 하였다. 구조동정에서는 HCM에 의한 초기 정점에 대하여 입력변수 수, 입력변수의 선택, 선택된 입력변수에 대한 멤버쉽함수의 수, 후반부 구조의 선택에 적용되었고, 선택된 구조에 대하여 멤버쉽함수의 정점을 최적화하는 전반부 파라미터 동정에 UNDX 연산자를 사용한 HFCGA는 기존의 단순 실수 코딩 유전자알고리즘(RCGA)을 사용한 동정보다 우수한 성능을 보인다.

### 감사의 글

이 논문은 2006년도 교육인적자원부의 재원으로 한국학술진흥재단(KRF-2006-311-D00194)의 지원을 받아 수행된 연구임

### 참고 문헌

- [1] L. A Zadeh, "Fuzzy sets," *Inf. Control* 8, pp.338-353, 1965.
- [2] W. Pderyca and G. Vukovich, "Granular neural networks," *Neurocomputing*, Vol. 36, pp. 205-224, 2001.
- [3] P. R. Krishnaiah and L. N. Kanal, editors. *Classification, pattern recognition, and reduction of dimensionality*, volume 2 of Handbook of Statistics. North-Holland, Amsterdam, 1982.
- [4] S.-K. Oh and W. Pedrycz, "Identification of Fuzzy Systems by means of an Auto-Tuning Algorithm and Its Application to Nonlinear Systems," *Fuzzy Sets and Syst.*, Vol. 115, No. 2, pp. 205-230, 2000.
- [5] Lin, S.C., Goodman, E., Punch, W.: Coarse-Grain Parallel Genetic Algorithms: Categorization and New Approach. IEEE Conf. on Parallel and Distrib. Processing. Nov. (1994)
- [6] Hu, J.J., Goodman, E.: The Hierarchical Fair Competition (HFC) Model for Parallel Evolutionary Algorithms. Proceedings of the 2002 Congress on Evolutionary Computation: CEC2002. IEEE. Honolulu. Hawaii. (2002)
- [7] Ono, I and Kobayashi, S : A Real-coded Genetic Algorithm for Function Optimization Using Unimodal Normal Distribution Crossover, Proceedings of the Seventh International Conference on GAs, pp. 246-253.