

# HCM 클러스터링과 유전자 알고리즘을 이용한 다중 퍼지 모델 동정

## Identification of Multi-Fuzzy Model by means of HCM Clustering and Genetic Algorithms

°박 호 성\*, 오 성 권\*

\* 원광대학교 전기전자공학부(Tel:82-653-850-6342; Fax:82-653-852-2196; E-mail:ohsk@wonnms.wonkwang.ac.kr)

**Abstract** : In this paper, we design a Multi-Fuzzy model by means of HCM clustering and genetic algorithms for a nonlinear system. In order to determine structure of the proposed Multi-Fuzzy model, HCM clustering method is used. The parameters of membership function of the Multi-Fuzzy are identified by genetic algorithms. A aggregate performance index with a weighting factor is used to achieve a sound balance between approximation and generalization abilities of the model. We use simplified inference and linear inference as inference method of the proposed Multi-Fuzzy model and the standard least square method for estimating consequence parameters of the Multi-Fuzzy. Finally, we use some of numerical data to evaluate the proposed Multi-Fuzzy model and discuss about the usefulness.

**Keywords** : Multi-Fuzzy, HCM clustering, genetic algorithms, nonlinear system.

### 1. 서론

많은 연구자들은 자연 현상을 이해하고 그것을 이용하기 위한 연구가 이루어지고 있다. 그 중 1965년에 Zadeh[1]에 의해 창안된 "퍼지집합"이론은 비선형적이고 복잡한 실 시스템의 특성을 해석하는데 적용함으로써 수학적 모델보다 좋은 결과를 가져왔다. 설계할 시스템의 성능 및 기능의 요구조건에 따라 퍼지 모델은 애매 모호한 언어적 변수를 수치적으로 표시할 수 있어서 융통성 있는 시스템 설계를 가능하게 하고 시스템의 기능을 향상시키며 설계를 간단하게 해주는 장점이 있다. 그러나 데이터 특성에 맞는 퍼지 모델을 구축하는데 있어서는 전문가와 시행착오에 의존해야 하는 어려움이 있어 동적으로 변화하는 환경에서 적용적으로 대처할 수 있는 퍼지 모델을 구축하기가 힘들다. 이러한 문제를 해결하기 위한 방법으로 본 논문에서는 다중 퍼지 모델을 제안한다.

제안된 다중 퍼지 모델은 시스템의 데이터 특성을 해석하기 위해 HCM 클러스터링 방법[2]을 이용하여 입력력 데이터를 서로 유사한 특성을 가진 그룹으로 분류하는 전처리 과정을 수행하고, 분류된 그룹은 다중 퍼지 모델 각각의 모델에 입력이 되어 출력하게 된다. 또한 데이터 특성 분류에 따른 각 퍼지 모델의 최적화를 위해 유전자 알고리즘[3,4]을 이용하여 멤버십 함수의 정점을 최적으로 동조한다. 더 나아가 학습 데이터 및 테스트 데이터를 고려한 퍼지 모델 성능 즉 근사화 및 일반화 능력 모두를 고려한 성능결과를 위해 하중 계수를 가진 합성 목적함수[5]를 사용한다. 이는 모델의 성능지수가 데이터의 개수, 비선형성의 정도(입출력 데이터 분포)등에 의해 근사화 및 예측의 정확도에 차이가 있으므로 상호 균형과 의존성을 고려한 합성 목적함수의 하중계수의 선택과 조절을 통하여 시스템의 최적모델을 구축하고자 한 것이다.

본 논문에서 제안된 퍼지 모델은 후반부가 간략추론과 선형 추론을 사용한 두 가지 형태를 사용하였으며, 멤버십 함수의 형태로는 삼각형 형태를 사용한다. 제안된 다중 퍼지 모델의 성능을 평가하기 위해서 Box와 Jenkins가 사용한 시계열 데이터[6]와 가스 터빈 발전소의 데이터[7]를 이용하였다.

### 2. 다중-퍼지 모델

#### 2.1 퍼지 모델의 구조와 동작

퍼지 모델링이란 if-then 형식으로 플랜트를 기술하는 것으로, 구체적으로 입력력 데이터의 상호관계에 의해 설정된 입력력 변수로부터 확립되는 것이다. 규칙베이스 퍼지 모델의 동정은 전반부와 후반부의 동정으로 나누어진다. 전반부 동정은 입력 정보, 즉 퍼지 규칙의 전반부 변수의 선택과 입력변수 공간의 퍼지분할 결정, 그리고 입력공간의 파라미터 결정이 필요하다. 후반부 동정은 후반부 변수의 선택과 후반부 변수의 파라미터를 결정하는 것이다. 전반부의 입력변수와 후반부 변수의 멤버십함수는 입력변수의 공간을 규칙 수만큼의 퍼지분할된 부분공간으로 나누는 것과 관련되고, 후반부의 파라미터는 각각 퍼지분할된 부분공간에서 입력력 관계를 묘사한다. 본 논문에서는 전반부 구조로 결정된 입력 변수와 규칙을 가지고, 유전자 알고리즘을 이용하여 전반부 파라미터를 동정한다. 후반부 구조는 두형태의 구조로 표현하며, 후반부 파라미터 동정으로는 표준 최소자승법을 사용한다.

#### 2.1.1 전반부 동정

퍼지 모델링에서 전반부 동정, 즉 구조 동정 및 파라미터 동정은 비선형 시스템을 표현하는데 있어서 매우 중요하다. 왜냐하면 전반부 입력 변수의 선택과 선택된 입력 변수의 공간 퍼지 분할 그리고 파라미터 동정은 비선형 시스템의 성능을 결정하는데 많은 영향을 미치기 때문이다. 전반부에서 멤버십 함수의 형태로는 삼각형 형태를 사용한다. 기존의 방법은 멤버십 함수를 그림 1처럼 입력 변수의 최소값과 최대값 사이를 임의의 개수로 등분해서 일률적으로 정의하였으나 이는 데이터들이 가지고 있는 특성을 제대로 반영하지 못하는 단점이 있다. 그래서 전반부 파라미터 동정을 위해 유전자 알고리즘을 이용한다. 유전자 알고리즘을 이용하면 위에서 언급한 문제점을 해결할 수 있고 멤버십 함수의 정점과 같은 파라미터들을 최적으로 동정할 수 있다.