

# Bacterial Foraging Algorithm과 FCM 기반 퍼지 시스템을 이용한 비선형 시스템 모델링 Nonlinear System Modeling Using Bacterial Foraging and FCM-based Fuzzy System

조재훈<sup>1</sup>, 전명근<sup>1</sup>, 김동화<sup>2</sup>

<sup>1</sup> 충북 청주시 충북대학교 전기.전자.컴퓨터 공학부  
E-mail: [mmi8988@lycos.co.kr](mailto:mmi8988@lycos.co.kr), [mgchun@chungbuk.ac.kr](mailto:mgchun@chungbuk.ac.kr)

<sup>2</sup> 대전광역시 한밭대학교 전기.전자.제어 공학부  
E-mail : [kimdh@hanbat.ac.kr](mailto:kimdh@hanbat.ac.kr)

## 요 약

본 논문에서는 Bacterial Foraging Algorithm과 FCM(fuzzy c-means)클러스터링을 이용하여 TSK(Takagi-Sugeno-Kang)형태의 퍼지 규칙 생성과 퍼지 시스템(FCM-ANFIS)을 효과적으로 구축하는 방법을 제안한다. 구조동정에서는 먼저 PCA(Principal Component Analysis)을 이용하여 입력 데이터 성분간의 상관관계를 제거한 후에 FCM을 이용하여 클러스터를 생성하고 성능지표에 근거해서 타당한 클러스터의 수, 즉 퍼지 규칙의 수를 얻는다. 파라미터 동정에서는 Bacterial Foraging Algorithm을 이용하여 전제부 파라미터를 최적화 시킨다. 결론부 파라미터는 RLSE(Recursive Least Square Estimate)에 의해 추정되어진다. PCA(Principal Component Analysis)와 FCM을 적용함으로써 타당한 규칙 수를 생성하였고 Bacterial Foraging Algorithm을 이용하여 최적의 전제부 파라미터를 구하였다. 제안된 방법의 성능을 평가하기 위하여 Box-Jenkins의 가스르 데이터와 Rice taste 데이터의 모델링에 적용하였고 우수한 성능을 보임을 알 수 있었다.

**Key Words** : Bacterial Foraging Algorithm, FCM, PCA, TSK fuzzy model, Non-linear system

## 1. 서 론

비선형 시스템 모델링은 여러 산업현장에서 쉽게 접할 수 있으나 수학적 방법으로 모델링하기에는 그 한계가 있다. 수학적 방법의 어려움을 극복하기 위하여 퍼지 시스템, 신경회로망 등과 같은 지능형 알고리즘을 이용한 방법들이 많이 제안되어져 왔으며, 최근에는 다양한 지능형 알고리즘의 장점만을 융합하여 그 성능을 개선하려는 연구가 많이 진행되어지고 있다.[1][2]

Bacterial Foraging Algorithm은 Bacteria의 생존을 위한 먹이찾기 행동을 모방한 알고리즘으로서 최근 최적화 분야에서 많이 연구되어지고 있다.

Bacteria를 모방한 최적화 방법은 Bremermann에 의해 제안된 bacteria의 주화성반응(chemotaxis)을 이용한 방법과 최근 Passino에 의해 제안된 Bacterial Foraging 방법이 있

다.[3],[4]. Bremermann에 의해 제안된 방법은 박테리아의 주화성 반응만을 기반으로 하고 Passino가 제안한 방법은 박테리아의 주화성 반응, 재생산, 제거-분산, Swarming 반응들을 포함한다.

본 논문에서는 Passino가 제안한 Bacterial Foraging Algorithm과 FCM클러스터링을[5] 이용하여 TSK(Takagi-Sugeno-Kang)형태의 효과적인 퍼지시스템(FCM-ANFIS)을 구축하는 방법을 제안한다. Jang[1]에 의해 제안된 ANFIS(Adaptive Neuro-Fuzzy Inference System)는 규칙의 수를 최소로 하면서 복잡한 비선형 시스템의 표현에 뛰어난 능력을 보이거나 입력공간의 증가에 따른 계산량의 급증과 규칙의 수가 지수적으로 증가하는 문제점을 가지고 있다. 이런 문제점을 해결하기 위해 개선된 ANFIS을 구축하고자 한다. 먼저 구조동정(Structure Identification)에서는 입력 데이터 성분간의 상관관계를 제거하기 위해

PCA(Principal Component Analysis)[6]을 이용한 후에 FCM 클러스터링을 이용하여 클러스터를 생성하고 성능지표에 근거해서 타당한 클러스터의 수, 즉 퍼지 규칙의 수를 얻는다. 파라미터 동정에서는 Bacterial Foraging Algorithm을 이용하여 전제부 파라미터를 최적에 가깝도록 탐색을 시도한다. 결론부 파라미터는 선형 방정식을 이루고 있기 때문에 전제부 파라미터가 결정된 후에 RLSE(Recursive Least Square Estimate)에 의해 추정되어진다. 이렇게 함으로서 구조동정에서 타당한 규칙 수와 효율적인 퍼지 규칙을 얻을 수 있으며, 파라미터 동정에서 최적에 가깝도록 전제부 및 결론부 파라미터를 구할 수 있다. 제안된 방법의 유용성을 보이기 위해 Box-Jenkins의 가스로 데이터에 적용하여 제안된 방법이 이전의 연구보다 좋은 결과를 보임을 알 수 있었다.

## 2. 박테리아 먹이찾기(Bacterial Foraging)

자연 선택은 동물들이 세대 번식을 하는데 있어서 성공적인 전략과 친척들의 공격으로부터 살아남는 것들만이 존재하도록 선택되어지고 유지되어진다. 이러한 여러 가지 동물들의 전략들은 기본적으로 먹이를 찾는 것으로부터 시작된다. 먹이 찾기는 동물들이 동물 자신들을 유지하는데 있어서 가장 기본적이고 근본적인 전략 중에 하나이다.

많은 세대 후에 이러한 전략들 중 우수한 전략을 가진 동물들만이 번식에 성공하고 진화하여 자신의 후손들을 유지하게 된다. 이러한 이론들은 공학에 있어 최적화 분야에 새로운 방법들을 제시하는데 바탕이 되어져 왔다.[4]

### 2.1 박테리아의 주화성, 집단이동, 재생산, 제거 그리고 분산 단계

실제적인 박테리아 먹이찾기 알고리즘을 최적화 문제에 적용하기 위하여 아래와 같이 정의하기로 한다.

$j$ 는 주화성 단계를 표시하고  $k$ 는 재생산 단계, 그리고  $l$ 을 제거-분산 동작으로 정의하면 박테리아의 행동은 아래의 식(1)로 표현될 수 있다.

$$P(j, k, l) = \{\theta^i(j, k, l) \mid i = 1, 2, \dots, S\} \quad (1)$$

식(1)은 박테리아 집단  $S$ 에서  $l$ 번째 제거-

분산 동작,  $k$ 번째 재생산,  $j$ 번째 주화성 단계를 가지는  $i$ 번째 박테리아의 위치( $\theta$ )를 말한다.

여기서 우리는  $i$ 번째 박테리아의 위치  $\theta^i(j, k, l) \in R^p$ 에서의 평가된(예, 좋은 환경이나 나쁜환경이나) 값을  $J(i, j, k, l)$ 으로 표현할 수 있다.

최적화 문제에서는 비용함수로 생각할 수 있다.

$C(i) > 0, i = 1, 2, \dots, S$ 는 주화성 단계에서 한 번 이동할 때의 이동 길이로 정의한다.

$\phi(j)$ 는 주화성 단계 동안 박테리아가 방향 전환을 할 때 선택되어지는 무작위적인 방향을 나타낸다. 위의 기호들을 사용해서 아래에 식(2)처럼 다음 단계의 박테리아의 행동을 표현한다.

식(2)는 어떤 임의의 박테리아가 전 단계에서 ( $j$ )에서 다음 단계 ( $j+1$ )에는 무작위 방향  $\phi(j)$ 에서  $C(i)$ 만큼 이동 된다는 것을 의미한다.

$$\theta^i(j+1, k, l) = \theta^i(j, k, l) + C(i)\phi(j) \quad (2)$$

만약  $\theta^i(j+1, k, l)$ 에서  $J(i, j+1, k, l)$ 가  $\theta^i(j, k, l)$ 에서 보다 더 좋은 값을 가지면 이 같은 방향에서 크기  $C(i)$ 의 길이를 작게 하여 더 정밀한 탐색을 할 수 있고, 더 나쁜 값을 가진다면  $C(i)$ 의 길이를 크게 하여 좀 더 광범위한 탐색을 할 수 있도록 조정할 수 있다.

이러한 주화성 반응은 정해진 주화성 단계에 도달 하거나 사용자가 원하는 목적을 달성하면 주화성 반응은 멈추게 된다.

## 3. Bacterial Foraging과 FCM 클러스터링에 의한 퍼지 시스템

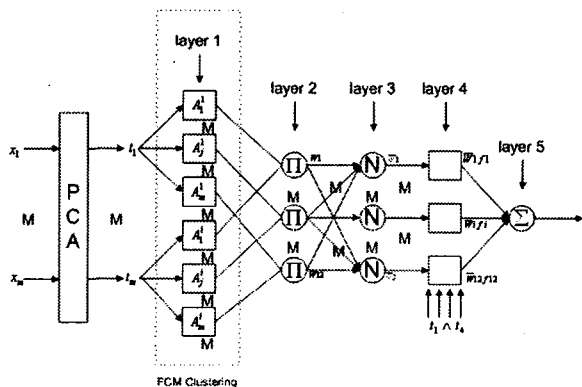


그림 1. FCM-ANFIS의 구조

퍼지 시스템을 구축하기 위해 구조동정과 파라미터 동정에 대해 설명하고자 한다. 먼저, PCA를 이용한 입력 데이터 성분간의 상관관계를 제거하여 입력공간을 변환한다.

PCA는 통신이론과 영상처리분야에서 K-L 변환[6]으로서 알려진 주성분분석기법으로 입력을 선형적으로 결합하고 입력공간의 차원을 줄이기 위한 접근방법으로 알려져 있다.

우선, 입력 데이터 집합을 식(3)과 같이 공분산행렬(covariance matrix)을 계산하고 고유치와 고유벡터를 구한다.

$$C = \frac{1}{N} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \quad i=1,2,\dots,n \quad (3)$$

PCA에 의해 변환된 입력을 0과 1사이에 정규화 시킨 다음 입력공간을 스캐터(scatter) 분할하여 적은 규칙의 수를 적용적으로 생성하도록 하는 FCM 클러스터링에 의한 퍼지 시스템(FCM-ANFIS)을 구축한다.

FCM-ANFIS을 구축한 후에 Bacterial Foraging을 이용하여 파라미터 동정을 수행한다. 이때 구축된 FCM-ANFIS의 전체부 파라미터를 초기치 집단으로 발생시켜 Bacterial Foraging Algorithm에 의해 최적에 가깝도록 탐색되어진다. 결론부 파라미터는 선형방정식을 이루기 때문에 전체부 파라미터가 결정된 후에 의해 추정되어진다. 그로 인해 Bacterial Foraging Algorithm의 탐색공간이 줄어들어 최적에 가깝도록 원하는 해를 찾아낼 수 있다.

Bacterial Foraging Algorithm에 의한 수행 순서는 다음과 같다.

**[step 1]** 기본적인 Bacterial Foraging Algorithm의 초기의 파라미터 값들을 정의하고 FCM-ANFIS에서 얻어진 전체부 파라미터를 초기치 집단으로 발생시킨다.

$$(n, S, N_c, N_{re}, N_{ed}, P_{ed}, C(i), \phi^i)$$

$n$ : 문제 해공간의 차원.

$S$ : Bacteria의 집단 수.

$N_c$ : 주화성 단계의 반복 횟수.

$N_{re}$ : 재생산 단계의 반복 횟수.

$N_{ed}$ : 제거-분산 반응의 반복 횟수.

$P_{ed}$ : 제거-분산의 비율

$C(i)$ : 주화성 반응에서의 단위 이동 크기.

**[Step 2]** 제거-분산 loop:  $l=l+1$

**[Step 3]** 재생산 loop:  $k=k+1$

**[Step 4]** 주화성 loop:  $j=j+1$

**[substep a]** 개체에 대한 적합함수  $F$ 를 식(4)과 같이 계산한다. 이는 학습데이터와 출력 데이터에 대하여 퍼지 모델의 출력과 실제 출력 값 사이의 MSE(Mean Square Error)에 대한 역수로 나타내어진다. 이 적합 함수는 학습 데이터에 편중되지 않고 새로운 데이터에 대해서도 타당한 결과를 내는 모델인가를 평가할 수 있다.

$$F = \frac{1}{1 + MSE_{tm}} \quad (4)$$

**[substep b]** 전 단계의 적합함수와 비교하여 전 단계보다 더 우수하면 방향을 유지하고 그렇지 않으면 새로운 무작위적 방향을 선택하여 탐색 방향을 재설정한다.

**[step 5]** If,  $j < N_c$ , go to step 5.

설계자에 의해 정해진 주화성 횟수나 목적함수 값을 만족하면 반복을 멈추고 그렇지 않으면 step5로 가서 주화성 반응을 반복한다.

**[step 6]** If  $k < N_{re}$ , go to step 4.

설계자에 의해 정해진 재생산 횟수 만큼 반복.

**[step 7]** If  $l < N_{ed}$ , then go to step3

설계자에 의해 정해진 제거-분산 횟수 만큼 반복.

end

## 4. 시뮬레이션 및 결과 고찰

본 논문에서 구축된 FCM-ANFIS와 Bacterial Foraging Algorithm을 이용하여 Box-Jenkins의 가스공정 모델링에 적용하여 이전에 연구된 방법들에 비하여 보다 나은 결과를 얻을 수 있음을 보이고자 한다. [1]

### 4.1 Box와 Jenkins의 가스 데이터의 모델링

Box-Jenkins의 가스 데이터는 가스에서의 가스 흐름을  $u(k)$ 와 이산화탄소 농도  $y(k)$ 의 시계열 데이터인 296  $\{y(k), u(k)\}$ 개의 쌍이 사용된다. 기존의 논문에서 이 예제에 대해 퍼지 시스템을 구축할 경우에 전체적인 데이터를 학습데이터로 사용한 경우와 학습데이터와 검증데이터로 나누어 사용한 경우가 있다. 본 논문에서는 전체적인 데이터를 학습데이터로 하고 두 개의 입력  $y(k-1), u(k-4)$ 으로 선택하고 출력 변수는  $y(t)$ 를 선택하였다.

그림(2)은 입력 데이터  $y(k-1), u(k-4)$ 의 분포도이며 상관관계가 존재함을 알 수 있고, 그림(3)은 데이터 전처리로서 PCA를 이용하여 입력 데이터 성분간의 상관관계를 제거한 분포도이다.

PCA에 의해 변환된 입력데이터에 대해 정

규화 과정을 거친 후에 FCM 클러스터링을 수행하였으며, 성능지표에 근거해서 S(c)의 값 값이 제일 적은 6개의 클러스터, 즉 퍼지 규칙을 얻을 수 있었다. 이렇게 전제부 파라미터가 결정된 후에 RLSE에 의해 결론부 파라미터를 추정하여 그림(1)과 같이 FCM-ANFIS을 구축하였다.

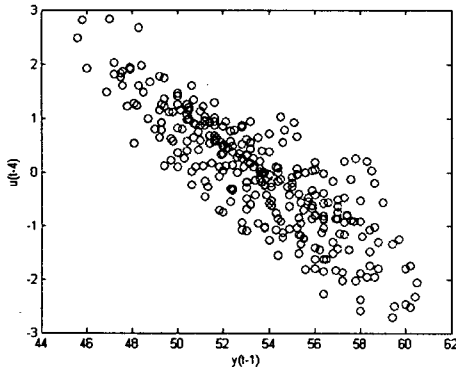


그림.2 상관관계가 존재하는 입력  $y(k-1), u(k-4)$ 의 분포도

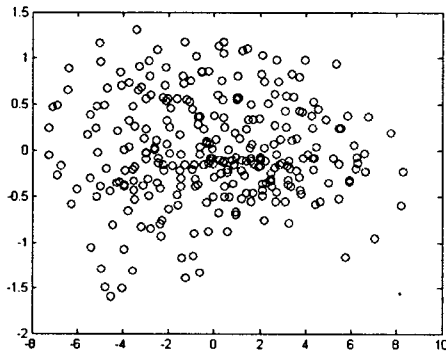


그림.3 PCA를 통해 상관관계가 제거된 입력  $y(k-1), u(k-4)$ 의 분포도

### 5. 결론

본 논문에서는 비선형 시스템 모델링을 위해서 bacterial foraging 과 FCM 클러스터링을 이용하여 TSK 형태의 효과적인 퍼지 시스템인 FCM-ANFIS을 구축하는 방법을 제안하였다. 제안된 방법을 Box-Jenkins 가스로 공정 모델에 적용하여 실험해 본 결과 이전의 다른 논문에 비해 적은 규칙과 파라미터로 좋은 모델링 특성을 보임을 알 수 있었다.

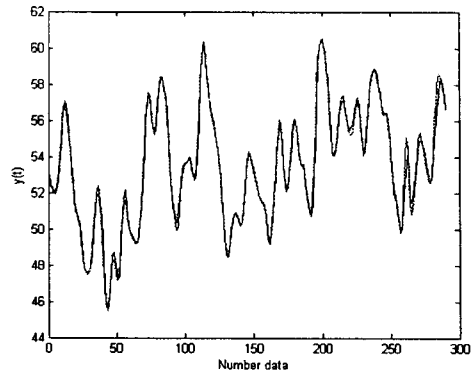


그림. 4 제안된 모델의 추론 값과 실제 데이터 출력과의 비교(입력2)

표.1 규칙 수와 오차의 비교

	입력수	규칙수	평균제곱오차
Tong	2	19	0.467
Pedrycz	2	20	0.776
Xu	2	25	0.328
Sugeno	3	6	0.19
Joo	2	6	0.166
Kwak	2	6	0.1176
제안된방법	2	6	0.1173

### 참 고 문 헌

- [1] Jyh-Shing Roger Jang, "ANFIS: Adaptive-Neuro-based Fuzzy Inference System", IEEE Trans. on Syst., Man, Cybernetics, vol., 23, No.3, pp.665-685, 1993.
- [2] C. T. Lin, C. S. G. Lee, "Neural fuzzy Systems : A Neuro-Fuzzy Synergism to Intelligent Systems", Prentice Hall, 1996.
- [3] H. J. Bremermann, " Chemotaxis and optimization," J. Franklin Inst., vol.297, pp. 397-404, 1974.
- [4] PASSINO, K. M., "Biomimicry of Bacterial foraging for Distributed Optimization", University Press, Princeton, New Jersey, 2001.
- [5] J.c. Bezdek, "Pattern Recognition with fuzzy objective Function Algorithms," Plenum Press, New York, 1981.
- [6] M. Loeve, "Probability theory", Van Nostrand, 3rd edition, 1963.