

# 바이러스진화 유전 알고리즘을 이용한 퍼지 모델링 Fuzzy Modeling Using Virus-Evolutionary Genetic Algorithm

이승준 · 주영훈\* · 박진배\*\*

Seung-Jun Lee, Young-Hoon Joo\* and Jin-Bae Park\*\*

延世大 電氣 · 컴퓨터工學科

\* 群山大 電子情報工學部

\*\* 延世大 電氣 · 컴퓨터工學科

## 요 약

본 논문은 기존의 수학적인 모델링으로는 만족스러운 결과를 얻기 어려운 복잡하고 불확실한 비선형 시스템에 대한 퍼지 모델링 기법을 다룬다. 유전 알고리즘은 어느 정도 최적해를 전역적으로 찾을 수 있기 때문에 퍼지 모델링시에 파라미터와 구조를 동정하기 위하여 사용되었다. 하지만, 유전 알고리즘은 개체군이 유전적 다양성을 잃었을 경우 조기 수렴한다는 문제점이 있으며 바이러스진화 유전 알고리즘은 이러한 지역수렴에 대한 방안이 될 수 있다. 따라서, 본 논문에서는 바이러스 이론이 적용된 VEGA를 퍼지 모델링할 때 이용할 수 있는 방법을 제안한다. 이 방법에서는 지역정보가 개체군 내에서 교환됨으로써 유전적 다양성을 유지하게 된다. 마지막으로, 본 논문에서 제안한 방법의 우수성과 일반성을 평가하기 위해 몇 가지의 수치적 예제를 제공한다.

## ABSTRACT

This paper deals with the fuzzy modeling for the complex and uncertain nonlinear systems, in which conventional and mathematical models may fail to give satisfactory results. Genetic algorithm has been used to identify parameters and structure of fuzzy model because it has the ability to search optimal solution somewhat globally. The genetic algorithm, however, has a problem, which optimization process can be premature convergence in the case of lack of genetic divergence of population. Virus- evolutionary genetic algorithm(VEGA) could be a strategy against this local convergence. Therefore, we use VEGA for fuzzy modeling. In this method, local information is exchanged in population so that population can sustain genetic divergence. finally, to prove the theoretical hypothesis, we provide numerical examples to evaluate the feasibility and generality of fuzzy modeling using VEGA.

## 1. 서 론

복잡한 비선형 시스템을 수학적으로 모델링할 경우 고차의 모델을 이용하거나 모델 근사화를 이용하여 낮은 차수로 표현하지만 결과는 만족하지 않다고 알려져 있다. 이러한 복잡한 비선형 시스템을 표현하기 위해 퍼지 모델링 기법이 활발히 연구되어 왔다[1-6]. Zadeh[7]에 의해 최초로 도입된 퍼지 논리 개념은 각 계에서 새로운 가능성을 보여주었다. Tong[1]은 퍼지 논리를 바탕으로 시스템을 묘사하는 퍼지 모델링 기법을 시도하여 복잡한 비선형 시스템의 모델링에 새로운 접근 방법을 제시하였다. 그 후, Pedrycz[2]는 상대 퍼지 집합과 Zadeh의 조건 확률 분포에 기초하여 퍼지 시스템의 새로운 동정 알고리즘을 제안하였고, Xu[3]는 퍼지 규칙의 구조와 파라미터 동정과 연관된 자기 학습 알고리즘을 포함하는 일반적인 퍼지

시스템 동정 기법을 제안하였다. Sugeno[4-6]는 선행적 지식이 없는 시스템의 입출력 데이터만으로 시스템의 정성적 모델을 구하는 방법과 규칙의 후반부를 선형식으로 대체하는 방법으로 퍼지 모델 동정의 일반적 기법을 제안하였다.

퍼지 모델링의 가장 주된 목적은 주어진 시스템을 퍼지 규칙과 퍼지 집합으로 정확히 묘사하는데 있다. 그러나, 기존의 퍼지 모델은 전문가의 지식에 기초하고 있지만 이러한 방법에 의해 퍼지 모델을 구하는 것은 많은 시간이 필요하다. 최근에는 유전 알고리즘(genetic algorithm; GA)과 클러스터링 기법을 이용해 퍼지 모델의 파라미터와 구조를 동정하는 자동 동정 기법이 연구되어 왔다[8,9]. 이 중 GA는 주요한 현대 생물학적 사고의 하나로서 컴퓨터를 이용해 선택, 교배와 돌연변이로 구성된 생물학적 진화를 모방하는 최적화 작업이다. 즉, 퍼지 모델링시 선택적인 지식으로

부터 결정해야하는 소속함수의 파라미터와 규칙의 구조 등을 결정하기 위해 후보 해 공간을 설정하여 개체군을 생성하고 GA의 연산자를 이용해서 다음 세대의 개체군을 구하는 방법에 의해 최적해를 향해 진화해 나가게 된다. 여기에서 GA의 연산자란 교배와 선택으로 이 연산자들은 효과적인 스키마타를 증가시키기도 하지만 스키마타를 직접적으로 다루지는 않는다. 그리고, 비례적 선택은 동 개체군내 다른 개체의 적합도보다 훨씬 높은 적합도를 가지는 개체가 출현할 경우 비례적으로 선택되는 GA의 특성에 기인하여 해당 개체의 유전 정보가 전체 개체군내로 전파되기 때문에 조기 지역 수렴의 원인이 될 수 있을 뿐 만 아니라 효과적인 스키마타를 증가시키면서 동시에 비효과적인 스키마타도 함께 증가시킨다는 문제점이 있다. 이에 의해 일정한 수준의 적합도에 도달했을 경우나 같은 세대에서의 다른 개체보다 우수한 형질을 보이는 개체가 출현한다면 더 이상 적합도의 증가를 보장하기 힘들다. 이를 막기 위해 생태학적인 모델[10,11], 평행 GA(parallel GA)[12], 혼성 GA(hybrid GA)[13], 바이러스-진화 유전 알고리즘(virus-evolutionary genetic algorithm; VEGA)[14] 같은 다양한 기법이 제안되어 왔다.

본 논문에서는 VEGA를 이용하여 퍼지 규칙의 소속함수의 파라미터와 규칙의 구조를 동시에 자동 동조할 수 방법을 제안한다. 일반적으로, VEGA는 호스트와 바이러스 개체 사이의 수평적 진화와 세대간 유전적 정보의 수직적 진화 모두를 기반으로 하여 공동 진화를 모의한다는 특징이 있다. 본 논문에서는 이러한 특징을 이용하여 복잡한 비선형 시스템의 퍼지 모델을 구하는 방법을 제안한다. 즉, 퍼지 모델의 파라미터와 규칙의 구조를 임의의 수로 암호화하여 호스트 개체를 구성하고, 미리 정해 놓은 수의 감염될 개체와 각 개체의 부분자열을 결정하여 바이러스 개체를 생성한 후, 각 호스트 개체의 적합도와 각 바이러스 개체의 적합도를 각각 적합도 함수에 의해서 감염된 호스트 개체의 적합도의 차의 합에 의해서 구한다. 그리고, 동일한 세대내에서는 바이러스 개체가 감염률에 의해 역전사 연산자(reverse transcription operator)가 작용하게 되며, 다음 세대로 진화할 때에는 호스트 개체는 교배, 재생, 돌연변이에 의해서 다음 세대의 개체를 구하고 바이러스 개체는 생명력에 의해서 존속과 비존속을 결정하게 된다. 마지막으로, 본 논문에서 제안한 VEGA를 이용한 퍼지 모델링 방법의 효율성과 우수성을 검증하기 위해 비선형 시스템과 가스로의 퍼지 모델링에 응용한다.

## 2. 퍼지 모델

퍼지 모델링이란 언어 규칙인 IF-THEN을 사용하여 시스템을 묘사하기 위한 접근 방법으로 퍼지 집합을 이용하여 시스템을 모델링하는 것이다. 따라서, 퍼지 모델링의 초기 작업은 모델링 과정을 통하여 동정될 퍼지 모델의 형태를 정하는 것이며 본 논문에서는 다음과 같은 모델로 정의한다.

$$\text{Rule } i : \text{IF } x_i \text{ is } F_i^i, \dots, \text{ and } x_n \text{ is } F_n^i, \text{ THEN } y_i \text{ is } O_i \quad (1 \leq i \leq c) \quad (1)$$

여기서, Rule  $i$ 는  $i$ 번째 규칙( $1 \leq i \leq c$ )을 말하며,  $x_j$  ( $1 \leq j \leq n$ )은  $j$ 번째 입력 변수,  $y_i$ 는  $i$ 번째 규칙의 출력으로 실수치  $O_i$ 를 갖으며,  $F_j^i$  ( $1 \leq j \leq n$ )는 퍼지 집합을 나타낸다. 또한, 최종 출력은 식 (2)와 같이 정의된다.

$$y = \sum_{i=1}^c \mu_i O_i \quad (2)$$

$$\text{여기서, } \mu_i = \frac{\omega_i}{\sum_{i=1}^c \omega_i}, \quad \omega_i = \prod_{j=1}^n F_j^i(x_j)$$

또한, 퍼지 집합  $F_j^i(x_j)$ 는 실제의 값을 퍼지 집합으로 대응시키는 함수로서  $i$ 번째 규칙에서  $j$ 번째 입력 변수의  $j$ 번째 소속함수에 대한 값을 구하는 함수로 본 논문에서는 삼각형의 소속함수를 이용한다.

본 논문에서는 복잡한 비선형 시스템의 입출력 데이터만을 이용하여 식 (1)로 표현되는 퍼지 모델을 구하기 위하여 각 소속함수  $F_j^i$ 의 파라미터와 후반부  $O_i$ , 그리고 규칙의 수를 구하는 방법을 제안한다.

## 3. 바이러스 - 진화 유전 알고리즘을 이용한 퍼지 모델링

### 3.1 유전 알고리즘(GA)와 바이러스-진화 유전 알고리즘(VEGA)

유전 알고리즘은 자연 진화와 개체군의 유전 정보라는 원리에 기반한 최적화 과정이며 Holland[15]에 의해 제안되었다. 생명체의 역사는 개체군과 종 사이에서 작용하는 몇 개의 통계적인 진화 과정에 의해 설명될 수 있다. 최근, 진화라는 개념은 생명체 연구의 범위를 넘어서 컴퓨터나 다른 기기를 이용해 모의 실험할 수 있는 최적화 과정으로 이용되고 있다. 이 기술의 응용분야가 전력시스템, 패턴 인식, 제어 시스템 등의 기존 기술과 결합됨에 따라서 모의실험에 대한 관심이 최근 몇 년 사이에 큰 폭으로 증가해왔다. 진화 과정을 수행하기 위한 연산자에는 재생, 교배 그

리고 돌연변이를 사용하게 된다. 재생을 통해서 낮은 적합도를 가진 개체는 다음 세대에 몇 개의 복제를 갖거나 아무런 유전 형질도 전달하지 못하는데 반해 높은 적합도를 가진 개체는 많은 복제를 하게 된다. 교배 연산자는 두 단계에 의해 두 부모 개체로부터 유전 정보를 합성하여 새로운 자식 개체를 생성한다. 돌연변이는 개체군의 다양성을 유지하기 위해 임의의 개체, 임의의 염색체를 변화시키는 것이다.

그러나, 기존의 GA는 일반적으로 강인하다고 알려져 있지만, 어떤 특수한 영역하에서는 성공적인 최적화 알고리즘이 되지 못한다고 알려져 있다. 즉, 개체군내에서 진화할 때 월등히 우수한 개체의 출현은 비례 선택적인 재생 연산자의 사용에 따라 우수한 개체의 염색체를 후보 개체들에 확산하게 된다. 따라서, 유전적인 다양성을 상실하게 되고 그 개체의 진화 방향으로 전 개체군의 진화 방향이 결정되게 된다. 이럴 경우 조기 지역 수렴이 발생하게 되고 돌연변이 연산자에 의해 현재의 진화 방향에서 탈피할 수 있을 때까지 더 이상의 적합도의 향상을 기대할 수 없다. 그 이유는 기존의 GA에서는 스키마타를 직접적으로 다루지 않고 간접적으로만 다루기 때문이다. 이에 따라 바이러스-진화 유전알고리즘의 사용이 요구된다.

분자 생물학의 발전에 힘입어 공생 이론이나 바이러스 이론 같은 다양한 진화 이론이 진화 메커니즘을 설명하기 위해 제안되었다. 특히, 바이러스는 그 호스트의 적합도를 감소시킬 수도 있다는 이유 때문에 중요성이 낮게 평가되어 왔으나 최근의 논문에서 증명된 것처럼 바이러스는 생태학적 그리고 진화론적 모두에 중요한 역할을 수행하고 있다[16]. 진화론적 관

점에서 바이러스 이론은 바이러스 형질도입이 종간의 DNA의 본질을 이동시키는 중요 메커니즘이라는 사실에 기초하고 있다. VEGA는 위와 같이 진화적 바이러스 이론에 바탕을 둔 유전 정보의 수평적 전파와 수직적 상속의 진화를 모방한 알고리즘이다. VEGA에서는 후보 해의 개체군의 현재 상태에 따라 지역과 전역 사이의 탐색을 적응적으로 변화시킬 수 있다. 이를 위해 새로운 탐색 연산자로서 두 개의 감염 연산자 즉, 역전사 연산자(reverse transcription operator)와 형질도입 연산자(transduction operator)가 GA에 도입된다[17]. 역전사 연산자는 바이러스 개체의 우수성이 입증된 유전 형질을 주 개체군에 전달함으로써 효과적인 스키마타를 증가시킬 수 있도록 해주고, 형질도입 연산자는 세대간의 진화시에 바이러스 개체가 주 개체군으로부터 새로운 유전형질을 공급받을 수 있도록 함으로써 바이러스 개체의 유전적 다양성을 보장해 준다. 여기서, 효과적인 스키마타의 직접적인 수직적, 수평적 진화는 스키마타와 같은 지역정보가 최적화 문제를 푸는데 유용하며 호스트 개체군과 바이러스 개체군의 공동 진화는 최적화 문제의 해를 빨리 찾을 수 있도록 보장한다.

VEGA는 그림 1과 같이 두 개체군 즉, 호스트 개체군과 바이러스 개체군을 갖는다. 여기에서 호스트 개체군과 바이러스 개체군이란 각각 후보 해의 집합과 호스트 개체의 부문자열의 집합으로 정의할 수 있다.

GA는 스키마타를 간접적으로 다루는데 반해 VEGA는 스키마타를 직접 다룬다는 특징을 갖는다. 즉, GA의 개체군은 세대간 수직적으로만 진화하는 반면에, VEGA의 개체군은 역전사 연산자와 형질도입

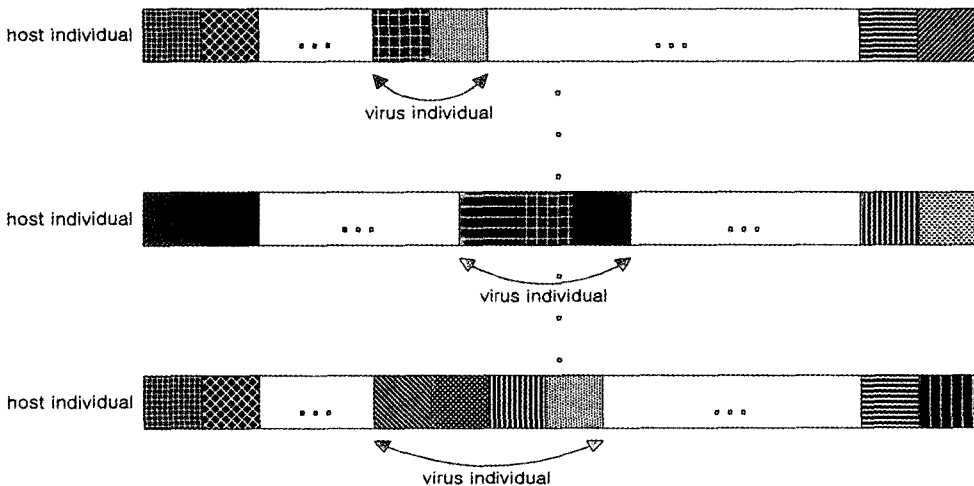


그림 1. 감염된 호스트 개체  
Fig. 1. Infected host individuals

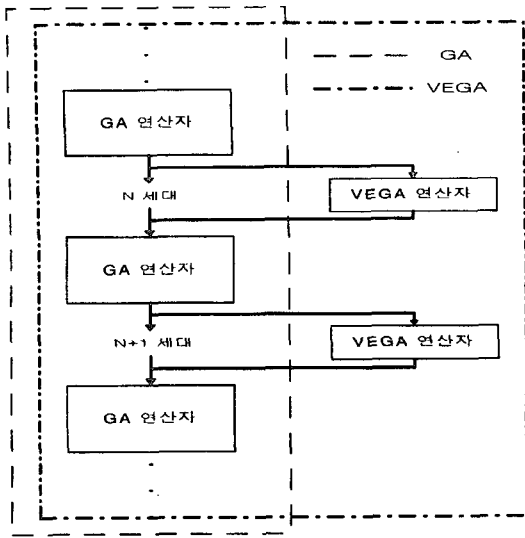


그림 2. GA 연산자와 VEGA 연산자  
Fig. 2. GA operator and VEGA operator

연산자에 의해 수직적으로 진화할 뿐만 아니라 한 세대내에서 수평적으로도 진화한다는 특징을 가지고 있다. 따라서, VEGA는 그림 2와 같이 GA의 연산자에 추가적인 연산자를 사용하게 된다.

일반적으로 지역 정보는 최적해를 찾는 데 매우 중요하다. 따라서, GA에서는 무시되었던 지역 정보를 VEGA에서는 바이러스 감염에 의해 스키마타를 효과적으로 사용함에 따라 더 좋은 해를 구할 수 있다. 위에서 언급한 역전사 연산자와 형질도입 연산자는 GA에서 사용되는 연산자 이외의 추가된 연산자이다. 그림 3은 역전사 연산자와 형질 도입 연산자의 연산 작용을 나타낸다.

**역전사 연산자:** 생물학에서 바이러스가 인체를 감염 시키듯 바이러스 개체가 호스트 개체에 바이러스 개

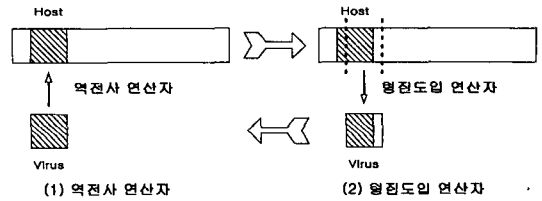


그림 3. 역전사 연산자와 형질도입연산자  
Fig. 3. Reverse transcription operator and transduction operator

체의 유전 형질을 전해주는 연산자로서 한 바이러스 개체가 모든 호스트 개체를 감염시키는 것이 아니라 정해진 비율의 임의로 선택된 호스트 개체를 감염시키게 된다.

**형질도입 연산자:** 바이러스 개체의 생명력 계수 가 음의 값, 또는 일정한 값보다 작을 경우 현재의 바이러스 개체는 사라지고 새로운 바이러스 개체가 생겨나며, 바이러스 개체의 적합도에 따라 바이러스 개체는 호스트 개체로부터 부분자열의 길이를 증가하거나 감소하여 형질 도입하게 된다.

그림 4는 바이러스 감염과 선택을 도시한 그림이다. 그림에서 같이 각 바이러스 개체는 감염률에 따라 호스트 개체를 감염시키게 되는데 바이러스에 감염된 호스트 개체들의 적합도가 향상된다면 바이러스가 감염된 상태의 호스트 개체들이 선택되고 그렇지 않다면 원래의 호스트 개체가 선택된다.

VEGA는 바이러스의 적합도에 따라 연산을 수행하므로 감염률, 바이러스 적합도, 생명력 등의 파라미터들을 갖게 되는데 다음의 식과 정의를 따른다.

$$fitvirus_j = fithost'_j - fithost_j \tag{3}$$

$$fitvirus_i = \sum_{j \in S} fitvirus_j \tag{4}$$

여기서,  $fithost_j$ 와  $fithost'_j$ 는 각각  $j$ 번째 호스트 개체에

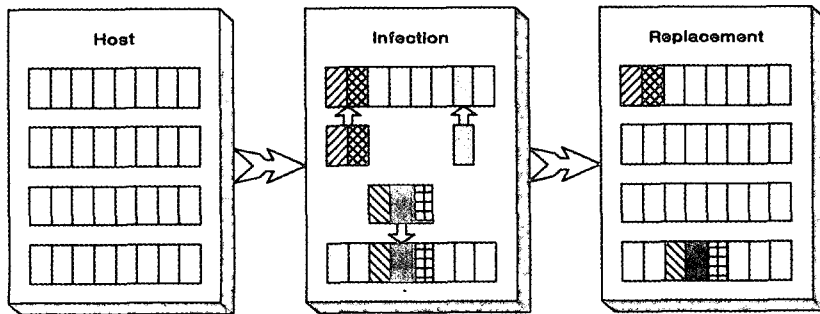


그림 4. 바이러스 감염과 선택  
Fig. 4. Virus infection and selection

감염하기 전과 후의 적합도,  $fitvirus_j$ 는  $i$ 번째 바이러스가  $j$ 번째 호스트를 감염시키기 전과 감염시킨 후의 호스트 개체의 적합도의 차이, 그리고  $fitvirus_i$ 는  $i$ 번째 바이러스의 적합도를 나타낸다. 또한,  $S_i$ 는 바이러스 개체의 감염률에 의해 감염된 호스트 개체의 후보 해 공간을 나타낸다. 식 (3)에 의하면 바이러스 적합도란 현재의 세대에서 바이러스 개체의 적합도를 나타내는 지표로서 감염시키기 전과 후의 적합도의 차이의 합으로 정의된다.

감염률은 식 (5)와 같이 정의된다.

$$\begin{aligned} \text{if } fitvirus_i \geq 0, \text{ then } infrate_i &= infrate_i + \beta \\ \text{if } fitvirus_i < 0, \text{ then } infrate_i &= infrate_i - \beta \end{aligned} \quad (5)$$

즉,  $infrate_i$ 로 표현되는 감염률은 바이러스 개체가 호스트 개체를 감염시키는 것을 정의하는 계수로서 식 (5)와 같이 초기에 정해진 값으로 시작해서 각 세대마다 바이러스 개체의 적합도가 양수면  $\beta$ 만큼 증가시키고 음수이면  $\beta$ 만큼 감소된다.

마지막으로, 생명력은 식 (6)과 같이 정의된다.

$$life_{i,t+1} = r \times life_{i,t} + \alpha \times fitvirus_i \quad (6)$$

여기서,  $life_{i,t+1}$ 는  $t+1$ 세대에서의 생명력이고,  $life_{i,t}$ 은  $t$ 세대에서의 생명력,  $r$ 은 감소율,  $\alpha$ 는 생명도 계수를 나타낸다. 만약 생명력의 값이 양의 값을 가지면 바이러스 개체는 다음 세대에 살아 남고, 음의 값을 갖으면 새로운 부분자열을 형질 도입한다. 그러나, 현재 세대의 바이러스 개체의 적합도가 낮더라도 바이러스 개체의 적합도가 전체의 호스트 개체를 감염시켜 얻은 값이 아니고 임의로 선택된 호스트 개체만을 감염시켜 얻은 값이면 현재의 바이러스 개체의 적합도가 좋지 않더라도 이전까지 좋은 생명도 값을 유지했을 경우 현재 세대의 생명도 계수가 다음 세대의 생명력

에 영향을 미칠 수 있어야만 한다.

역전사 연산자, 형질도입 연산자, 바이러스 개체의 적합도, 감염률, 생명력을 GA의 연산자에 추가하여 VEGA에서 사용하는 전체 연산자를 구성할 수 있다. 여기서, 초기 바이러스 개체의 길이, 감염률의 초기값과 변화값, 생명도 계수를 구하는데 필요한  $r$ 과  $\alpha$ 를 적용 대상에 따라 적절히 선택하여야만 빠른 최적화를 이룰 수 있다.

### 3.2 퍼지 모델의 유전 암호화

본 절에서는 퍼지 모델의 파라미터와 구조를 동시에 동정하기 위해서 퍼지 규칙의 전반부와 후반부로 나누어 그림 7과 같이 부호화하는 방법을 제안한다. 즉, 전반부의 소속함수의 파라미터를 폭과 중심값으로 나타낸 쌍이 최대 규칙의 수만큼 되도록 배열하고, 그 뒤로 후반부 값을 배열하였으며, 구조를 동정하기 위해서 0 또는 1로 규칙의 필요 유무를 탐색하도록 구성된다.

개체군의 초기화를 한 후에는 각 개체를 평가하기 위한 적합도 함수를 정하는 것이 필요하다. 호스트 개체와 바이러스 개체는 각각의 적합도를 가지고 VEGA의 연산자들에 의해 진화하게 된다. 퍼지 모델의 평가는 실제 출력과 모델 출력의 평균 제곱 오차로 구하며 식 (7)과 같이 정의한다.

$$E = \frac{1}{N} \sum_{i=1}^N (y_d - y)^2 \quad (7)$$

여기서,  $N$ 은 데이터 수,  $y_d$ 는 교사 데이터의 출력이고,  $y$ 는 퍼지 모델로부터 추론된 값이다. 퍼지 모델의 선정과정에서의 필요조건은 전체 입력 집합내에서 퍼지 모델의 유효성 여부이다. 식 (2)에서  $i$ 번째 규칙의 적합도  $\mu_i$ 를 구할 때에 각 입력 값에 대한 소속함수의 값 중에 한 개라도 소속함수의 범위에서 벗어나게

	Premise string				Consequent string				Rule number			
Individual 1	width_11	center_11	...	width_1c	center_1c	position_11	...	position_1c	rule_11	...	rule_1c	
Individual 2	width_21	center_21	...	width_2c	center_2c	position_21	...	position_2c	rule_21	...	rule_2c	
...	...	...	...	...	...	...	...	...	...	...	...	
Individual n	width_n1	center_n1	...	width_nc	center_nc	position_n1	...	position_nc	rule_n1	...	rule_nc	

그림 5. 퍼지 모델의 유전 암호화  
Fig. 5. Genetic coding of fuzzy model

되면 0이 된다. 따라서, 모든 규칙에서 각 입력값에 대해 0을 갖는 소속함수가 적어도 한 개 이상 존재한다면  $\mu_i$ 의 분모가 0이 되고 이는  $\mu_i$ 의 값이 무한대가 됨을 의미한다. 따라서, 본 논문에서는 위와 같은 개체가 다음 세대에서 살아남지 않게 하기 위해서는 해당 적합도에 작은 수를 곱해 줌으로써 후보 해 선정 과정에서 배제한다.

일반적으로 퍼지 모델링의 최적화는 식 (7)의 평가 함수값  $E$ 를 최소화시키는 최적의 퍼지 규칙과 파라미터값을 찾아내는 것이다. 그러나, GA에서는 적합도 함수를 최대화시키도록 진화가 이루어져야 하므로 최소화 문제를 최대화 문제로 변경하여야 한다. 본 논문에서 사용한 적합도 함수는 모델 오차의 최소화와 필요한 규칙 수의 최소화를 동시에 수행하기 위하여 식 (8)과 같은 적합도 함수를 이용하며 유용한 퍼지 모델의 선정을 위해 공집합 형성시 벌점 부과가 있도록 한다[18].

$$\begin{cases} \sum_{i=1}^c \omega_i = 0, f = \left( \frac{C_e}{E + \varepsilon} + \frac{C_r}{R + \rho} \right) \times \Pi \\ \sum_{i=1}^c \omega_i \neq 0, f = \frac{C_e}{E + \varepsilon} + \frac{C_r}{R + \rho} \end{cases} \quad (8)$$

여기서,  $C_e$ ,  $C_r$ 과  $\rho$ ,  $\varepsilon$ 은 경쟁 수준 조정을 위한 상수,  $E$ 는 식 (7)에서 정의한 평가함수,  $\Pi$ 는 퍼지 모델의 입력 분할에 대하여 공집합을 가지는 것을 방지하기 위해 적합도 함수에 곱해지는 상수, 그리고  $\varepsilon$ 은  $E=0$ 이 될 경우 적합도 함수의 발산을 방지하기 위한 값이며  $R$ 은 해당 염색체에서 추출된 퍼지 모델의 규칙 수이다. 식 (8)에서,  $C_e$ 는 퍼지 모델의 정확성에 관계되고,  $R$ 과  $C_r$ 은 퍼지 모델의 규칙 수에 관계된다. 그러므로, 설계자는 이러한 상수들의 값을 조정함으로써 원하는 방향의 퍼지 모델링을 행할 수 있다. 즉,  $C_e$ ,  $C_r$ 의 값에 따라 퍼지 모델의 정확성을 중요시한다면 그 비를 크게 두고, 규칙 수가 적은 모델을 원한다면 그 비를 작게 둔다.

GA에서 높은 적합도를 갖는 개체는 재생 단계에서 보다 높은 확률로 선택된다. 그러나, 이 경우 초기에 상대적으로 적합도가 너무 높은 우수 개체의 출현은 조기 수렴의 결과를 가져올 수 있다. 본 논문에서는 조기 수렴 방지를 위해 원래의 적합도 값을 크기대로 정렬하고 각각에 새로운 적합도 값을 식 (9)와 같이 배당하여 이를 재생 단계에서 이용한다[18].

$$f(1) = \text{Max}, \dots, f(k) = \alpha^{k-1} \times \text{Max}, \dots, f(N) = \alpha^{N-1} \times \text{Max} \quad (9)$$

$\alpha(0 < \alpha < 1)$ 는 적합도 조정 계수이다.

경험에 의하면  $\alpha$ 값은 0.7~0.9로 정할 수 있으며, 식 (9)과 같이 행하여도 적합도 값은 바뀌지만 그 차례에는 변함이 없으면서 동일한 세대내에 적합도가 월등히 좋은 개체가 출현하더라도 비례 재생시에 개체군의 다양성을 보장해 줄 수 있게 된다.

### 3.3 퍼지 모델링 알고리즘

본 논문에서 제안한 VEGA에 의한 퍼지 모델링은 임의로 초기화된 유전 개체들을 기본적인 유전 연산자들인 재생, 교차와 돌연변이 연산자에 의해 자식 개체들을 생성하며 진화할 뿐 만 아니라 역전사 연산자, 형질도입 연산자에 의해 한 세대 내에서 진화하게 된다. 유전 알고리즘과 같이 기본적인 유전 연산자들을 통한 세대 교체시 최고의 적합도를 갖는 개체는 다음 세대에서 보존하는 이론을 적용하여 적합도 함수의 수렴을 보장하게 된다. 본 논문에서 제안한 퍼지 모델링 알고리즘은 다음과 같다.

단계 1 : 교차율, 돌연변이율, 감염률과 생명도 계수를 결정하고 최대 세대수, 개체군 수와 바이러스 개체의 문자열의 길이, 감염률, 생명력을 초기화한다. 또한, 퍼지 모델의 최대 규칙수와 후보 소속 함수의 수를 정한다.

단계 2 : 동정될 입출력 데이터들을 정규화한다.

단계 3 : 각 개체의 염색체에서 퍼지 모델을 추출하여 식 (8)에 의한 평가를 거쳐 식 (9)에 의해 적합도 함수 값을 각 개체에 부여한다.

단계 4 : 역전사 연산자에 의해 바이러스 개체로부터 호스트 개체로 감염이 일어나고 이로부터 다음 세대에 이용할 감염률과 생명력을 구한다.

단계 5 : 적합도 값에 비례하여 선택이 이루어지고 교차와 돌연변이를 거쳐 새로운 세대를 구성한다. 돌연변이의 경우 유전자 값이 정상적인 범위를 벗어나는 경우는 이를 취소한다. 이 과정에서 최고의 적합도 값을 갖는 개체는 보존된다.

단계 6 : 만족할 만한 결과가 나오거나 최대 세대수가 지날 때까지 단계 3-5를 반복한다.

## 4. 시뮬레이션 및 결과 고찰

### 4.1 비선형 시스템

식 (10)은 2개의 입력과 1개의 출력으로 이루어진 Sugeno가 사용한 비선형 시스템이다. 식 (10)의 비선형 시스템의 입출력 데이터만을 이용하여 퍼지 모델을 구하기 위하여 Sugeno가 사용한 50개의 입출력

데이터가 이용되었다[6].

$$y = (1 + x_1^2 + x_2^{1.5})^2, 1 \leq x_1, x_2 \leq 5 \quad (10)$$

본 모의실험을 위해 사용한 파라미터값을 표 1에 나타내었다.

그림 6은 GA와 VEGA의 적합도의 변화를 도시한 그림이다. 그림 6을 살펴보면, GA는 일정한 세대 이후에는 더 이상의 적합도 값의 증가가 이루어지지 않지만, VEGA는 스키마타의 증가로 적합도 값이 계속 해서 증가하는 것을 나타낸다. 따라서, VEGA는 유전적 다양성을 보장하고 효과적인 스키마타를 보존하므로써 조기 지역수렴을 방지하는 것을 알 수 있다. 그림 7과 8은 각각 GA와 VEGA의 최소 제곱 오차(MSE)의 변화율과 규칙수의 변화를 도시한 그림이다. 약 550세대 이후의 GA는 MSE값의 변화가 거의 안 일어나지만 VEGA는 계속 감소하고 있음을 알 수 있다. 이는 바이러스 감염에 의해 개체군의 유전 정보의 다양성이 보장될 수 있다는 것을 의미한다. 그림 8에서는 VEGA의 규칙수가 GA보다 크지만 식 (9)에서  $C_e$ 와  $C_r$  값의 조정에 따라 규칙수를 쉽게 줄일 수 있다. 그림 9는 본 논문에서 제안한 퍼지 모델링 과정

표 1. 비선형 시스템 모델링을 위한 파라미터들  
Table 1. Parameters for modeling of nonlinear system

Parameter's name	Parameter's value
Maximum generation number	1,000
Population number	100
Crossover rate	0.7
Mutation rate	0.2
max_c	6
$\alpha$	0.8
$\epsilon, C_e, C_r, \rho, \Pi$	0, 1, 0.3, 3, 0.0001
Reduction rate, Life coefficient	0.1, 0.9

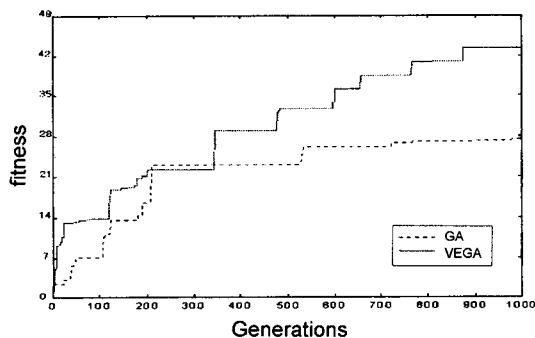


그림 6. GA와 VEGA의 적합도의 변화  
Fig. 6. Changes of fitness of GA and VEGA

을 거쳐 최종적으로 동정된 6개 퍼지 제어 규칙을 도시한 그림이다.

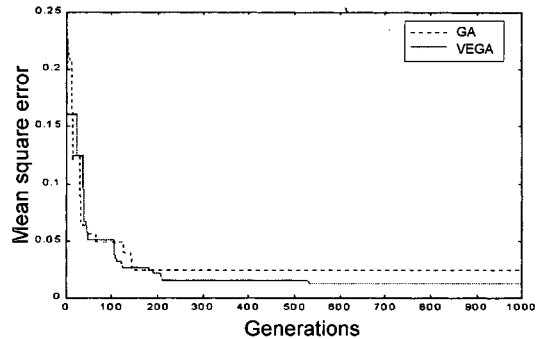


그림 7. GA와 VEGA의 MSE의 변화  
Fig. 7. Changes of MSE in GA and VEGA

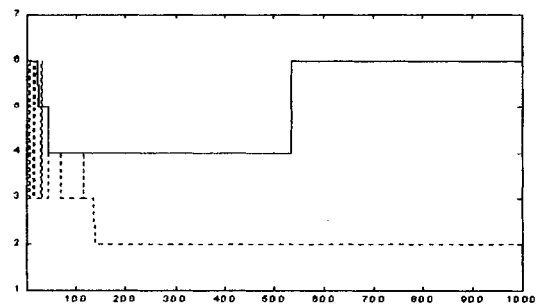


그림 8. GA와 VEGA의 규칙수의 변화  
Fig. 8. Changes of the number of rules in GA and VEGA

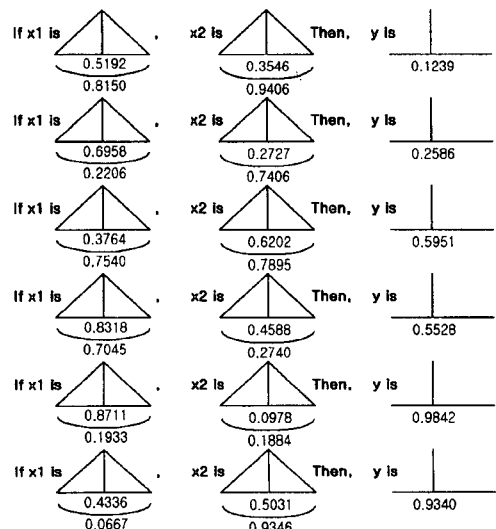


그림 9. 동정된 퍼지 모델  
Fig. 9. Identified fuzzy model

표 2. 다른 모델과의 비교

Table 2. Comparisons of our model with other model

Model name	Number of rules	MSE
Sugeno's [1]	6	0.079
Chang's [18]	4	0.032
GA	2	0.039
VEGA	6	0.021

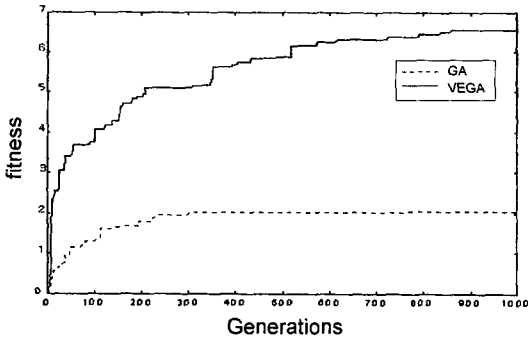


그림 10. GA와 VEGA의 적합도의 변화  
Fig. 10. Changes of fitness in GA and VEGA

표 2는 본 논문에서 제안한 모델링 기법과 다른 모델링 기법들과의 MSE의 비교를 보여준다. 표 2에서 알 수 있듯이 본 논문에서 제안한 방법이 다른 방법에 비해 정확도면에서 더욱 우수함을 알 수 있다.

#### 4.2 가스로의 퍼지 모델링

본 절에서는 Box와 Jenkins [19]가 제시한 가스로의 시계열 입출력 데이터를 사용하여 가스로의 연소된 이산화탄소의 농도를 퍼지 모델링한다. 이 데이터들은 입력으로 가스 유입률( $u(t)$ )과 출력으로  $CO_2$  밀도( $y(t)$ )로 나타내는 296개의 데이터 쌍으로 구성되어있다. 본 논문에서는 시뮬레이션을 위해 현재의 출력  $y(t)$ 에 영향을 미치는 입력들로서는  $u(t-4)$ 와  $y(t-1)$ 의 두 개를 택하였다.

시뮬레이션을 위한 초기 파라미터 값들은 표 1과 같이 하였다. 그림 10-12은 GA와 VEGA로 모의실험한 후의 적합도, 최소 제곱 오차, 규칙 수등을 도시한 것이다. 그림 10과 11에서도 4.1절과 같이 GA는 일정 세대이후 변화율이 0에 가까워지는 반면, VEGA는 변화율이 작아지기는 했으나 계속해서 적합도 값은 증가, MSE 값은 감소하는 것을 나타낸다.

표 3은 다른 기법에 의한 퍼지 모델링의 MSE 값들을 나타낸 표이다. 표에서 알 수 있듯이 본 논문에서 제안한 방법에 다른 방법에 비해 정확도면에서 우

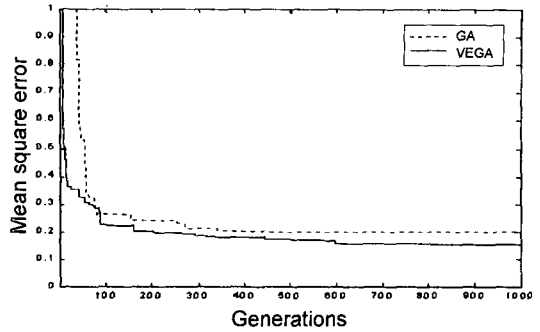


그림 11. GA와 VEGA의 MSE의 변화  
Fig. 11. Changes of MSE in GA and VEGA

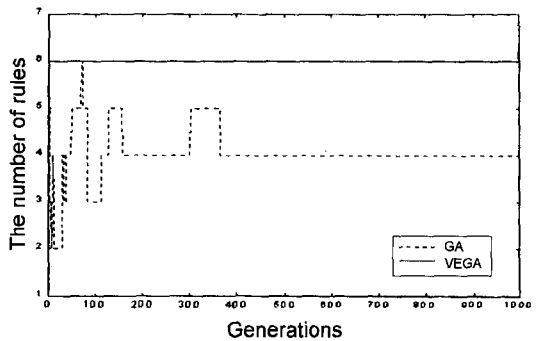


그림 12. GA와 VEGA의 규칙수의 변화  
Fig. 12. Changes of rule number in GA and VEGA

표 3. 다른 모델들과의 비교

Table 3. Comparisons of our model with other models

Model name	Inputs	Number of rules	MSE
Tong's [1]	$y_{k-1}, u_{k-4}$	19	0.469
Pedrycz's [2]	$y_{k-1}, u_{k-4}$	81	0.320
Xu's [3]	$y_{k-1}, u_{k-4}$	25	0.328
Sugeno's [6]	$y_{k-1}, u_{k-3}, u_{k-4}$	6	0.190
GA	$y_{k-1}, u_{k-4}$	4	0.196
VEGA	$y_{k-1}, u_{k-4}$	6	0.154

수함을 나타낸다. 그림 13은 최종 동정된 6개 규칙의 소속함수와 후반부의 파라미터 값을 도시한 그림이다.

상기 시뮬레이션의 결과, GA를 이용했을 때에는 높은 적합도를 가진 개체가 출현했을 경우 지역해로 수렴해 가는 반면에 VEGA는 바이러스 개체들의 도움으로 꾸준히 증가하고 있음을 검증할 수 있었다. 비록 규칙수의 동정에서는 GA가 4개로 6개인 VEGA보다 우수하였으나, 이는 3.2절에서도 언급했듯이 식 (9)의 정확성과 효율성의 계수를 적절히 바꿔주면 쉽게 개



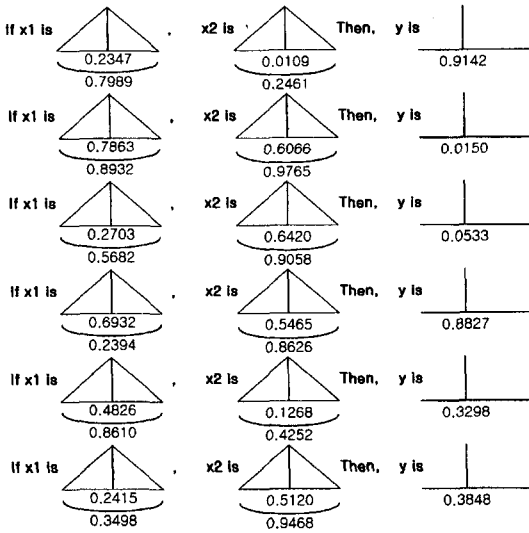


그림 13. 동정된 퍼지모델  
Fig. 13. Identified fuzzy model

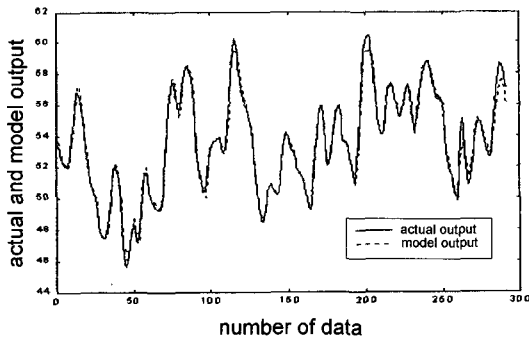


그림 14. VEGA를 이용한 가스의 실제 출력과 모델출력  
Fig. 14. Actual output and model output using VEGA

선될 수 있다. 그림 14는 최종 동정된 퍼지 모델을 이용해서 가스로 시계열 데이터의 실제 출력과 비교한 그림이다. 그림에서 알 수 있듯이 본 논문에서 제안한 VEGA를 이용한 모델의 출력을 구한 결과 실제 출력을 근사하게 모델링 했음을 알 수 있다.

## 5. 결 론

본 논문에서는 VEGA를 이용하여 퍼지 모델링시 소속함수의 파라미터와 규칙의 구조를 동시에 자동 동조할 수 알고리즘과 새로운 염색체를 제안하였다. 일반적으로, VEGA는 호스트와 바이러스 개체 사이의 수평적 진화와 세대간 유전적 정보의 수직적 진화 모듈을 기반으로 하여 공동 진화를 모의한다는 특징이

있다. 제안한 알고리즘은 먼저 퍼지 모델의 파라미터와 규칙의 구조를 임의의 수로 암호화하여 호스트 개체를 구성하고, 미리 정해 놓은 수의 감염될 개체와 각 개체의 부분자열을 결정하여 바이러스 개체를 생성한다. 다음, 각 호스트 개체의 적합도와 각 바이러스 개체의 적합도를 구하여 감염된 호스트 개체의 적합도의 차의 합에 의해 감염률과 생명력을 구한다. 마지막으로, 동일한 세대내에서는 바이러스 개체가 감염률에 의해 역전사 연산자가 작용하게 되며, 다음 세대로 진화할 때에는 호스트 개체는 교배, 재생, 돌연변이에 의해서 다음 세대의 개체를 구하고 바이러스 개체는 생명력에 의해서 존속과 비존속을 결정하게 된다. 본 논문에서 제안한 VEGA를 이용한 퍼지 모델링 방법의 효율성과 우수성을 검증하기 위해 비선형 시스템과 가스의 퍼지 모델링에 응용하여 우수한 결과를 얻었다.

## 감사의 글

본 논문은 2000년 두뇌한국사업21에 의해서 지원되었음.

## 참고문헌

- [1] R. M. Tong, "The Evaluation of Fuzzy Models Derived from Experimental Data", *Fuzzy Sets and Systems*, Vol. 4, pp. 1-12, 1980.
- [2] W. Pedrycz, "An Identification Algorithm in Fuzzy Relational Systems", *Fuzzy Sets and Systems*, Vol. 13, pp. 153-167, 1984.
- [3] C. W. Xu, "Fuzzy Model Identification and Self-learning for Dynamic Systems", *IEEE Trans. Syst. Man. Cybern.*, Vol. 17, No. 4, pp. 683-689, 1987.
- [4] H. Takagi and M. Sugeno, "Fuzzy Identification of Systems and Its Application to Modeling and Control", *IEEE Trans. on Sys. Man and Cybern.*, Vol. 15, pp. 116-132, 1985.
- [5] M. Sugeno and G. T. Kang, "Structure Identification of Fuzzy Model", *Fuzzy Sets and Systems*, Vol. 28, pp. 15-33, 1988.
- [6] M. Sugeno and T. Yasukawa, "A Fuzzy-Logic-Based Approach to Qualitative Modeling", *IEEE Trans. on Fuzzy system*, Vol. 1, No. 1, pp. 7-31, 1993.
- [7] L. A. Zadeh, "Fuzzy Sets", *Information Control*, Vol. 8, pp. 338-353, 1965.
- [8] Y. H. Joo, H. S. Hwang, K. B. Kim and K. B. Woo, "Linguistic Model Identification for Fuzzy System", *Electronics Letters*, Vol. 31, No. 4, pp. 330-331, 1995.
- [9] Y. H. Joo, H. S. Hwang, K. B. Kim and K. B. Woo, "Fuzzy System Modeling by Fuzzy Partition and GA Hybrid Schemes", *Fuzzy Sets and Systems*, Vol. 86, No. 3, pp. 279-288, 1997.

- [10] Y. Davidor, "A Natural Occurring Niche & Species Phenomenon: The Model and First Results", *Proc. of The Fourth Int. Conf. on Genetic Algorithms*, 257-263, 1991.
- [11] B. Manderick, and P. Spoessens, "Fine-Grained Parallel Genetic Algorithms", *Proc. of The Fourth Int. Conf. on Genetic Algorithm*, pp. 428-433, 1991.
- [12] S. Mahfoud, and D. E. Goldberg, "A Genetic Algorithm for Parallel Simulated Annealing, Parallel Problem Solving from Nature 2", Elsevier Science Publishers B. V., 1992.
- [13] J. Renders, and H. Bersini, "Hybridizing Genetic Algorithms with Hill-Climbing Methods for Global Optimization: Two Possible Ways", *Proc. of The First IEEE Conf. on Evolutionary Computing*, Vol. 1, pp. 312-317, 1994.
- [14] N. Kubota, T. Fukuda and K. Shimojima, "Virus-Evolutionary Genetic Algorithm for a Self-Organizing Manufacturing System", *Computers & Industrial Engineering*, Vol. 30, Issue 4, pp. 1015-1026, Sep. 1996.
- [15] J. H. Holland, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, MA, 1989.
- [16] M. Begon, J. L. Harper and C. R. Townsend, "Ecology: Individuals, Populations and Communities", Blackwell Scientific Publications, Oxford, 1996.
- [17] N. Kubota, K. Shimojima, T. Fukuda, "The Role of Virus Infection in Virus- Evolutionary Genetic Algorithm", *Proc. of IEEE Int. Conf. on*, pp. 182 -187, 1996.
- [18] 장 옥, 손유석, 주영훈, 박진배, "비선형 시스템의 퍼지 모델링에 관한 연구", 대한 전기학회 논문집, 47권, 6호, pp. 826-833, 1998, 6월.
- [19] G. E. P. Box and G. M. Jenkins, "Time Series Analysis, Forecasting and Control", Holden Day, San Francisco, 1970.

**이 승 준 (Seung-Jun Lee)**



1999년 : 연세대 공대 전기공학과 졸업  
 1999년~현재 : 연세대 대학원 석사과정 재학중  
 관심분야 : 퍼지제어, 지능제어, 유전알고리즘, 신경회로망

**주 영 훈 (Young-Hoon Joo)**



1978년 : 연세대 전기공학과 졸업  
 1984년 : 연세대 대학원 전기공학과 졸업  
 1995년 : 동 대학원 전기공학과 졸업 (공학박)  
 1986년~1995년 8월 : 삼성전자(주) 자동화연구소(선임 연구원)  
 1998년~1999년 : 미국 휴스턴대학 전기 및 컴퓨터공학과 교환교수

1995년 9월~현재 : 군산대 공대 전자정보공학부 조교수  
 2000년~현재 : 한국퍼지 및 지능시스템학회 이사 및 편집위원  
 관심분야 : 퍼지제어, 지능제어, 유전알고리즘, 신경회로망, 지능형 로봇

**박 진 배 (Jin-bae Park)**



1977년 : 연세대 전기공학과 졸업  
 1990년 : Kansas State University 전기 및 컴퓨터 공학과 졸업(공학박)  
 1990년~1991년 : Kansas State University 전기 및 컴퓨터 공학과 조교수  
 1998년~1999년 : 연세대 창업보육센터 (센터장)

1998년~1999년 : 연세대 연세공학원(부원장)  
 1994. 9~현재 : 연세대 공과대학 전기 및 컴퓨터공학과(부교수)  
 현재 : 연세대 공대 전기 · 컴퓨터공학과 부교수  
 관심분야 : 자동제어, 강인제어, 지능제어, 예측제어, 필터링 이론 및 시스템 동정