

論文94-31B-6-6

퍼지추론 방법에 의한 퍼지동정과 하수처리공정시스템 응용

(Fuzzy Identification by means of Fuzzy Inference Method and its Application to Waste Water Treatment System)

吳聖權*, 周永焄**, 南宜錫**, 禹廣芳**

(Sung Kwun Oh, Young Hoon Joo, Eui Suck Nam and Kwang Bang Woo)

要約

규칙베이스 퍼지모델링의 설계 방법을 복잡하고 비선형적인 시스템의 모델동정을 위하여 제안한다. 제안된 규칙베이스 퍼지모델은 최적화이론, 퍼지구현규칙과 퍼지 c-means 클러스터링을 사용하여 "IF..., THEN..."의 효율적인 형태로 시스템 구조와 파라미터 동정을 시행한다. 이 논문에서 간략추론, 선형추론, 변형된 선형추론의 3가지 종류의 퍼지모델링 방법을 제시한다. 퍼지구칙의 전반부 구조와 파라미터를 동정하기 위하여 퍼지 클러스터링, 수정된 컴플렉스법이 사용되고, 최적 후반부 파라미터를 동정하기 위해 최소자승법이 사용된다. 가스공정의 시계열 데이터와 하수처리공정의 데이터가 제안된 규칙 베이스 퍼지 모델링을 성능평가 하기위해 사용된다. 제안된 모델링 방법이 다른 논문과 비교시 더 높은 정확도를 가진 퍼지모델을 생성함을 보인다.

Abstract

A design method of rule-based fuzzy modeling is presented for the model identification of complex and nonlinear systems. The proposed rule-based fuzzy modeling implements system structure and parameter identification in the efficient form of "IF..., THEN...", using the theories of optimization theory, linguistic fuzzy implication rules and fuzzy c-means clustering. Three kinds of method for fuzzy modeling presented in this paper include simplified inference (type 1), linear inference (type 2), and modified linear inference (type 3). In order to identify premise structure and parameter of fuzzy implication rules, fuzzy c-means clustering and modified complex method are used respectively and the least square method is utilized for the identification of optimum consequence parameters. Time series data for gas furnace and those for sewage treatment process are used to evaluate the performance of the proposed rule-based fuzzy modeling. Comparison shows that the proposed method can produce the fuzzy model with higher accuracy than previous other studies.

I. 서론

* 正會員, 圓光大學校 制御計測工學科

(Dept. of Cont. & Inst. Eng., Wonkwang Univ.)

** 正會員, 延世大學校 電氣工學科

(Dept. of Elec. Eng., Yonsei Univ.)

接受日字 : 1993년 5월 10일

퍼지 집합이론이 1965년 Zadeh에 의해 창안된 이래, 퍼지시스템 동정이 최근 퍼지시스템 이론의 주요 연구대상이 되었다.^{1,2} 비선형적이고 복잡한 실 시스템의 특성을 해석하는데 수학적 모델은 만족스러운

결과를 주지 못한다. 이러한 어려움을 극복하기 위한 방법으로 이들 시스템의 정적 혹은 동적 특성을 묘사하기 위해 퍼지모델을 사용한다.^{1,3,4} 퍼지시스템 이론의 발전으로 퍼지모델 동정 알고리즘의 접근 방식도 거듭된 연구 결과로 향상되었다. 초기 퍼지모델의 동정연구로 언어적 접근방식^{1,2} 과 퍼지관계 방정식에 기초한 접근방식^{3,4}의 두 가지 방법이 제안되었다. 언어적 접근방식에서, Tong은 논리적 조사(logical examination) 방법에 의해 가스로 공정을 동정하였고⁵, B. Li 등은 Tong의 방법을 수정함으로써 더 좋은 결과를 얻었으며⁶, 그들은 이외에 결정테이블(decision table)에 기초한 적응모델 수정 알고리즘을 제안하였으나, 제안된 알고리즘은 필요한 컴퓨터 용량 및 계산시간 때문에 고계 다변수 시스템에의 적용에는 문제점이 발생하였다.⁵ Pedrycz는 언어적 모델링에서 기준퍼지집합(referential fuzzy set) 개념에 의해 해석하였고² T. Li 등은 간단한 SISO 퍼지모델에 대한 자기학습 알고리즘을 제안하였다.⁵ 퍼지관계 방정식에 기초한 접근방식에서, Pedrycz는 퍼지관계 방정식에 의해 새로운 합성규칙인 기준퍼지집합(referential fuzzy set)과 Zadeh의 제한적 가능성분포(conditional possibility distribution)의 개념에 기초해서 동정하였다.⁵ 그리고, Xu는 이 두가지 방법에 기초한 기준퍼지집합 이론과 자기학습 알고리즘에 의해 퍼지관계 모델을 구성하여 동정하였다.^{5,6} 이들 두 방식에서 사용되는 직접추론방식은 혼합(선형)추론방식 보다 그 성능이 떨어진다. Sugeno는 자신이 제안한 선형퍼지추론식에 경험적 방법을 통한 구조동정으로 컴플렉스법과 최소자승법에 의해 시스템을 동정하였으나¹⁰, 전반부 구조 동정에서 반복적인 퍼지분할에 의한 경험에 의존하였다. 그리고 그는 퍼지 클러스터링방법에 의해 가스로 공정을 퍼지 동정하였으나⁸, 이 또한 직접 추론법에 의한 것으로 그 동정성능이 떨어진다. 본 논문에서는 간략추론(구조 1), 선형추론(구조 2)과 변형된 선형추론(구조 3)의 세가지 퍼지추론 방식에 퍼지 클러스터링방법과 자동동조 알고리즘인 최적 수렴의 정도를 보다 향상시킨 수정된 컴플렉스법을 이용하여 개선된 알고리즘을 제시하였고, 이의 평가를 위해 Box와 Jenkins가 사용한 가스로 시계열 입출력 데이터⁹와 하수처리공정 데이터¹⁴를 모델링함으로써 기존의 퍼지모델링방법^{2,5,7,8,14}과 동정오차면에서 비교하였다.

II. 퍼지 추론에 의한 모델링

규칙베이스 퍼지 모델의 동정 알고리즘은 전반부와

후반부의 동정으로 분류된다. 전반부에서는 1)퍼지구현규칙의 전반부 입력변수 x_1, x_2 등의 선택, 2)조건 설정 즉 전반부 변수공간의 퍼지 분할(small, large)의 결정이다. 즉 입력공간의 최적 퍼지공간분할의 수를 결정하는것으로 퍼지 부분공간의 수는 구현규칙수와 동일하게 된다. 그리고 3)퍼지변수의 멤버십 파라미터 결정이 필요하다. 후반부에서는 1)구현규칙의 구조 즉 후반부의 변수의 선택과, 2)후반부의 파라미터의 계수 결정이 필요하다. 전반부의 모든 입력변수, 퍼지집합의 멤버십함수는 입력 변수의 공간을 규칙수만큼의 퍼지분할된 부분공간으로 나누는 것과 관련되고, 후반부의 파라미터는 각각 퍼지 분할된 부분공간에서 입,출력 관계를 묘사한다. 그러므로 퍼지모델링에서는 규칙수의 결정을 위해 구현규칙의 구조동정 및 파라미터들의 동정이 중요한 문제이다. 본 논문에서는 전반부 구조동정으로 퍼지 C-means 클러스터링 알고리즘을 사용하고, 전반부 파라미터 동정에는 수정된 컴플렉스 알고리즘을 사용한다. 그리고 후반부 구조는 3가지 구조로 표현한다. 즉 구조 1, 구조 2, 구조 3으로 각각 단일의 상수, 일차선형식, 본 논문에서 제안한 변형된 일차선형식의 구조를 사용한다. 후반부 파라미터동정으로 최대 피벗팅(maximal pivoting)알고리즘을 가지는 가우스 소거법에 의한 최소자승법을 사용한다. 퍼지 모델링의 전체 구성도는 그림 1.과 같다.

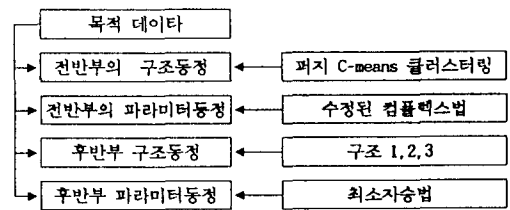


그림 1. 퍼지 모델링의 전체 구성도

Fig. 1. Overall scheme of fuzzy modeling.

III. 퍼지 동정 알고리즘

1. 퍼지 모델의 전반부 동정

기존의 전반부 구조 동정은 변수증가법에 따라 각 변수 선택에 의해 퍼지분할의 수를 하나씩 증가시키면서 최적구조를 찾아내는 경험적인 추적방법¹⁰이 사용되었으나, 본 논문에서는 퍼지 클러스터링 알고리즘을 사용하여 구조동정을 행한다. 퍼지 클러스터링은 각 클러스터의 중심벡터와 입력벡터와의 기하학적 거리 곱하기 각 입력벡터가 클러스터에 속하는 멤

버섯함수 값으로 표시된 목적함수를 최소로하는 클러스터 수와 클러스터 중심벡터를 구하기 위한 것으로 목적함수의 반복적인 최소화를 통해 입력벡터 공간을 퍼지분할하는 알고리즘이다. 따라서 입력벡터 공간의 퍼지분할 기능을 갖는 퍼지 클러스터링을 이용하여 퍼지규칙을 구성하는 퍼지조건문의 입력변수 공간을 퍼지분할하고 최적의 분할 수로부터 퍼지조건문 수를 얻을 수 있다. 본 논문에서는 퍼지 클러스터링 알고리즘으로 기존의 SCM(Soft C-Means) 클러스터링이 사용된다. [12,13]

본 논문에서는 퍼지분할의 압축성과 분리성을 측정하는 퍼지 클러스터링 평가규범인 식(1)의 S를 사용한다.

$$S = \frac{\sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|X_k - v_i\|^2}{n \cdot \min_{i,j} \|v_i - v_j\|^2} \quad (1)$$

여기서 v_i 와 v_j 는 클러스터 중심 벡터이다.

전반부 파라미터 동정은 컴플렉스 개념에 기존의 최적화 이론인 심플렉스 개념을 확장, 융합함으로써 다음과 같은 수정된 컴플렉스법을 제안한다. 이 수정된 컴플렉스법은 목적함수를 최소화시키는 방법인 심플렉스법에 제약조건이 주어지는 경우로 식 (2)의 형태의 수식을 풀기위한 방법으로 반사, 확장, 축소 방법의 각 단계를 반복 수행함으로써 수렴성을 보다 향상시킨다.

$$\begin{aligned} & \text{Minimize } f(x) \\ & \text{Subject to } g_j(x) \leq 0, \quad j = 1, 2, \dots, m \\ & x_i^{(l)} \leq x_i \leq x_i^{(u)} \quad i = 1, 2, \dots, n \end{aligned} \quad (2)$$

여기서, l은 하한(lower bound), u는 상한(upper bound)을 나타낸다.

〈단계 1〉

모든 제약조건을 만족시키는 K개의 해집합($X_1, X_2, \dots, X_k, k \geq n$)을 대략 $K=2n$ 개를 선정한다. (n은 초기 정점의 수)

〈단계 2〉

심플렉스의 기본개념에서 반사(Reflection), 확장(Expansion), 축소(Contraction)의 기본개념을 이용해서, α, β, γ 의 초기값을 설정한다. 즉,

i) Reflection : $X_r = X_0 + \alpha(X_n - X_0)$ (3)

ii) Expansion : $X_e = X_0 + \gamma(X_r - X_0)$ (4)

iii) Contraction : $X_c = X_0 + \beta(X_n - X_0)$ (5)

여기서, X_0 는 모든 정점 X_i 의 중심, X_n 는 목적 함수 값이 최대로 되는 정점이고, X_r, X_e, X_c 는 각각 식 (3)-(5)의 반사, 확장 축소방법에 의해 계산된 값이다.

〈단계 3〉

초기 평가치로써 최대, 최소값($f(X_n), f(X_0)$)을 구하고, 주어진 모든 조합(X_k)의 중심값을 구한다. 반사방법을 이용하여 X_r 를 구한다. 즉, $X_r = X_0 + \alpha(X_n - X_0)$. 여기에서 X_r 의 제약조건 및 그외의 조건을 체크한 후 만족치 않을 때 $X_r = (X_0 + X_r)/2$ 을 구한 후 반복하여 X_r 조건을 체크한다.

〈단계 4〉

반사방법(reflection)에 의한 $f(X_r)$ 를 초기 최소 평가치와 비교하여 최소값보다 적을 경우 확장방법(expansion)을 이용하여 X_e 를 구한다. 즉, 여기서 X_e 의 제약조건 및 그외의 조건을 체크한 후 만족치 않을 때 $X_e = (X_0 + X_e)/2$ 을 구한 후 반복하여 X_e 조건을 체크한다. $X_e = X_0 + (X_r - X_0)$. 이 X_e 의 평가치 $f(X_e)$ 를 구한다음 $f(X_r)$ 및 $f(X_0)$ 를 초기 최소 평가치와 비교하여 초기 최대 평가치를 가지는 X_n 에 치환한다.

〈단계 5〉

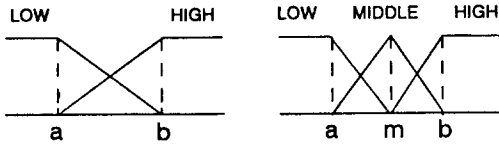
반사방법(reflection)에 의한 $f(X_r)$ 이 평가치 최대 값을 제외한 모든 값과 비교하여 큰 경우에 축소방법(contraction)을 이용하여 X_c 를 구한다. 즉, $X_c = X_0 + (X_n - X_0)$. 여기서 X_c 의 제약조건 및 그외의 조건을 체크한 후 만족치 않을 때 $X_c = (X_0 + X_c)/2$ 를 구한 후 반복하여 X_c 조건을 체크한다. 이 X_c 의 평가치 $f(X_c)$ 를 구한다음 $f(X_c)$ 가 $f(X_n)$ 보다 적으면 X_n 는 X_c 로 치환되나 그렇지 않으면 $X_1 = (X_k + X_1)/2$ 로 치환한다.

〈단계 6〉

$$Q = \left(\sum_{i=1}^{n+1} \frac{[f(X_i) - f(X_0)]^2}{n+1} \right)^{1/2} \leq \epsilon \quad (6)$$

인 것을 검정한 후 부등식을 만족치 않으면 〈단계 3〉으로 간다.

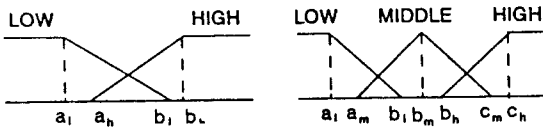
멤버쉽 함수로 본 동조 알고리즘 적용시 우수한 특성을 가지는 삼각형 함수를 사용한다. 삼각형 형태인 멤버쉽함수를 갖는 퍼지변수가 2종류 및 3종류의 경우에 2가지 구조의 형태를 제안한다. 그림 2는 멤버쉽함수의 각 퍼지 변수의 정점이 고정인 상태에서 조정되는 경우를 나타내고, 그림 3은 멤버쉽 함수의 각 퍼지변수의 정점이 달리 조정되는 경우를 나타낸다.



(a) 퍼지 변수 2인 경우 (b) 퍼지 변수 3인 경우

그림 2. 튜닝하기 위한 퍼지변수들의 파라미터 a, m, b (구조 a)

Fig. 2. Parameters a, m, b for tuning in the fuzzy variables. (Type a)



(a) 퍼지변수 2인 경우 (b) 퍼지변수 3인 경우

그림 3. 튜닝하기 위한 퍼지변수들의 파라미터 a1, bm, am, bm, cm, ah, bh, ch (구조 b)

Fig. 3. Parameter a1, bm, am, bm, cm, ah, bh, ch for tuning in the fuzzy variables. (Type b)

퍼지모델링을 하기 위해 구조 a, 구조 b의 선택은 시스템의 입출력관계, 입출력 변수의 수, 비선형성등을 고려하여 선택하도록 한다.

2. 퍼지 모델의 후반부 동정

퍼지모델의 구조동정에서 후반부 구조로는 서로 다른 3가지의 퍼지추론에 의해 구별되는 구조 1,2,3을 사용한다. 3가지 구조에의한 퍼지규칙추론과 파라미터동정을 위한 매개변수 추정은 다음과 같다.

구조 1(후반부 : 상수)

후반부가 단일의 상수항만을 가지는 것으로 이와같은 추론법을 간략추론법이라한다. 이 퍼지모델은 식 (7)의 형태를 가지는 구현규칙들로 구성된다.

Rⁱ : If x₁ is A_{i1}, ..., and x_k is A_{ik}, then y=a_i

$$y^o = \frac{\sum_{i=1}^n w_i a_i}{\sum_{i=1}^n w_i} \tag{7}$$

여기서 Rⁱ는 i(i=1, ..., n)번째 규칙, x_j(j=1, ..., k)는

입력변수, A_{ij}(i=1, ..., n; j=1, ..., k)는 퍼지집합의 멤버쉽 함수, a_i는 상수이고, n은 퍼지규칙수, 그리고 y^o는 추론된 값이다. 후반부 파라미터동정에서 전반부 입력변수 및 파라미터가 주어지면, PI(Performance index)를 최소화하는 최적 후반부 파라미터를 결정할 수 있다. 이 PI는 성능지수로 식 (8)로 정의된다.

$$PI = \sum_{i=1}^m \frac{(y_i - y_i^o)^2}{m} \tag{8}$$

여기서 y_i^o(=∑_{i=1}ⁿ w_i a_i / ∑_{i=1}ⁿ w_i)는 입력 X^o(=(x₁^o, x₂^o, ..., x_k^o))일때의 출력이며 w^o=∏_{i=1}^k A_i(x_i) n은 구현규칙 수, k는 입력변수의 수, m은 데이터 수이다.

후반부의 파라미터는 a_i로서 입출력 데이터가 x_{1i}, x_{2i}, ..., x_{ki}-y_i(i=1, 2, ..., m)의 집합으로 주어졌을때 최소자승법에 의해 구해진다. 구조 1의 퍼지모델에 있어서 최소자승법에 의한 매개변수 추정은 식 (10)과 같다.

$$\begin{aligned} V(a, m) &= 1/2 \cdot \sum_{i=1}^m \epsilon_i^2 = 1/2 \cdot \sum_{i=1}^m (y_i - y_i^o)^2 = 1/2 \cdot \sum_{i=1}^m \left[y_i - \sum_{j=1}^m a_j w_{ji} \right]^2 \\ &= 1/2 \cdot \sum_{i=1}^m [y_i - x_i^T a]^2 = 1/2 \cdot \|E\|^2 \end{aligned} \tag{9}$$

여기서

$$w_{ji} = \frac{A_{j1}(x_{1i}) * \dots * A_{jk}(x_{ki})}{\sum_{j=1}^m A_{j1}(x_{1i}) * \dots * A_{jk}(x_{ki})}$$

j:규칙수 i:데이터수 m:데이터 총수 n:규칙 총수이다.

그리고, x_i^T = [w_{1j}, ..., w_{nj}] , a^T = [a₁, ..., a_n] ,

$$Y = [y_1, y_2, \dots, y_m]^T,$$

$$E = [\epsilon_1, \dots, \epsilon_m]^T,$$

$$X = [x_1 \ x_2 \ x_m]^T$$

식 (9)에서 E = Y - Xa 이므로 최소자승법에 의한 최소값은 식 (17)에 의해 구해진다.

$$\hat{a} = (X^T X)^{-1} X^T Y \tag{10}$$

구조 2(후반부 : 일차 선형식)

후반부가 일차 선형식인 경우의 추론법을 선형추론법 혹은 복합추론법이라 한다. 이 퍼지 모델은 식 (11)의 형태를 가지는 구현규칙들로 구성된다.

$$\begin{aligned} R^i : \text{If } x_1 \text{ is } A_{i1}, \dots, \text{ and } x_k \text{ is } A_{ik}, \text{ then } y &= f_i(x_1, \dots, x_k) \\ f_i(x_1, \dots, x_k) &= a_{i0} + a_{i1}x_1 + \dots + a_{ik}x_k \\ y^o &= \frac{w_1 f_1(x_1^o, \dots, x_k^o) + \dots + w_n f_n(x_1^o, \dots, x_k^o)}{w_1 + w_2 + \dots + w_n} \end{aligned} \tag{11}$$

여기서 Rⁱ는 i번째 규칙, x_j는 입력변수, A_{ij}는 퍼지

집합의 멤버십함수, 그리고 $a_{ij}(i=1, \dots, n; j=0, \dots, k)$ 는 후반부의 파라미터 y^o 는 추론된 값이다. 후반부 파라미터는 최소자승법에 의한 최소값으로 식 (12)로 구해진다.

$$\hat{a} = (X^T X)^{-1} X^T Y \quad (12)$$

$$\begin{aligned} x_i^T &= [w_{i1}, \dots, w_{in}, x_{i1}, w_{i1}, \dots, x_{ik}, w_{ik}, \dots, x_{ik}, w_{ik}] \\ \hat{a}^T &= [a_{i0}, \dots, a_{in}, a_{i1}, \dots, a_{i1}, \dots, a_{ik}, \dots, a_{ik}] \\ Y &= [y_1, y_2, \dots, y_m]^T \\ X &= [x_1, x_2, x_m]^T \end{aligned}$$

구조 3 (후반부 : 변형된 일차선형식)

후반부가 전반부 적합도의 함수로 표현되는 일차선형식인 경우, 이와같은 추론법을 변형된 선형 추론법이라 한다. 이 퍼지모델은 식 (13)의 형태를 가지는 구형규칙들로 구성된다.

$$\begin{aligned} R^i : \text{If } x_i \text{ is } A_{i1}, \text{ and } x_k \text{ is } A_{ik}, \text{ then } y &= f_i(w_i) \\ f_i(w_i) &= a_{i0} + a_{i1}w_i \end{aligned} \quad (13)$$

여기서, R_i 는 i 번째 규칙, x_j 는 입력변수, A_{ij} 는 퍼지집합의 멤버십함수, 그리고 $a_{ij}(i=1, \dots, n; j=0, 1)$ 는 후반부의 파라미터이다. 식 (13)의 입력 $x_1^o, x_2^o, \dots, x_k^o$ 에 대한 전반부 적합도 w_1, w_2, \dots, w_n 는 대수곱에 의해 구해진다. 후반부의 출력 $f_i(w_i)$ 는 입력이 w_i 인 후반부 선형식에 의해 결정되며 전체의 추론결과는 식 (14)와 같다.

$$\begin{aligned} y^o &= \frac{w_1 f_1(w_1) + \dots + w_n f_n(w_n)}{w_1 + w_2 + \dots + w_n} \\ &= \frac{w_1 f_1(a_{10} + a_{11}w_1) + \dots + w_n (a_{n0} + a_{n1}w_n)}{w_1 + w_2 + \dots + w_n} \end{aligned} \quad (14)$$

후반부 파라미터는 최소자승법에 의해 식 (15)에 의해 구해진다.

$$\hat{a} = (X^T X)^{-1} X^T Y \quad (15)$$

여기서,

$$\begin{aligned} x_i^T &= [w_{i1}, \dots, w_{in}, w_{i1}, w_{i1}, \dots, w_{in}, w_{in}] \\ \hat{a}^T &= [a_{i0}, \dots, a_{in}, a_{i1}, \dots, a_{i1}] \\ Y &= [y_1, y_2, \dots, y_m]^T \\ X &= [x_1, x_2, x_m]^T \end{aligned}$$

IV. 시뮬레이션 및 결과고찰

1. 가스공정의 경우

본 절에서는 II, III절에서 제안한 규칙베이스 퍼지 모델링 방법을 적용함으로써 Box 와 Jenkins가 사용한 가스로 시계열 데이터를 이용하여 그림 4.와 같이 입출력 데이터인 가스 흐름율과 연소된 이산화탄소의 농도의 가스로 공정을 퍼지 모델링한다. 메탄 가스 흐름율($X(t)$)은 실험에서 -2.5 에서 2.5로 변하는 것을 사용하나 실제 공정에서 메탄 가스 흐름율($X_m(t)$)은 식(16)과 같이 0.5에서 0.7로 변한다.^[9]

$$X_m(t) = 0.60 - 0.048 * X(t) \quad (16)$$

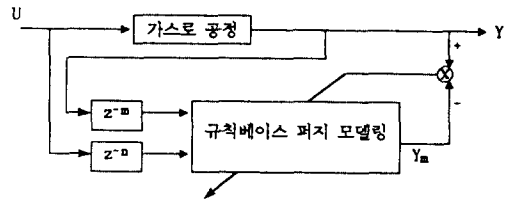


그림 4. 가스로 공정에서 규칙베이스 퍼지모델링을 위한 구조

Fig. 4. Structure for rule-based fuzzy modeling in the gas furnace.

그림 4에서 U는 입력 데이터로 가스 흐름율, Y는 출력데이터로 연소된 이산화탄소 농도, Ym은 모델 출력을 나타낸다. 가스로 시계열 입출력 데이터 296쌍을 사용하여 퍼지 추론에 의한 모델링을 한다. 이 퍼지 동정의 단계별 수행절차는 다음과 같다.

<단계 1>

이 단계에서는 전반부 구조 동정을 하게되며, 전반부 입력 변수로 $y(t-1) \sim y(t-3)$, $u(t-1) \sim u(t-6)$ 의 9개의 입력을 고려한다. 입력 변수 $y(\cdot)$ 와 $u(\cdot)$ 가 서로 독립적이어서 각각의 입력 변수에 대한 입출력 관계로부터 전체 입출력 관계를 얻는다. 즉, 가스로 입출력 $y(t-1) \sim y(t-3)$, $u(t-1) \sim u(t-6)$, $y(t)$ 의 데이터 집합을 각 입력변수와 출력 변수의 관계 $(u(\cdot), y(t))$ 와 $(y(\cdot), y(t))$ 로 된 데이터 집합을 분할한다. 각 분할된 데이터 집합을 사용하여 2개의 클러스터 수로부터 하나씩 증가시키면서 클러스터링하고 퍼지분할 수 결정을 위해 식(1)에 정의된 평가 규범 S를 계산한다. 보다 적은 S값이 입출력 공간의 적절한 퍼지 분할수를 표시한다. 표 1은 $(u(t-4), y(t))$, $(u(t-3), y(t))$, $(y(t-1), y(t))$ 로 된 데이터 집합에서 클러스터링에 의한 각 분할수에서의 평가 규범을 나타낸다.

표 1. 각 데이터집합의 분할 수에서의 평가규범
Table 1. Evaluation criterion in the partition number of each data set.

데이터 집합	분할수					
	2	3	4	5	6	7
$((u(t-4),y(t)))$	0.118	0.124	0.118	0.102	0.149	0.141
$((u(t-3),y(t)))$	0.119	0.125	0.120	0.103	0.152	0.148
$((y(t-1),y(t)))$	0.112	0.114	0.123	0.114	0.148	0.148

표 1에 나타난 데이터 집합이 동정에 사용시 가장 우수한 성능 특성을 나타낸다. 표 1에서 $(u(t-4),y(t))$ 와 $(y(t-1),y(t))$ 데이터 집합에 대해 각 클러스터 수에 대한 퍼지 분할 평가규범 S를 보면 $(y(t-1),y(t))$ 는 클러스터 수가 2일때 S값이 가장 작았으나 $(u(t-4),y(t))$ 는 클러스터 수가 5 일때 S값이 가장 작다. 구조동정의 궁극적인 목적이 클러스터링 목적 함수를 최소화시키는 것이 아니라 식 (17)로 정의된 PI 즉 실제 데이터 출력과 퍼지모델 출력사이의 오차를 최소화시키는 것이므로 가장 적은 S를 갖는 분할수가 PI(performance index)를 최소화시킨다는 보장이 없다.

$$PI = (1/N) \cdot \sum_{k=1}^N (y(k) - y^o(k))^2 \quad (17)$$

여기서, N은 데이터 수, $y^o(k)$ 는 실제 데이터 출력 값이고, $y_o(k)$ 는 동정된 모델로부터 추론된 값이다. 따라서 적절한 퍼지 분할 수에 대한 몇개의 해를 얻는데 평가규범이 사용되므로 $(u(t-4),y(t))$ 의 가능한

분할 수로 클러스터 수 2,3,4,5가 선정되고 퍼지 규칙 수의 최소화과 식(17)의 PI의 최소화를 고려하여 클러스터수를 최종 선정하므로 선정된 최종 분할수는 $(u(t-4),y(t))$ 는 2 혹은 3이며, $(y(t-1),y(t))$ 의 경우는 2이다.

<단계 2>

전반부 멤버쉽함수의 파라미터를 결정하기 위해서 그림 2의 구조 a와 그림 3의 구조 b의 두 경우에서 PI를 최소화시키는 파라미터를 수정된 컴플렉스법인 자동동조 알고리즘에 의해 구한다. 그림 5(a),(c)는 입력변수 $u(t-4)$ (혹은 $u(t-3)$)가 각각 2,3개로 퍼지 분할된 경우이고 (b),(d)는 입력변수 $y(t-1)$ 가 2개로 퍼지 분할된 경우로써 튜닝할 멤버쉽함수의 파라미터를 가지는 전반부 구조 a를 나타낸다. 그림 6은 입력 변수를 $u(t-4)$ (혹은 $u(t-3)$), $y(t-1)$ 의 튜닝할 멤버쉽함수의 파라미터를 가지는 전반부 구조 b를 나타낸다.

각 멤버쉽 함수의 초기값은 클러스터 중심값 혹은 각 입력 데이터의 최대, 최소값을 사용한다. 전반부 각 입력 퍼지변수의 멤버쉽 함수의 파라미터를 동정하기위해 수정된 컴플렉스법을 사용한다. 투영(reflection), 확장(expansion)과 축소(contraction) 방법의 파라미터의 초기값 $\alpha=1.0$ $\gamma=2.0$ $\beta=0.5$ 로 설정한다.

<단계3>

후반부는 3가지 구조인 구조 1,2,3을 사용한다. 후반부 파라미터 동정으로 최대 퍼벗팅(maximal pivoting) 알고리즘을 가지는 가우스 소거법에 의한

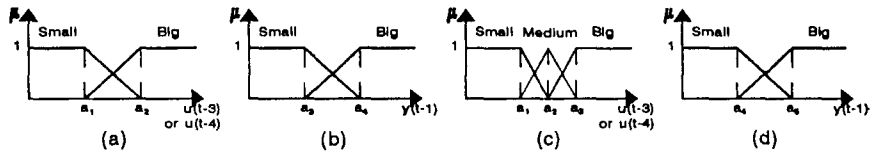


그림 5. 구조 a에서 동조할 파라미터를 가지는 입력 변수들의 멤버쉽함수
Fig. 5. Membership functions of input fuzzy variables with parameters to tune in the type a.

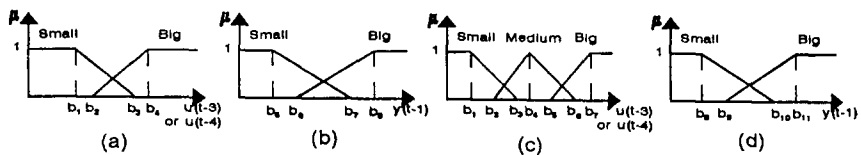


그림 6. 구조 b에서 동조할 파라미터를 가지는 입력 변수들의 멤버쉽함수
Fig. 6. Membership functions of input fuzzy variables with parameters to tune in the type b.

최소 자승법을 사용한다. 표 4는 각 후반부 구조 1, 2,3과 전반부 구조 a,b로 구성된 각 모델에 대하여 퍼지 동정으로 얻어진 최적 구현규칙들의 성능지수를 나타낸다.

표 2. 구조 1,2,3을 가지는 각 모델에 대한 최적 성능지수

Table 2. Optimal performance index for each model with type 1, 2, 3.

Model No.	후반부구조	전반부구조	입력 변수	규칙수	PI
1	Type 1	Type a	$u(t-3), y(t-1)$	4	0.164
2	Type 1	Type b	$u(t-3), y(t-1)$	4	0.137
3	Type 1	Type a	$u(t-3), y(t-1)$	6	0.147
4	Type 1	Type b	$u(t-3), y(t-1)$	6	0.134
5	Type 2	Type a	$u(t-3), y(t-1)$	4	0.129
6	Type 2	Type b	$u(t-4), y(t-1)$	4	0.125
7	Type 2	Type a	$u(t-3), y(t-1)$	6	0.106
8	Type 2	Type b	$u(t-4), y(t-1)$	6	0.098
9	Type 3	Type a	$u(t-3), y(t-1)$	4	0.142
10	Type 3	Type b	$u(t-3), y(t-1)$	4	0.133
11	Type 3	Type a	$u(t-4), y(t-1)$	6	0.137
12	Type 3	Type b	$u(t-3), y(t-1)$	6	0.136

본 논문에서 제시한 퍼지동정 알고리즘에 의해 구조 1에서 최적퍼지모델은 Model 4로 성능지수 0.134이고, 구조 2에서 최적퍼지모델은 Model 8로 성능지수 0.098이고, 구조 3에서 최적 퍼지모델은 Model 10으로 성능지수 0.133 이다. 이상에서 구한 각 구조의 최적 퍼지 규칙중에서 가장 우수한 성능 특성은 후반부가 선형식인 구조 2인 경우인 표 2의 모델 8이다. 이 모델 8의 출력 데이터와 가스로의 실 데이터의 비교는 그림 7과 같다.

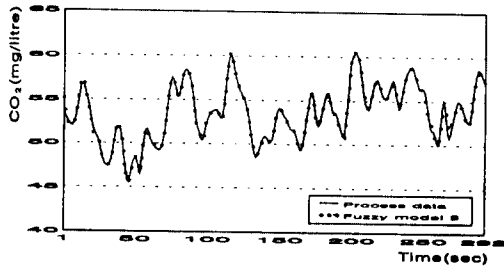


그림 7. 실 데이터와 모델 8의 출력 데이터의 비교
Fig. 7. Comparison of original data and output data for model 8.

표 2에서의 구조 1,2,3에서 각 규칙수가 4개 혹은 6개인 경우의 동정오차(혹은 성능지수)를 표 3에서

다른 퍼지 모델의 동정오차와 비교한다.

표 3. 다른 퍼지모델링 방법과 동정오차 비교
Table 3. Comparison of identification error with other fuzzy modeling methods.

Model name	Mean square error	Number of rules
Tong's model[7]	0.469	19
Pedrycz's model[2]	0.776	20
Xu's model[5]	0.328	25
Sugeno's model[8]	0.355	6
Our model	Type 1 (Model 2)	0.137
	Type 1 (Model 4)	0.134
	Type 2 (Model 6)	0.125
	Type 2 (Model 8)	0.098
	Type 3 (model 10)	0.133
Type 3 (Model 12)	0.136	6

2. 하수처리공정의 경우

하수처리시스템의 개요도는 그림 8과 같다.

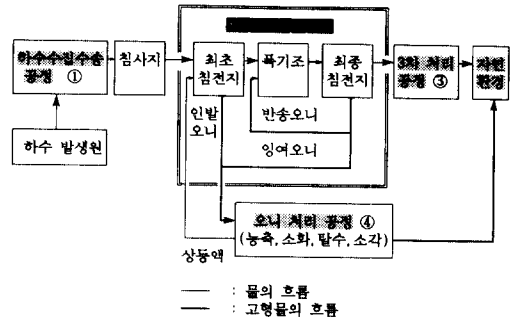


그림 8. 하수처리시스템의 개요도

Fig. 8. Configuration of sewage treatment system.

하수처리에는 활성오니공정이 일반적으로 사용되고 있다. 활성오니공정은 침사지, 최초침전지, 폭기조 및 최종침전지를 포함한다. 침사지와 최초침전지에서는 하수중에 포함된 현탁물질의 입자가 중력에 의해 침강된다. 폭기조에서는 수시간 연속적으로 하수중에 공기가 흡입된다. 유기물 제거의 중심을 이루는 미생물은 폭기조의 과정에서 자연발생하는 다종류의 미생물 군이고 이를 활성오니라 한다. 응집한 미생물의 덩어리를 플록(floc)이라 한다. 활성오니는 폭기조내 공기로부터 용해한 산소에 의해 유기물을 생물화학적 산화해서 증식하고 대사과정에서 유기물을 물과 탄산가스로 분해한다. 최종침전지에서 플록(floc)은 침강 제거되고 맑은 물이 3차처리 공정으로 유입된다. 한편 최종 침전지에서 침강한 미생물은 폭기조

입구에서 재순환 되어 다시 유기물 제거에 사용된다.

활성오니공정은 폭기조 및 최종침전지를 하나의 단일 모델로 선정하였다. 실제 최초침전지로 유입되는 생물학적 산소요구량(Biological Oxygen Demand : BOD) 및 유입수의 부유물(Suspended Solid : SS) 농도를 측정하고, 최종침전지에서 유출되는 생물학적 산소요구량(Effluent Boiological Oxygen Demand : EBOD) 및 부유물 농도(Effluent Suspended Solid : ESS)가 주요 측정 대상이 된다. 활성오니공정의 용존산소 설정치(Dissolved Oxygen Set-Point : DOSP) 및 반송율 설정치(Recycle Sludge Ratio Set-Point : RRSP)는 유입수의 생물학적 산소요구량 및 부유물 농도에 따라 유출수의 생물학적 산소요구량(EBOD), 부유물(ESS) 농도가 영향을 받으므로 조작자가 이 유출수의 농도를 소정치 이하로 낮추기 위해 사용되는 데이터이다. 유출수의 생물학적 산소요구량(EBOD) 및 부유물(ESS) 농도는 측정된 데이터의 혼합액 부유물(Mixed Liquid Suspended Solid : MLSS) 농도, 잉여오니흐름율(Waste Sludge Ratio : WSR), 반송율 설정치(RRSP), 용존산소 설정치(DOSP)에 의존하며 BOD, SS 농도는 서로 일정한 상관관계를 가지고 있다.

본 논문에서 수도권 하수처리장 중의 하나를 모델로 선정하여 이 처리장의 1년치 수질 데이터를 토대로 활성오니공정을, 절에서 제안한 퍼지 추론에 의해 규칙베이스 퍼지 모델링 한다. 이 활성오니공정의 동일한 입력력 데이터를 이용하여 기존의 방법에 의한 퍼지모델링¹⁴ 과 성능을 비교한다.

활성오니공정에서 규칙베이스 퍼지 모델링을 위한 구조는 그림 9와 같다.

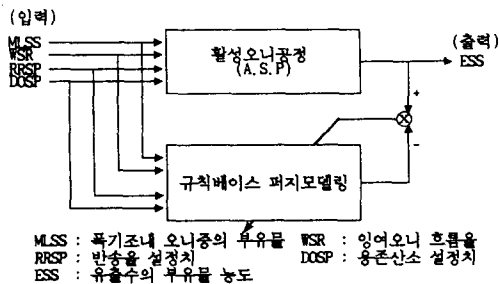


그림 9. 활성오니공정에서 규칙베이스 퍼지 모델링을 위한 구조

Fig. 9. Structure for rule-based fuzzy modeling in the A.S.P.

활성오니공정 입출력 데이터 52쌍을 사용하여 퍼지

추론에 의한 모델링을 한다. 퍼지모델의 전반부 구조 동정은 전반부 파라미터 동정은 그림 2와 그림 3에서 표시한 구조 a와 구조 b의 방법을 사용한다. 후반부 동정도 구조동정 및 파라미터 동정으로 구분된다. 후반부 구조로는 구조 2를 사용한다. 제안한 퍼지 동정 절차는 다음과 같다.

전체 입력변수 MLSS, WSR, RRSP, DOSP의 각 변수와 출력 변수 ESS에 대해 퍼지 클러스터링에 의한 평가규범은 표 4와 같다.

표 4. 각 데이터의 퍼지 분할수에서의 평가규범
 Table 4. Evaluation criterion in the fuzzy partition number of the each data set.

데이터집합	분할수			
	2	3	4	5
(MLSS, ESS)	0.116	0.151	0.252	0.375
(WSR, ESS)	0.121	0.266	0.223	0.319
(RRSP, ESS)	0.120	0.140	0.186	0.392
(DOSP, ESS)	0.145	0.213	0.242	0.248

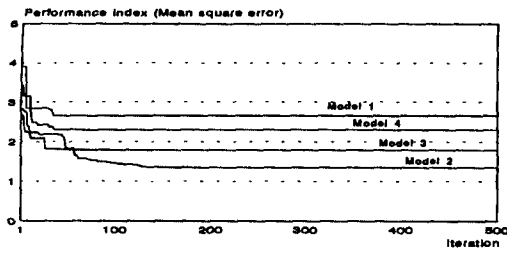
표 5. 각 모델의 성능지수
 Table 5. Performance index of each model.

모델 No.	입력변수	Type a	Type b
Model 1	MLSS_WSR_RRSP	2.65	*1.40
Model 2	MLSS_WSR_DOSP	*1.34	1.85
Model 3	MLSS_RRSP_DOSP	1.78	2.26
Model 4	WSR_RRSP_DOSP	2.29	2.01

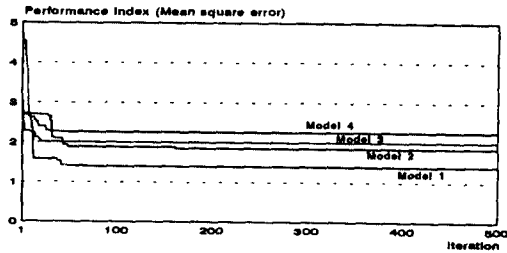
*최적모델

이들 입력변수 4개로부터 PI(동정오차)와 퍼지 규칙수의 최소화를 고려하여 입력변수는 3개로하고 각 입력변수와 출력변수의 데이터 쌍에 의한 퍼지 분할수는 2개 즉 "Big"과 "Small"로 한다. 2차원 공간상의 멤버쉽함수를 각 입력변수축인 1차원 공간으로 사상 시킨다. 그림 10은 퍼지 모델링의 동정오차(성능지수)가 동조 알고리즘에 의한 최적값으로의 수렴되는 과정을 나타내고, 표 5는 이들 최적값에 의한 각 모델에 의한 동정오차를 나타낸다.

구조 2에서 추출된 최적 퍼지규칙으로 이들의 동정 오차(성능지수)와 기존의 퍼지모델의 동정오차와 비교하면 표 6과 같다.



(a) 구조 a



(b) 구조 b

그림 10. 각 퍼지모델 성능지수의 최적값으로의 수렴과정

Fig. 10. Covergence procedure to optimal values of PI for each fuzzy model.

표 6. 기존의 퍼지모델링 방법과 동정오차 비교
Table 6. Comparison of identification error with conventional fuzzy modeling method.

Model name		Mean square error	Number of rules
Conventional model[14]		5.21	8
Our model	Model 2(Type a)	1.34	8
	Model 1(Type b)	1.40	8

본 논문에서 제안한 퍼지추론에 의한 규칙베이스 퍼지모델링이 기존의 퍼지모델링 방법보다 우수함을 보인다.

V. 결론

본 논문에서 간략추론(구조 1), 선형추론(구조 2) 과 변형된 선형추론(구조 3)의 세 가지 퍼지추론방식으로 구조동정에 퍼지 클러스터링 방법과 파라미터 동정에 제안된 자동동조 알고리즘인 수정된 콤플렉스법을 사용함으로써 효율적인 퍼지규칙의 동정기법을 제시하였다. 가스공정에서 각 퍼지추론에 의한 모

델링의 동정오차는 구조 1에서 0.134, 구조 2에서 0.098, 구조 3에서 0.133으로 기존의 퍼지 모델링의 동정오차 0.355(규칙수:6)보다 성능이 우수함을 보였다. 또한 하수처리공정의 선형추론에 의한 퍼지 모델링의 동정오차는 1.34로 기존의 퍼지 모델링의 동정오차 5.21보다 1/4로 감소하였다. 가스공정 시계열 입출력 데이터 및 하수처리공정 데이터를 사용하여 제안된 퍼지 모델링 방법이 기존의 모델링 방법과 비교하여 성능이 우수하며 더 적은 규칙수로 최적 구현 규칙을 얻을 수 있었다.

參考文獻

- [1] R. M. Tong, "Synthesis of fuzzy models for industrial processes", *Int. J. Gen. Syst.*, Vol. 4, pp.143-162, 1978.
- [2] W. Pedrycz, "An identification algorithm in fuzzy relational system", *Fuzzy Sets Syst.*, Vol.13, pp.153-167, 1984.
- [3] W. Pedrycz, "Numerical and application aspects of fuzzy relational equations", *Fuzzy Sets Syst.*, Vol.11, pp.1-18, 1983.
- [4] E. Czogala and W.Pedrycz, "On identification in fuzzy systems and its applications in control problems", *Fuzzy Sets Syst.*, Vol.6, pp.73-83, 1981.
- [5] C. W. Xu, and Y. Z. lu, "Fuzzy model identification self-learning for dynamic system", *IEEE Trans. on Syst. Man, Cybern.*, Vol.SMC-17, No.4, pp.683-689, 1987.
- [6] C. W. Xu, "Fuzzy systems identification", *IEE Proceeding* Vol.126, No.4, pp.146-150, 1989.
- [7] R. M. Tong, "The evaluation of fuzzy models derived from experimental data", *Fuzzy Sets Syst.*, Vol.13, pp.1-12, 1980.
- [8] M. Sugeno and T. Yasukawa, "Linguistic Modeling Based on Numerical Data", *IFSA '91 Brussels, Computer, Management & System Science*, pp. 264-267, 1991.
- [9] Box and Jenkins, "Time Series Anal-

ysis, Forcastig and Control". Holden Day, Sanfransisco.

[10] 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.

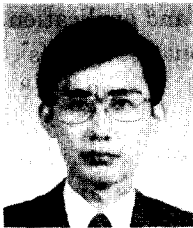
[11] M. Sugeno and G. T. Kang, "Structure identification and fuzzy model". Fuzzy Sets and Systems, pp.15-33, 1988.

[12] M. A. Ismail, "Soft Clustering Algorithm and Validity of Solutions". In: Fuzzy Computing Theory, Hardware and Applications, edited by M. M. Gupta, North Holland, pp.445-471, 1988.

[13] Xuanli, L. X. and Gerado, B., "A Validity Measure for Fuzzy Cluster". IEEE Trans. Pattern Anal. Machine Intell., Vol.PAMI-13, No.8, pp.841-847, August, 1991.

[14] 오성권, 김현기, 우광방, "활성오니공정의 퍼지 모델링과 제어시스템", 대한전기학회 논문집 40권 10호 pp.1072-1083, 1990년 10월.

著者紹介



吳 聖 權(正會員)

1957年 2月 13日生. 1981年 연세대학교 전기 공학과 졸업. 1983년 ~ 89년 금성 산전 연구소(선임 연구원). 1993년 연세대 대학원 전기공학과 졸업(공학박). 현재 원광대 공대 제어계측 공학과 전임강사. 주관심 분야는 시스템 자동화, 퍼지 논리, 신경회로망, 모델링 및 지능형 제어 등임.

周 永 燾(正會員) KITE of Journal of Electronics Engineering, Vol.4, No. 1, OCTOBER 1993. 參照

1958年 6月 26日生. 1882年 연세대학교 전기공학과 졸업(공학사). 1984년 연세대학교 대학원 전기공학과 졸업(공학석사). 1986년 ~ 현재 삼성전자(주)생산기술센터 근무. 1991년 ~ 현재 연세대학교 대학원 전기공학과 박사과정. 주관심 분야는 이동로봇 및 AGV의 지능형 제어, 퍼지논리, 신경회로망등의 산업적 응용 등임.



南 宜 錫(正會員)

1968年 2月 1日生. 1991年 연세대학교 전기공학과 졸업(공학사). 1993년 연세대학교 대학원 전기공학과 졸업(공학석사). 현재 연세대학교 대학원 전기공학과 박사과정. 주관심 분야는 시스템 자동화, 퍼지논리, 신경회로망, 모델링 등임.

禹 廣 芳(正會員) 第 28卷 B 編 第 7號 參照