

# 유전자 알고리즘을 이용한 퍼지 시스템의 최적화

## Optimization of fuzzy systems by means of GA

박병준\*, 박춘성\*, 오성권\*, 김현기<sup>o</sup>

\*원광대학교 제어계측공학과, 전라북도 익산시 신용동 344-2 ☎570-749

<sup>o</sup>수원대학교 전기전자정보통신공학부, 경기도 화성군 봉삼면 와우리 ☎445-743

Byoungjun Park\*, Chunseong Park\*, Sungkwun Oh\*, Hyunki Kim<sup>o</sup>

\*Dept. of Control & Instrumentation Engineering, Wonkwang Univ., Iksan, KOREA

<sup>o</sup>Dept. of Electrical Engineering, Suwon Univ., Whasung-gun, KOREA

### ABSTRACT

본 논문은 퍼지 추론 시스템 모델의 최적화를 제시한다. 비선형적이고 복잡한 실시스템의 특성을 해석하는 방법으로써 시스템의 정적 혹은 동적 특성을 묘사하기 위해 퍼지 모델이 사용된다. 그러나 퍼지시스템의 동정은 경험적 방법에 의해 규칙을 추출하기 때문에, 보다 논리적이고 체계적인 방법에 의한 추출 방법의 고찰이 필요하다. 제안된 규칙베이스 퍼지모델은 GA 및 퍼지 규칙의 이론을 이용한 시스템 구조와 파라미터 동정을 시행한다. 두 형태의 퍼지모델 방법은 간략추론 및 선형추론에 의해 시행된다. 본 논문에서는 퍼지 추론 시스템의 전반부 파라미터 동정을 통해 퍼지 입력공간을 정의함으로써 비선형 시스템을 표현한다. 전반부 파라미터의 동정에는 유전자 알고리즘을 사용하고, 후반부는 표준 가우스 소거법을 사용하여 동정한다. 최적화는 유전자 알고리즘에 기초한 자동-동조 방법이며, 학습 및 테스트 데이터의 성능결과의 상호균형을 얻기 위한 하중값을 가진 성능지수가 제시된다.

### 1. 서론

많은 연구자들은 자연 현상을 이해하고, 그것을 이용하기 위해 연구가 이루어지고 있다. 그 중 1965년에 Zadeh에 의해 창안된 "퍼지집합" 이론은 비선형적이고 복잡한 실 시스템의 특성을 해석하는데 적용함으로써 수학적 모델보다 좋은 결과를 가져왔다. 수리적인 모델 방법을 대신하여 Zadeh가 제안한 방법은 수치가 아니라 언어적으로 애매한 변수를 이용하여 시스템을 표현하였고, 조건부 명제에 의해 애매한 변수간의 관계를 나타냈으며, 시스템을 퍼지 알고리즘방식을 이용하여 표현했다. 비선형적이고 다변수인 시스템을 대상으로한 퍼지 제어기와 퍼지 모델링 기법의 유용성은 이미 알려진 사실이며 이들은 퍼지 추론 시스템에 기초하고 있다. 퍼지 제어기와 퍼지 모델의 성능은 퍼지 규칙의 구성 방법에 의존하며 보다 좋

은 성능을 위해서는 퍼지 규칙의 동정이 필연적이다.

퍼지모델의 동정 알고리즘을 본 논문에서 제시한다. 기존의 퍼지시스템의 동정은 경험적 방법에 의해 규칙을 추출하기 때문에, 보다 논리적이고 체계적인 방법에 의한 추출 방법의 고찰이 필요하다. 따라서, 제안된 규칙베이스 퍼지모델은 유전자 알고리즘 및 퍼지규칙의 이론을 이용한 시스템 구조와 파라미터 동정을 시행한다. 두 형태의 퍼지모델 방법은 간략추론 및 선형추론에 의해 시행된다. 퍼지추론 시스템의 전반부 파라미터의 동정은 유전자 알고리즘을 사용하여 최적의 파라미터 값을 얻음으로써 모델의 입력공간을 정의하고 비선형 시스템을 표현한다. 후반부는 표준 가우스 소거법을 사용하여 동정한다. 이의 평가를 위해 Box와 Jenkins가 사용한 가스로 시계열 데이터[3]와 하수처리공정 시스템[7]을 모델링한다. 또

한, 학습 및 테스트 데이터의 성능결과에 대한 상호균형을 얻기 위해 하중값을 가진 성능지수가 제시된다.

## 2. 퍼지 동정 알고리즘

### 2.1 전반부동정

전반부 파라미터 동정에 따른 비교를 위해 구조를 정하고, 기존의 퍼지모델과 제안된 퍼지모델의 성능지수를 비교한다. 전반부 구조로는 입력이 두 개일 때, 그리고 각각의 입력에 대해 소속함수를 정의하였으며, 유전자 알고리즘을 이용하여 전반부 파라미터를 동정한다. 퍼지규칙의 전반부에는 유전자 알고리즘에 의해 자동동정 되는 파라미터(a, b)를 가진(그림 1) 삼각형 형태의 소속함수를 사용한다.

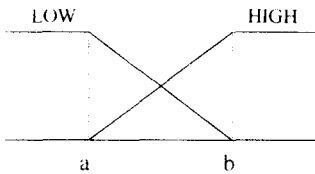


그림 1. 2개의 파라미터를 가진 퍼지변수

### 2.2 후반부 동정

퍼지 모델의 후반부 동정은 구조의 선택과 그들의 파라미터를 동정하는 문제이다. 후반부 구조로는 퍼지추론에 의해 구별되는 구조 1, 2를 사용한다. 구조 1은 후반부가 단일 상수항만을 가지고, 구조 2는 후반부가 일차 선형식 형태를 가진다. 두 가지 구조에 의한 퍼지규칙 추론과 파라미터 동정은 다음과 같다.

#### 구조 1(후반부 : 상수)

후반부가 단일 상수항만을 가지는 것으로, 이와 같은 추론법을 간략추론법이라 한다. 이 퍼지모델은 아래와 같은 형태를 가지는 구현규칙들로 구성된다.

$$\begin{aligned} R^1 \text{ IF } x_1 \text{ is } A_{11} \text{ and } \dots \text{ and } x_k \text{ is } A_{1k} \text{ Then } a_{1j} \\ \vdots \\ R^r \text{ IF } x_1 \text{ is } A_{r1} \text{ and } \dots \text{ and } x_k \text{ is } A_{rk} \text{ Then } a_{rj} \\ \vdots \\ R^n \text{ IF } x_1 \text{ is } A_{n1} \text{ and } \dots \text{ and } x_k \text{ is } A_{nk} \text{ Then } a_{nj} \end{aligned} \quad (1)$$

추론된 값  $y^*$ 는 다음과 같다.

$$y^* = \frac{\sum_{i=1}^n w_i a_{ij}}{\sum_{i=1}^n w_i} \quad (2)$$

(i : i(i=1, ..., m)번째 데이터)

$$w_{ij} = \frac{(A_{j1}(x_{1i}) \wedge \dots \wedge A_{jk}(x_{ki}))}{\sum_{i=1}^m (A_{j1}(x_{1i}) \wedge \dots \wedge A_{jk}(x_{ki}))} \quad (3)$$

후반부 파라미터 동정에서 전반부 입력변수 및 파라미터가 주어지면, 성능지수를 최소화하는 최적 후반부 파라미터를 결정할 수 있다. 시스템의 출력 데이터와 모델 데이터간의 차이를 의미하는 성능지수(PI: Performance Index)는 식(4)로 정의된다.

$$PI = \sum_{i=1}^m \frac{(y_i - y_i')^2}{m} \quad (4)$$

(m : 데이터의 수)

최소자승법에 의한 후반부 파라미터의 동정은 식(5)에 의해 구해진다.

$$\hat{a} = (X^T X)^{-1} X^T Y \quad (5)$$

#### 구조 2(후반부 : 일차 선형식)

후반부가 일차 선형식으로 표현된 것으로 선형추론법 또는 혼합추론법이라 한다.

$$\begin{aligned} R^1 \text{ IF } x_1 \text{ is } A_{11} \text{ and } \dots \text{ and } x_k \text{ is } A_{1k} \\ \text{ Then } y = a_{10} + a_{11}x_1 + \dots + a_{1k}x_k \\ \vdots \\ R^r \text{ IF } x_1 \text{ is } A_{r1} \text{ and } \dots \text{ and } x_k \text{ is } A_{rk} \\ \text{ Then } y = a_{r0} + a_{r1}x_1 + \dots + a_{rk}x_k \\ \vdots \\ R^n \text{ IF } x_1 \text{ is } A_{n1} \text{ and } \dots \text{ and } x_k \text{ is } A_{nk} \\ \text{ Then } y = a_{n0} + a_{n1}x_1 + \dots + a_{nk}x_k \end{aligned} \quad (6)$$

추론된 값  $y^*$ 는 다음과 같다.

$$\begin{aligned} y^* &= \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i} \\ &= \frac{\sum_{i=1}^n w_i (a_{i0} + a_{i1}x_{1i} + \dots + a_{ik}x_{ki})}{\sum_{i=1}^n w_i} \end{aligned} \quad (7)$$

최소자승법에 의한 후반부 파라미터의 동정은 구조 1과 같이 식(8)에 의해 구해진다

$$a = (X^T X)^{-1} X^T Y \quad (8)$$

## 3. 최적화를 위한 유전자 알고리즘

### 3.1 유전자 알고리즘

유전자 알고리즘은 3개의 기본 연산자 재생산(reproduction), 교배(crossover), 돌연변이(mutation)를 통해 문자열을 생성하고, 만들어진 문자열의 군을 가지고 3단계-평가(evaluation), 선택(selection), 재생산(reproduction)를 통해 반복 수행된다. (그림 2)

본 논문에서는 퍼지추론의 방법, 삼각형 소속함수의 특성, 그리고 하중값을 바탕으로 유전자 알고리즘을 이용하여 퍼지모델을 최적화한다. 동정대상은 앞절에서 제시한 것처럼 퍼

지추론의 전반부 파라미터이다.

### 3.2 하중값을 가진 목적함수

하중값을 가진 목적함수를 논문에서는 제시

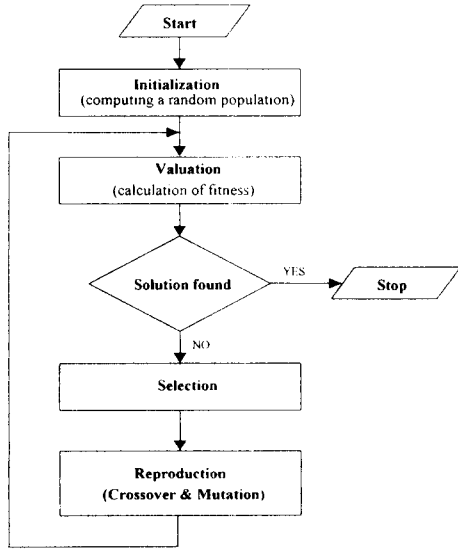


그림 2. 유전자 알고리즘

한다. 학습과 테스트 데이터에 대해 하중값을 가진 목적함수를 다음과 같이 정의한다.

$$f = \theta \times PI + (1 - \theta) \times E\_PI \quad (8)$$

$\theta$ 는 PI 와 E\_PI에 대한 하중값이고, PI는 학습에 대한 성능지수를, E\_PI는 테스트에 대한 성능지수 값을 나타낸다. 위에서 정의한 목적함수는 하중값의 설정에 따라 다음과 같은 특징을 가진다.

- $\theta=1$ 이면 모델은 테스트는 고려하지 않고 학습을 바탕으로 최적화된다.
- 반대로,  $\theta=0$ 일 경우 퍼지모델은 테스트을 바탕으로 전반부 파라미터는 최적화된다.
- $\theta=0.5$ 는 학습 및 테스트 모두가 같은 비중을 가지고 평가되는 경우이다.
- $\alpha \in [0, 1]$ 에 대해  $\theta = \alpha$  이면,  $\alpha$ 의 선택은 퍼지모델의 근사화와 일반화 사이에서 최적화에 대한 방향을 설정한다.

## 4. 시뮬레이션 및 결과고찰

### 4.1 가스로 공정

가스로 시계열 입출력 데이터를 학습과 테스트 데이터로 나누어 사용한다. 학습과 테스트 데이터의 성능지수를 유전자 알고리즘의 목적함수(식8)로 사용하여 모델을 최적화한다. 표1은 기존의 경험에 의해 퍼지 입력공간의 파라미터값을 동정한 성능지수와 유전자 알고리즘

에 의한 성능지수이다.  $\theta$ 의 선택에 따른 성능지수는 표2에 보였다.

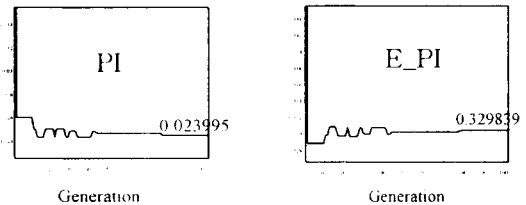
표1. 모델의 성능지수

구조	모델	PI	E_PI
1	기본모델	0.022855	0.335439
	최적화모델 ( $\theta = 0.5$ )	0.023456	0.344973
2	기본모델	0.024739	0.358243
	최적화모델 ( $\theta = 0.5$ )	0.018809	0.264046

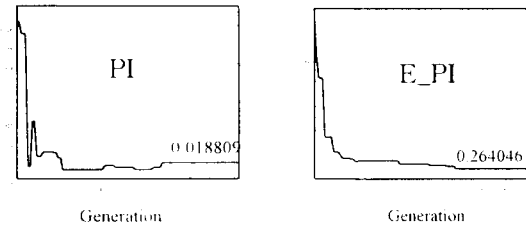
표2. 하중값을 가진 목적함수에 의한 성능지수

구조	Weight ( $\theta$ )	PI	E_PI
구조 1	1.0	0.022855	0.335407
	0.25	0.056886	0.316511
	0.1	0.025553	0.328181
	0.6	0.023995	0.329839
	0.75	0.022856	0.335137
구조 2	1.0	0.018360	0.350625
	0.25	0.029384	0.283629
	0.1	0.019560	0.278573
	0.6	0.018944	0.279414
	0.75	0.018859	0.279530
	0.0	0.031077	0.278938

그림3은 유전자 알고리즘에 의한 모델의 최적화 과정을 보여준다.



a) 구조 1의 최적화 ( $\theta = 0.6$ )



b) 구조 2의 최적화 ( $\theta = 0.5$ )

그림 3. 유전자 알고리즘에 의한 모델의 최적화 과정

### 4.2 하수처리 공정

수도권 하수처리장 중의 하나를 모델로 선정하여 1년치 수질 데이터를 토대로 퍼지 추론에 의한 모델링을 하고, 유전자 알고리즘으로 모델을 최적화한다. 시뮬레이션을 위해 2입력 1출력 구조를 사용하며, MLSS는 혼합액 부유물, WSR은 잉여오니 흐름율, RRSP는 반응율 설정치, DOSP는 용존산소 설정치를 의미한다. 표3은 기존모델과 유전자 알고리즘에 의한 성능지수이고,  $\theta$  값에 따른 모델의 성능지수는 표4와

같다.

표3. 구조 2의 성능지수

모델	입력1	입력2	PI	E.PI
기존모델	MLSS	WSR	6.396533	54.529442
	MLSS	RRSP	4.328988	222.968689
	MLSS	DOSP	8.332672	476.234711
	WSR	RRSP	7.798279	225.496414
	WSR	DOSP	8.622758	57.188206
최적화 모델 ( $\theta=0.5$ )	MLSS	WSR	7.087610	32.970261
	MLSS	RRSP	13.319167	65.584373
	MLSS	DOSP	7.556438	26.155487
	WSR	RRSP	8.657568	26.561729
	WSR	DOSP	6.871958	28.599873
	RRSP	DOSP	4.135979	66.810814

표4. 하중값을 가진 목적함수에 의한 성능지수

구조	입력1	입력2	Weight ( $\theta$ )	PI	E.PI
구조 1	MLSS	RRSP	1.0	9.917910	117.707330
			0.25	12.913282	13.499081
			0.4	12.913285	13.499073
			0.5	12.873838	13.531213
			0.6	12.210629	14.308290
			0.75	11.730998	15.339379
구조 2	WSR	DOSP	1.0	3.214219	822.611267
			0.25	6.108916	26.913485
			0.4	6.103705	27.942829
			0.6	6.269360	26.920230
			0.75	6.105598	27.116993
			0.0	6.136031	30.676236

구조 2의 동정된 파라미터는 그림4에 보이고, 그림5는 구조 1의 유전자 알고리즘에 의한 파라미터 탐색과정을 보여준다.

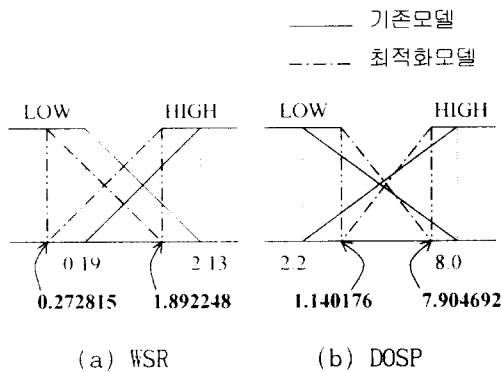


그림 4. 구조 2의 동정된 파라미터 ( $\theta=0.25$ )

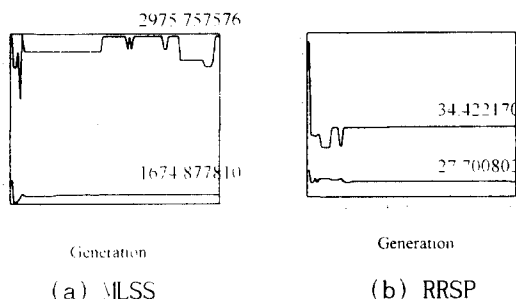


그림 5. 유전자 알고리즘에 의한 구조 1의 파라미터 탐색과정 ( $\theta=0.5$ )

## 5. 결론

본 논문에서 모델의 최적화를 위해 유전자 알고리즘을 사용하였으며, 퍼지추론 시스템 전 반부 파라미터를 동정함으로써 퍼지 입력공간을 정의하였고 비선형 시스템을 표현했다. 또한 하중값을 가지는 목적함수에 의해 퍼지모델의 근사화와 일반화 사이에서 상호 최적화에 대한 방향을 제시하였다. 하중값 변화에 따라 학습과 테스트 사이의 관계를 표현함으로써, 그에 따른 모델의 최적화를 이룰 수 있었다. 시뮬레이션에서 볼 수 있듯이 하중값이 가지는 특성에 따라 학습과 테스트 사이의 최적화 방향을 설정하였고, 복잡하고 비선형이 강한 공정에 적용함으로써 보다 향상된 성능을 가진 모델을 구축하였다.

## 6. 참고문헌

- [1] R.M. Tong, "Synthesis of fuzzy models for industrial processes", *Int. J. Gen. Syst.*, Vol. 4, pp.143-162, 1978.
- [2] C.W. Xu, "Fuzzy system identification", *IEE Proceeding* Vol. 126, No. 4, pp.146-150, 1989.
- [3] Box and Jenkins, "Time Series Analysis, Forecasting and Control", Holden Day, San Francisco, CA.
- [4] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control", *IEEE Trans. Syst. Cybern.*, Vol. SMC-15, No. 1, pp.116-132, 1985.
- [5] David E. Goldberg, "Genetic Algorithms in search, Optimization & Machine Learning", Addison-wesley.
- [6] Zbigniew Michalewicz, "Genetic Algorithms + Data Structure = Evolution Programs", Springer-Verlag.
- [7] 오성권, 김현기, 우광방, "활성오니공정의 퍼지 모델링과 제어시스템", 대한 전기학회 논문집 40권 10호 PP.1072-1083, 1990
- [8] 오성권, 우광방, "퍼지추론 방법에 의한 퍼지동정과 하수처리공정시스템 응용", 대한전자공학회 논문집 제 31권 B편 제 6호, PP.43-52, 1994.