

# 퍼지 클러스터링을 이용한 사례기반 추론의 성능 개선

## Performance Improvement of Case-based Reasoning Using Fuzzy Clustering

현우석

한국성서대학교 정보과학부

Woo-Seok Hyun

Dept. of Information and Science, Korean Bible University

E-mail : wshyun@bible.ac.kr

### ABSTRACT

사례 기반 추론(case-based reasoning)은 과거에 유사하게 수행된 적이 있는 사례를 유추하고, 유추된 사례의 해를 이용하여 현재의 문제를 해결하는 기법으로서 규칙 기반 추론과 함께 여러 분야에 이용되고 있다. 하지만 사례기반 추론시 사례베이스로부터의 유사성에 근거한 검색을 해야 하므로 사례베이스의 크기가 증가하게 되면 검색시간이 길어지게 되거나 적절하지 못한 사례가 조회될 수 있다. 특히 사례베이스 내의 모든 사례에 대하여 유사도를 계산하게 되기 때문에 수행속도가 현저히 저하되는 문제점을 지니고 있다.

본 논문에서는 규칙 및 퍼지 클러스터링에 의한 사례기반추론을 이용한 E-FFIS(Enhanced-Fire Fighting Intelligent System)를 제안한다. 제안하는 시스템은 기존의 H-FFIS(Hybrid-Fire fighting Intelligent System)와 비교해 보았을 때 수행시간을 감소시키면서 정확성을 높이게 되었다.

**Key Words:** Case-based Reasoning, Fuzzy Clustering

### 1. 서론

지금까지 개발된 지능시스템들은 전문지식을 생성 규칙(production rule)[1]으로 표현한 규칙 기반 시스템이 대부분이다. 퍼지논리를 이용한 선박의 화재진압통제 지능시스템(FFIS)[2]은 선박에서 화재발생시 선박이 처한 상황에 따라 가장 적절한 후속조치를 취하게 하는 시스템으로서 화재관련 전문지식을 생성규칙으로 표현하는 규칙기반 시스템이다. 실제로 화재를 탐지하는데 필요한 지식은 정형화된 규칙만으로 표현하기 어려우며, 과거의 사례 발생을 기초로 탐지하는 경우가 적지 않다. 또한 시스템의 성능 향상을 위해 규칙을 계속 수정하고 추가해야 하는 단점이 있으며, 예외적인 상황에서 화재가 발생시 적시에 화재를 탐지하는데 문제점을 지니고 있다. 이

런 문제점을 해결하고자 사례기반 추론[3-6]에 의해 확장된 하이브리드 화재진압통제 지능시스템(H-FFIS:Hybrid - Fire Fighting Intelligent System)[7]이 제안되었다. 그런데 사례기반 추론시 사례베이스로부터 유사성에 근거한 검색을 해야 하므로 사례베이스의 크기가 증가하게 되면 사례베이스 안의 모든 사례들을 검색해야 하기 때문에 검색시간이 길어지게 되는 문제점이 발생하게 된다[8].

본 논문에서는 규칙 및 퍼지 클러스터링에 의한 사례기반추론을 이용한 E-FFIS(Enhanced-Fire Fighting Intelligent System)을 제안한다. 제안하는 시스템은 기존의 H-FFIS 와 비교해 보았을 때 수행시간을 감소시키면서 정확성을 높이게 되었다.

## 2. 퍼지 클러스터링

본 논문에서는 예외상황 사례베이스에 들어있는 사례들을 클러스터링하기 위하여 Bezdek의 FCM(Fuzzy C-Means)[9-11] 알고리즘을 사용하였으며, 다음과 같다.

사례베이스를  $c$ 개의 클러스터로 분류할 때 각 클러스터의 중심 벡터  $v_i(i=1,2, \dots, c)$ 와 데이터  $x_k$ 와의 비유사도(dissimilarity)  $d_{ik}$ 는 유클리디언(Euclidean) 거리로 구하며 식 (1)과 같다.

$$d_{ik} = d(x_k, v_i) = \|x_k - v_i\| = \left\{ \sum_{j=1}^p (x_{kj} - v_{ij})^2 \right\}^{(1/2)} \quad (1)$$

FCM 알고리즘은 식 (2)의 목적함수를 최소화 하는데 목적이 있다.

$$J(U, v) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^m \|x_k - v_i\|^2 \quad (2)$$

이 때, 클러스터 중심( $v_i$ )과 소속도 함수 값( $u_{ik}$ )은 식 (3)과 (4)를 이용하여 구하며 다음과 같다.

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}, 1 \leq i \leq c \quad (3)$$

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}} \quad (4)$$

FCM 알고리즘은 다음과 같이 4 단계로 표현할 수 있다.

1단계: 클러스터의 수  $c$ 와 지수가중치  $m$  값을 선정하고, 퍼지  $c$ 분할행렬 ( $U^{(1)}$ )을 초기화한다.

$$2 \leq c \leq n$$

$$1 < m < \infty$$

2단계: 1단계에서 구한 ( $U^{(1)}$ )과 식(3)을 이용하여 새로운 클러스터의 중심  $v_i^{(1)}(i=1,2,\dots,c)$ 을 구한다.

3단계:  $x_k \neq v_i^{(1)}$  일 때  $U^{(1)}$ 를 식 (4)에 의해  $U^{(1+1)}$ 로 업데이트하고, 그 외에는

$$u_{ik}^{(1+1)} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \text{로 둔다.}$$

4단계:

$$\|U^{(1+1)} - U^{(1)}\|_G \leq \epsilon \text{의 조건이 만족되면 클러스}$$

터링을 종결하고 그렇지 않으면  $l$ 을 1 증가시킨 후 2단계로 되돌아가서 반복 수행한다.

## 3. E-FFIS(Enhanced-Fire Fighting Intelligent System)

선박의 각 컴파트먼트에는 열 감지기, 화염 감지기, 연기 감지기 등이 설치되어 있어서 화재를 탐지하게 된다. 규칙기반 FFIS[2]에서는 시스템의 성능 향상을 위해 규칙을 계속 수정하고 추가해야 하는 단점이 있으며, 예외적인 상황에서 화재가 발생시 적시에 화재를 탐지하는데 문제점을 지니고 있다. 이런 문제점을 해결하고자 사례기반 추론에 의해 확장된 H-FFIS[7]이 제안되었으나, 사례기반 추론시 사례베이스로부터 유사성에 근거한 검색을 해야 하므로 사례베이스의 크기가 증가하게 되면 수행시간이 증가되는 문제점을 지니고 있다. 본 논문에서는 사례베이스 크기의 증가에 따른 수행시간의 증가 문제를 해결하고자 퍼지 클러스터링을 사용하여 사례베이스를 클러스터링함에 의해서 수행시간을 감소시키면서 정확성을 높이게 되었다. 규칙 및 퍼지 클러스터링에 의한 사례기반추론을 이용한 개선된 화재진압 통제 지능시스템(E-FFIS)의 구조는 그림 1과 같으며, 본 시스템의 화재탐지 과정은 그림 2와 같다. 먼저 각 감지기의 감지값이 입력되어 규칙으로 표현된 화재제어 지식베이스를 기반으로 퍼지논리를 이용하여 화재탐지를 수행하고, 화재탐지에 실패했을 경우 퍼지 클러스터링을 이용한 예외상황 사례베이스를 기반으로 화재탐지를 재시도하게 된다.

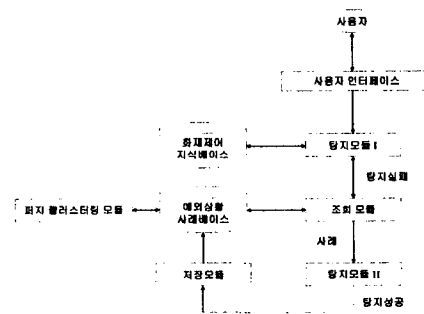


그림 1 E-FFIS의 구조

### 3.1 사용자 인터페이스

사용자가 시스템을 원활히 사용할 수 있도록 시스템과 사용자간을 연결해 주는 기능을 갖는다. 선박의 각 컴파트먼트마다 부착된 여러 개의 연기감지 센서, 화염감지 센서, 열감지 센서로부터의 정보를 받아들이고 화재탐지의 결과를 사용자에게 제시해 준다.

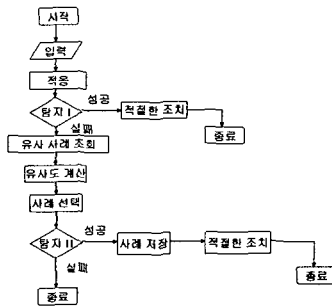


그림 2 화재탐지 과정

3.2 화재제어 지식베이스

화재 진압 통제를 제어하기 위한 사실이나 규칙으로써 구성되며, 다음의 예는 생성규칙을 이용한 화재탐지를 위한 기초적인 관련 지식을 표현한다[2].

규칙 1)

IF HR(t) is  $m_i$  AND HI(t) is  $m_j$   
 THEN HPF(t) is  $m_k$  [단,  $m_k = \min(m_i, m_j)$ ]  
 PR(t) is 0.9

3.3 탐지모듈 I

화재제어 지식베이스를 기반으로 하는 화재탐지는 전향추론(forward chaining) 기법을 사용한다.

3.4 퍼지 클러스터링 모듈

예외상황 사례베이스에 들어 있는 사례들을 2절의 FCM[9-11] 알고리즘을 사용하여 퍼지 클러스터링 한다. 퍼지 클러스터링 모듈이 실행되기 위해서는 '열감지 센서 반응정도', '화염감지 센서 반응정도', '연기감지 센서 반응정도', '열감지 센서 증가정도', '화염감지 센서 증가정도', '연기감지 센서 증가정도' 등의 속성이 필요하다. 모듈의 실행이 끝나면 그 결과로 c개의 클러스터와 클러스터별로  $v_i$ 개의 중심이 도출된다. 이 결과는 클러스터별로 인덱싱 되어 예외상황 사례베이스에 저장되며 이 정보는 조회모듈에서 사례들을 조회할 때 사용된다. 본 연구에서는 c 값은 3,  $v_i$  값은 1로 정했다.

3.5 예외상황 사례베이스

예외적인 경우 화재가 탐지되는 각 사례들은 퍼지 관계형 데이터베이스에서 하나의 테이블 형태로 저장되도록 설계하였다.

<표 1>는 초기 사례베이스의 일부분을 보여준다.

표 1 A-FFIS의 사례 베이스

열감지 센서 반응정도	화염감지 센서 반응정도	연기감지 센서 반응정도	열감지 센서 증가정도	화염감지 센서 증가정도	연기감지 센서 증가정도	유사도	참조횟수
0.59	0.62	0.53	0.62	0.68	0.59		0
0.62	0.68	0.41	0.65	0.76	0.47		0
0.69	0.54	0.35	0.78	0.51	0.39		0
⋮	⋮	⋮	⋮	⋮	⋮		⋮

3.6 조회모듈

탐지모듈 I에서 화재탐지에 실패할 경우, 퍼지 클러스터링에 의하여 클러스터별로 인덱싱되어 있는 예외상황 사례베이스에 있는 사례들을 조회하여 현 상황과 유사한 사례를 찾아낸다. 유사한 사례를 찾기 위해서는 유사도를 사용하게 되는데, 퍼지 근접관계[12]를 사용하였으며 다음과 같다.

$$SM(case_i, case_j) = 1 - \frac{\sum_{k=1}^n |attr_{ik} - attr_{jk}|}{|n|}$$

n: 속성 수

case<sub>i</sub>: 현재 사례

case<sub>j</sub>: 과거 사례

attr<sub>ik</sub>: 현재 사례를 구성하는 k번째 속성을 나타내는 퍼지값 (1 ≤ k ≤ n)

attr<sub>jk</sub>: 과거 사례를 구성하는 k번째 속성을 나타내는 퍼지값 (1 ≤ k ≤ n)

3.7 탐지모듈 II

시스템이 제시하는 사례를 기초로 화재발생 유무를 탐지한다.

3.8 저장 모듈

조회모듈에서 제시한 사례를 기초로 화재탐지에 성공하고, 조회된 사례와 현재의 상황 사례가 완전히 동일하지 않으면 현재 상황 데이터를 예외상황 사례베이스에 새로운 사례로 저장시킨다. 이 때 클러스터링 모듈로 제어가 넘어가 새로운 사례가 어느 클러스터에 속하는지 결정하여 해당 클러스터에 저장시키게 된다.

4. 평가

본 시스템의 성능을 평가하기 위하여 기존의 H-FFIS와 제안하는 E-FFIS을 비교하였다. 시뮬레이션 환경에서는 화재가 발생하게 되는 200가지 사건을 수집하여 100가지 사건을 사례베이스

에 저장하고 나머지 100가지 사건은 테스트용으로 사용하였다. 여기서 사건이란 시스템에 일정 시간에 입력되게 되는 30개의 센서의 감지값을 모두 포함하는 것이다. 두 시스템에서 100가지 사건을 10 가지 test set으로 나누어 실험을 해서 평균 수행시간과 평균 화재탐지율을 비교하였는데 그림 3, 4 와 같다.

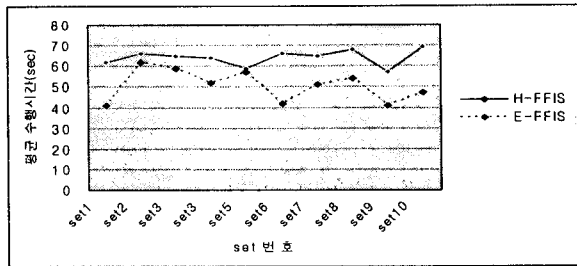


그림 3 시스템에 따른 평균 수행시간 비교

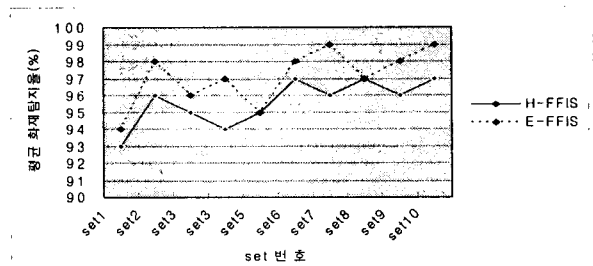


그림 4 시스템에 따른 평균 화재탐지율 비교

### 5. 결론 및 향후과제

퍼지 클러스터링에 의한 사례기반 추론을 하는 제안하는 E-FFIS는 기존의 H-FFIS와 비교해 볼 때 평균 수행시간이 감소되었다. 또한 기존의 H-FFIS보다 평균 화재탐지율도 향상시켰다.

향후 연구과제로는 퍼지 클러스터링에서 클러스터나 중심의 개수 선정에 따라 결과가 다르게 나타날 수 있으므로 최적의 클러스터나 중심의 개수를 선정하는 연구가 남아있다.

### 6. 참고문헌

[1] J. Giarratano and G. Riley, 'Expert Systems Principles and Programming', 2nd Ed., PWS, 1994.  
 [2] 현우석, 김용기, "통합플랫폼관리체제에서 퍼지논리를 이용한 화재진압통제 지능시스템의 설계 및 구현". 한국정보처리학회 논문지, Vol. 7 No. 7, pp.2095-2108, 2000.7.  
 [3] R. Barletta, "Case-based reasoning and information retrieval: Opportunities for

technology sharing", *IEEE Expert*, Vol.8, No.6, pp.2-3, 1993.

[4] M. P. Feret and J. I. Glasgow, "Hybrid Case-Based Reasoning for the Diagnosis of Complex Devices", *Proc. of the National Conf. on Artificial Intelligence(AAAI-93)*, pp.168-175, 1993.

[5] J. L. Kolodner, "Improving human decision making through Case-base decision aiding", *AI Magazine*, Vol.12, No.2, pp.52-68, 1991.

[6] A. Aamodt and E. Plaza, "Foundational Issues, Methodological Variations, and System Approaches", *AI Communications*, Vol.7, No.1, pp.65-79, 1994.

[7] Woo-Seok Hyun, Yong-Gi Kim, "Improving Accuracy by Combining Rule-based and Case-based Reasoning for Fire Detection in a Ship", *Proceedings of The 1st International Conference on EALPIIT*, pp.232-235, Shenyang, China, August, 2000.

[8] D. W. Aha and D. Wettschereck, "Case-Based Learning: Beyond Classification of Feature Vectors," *Proceedings of the European Conference on Machine Learning*, 1997.

[9] N. Pal and J. Bezdek, "On Cluster Validity for the Fuzzy C-Means Model," *IEEE Trans. on Fuzzy Systems*, Vol. 3, No. 3, pp.370-379, 1995.

[10] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981.

[11] R. L. Cannon, J. V. Dave and J. C. Bezdek, *Efficient Implementation of the Fuzzy C-Means Clustering Algorithms*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.8, No.2, 1986, pp.248-255.

[12] G. Klir and T. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall International Editions, 1992.