

# 자기-구성 클러스터링의 모델링 및 성능평가

정회원 유정웅\*, 김승석\*, 송창규\*, 김성수\*

## Modeling of Self-Constructed Clustering and Performance Evaluation

Jeong woong Ryu\*, Sung Suk Kim\*, Chang kyu Song\*, Sung Soo Kim\* *Regular Members*

### 요 약

본 논문에서는 퍼지 추론 시스템의 추론 정보를 이용하여 자율적으로 구조를 결정하는 클러스터링 기법을 제안한다. 제안된 기법은 주어진 입출력 데이터를 이용하여 자율적으로 클러스터의 수를 추정하고 동시에 이들 파라미터를 최적화한다. 일반적인 클러스터링 기법에서 볼 수 있었던 비교사학습을 교사학습으로 확장하여 클러스터 추정에 입출력 인과 관계를 고려한 학습을 실시하게 하여 전체 모델의 성능을 개선하고자 하였다. 출력 정보가 입력공간에서 클러스터링 학습에 적용됨으로써 클러스터링에서의 각 클래스의 구분 작업이 더 원활하게 이루어 질 수 있다. 모의실험을 통하여 기존의 연구 결과와 비교하여 제안된 기법의 유용성을 보인다.

Key Words : MLE, Fuzzy Clustering, Subtractive Clustering, TSK Fuzzy Model, Supervised Learning

### ABSTRACT

In this paper, we propose a self-constructed clustering algorithm based on inference information of the fuzzy model. This method makes it possible to automatically detect and optimize the number of cluster and parameters by using input-output data. The propose method improves the performance of clustering by extended supervised learning technique. This technique uses the output information as well as input characteristics. For effect the similarity measure in clustering, we use the TSK fuzzy model to sent the information of output. In the conceptually, we design a learning method that use to feedback the information of output to the clustering since proposed algorithm perform to separate each classes in input data space. We show effectiveness of proposed method using simulation than previous ones

### I. 서 론

비선형 데이터 공간에서 다양하게 존재하는 클래스나 패턴의 정보를 추출하고자 할 때 사용되는 여러 가지 기법 중 클러스터링 방법은 데이터의 특성을 고려한 클러스터 파라미터 추정기법이다<sup>[1][2]</sup>. 클러스터링은 선형회귀문제나 패턴분류문제에 대해서 독립적으로 적용할 수 있으나 다른 기법과 순차적으로 연결하여 전체 모델의 성능을 개선할 수도 있다<sup>[3]</sup>. 또 다른 관점에서 보면 주어진 비선형 데이터

공간을 각 클러스터 공간으로 분해하여 각각에 대해서 부분 선형화하는 작업으로 볼 수 있다<sup>[2]</sup>. 이러한 클러스터링 기법은 수행 목적에 따라 크게 두 가지로 나눌 수 있다. 먼저 주어진 데이터로부터 사전 정보를 알지 못하는 경우 임계값을 이용하여 클러스터나 패턴의 수를 추정하는 방법과 사전에 클러스터의 수를 알고 있는 경우 이들 파라미터를 최적화하는 방법 등이 있다<sup>[2][4][5]</sup>. 전자의 경우는 데이터의 분포에 따른 밀도 등을 이용하여 추정하는 방식이다. 이것은 일반적으로 각 클러스터 사이의 상

\* 충북대학교 전기전자컴퓨터공학부(jwryu, powerkim, sckyu, sungkim)@chungbuk.ac.kr

논문번호 : KICS2005-06-230, 접수일자 : 2005년 6월 3일

※ 이 논문은 2004년도 충북대학교 학술연구지원사업의 연구비지원에 의하여 연구되었음.

관관계 등을 정확하게 고려하지 않으므로 파라미터 최적화 특성이 상대적으로 미약하다. 후자는 주어진 데이터의 패턴과 클러스터의 수가 일치하는 경우 좋은 성능을 보이지만 사전 정보가 정확하지 않는 경우 원하는 성능을 얻지 못할 수 있다. 또한 클러스터링 기법은 입력 공간에 대해서 수행되는 전형적인 비교사학습(Unsupervised Learning)으로 입력력의 상관관계에 대한 정보를 포함하지 않는다. 이러한 특성들로 인해서 클러스터링 기법은 다른 주된 기법을 구현하기 전에 선행 처리되는 기법으로 많이 이용된다. Pedrycz 등은 입력 공간에서 실시한 기법을 다시 출력공간에서 생성된 정보를 이용하여 학습에 영향을 주는 형태의 교사학습(Supervised Learning)을 통하여 클러스터링의 성능을 개선했지만 이 경우에도 클러스터의 사전정보를 알아야 하는 문제점이 있다<sup>6)</sup>.

본 논문에서는 사전 임계값을 이용하여 자율적으로 클러스터의 수를 추정함과 동시에 이들 파라미터를 최적화하는 기법을 제안하였다. 이 기법은 Chen 알고리즘을 기반으로 클러스터의 수를 추정하고 변형된 Subtractive 클러스터링 기법을 이용하여 클러스터의 수에 따른 연산량 증가문제를 개선하였다<sup>7)</sup>. 모델의 1회 학습 중 클러스터에서 추정된 정보를 퍼지 모델을 통하여 최종 출력을 추정하고 이들 정보를 다시 다음 클러스터 학습에 적용하여 성능에 영향을 주도록 하였다. 클러스터링 학습 후의 파라미터를 다음 모델의 초기 파라미터로 구성하고 순차적으로 모델을 학습하는 기법과는 달리 클러스터링 과정에서 발생하는 추정결과를 특정 모델을 통하여 평가를 받으며 이 결과를 다시 모델의 학습에 이용하는 통합적인 클러스터링 모델을 구성하였다. 제안된 클러스터링 모델은 비선형 공간을 클러스터링을 통하여 부분 선형 공간으로 분할하여 결과를 추정하고 교사학습 기법을 적용하여 매회 학습의 출력에 영향을 주도록 하였다. 이때 클러스터의 파라미터를 이용하여 결과를 추정하는 퍼지 모델을 클러스터링의 조건부 평가 부분으로 고려하였다. 제안된 기법을 Benchmark 문제에 적용하여 유용성을 보였다.

## II. 자기-구성 클러스터링의 모델링

클러스터링의 기본개념은 유사도가 높은 데이터는 같은 클러스터에 속하게 하고 그렇지 않은 클러스터는 다른 클러스터에 속하게 하는 것이다. 이를

위해서 일반적으로 클러스터링 알고리즘에서는 유사도 측정함수를 이용한다. 식 (1)은 제안된 기법에서 이용한 일반화된 유사도 측정함수이다<sup>7)</sup>.

$$r_{ij} = \exp\left(-\frac{1}{2}(x_i - c_j)^T \Sigma_j^{-1}(x_i - c_j)\right) \quad (1)$$

여기서  $c_j$ 와  $\Sigma_j$ 는 각각  $j$ 번째 클러스터의 중심과 분산을 나타내는 파라미터이고 데이터의 수는  $N$ 개 로써  $i=1, 2, \dots, N$ 이며 클러스터의 수는  $c$  로써  $j=1, 2, \dots, c$ 이다. 이 함수는 클러스터의 중심에 데이터가 존재하는 경우 유사도는 1로 하고 거리가 멀어질수록 유사도는 지수함수형태로 감소한다. 현재의 클러스터에 대하여 영향을 받지 않는 데이터에 대해서도 그 유사도가 존재하므로 현재 클러스터 파라미터 추정에 영향을 미치게 된다. 따라서 사전에 지정한 특정 임계값  $\xi$ 를 이용하여 현재 클러스터 추정에 필요한 유사도의 크기를 식 (2)와 같이 제한한다.

$$r_{ij} = \begin{cases} r_{ij}, & \text{if } r_{ij} > \xi \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

이 경우  $\xi$ 의 크기에 따라 클러스터의 추정결과가 달라지는데 크기가 큰 경우 현재 추정하려 하는 클러스터의 정보까지 제거할 수 있으며 너무 작은 경우  $\xi$ 의 필요성이 없어진다.

따라서 이러한 문제점을 보완하기 위하여 식 (1)의 유사도 측정함수의 형태를 식 (3)과 같이 수정한다.

$$r_{ij} = \exp\left(-\frac{1}{2}(x_i - c_j)^T (\xi \cdot \Sigma_j^{-1})(x_i - c_j)\right) \quad (3)$$

사전 임계값을 유사도 제한조건과 유사도 측정함수 내에 역함수 형태로 추가함으로써 사전임계값의 변화에 따른 급격한 클러스터 추정결과 변화를 억제한다. 유사도를 이용하여 클러스터의 중심은 식 (4)를 이용하여 추정한다.

$$c_j = \frac{\sum_{i=1}^N r_{ij} x_i}{\sum_{i=1}^N r_{ij}} \quad (4)$$

클러스터의 중심이 추정되면 이의 공분산  $\Sigma_j$ 는 MLE(Maximum Likelihood Estimation) 기법을 이용하여 추정한다. 이를 추정하기 위한 사전확률(prior probability)은 다음과 같다<sup>5)</sup>.

$$Pr(x_i) = \frac{1}{N} \sum_{j=1}^N u_{ij} \quad (5)$$

여기서  $u_{ij}$ 는 식 (6)의 확률을 정규화(Normalization)한 후에 이전의 사전확률  $Pr(x_i)$ 을 곱한 분할행렬  $U$ 의 요소이며 식 (7)과 같은  $j$ 번째 데이터가  $j$ 번째 클러스터에 속할 확률이다.

$$b_{ij} = \frac{1}{(2\pi)^{d/2} |\Sigma_j|^{1/2}} \times \exp\left(\frac{1}{2}(x_i - c_j)^T \Sigma_j^{-1}(x_i - c_j)\right) \quad (6)$$

여기서  $d$ 는 입력공간의 차원을 말한다.

$$u_{ij} = \frac{b_{ij}Pr(x_i)}{\sum_{j=1}^N b_{ij}Pr(x_i)} \quad (7)$$

제안된 기법의 초기 클러스터의 수는 데이터의 수와 같다. 알고리즘이 수행되는 동안 사전 임계값에 의하여 데이터의 밀도가 밀집된 곳으로 수렴을 한다. 각 클러스터 중심은 모든 데이터에 대하여 유사도를 측정하고 이를 이용하여 클러스터의 수를 수렴시키거나 파라미터를 최적화한다. 데이터의 크기가 큰 경우에는 클러스터 추정 연산이 많아지며 클러스터의 수렴이 늦어지거나 원하는 수렴이 이루어지지 않는 문제점을 갖는다. 이를 해결하기 위해서 변형된 Subtractive 클러스터링 기법을 알고리즘에 도입하였다. 변형된 Subtractive 클러스터링 기법은 추정된 클러스터 중심의 누적 밀도 값을 이용하여 클러스터의 밀집 정도에 따라 밀집 클러스터를 하나의 대표 클러스터로 만들고 인접 클러스터는 소거한다. 밀집된 클러스터의 초기 대표 클러스터는 식 (8)과 같이 추정한다<sup>[24]</sup>.

$$c_{new} = c_j, \text{ if } D_j = \max\left(\sum_{i=1}^N u_{ij}\right) \quad (8)$$

추정된 클러스터 중심을 이용하여 식 (9)와 같이 제거 유사도 범위를 정하여 인접한 클러스터 중심을 소거한다.

$$s_{i,new} = \exp\left(-\frac{1}{2}(c_i - c_{new})^T \Sigma_{new}(c_i - c_{new})\right) \quad (9)$$

이때 식 (2)의 개념을 도입하여 인접한 또 다른 클러스터에 영향을 미치는 것을 제거하기 위해서 수렴 임계값  $\epsilon$ 을 이용하여 식 (9)를 다시 한 번 연산한다.

$$s_{i,new} = \begin{cases} 0, & \text{if } s_{i,new} < \epsilon \\ s_{i,new}, & \text{otherwise} \end{cases} \quad (10)$$

이러한 과정이 끝나면 새로운 대표 클러스터의 중심을 식 (11)과 같이 추정한다.

$$c'_{new} = c'_j, \text{ if } D'_j = \max\left(D_{new} - \sum_{i=1}^c s_{i,i}\right) \quad (11)$$

주어진 종결조건을 만족할 때까지 식 (8)에서 식 (11)을 반복하여 수행한다. 이러한 반복 알고리즘이 모두 종결되면 클러스터링 알고리즘은 1회 학습을 하며 전체 클러스터링 종결조건을 검사하여 만족하는 경우 알고리즘을 끝내고 그렇지 않는 경우 전체 알고리즘을 다시 반복 수행한다.

### III. 클러스터링을 위한 성능평가 모델

본 논문에서 제안된 클러스터링 방법의 성능평가를 위해서 클러스터링에 의해 부분 선형화된 데이터를 Tagaki-Sugeno-Kang(TSK) 퍼지모델에 적용한다<sup>[2]</sup>. TSK 퍼지모델은 소속함수로 구성된 비선형 입력공간으로 고려할 수 있는 전제부와 1차 선형방정식 형태로 구성된 결론부로 구성되어 있다. 따라서 이 모델은 클러스터링의 형태를 갖는 전제부를 그대로 이용하며 부분 선형화된 결과를 결론부에서 선형추론을 가능하게 한다. 또한 각 클러스터에 의한 출력은 퍼지 모델의 규칙부에서 간단하게 구할 수 있다. 그림 1은 기본적인 TSK 퍼지모델의 구조를 나타낸다. 그림에서와 같이 소속함수로 표현할 수 있는 전제부를 클러스터링으로 하고 각 입력차원과 클러스터와의 관계를 퍼지 모델의 규칙부에서 연결하여 결론부에서 선형 출력을 한다.

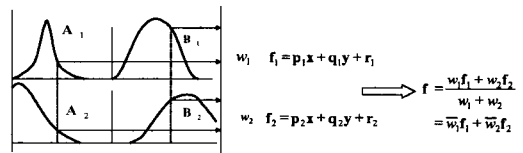


그림 1. TSK 퍼지모델의 개념

클러스터링에서 사용되는 함수와 퍼지 모델에서 사용되는 함수는 식 (12)와 같은 Gaussian 소속함수이므로 함수의 변환 없이 직접 이용한다.

$$u_i = \exp\left(-\frac{1}{2} \frac{(x - c_i)^2}{\sigma^2}\right) \quad (12)$$

퍼지모델에서 각 개별 클러스터링에 의한 부분 출력은 TSK 퍼지모델의 4층에서의 출력과 같다. 5층에서의 최종 출력은 단지 4층에서의 개별 출력을 가중 평균법에 의하여 정리한 것이다. 따라서 각 개별 클러스터의 성능은 4층에서의 개별 출력  $\bar{\omega}_{ij}$ 로 고려할 수 있다. 이것을 주어진 출력 데이터와 비교하여 식 (13)과 같이 개별 클러스터의 추정에 위한 오차를 발생시킨다.

$$e_{j,neurofuzzy} = y - \bar{\omega}_{ij} f_i \quad (14)$$

이를 이용하여 퍼지 모델의 오차가중치 소속도  $ef_{ij}$ 를 식 (15)와 같이 구성한다.

$$ef_{ij} = \exp\left(-\frac{1}{2} \frac{(e_{j,neurofuzzy} - \mu_{ei})^2}{\sigma_{ci}^2}\right) \quad (15)$$

여기서  $\mu_{ei}$ 는 결론부 소속 함수의 중심 파라미터이다. 또한 식 (15)를 이용하여 클러스터의 유사도 식 (3)을 식 (16)과 같이 재구성한다.

$$r_{ij} = r_{ij}(ef_{ij} \cdot (1-t) + t) \quad (16)$$

여기서  $t$ 는 결론부 변수이며 추정된 결과를 유사도 측정에 미치는 영향 범위를 결정한다.

그림 2는 클러스터링을 포함한 전체 구성을 나타낸다. 클러스터링 기법을 이용하여 비선형 입력데이터 공간을 부분 선형공간으로 분할하고 이를 TSK 퍼지 모델을 이용하여 선형 추론한다. 그리고 발생된 오차를 다시 개별 클러스터링 학습에 영향을 주도록 이용한다. 마지막으로 그림에서 볼 수 있듯이 TSK 퍼지 모델의 선형추론 특성과 각 규칙의 할당을 이용하여 클러스터링의 성능을 평가한다.

제안된 방법의 학습구조에서는 TSK 퍼지 모델의

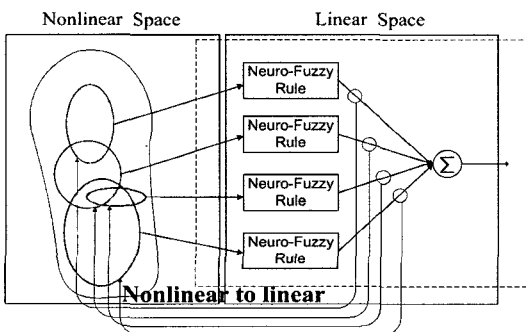


그림 2. 제안된 방법의 학습구조

성능을 클러스터링의 성능으로 볼 수 있으며 결론부의 파라미터들은 최소자승법을 이용하여 추정한다. 기존의 비선형모델의 파라미터 추정 방법보다 선형모델에 대해서 파라미터를 추정함으로써 더 쉽게 연산할 수 있는 장점을 갖는다. 또한 TSK 퍼지 모델을 변형하지 않고 원형대로 이용하고 전제부 학습 방식만 적용하기 때문에 이해가 편리하다.

#### IV. 모의실험 및 결과

제안된 방법의 타당성을 보이기 위해서 본 논문에는 Box-Jenkins의 가스르 데이터를 이용하였다. 이 데이터는 전형적인 비선형 시계열 데이터로써 가스르에서 9초 간격의 샘플링을 통하여 취득된 총 296개의 입출력 데이터쌍으로 시계열 데이터로 구성되어 있다<sup>[2]</sup>. 입력은 메탄 흐름률이고 출력은 이산화탄소의 비율이며 데이터의 구성은  $y(t-1)$ ,  $y(t-2)$ ,  $y(t-3)$ ,  $y(t-4)$ ,  $u(t-1)$ ,  $u(t-2)$ ,  $u(t-3)$ ,  $u(t-4)$ ,  $u(t-5)$ ,  $u(t-6)$ 로 되어 있다. 모델의 학습을 위해서 입력은 Jang<sup>[2]</sup>의 방법에 따라서  $y(k+1) = f(y(k), u(k-3))$ 를 학습 및 검증 데이터로 이용하였다. 또한 296개의 데이터쌍 중에서 홀수 데이터를 모델의 학습에 이용하였으며 짝수 데이터를 모델의 검증에 이용하였다.

먼저 사전 유사도  $\xi$ 를 0.2로 설정하고 결론부 변수를 0.4로 하여 50회 학습한 클러스터링의 추정 결과를 그림 3에 나타내었다.

표 1과 그림 4에 임계값을 가변하였을 때 추정된 규칙수의 변화를 나타냈다. 임계값이 작아질수록 클러스터의 수가 증가하는 것으로 볼 수 있다. 임계값이 0.05에서 2.5까지 크게 변화함에도 불구하고 클러스터의 수가 급변하지 않는 것을 확인할 수 있다.

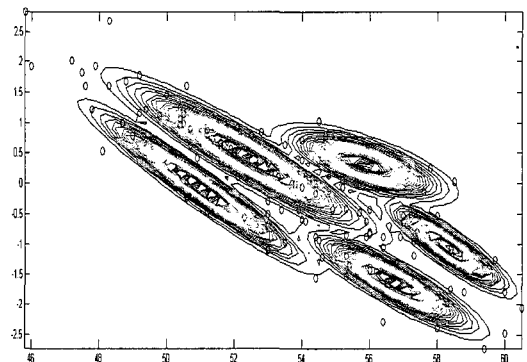


그림 3. Box-Jenkins 가스르 데이터의 클러스터링 결과

표 1. 학습 후 임계값에 의해서 생성된 규칙수

$\xi$ \ t	0.1	0.2	0.3	0.4	0.5	0.6	0.7
0.05	8	13	11	10	9	11	4
0.06	14	12	12	9	10	7	6
0.07	14	14	11	14	9	7	6
0.08	14	13	12	10	12	8	5
0.09	12	12	10	13	9	8	X
0.1	13	14	10	9	9	7	X
0.13	13	11	6	9	6	6	X
0.15	11	8	9	7	7	4	X
0.17	10	8	9	5	6	5	X
0.2	8	7	6	5	4	4	3
0.25	2	4	3	3	4	4	X

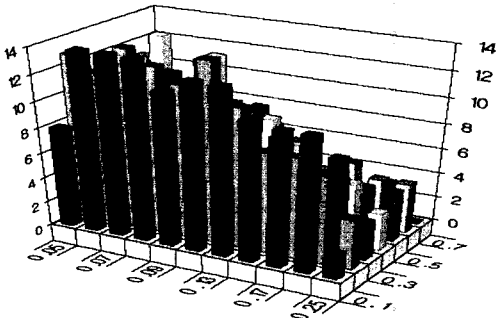


그림 4.  $\xi$ 와  $t$ 의 변화에 의한 규칙수 변화

여기서  $\times$ 로 표기된 부분은 클러스터링이 실시되지 않은 영역을 나타낸다.

또한 클러스터에 의한 전체 모델의 성능을 알아보기 위해서 표 2와 3에 각각 Box-Jenkins 가스로에 데이터에 대한 클러스터링 학습오차와 검증오차의 결과를 나타내었다.

표 2와 3을 보면 클러스터링에 의한 학습데이터의 오차뿐만 아니라 검증데이터에 대한 오차도 일정한 범위 내에서 동시에 감소하는 것을 알 수 있다. 표 4에서 제안된 기법의 모델 성능과 기존 연구 결과의 모델 성능을 비교하였다. 표 4를 보면 제안된 기법의 학습오차와 검증오차가 기존 연구결과에 비해서 크지 않음을 알 수 있다.

일반적인 모델에서 학습에 대한 오차 뿐만 아니라 검증모델에 대한 오차는 학습 모델의 과적합(overfitting) 문제로 인하여 중요한 관심대상이 된다. 주어진 학습데이터에 대해 지나친 학습을 실시하게 되면 학습데이터에 대한 성능은 우수한 반면 모델의 성능을 검증하기 위한 데이터에 대하여 적용하지 못하는 문제점들이 발생할 수 있다. 같은 시스템에서 취득된 입출력 데이터임에도 불구하고 과

표 2. Box-Jenkins 가스로의 학습오차(RMSE)

$\xi$ \ t	0.1	0.2	0.3	0.4	0.5	0.6	0.7
0.05	0.1294	0.0991	0.1333	0.1361	0.1412	0.1200	0.1496
0.06	0.0977	0.1132	0.1208	0.1331	0.1385	0.1393	0.1228
0.07	0.0972	0.1088	0.1237	0.0960	0.1429	0.1263	0.1390
0.08	0.0974	0.1213	0.0955	0.1279	0.1050	0.1219	0.1239
0.09	0.1018	0.1211	0.1346	0.1186	0.1263	0.1566	X
0.1	0.1157	0.1037	0.1081	0.1282	0.1250	0.1297	X
0.13	0.1158	0.1249	0.1567	0.1446	0.1180	0.1258	X
0.15	0.1307	0.1373	0.1367	0.1327	0.1288	0.1160	X
0.17	0.1208	0.1533	0.1245	0.1285	0.1287	0.1295	X
0.2	0.1294	0.1509	0.1270	0.1318	0.1526	0.1358	0.1652
0.25	0.1682	0.1666	0.1414	0.1381	0.1356	0.1311	X

표 3. Box-Jenkins 가스로의 검증오차(RMSE)

$\xi$ \ t	0.1	0.2	0.3	0.4	0.5	0.6	0.7
0.05	0.1762	0.2243	0.1669	0.1691	0.2071	0.2147	0.1873
0.06	0.2078	0.9866	0.2333	0.2181	0.1628	0.2312	0.1893
0.07	79.251	0.1690	0.2153	0.2963	0.1729	0.1769	0.1986
0.08	0.1510	0.2191	0.1919	0.3015	0.3153	0.1523	0.1544
0.09	0.2205	0.2005	920	2.3614	0.4344	0.1847	X
0.1	0.3615	0.1697	1.5345	0.1863	0.1894	0.2066	X
0.13	0.1827	0.1707	0.1707	0.2699	0.4005	0.2127	X
0.15	0.1560	0.2003	0.1675	0.1749	0.1624	0.1765	X
0.17	0.1703	0.1868	0.1634	0.1552	0.4352	0.1816	X
0.2	0.1762	0.1750	0.1529	0.1863	0.1745	0.1899	0.1762
0.25	0.1808	0.2028	0.1912	0.1994	0.1939	0.2087	X

학습이 발생하는 경우 검증오차에 대한 문제점이 발생할 수 있다. 클러스터링 기법에 의한 모델의 성과 학습 방법에서는 데이터의 형태를 고려한 학습을 실시함으로써 이러한 문제점을 개선할 수 있다. 제안된 클러스터링 기법은 검증데이터에 대한 오차 결과에서 볼 수 있듯이 과적합 문제를 개선하였다.

기존의 연구결과와 함께 제안된 모델의 성능을 비교하면 표 4에서와 같이 설명할 수 있다.

기존의 연구들에서 볼 수 있듯이 다양한 기법에서 주어진 입출력관계를 이용하여 모델링을 최적화하려는 시도들이 이어져 왔다. Babuska 등은 전체 부 소속함수를 클러스터링 기법을 이용하여 최적화 하였지만 클러스터링 과정에 입출력 상관관계를 고려하지 않았으며 Jang의 경우 모델을 오차 미분기반 학습 방법을 이용하였지만 일반적인 격자분할에

표 4. 이전 연구결과와의 모델성능 비교

Method	Error (RMSE)	rules	Training error	Checking error	Remark
Pedrycz[8]		81	0.320	None	
Xu[9]		25	0.328	None	
Sugeno[10]		2	0.359	None	
Abonyi[11]		16	0.154	None	
Babuska[12]		23	0.124	None	
Jang[13]		4	0.135	0.530	
Oh[14]		4	0.026	0.272	
Proposed $\xi=0.2, t=0.4$		9	0.1282	0.1863	50 epoch
Proposed $\xi=0.2, t=0.4$		5	0.1267	0.1523	100 epoch
Proposed $\xi=0.1, t=0.2$		14	0.1037	0.1697	50 epoch

의한 전체부 소속함수 생성 방식을 이용하여 입력 차원이 증가하거나 소속함수의 수가 증가할 경우 모델의 학습과 전체 크기에 대하여 문제점이 발생하였다<sup>[12][13]</sup>. Oh 등은 신경회로망과 퍼지 모델을 융합한 모델을 이용하여 모델링을 실시하고 유전 알고리즘을 이용하여 최적화를 실시하였지만 검증 오차가 증가하는 문제점이 있었다<sup>[14]</sup>. 반면 제안된 기법에서는 비교사 학습인 클러스터링을 교사학습으로 확장하였으며 전체 학습을 하나로 구성하는 학습 방식을 통하여 학습오차 뿐만 아니라 검증오차까지 감소하였다. 주어진 데이터의 입력출력관계를 이용하여 자율적으로 클러스터의 수를 수렴하고 최적화하는 기법을 통하여 주어진 시스템을 더욱 정확하게 묘사할 수 있는 모델을 생성하였다.

### V. 결론

본 논문에서는 자기-구성 클러스터링을 통하여 클러스터의 수와 이들 파라미터를 동시에 추정하고 최적화하는 기법을 제안하였다. 알고리즘을 수행하는 동안 연산량이 증가하는 문제를 개선하기 위해서 변형된 Subtractive 클러스터링 기법을 도입하였고 파라미터 최적화를 위해서 MLE를 이용하였다. 또한 출력의 정보를 이용한 교사학습을 통하여 클러스터링 학습 정보를 추가하였다. 클러스터링에 의해서 부분 선형화된 클러스터의 성능을 평가하기 위해서 TSK 퍼지 모델을 적용하였다. 그리고 모의 실험을 통하여 제안된 기법이 우수한 성능을 보임

을 검증하였다. 제안된 클러스터링 기법에서 연산량 증가문제는 어느 정도 개선됐지만 데이터의 크기가 증가할수록 연산량이 증가하는 문제를 해결하기 위한 지속적인 연구가 필요하다.

### 참고 문헌

- [1] Chin-Teng Lin, C. S. George. Lee, *Neural Fuzzy Systems : A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall, 1996.
- [2] J-S. R. Jang, C. T. Sun, E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall 1997.
- [3] 김승석, 김성수, 유정웅, “새로운 클러스터링 알고리즘을 적용한 향상된 뉴로-퍼지 모델링”, *대한전기학회 논문지*, Vol. 53D, No. 7, pp. 536-543, 2004.
- [4] R. R. Yager, D. P. Filev, “Generation of Fuzzy Rules by Mountain Clustering,” *Journal of Intelligent and Fuzzy System*, Vol.2, pp. 209-219, 1994.
- [5] Guorong Xuan, Wei Zhang, Peiqi Chai, “EM algorithm of Gaussian Mixture Model and Hidden Markov Model,” *Image Processing Proceedings, International Conference on*, Vol. 1, pp. 145-148. 2001.
- [6] Witold Pedrycz, “Conditional Fuzzy Clustering in the Design of Radial Basis Function Neural Network,” *IEEE Trans on. Neural Network*, Vol. 9, No. 4, pp. 601-612, 1998.
- [7] Ching-Chang Wong, Chia-Chong Chen, “A Hybrid Clustering and Gradient Descent Approach for Fuzzy Modeling,” *IEEE Trans on Systems, Man, and Cybernetics-Part B : Cybernetics*, Vol. 29, No.6, pp. 686-693, 1999.
- [8] Witold Pedrycz, “An Identification Algorithm in Fuzzy Relational Systems,” *Fuzzy Sets and Systems*, Vol. 13, pp. 153-167, 1984.
- [9] C. Xu and L. Lu, “Fuzzy model Identification and Self-Learning for Dynamic Systems,” *IEEE Trans on Systems, Man and Cyberne-*

