

# 점증적 입자 모델의 최적화 설계와 응용

염찬욱, 궤근창\*  
 조선대학교 제어계측공학과  
 \*e-mail: kwak@chosun.ac.kr

## An Optimization Design of Incremental Granular Model and Its Application

Chan-Uk Yeom, Keun-Chang Kwak  
 Dept of Control and Instrumentation Engineering, Chosun University

### 요 약

본 논문에서는 GA(Genetic Algorithm) 기반 점증적 입자모델(IGM: Incremental Granular Model)의 최적화 설계를 제안한다. IGM의 성능은 다양한 실세계 응용예제를 통해 성공적으로 연구되어져왔다. 그러나, IGM의 문제로 각 컨텍스트에서 동일한 클러스터 수가 사용되는 점과 전형적인 퍼지화 계수가 설정된다는 점이 있다. 이러한 문제를 해결하기 위해 IGM을 최적화하여 각 컨텍스트에서 클러스터 중심의 수와 퍼지화 계수를 최적화하는 설계 방법을 제시했다. 제안된 방법의 타당성을 확인하기 위해 Ecotect에서 시뮬레이션 한 12가지 건물 형태를 사용하여 에너지 효율 예측에 대한 실험을 수행하였고, 제안된 방법은 기존의 IGM보다 우수한 성능을 보이는 것을 확인했다.

### 1. 서론

현재 여러 분야에서 퍼지 모델에 대한 수많은 연구들이 진행되었다. 복잡하면서 비선형적인 특징을 가진 여러 문제들은 다양한 퍼지 모델들을 통합하는 하이브리드 지능 시스템 모델을 추구하고 있다. 이와 같은 모델은 개별적인 모델들의 융합하기 때문에 서로 다른 모델의 장점들을 얻을 수 있다. 대표적인 방법으로 뉴로-퍼지 추론 시스템이 있다. 뉴로-퍼지 추론 시스템은 변화하는 환경에 적응하고 학습을 할 수 있는 신경회로망과 전문가 지식인 퍼지 추론을 융합하여 만들어진다[1].

한편 입자 모델(Granular Model)을 설계하는데 사용되는 클러스터링의 근본적인 아이디어를 이용한 새로운 방법에 대한 연구들이 진행되고 있다. Pedrycz는 CFCM(Context-based Fuzzy C-Means) 클러스터링에 근거하여 RBFNN[2], 언어적 모델(Linguistic Model)[3], 동적 데이터 입자화 설계 방법[4] 등을 제안했다. Kwak[5]은 퍼지 입자화를 가진 GA 기반 언어적인 모델을 설계하는 방법을 제안하였다. 진화알고리즘을 통한 모델의 최적화에 관한 다양한 연구들이 진행되고 있다. 이러한 다양한 연구들이 수행되어졌음에도 불구하고, 모델링 오차를 예측하는 방법은 고려되어지지 않고 있다. Pedrycz는 모델의 오차를 예측하기 위해 전역적인 모델로 선형회귀(Linear Regression)와 국부적인 모델로 LM을 결합하여 점증적 입자 모델을 제안했다[6].

하지만, 점증적 입자 모델은 특성이 다른 컨텍스트 임에도 불구하고, 클러스터의 수가 동일하게 정해지는 문제

와 퍼지화 계수가 일정한 값으로 고정되는 문제점을 확인했다. 따라서, 본 논문에서는 GA 기반의 점증적 입자 모델을 제안하고, 각 컨텍스트 마다 클러스터의 수와 퍼지화 계수를 최적화하여 기존의 방법과 성능을 비교하고자 한다.

### 2. 유전알고리즘 기반 점증적 입자 모델의 설계

#### 2.1. 선형회귀와 언어적 모델의 설계

IGM은 전역적인 모델로 LR을 사용하고 국부적인 모델로 LM을 사용하도록 구성되며, 각각의 장점을 이용해 효과적으로 오차를 예측 할 수 있다. LR을 이용해서 최소자승법(LSE)으로 선형부분을 모델링하고, 비선형성을 갖는 오차를 CFCM 클러스터링을 이용하는 LM을 통해 보상된다.

CFCM 클러스터링은 입출력 데이터의 특성을 고려하기 때문에 보다 정확하게 클러스터를 찾을 수 있다. 소속행렬  $U$ 는 다음과 같이 계산되어진다[7].

$$u_{ik} = \frac{f_k}{\sum_{j=1}^c \left( \frac{\|x_k - c_i\|}{\|x_k - c_j\|} \right)^{\frac{2}{m-1}}} \quad (1)$$

여기서,  $m$ 은 퍼지화 계수를 나타내며,  $f_k$ 는 0과 1 사이의 소속 값을 나타낸다.  $f_k = W(d_k)$ 는 출력 공간에서 생성된 임의의 퍼지 집합에서  $k$ 번째 데이터의 포함 정도를 나타낸다. 출력 공간의 퍼지 집합은  $W: D \rightarrow [0,1]$ 로 나타낼 수

있으며,  $D$ 는 출력 변수의 전체 집합을 나타낸다. 이러한 특성에 의해 소속행렬  $U$ 를 나타낼 수 있다.

$$U(f) = \left\{ \begin{array}{l} u_{ik} \in [0, 1] \sum_{i=1}^c u_{ik} = f_k \forall k \\ \text{and } 0 < \sum_{k=1}^N u_{ik} < N \end{array} \right\} \quad (2)$$

출력 변수의 확률적인 분포에 의해서 각 컨텍스트를 생성하며, 컨텍스트에서 생성된 클러스터의 중심은 FCM 클러스터링과 같은 방법으로 계산된다.

**2.2. 유전알고리즘**

일반적인 IGM에서는 각 컨텍스트마다 고정된 클러스터의 수와 고정된 퍼지화 계수를 사용하고 있다. 하지만, 본 논문은 GA를 이용하여 위의 두 값을 최적화한다. GA는 각 파라미터 값을 초기 집단으로 사용하고 선택, 교차, 돌연변이, 엘리트즘 기법 등 연산자를 통해 새로운 집단을 생성한다.

최종적으로 적합도 함수(fitness function)를 이용하여 주어진 방법이 얼마나 적합한지를 세대를 반복하면서 아래와 같이 계산된다.

$$f = \frac{1}{F_t + F_c} \quad (3)$$

$$F_t = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \hat{Y}_k)^2} \quad (4)$$

$$F_c = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k' - \hat{Y}_k')^2} \quad (5)$$

**2.3. 유전알고리즘 기반 점증적 입자 모델의 설계**

본 논문에서 제안하는 GA 기반 점증적 입자 모델의 진행 과정은 다음과 같다.

**[단계 1]** 입출력 공간에서 전역적인 모델인 LR을 통해 선형예측을 하고, 실제 출력과 LR 모델 출력 사이의 오차를 구한다. 오차를 바탕으로 새로운 데이터를 생성한다.

**[단계 2]** 새롭게 생성된 입력 공간에서 오차 분포의 특성에 알맞게 컨텍스트들을 생성한다.

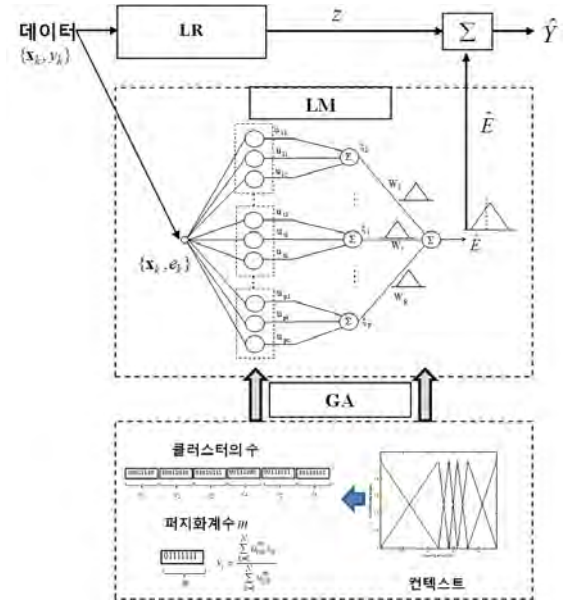
**[단계 3]** 각 클러스터 공간에서 CFM 클러스터링을 수행한다. 여기서, 각 컨텍스트마다 입력 데이터의 특성을 반영하여 클러스터를 결정한다. 결정된 클러스터는 GA를 통해 최적화된 클러스터의 수를 찾는다[2].

**[단계 4]** GA를 통해 최적화된 클러스터에서 활성화함수를 얻고, 컨텍스트를 통한 가중치에 의해 출력을 계산한다.

**[단계 5]** 전역적인 모델 LR의 출력과 국부적인 모델 LM의 출력을 결합하여 최종적인 출력을 계산한다.

$$\hat{Y} = z \oplus \hat{E} \quad (6)$$

그림 1은 제안하는 방법인 GA 기반 점증적 입자 모델의 구조를 보여주고 있다.



(그림 1) GA와 점증적 입자 모델의 구조

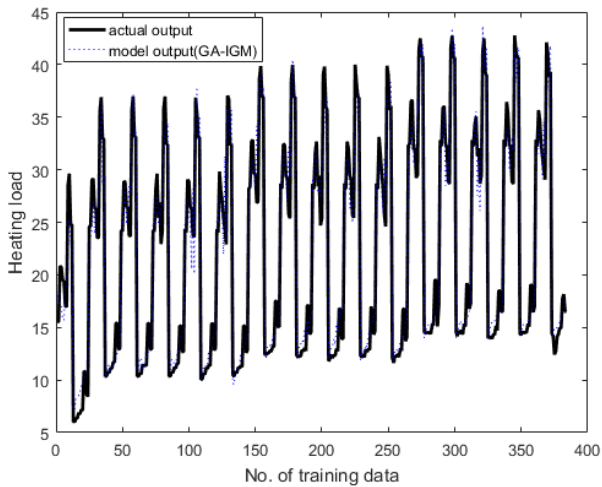
**3. 실험 및 결과분석**

본 장에서는 IGM과 제안된 방법의 성능을 비교 평가한다. 실험에는 빌딩 냉난방 부하 예측 데이터를 사용했다 [8]. 입력 변수는 건물 특성은 상대 간결성, 표면적, 벽으로 된 공간, 지붕으로 된 공간, 전체 높이, 방향, 유약 지역, 유약 지역 분포도로 이루어져 있으며, 출력 변수는 난방 부하와 냉방 부하로 되어있다. 본 논문에서는 난방 부하 출력에 대한 실험만 진행하였다. 학습 데이터와 검증 데이터는 각 50%와 50%로 나누고 0과 1 사이로 정규화 후 사용했다.

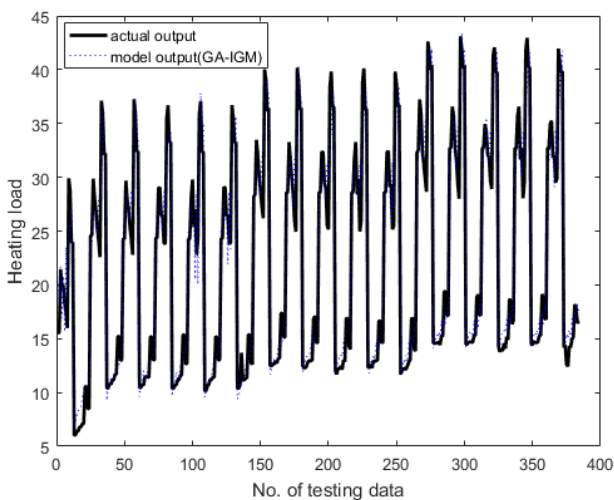
GA를 이용해서 최적화를 하기 위해서 클러스터의 수를 2개에서 9개 사이 공간에서 선택하고, 퍼지화 계수는 1.5에서 3까지 공간에서 선택하도록 설정했다. 기타 파라미터로는 세대수 30, 교차율 0.97, 돌연변이율, 0.01으로 설정하고 엘리트즘 기법을 이용해서 가장 우수한 적합도를 갖는 두 개의 개체를 유지했다.

그림 2와 3은 각각 빌딩 난방 부하에 대한 학습 데이터의 근사화 성능과 검증 데이터의 예측 성능을 보여준다. 그림 2와 3에서 보이는 바와 같이 제안된 방법이 우수한 성능을 보이는 것을 확인 할 수 있다. 표 1은 전체 성능 결과를 보여주고 있으며, 예측 성능은 RMSE를 통해 비교했다. GA 기반 점증적 입자 모델에서 구해진 최적의 파라미터 값은 컨텍스트가 6개 일 때, 클러스터가 8, 7, 6, 8,

9, 9이고, 퍼지화 계수는 1.5로 나온 것을 확인했다. 이로 인해, 기존의 방법보다 효율적인 규칙도 구해지는 것을 확인할 수 있었다.



(그림 2) GA 기반 점증적 입자 모델의 학습 데이터에 대한 근사화 성능



(그림 3) GA 기반 점증적 입자 모델의 검증 데이터에 대한 일반화 성능

<표 1> 빌딩 난방 부하 예측 성능 평가

알고리즘	p, c	규칙 수	RMSE (training)	RMSE (testing)
IGM [6]	p=c=6	36	2.46	2.63
	p=6,c=7	42	2.36	2.62
	p=6,c=8	48	2.33	2.56
	p=6,c=9	54	2.26	2.54
	p=6,c=10	60	2.21	2.53
제안된 방법	p=6,c=[8, 7, 6, 8, 9, 9], m=1.5	47	2.12	2.36

## 5. 결론

본 논문에서는 GA를 이용한 점증적 입자 모델의 최적화 방법을 제안했다. 기존의 IGM은 서로 다른 컨텍스트 내부 클러스터의 수가 동일한 문제와, CFCM 클러스터링 과정에 사용되는 퍼지화 계수가 고정된 값을 갖는 문제가 있다. 제안된 방법은 GA를 통해 클러스터의 수와 퍼지화 계수의 최적화된 값을 구해 위와 같은 문제를 해결했다. 빌딩 난방 데이터에 이용해서 실험한 결과 제안된 방법이 기존의 IGM보다 우수한 예측 성능을 보이는 것을 확인했다. 향후, 컨텍스트의 수를 최적화하는 방법과 GA가 아닌 PSO 방법을 이용한 최적화에 대한 연구를 진행 할 계획이다.

## 감사의 글

본 연구는 2017년도 산업통상자원부의 재원으로 한국에너지기술연구원(KETEP)의 에너지인력양성사업으로 지원받아 수행한 인력양성 성과입니다(No. 20174030201620). 또한, 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학지원사업의 연구결과로 수행되었음(2017-0-00137).

## 참고문헌

- [1] J. S. R. Jang, C. T. Sun, and E. Mizutani, "Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice Hall. 1997.
- [2] W. Pedrycz, "Conditional fuzzy clustering in the design of radial basis function neural networks," IEEE Trans.on Neural Networks, Vol.9, No. 4, pp. 601-612, 1998.
- [3] W. Pedrycz and K. C. Kwak, "Linguistic models as a framework of user-centric system modeling." IEEE Trans. on Systems, Man, and Cybernetics-Part A: Systems and Humans, Vol. 36, No. 4, pp. 727-745, 2006.
- [4] W. Pedrycz, "A dynamic data granulation through adjustable fuzzy clustering." Pattern Recognition Letters Vol. 29, No. 16, pp. 2059-206, 2008.
- [5] K. C. Kwak, "A design of genetically optimized linguistic models," IEICE Trans. on Information and Systems, Vol. E95D, No.12, pp. 3117-3120, 2012.
- [6] W. Pedrycz, K. C. Kwak, "The development of incremental models," IEEE Trans. on Fuzzy Systems. Vol. 15, No. 3, pp. 507-518, 2007.
- [7] Y. H. Byeon and K. C. Kwak, "A design of genetically oriented rules-based incremental granular models and its application," Symmetry, Vol. 9, No. 12, Article ID 324, 2017.
- [8] UCI Machine Learning Repository. Available online: <https://archive.ics.uci.edu/ml/datasets>