

바이러스-메시 유전 알고리즘에 의한 퍼지 모델링

The Fuzzy Modeling by Virus-messy Genetic Algorithm

최종일 · 이연우 · 주영훈 · 박진배*

군산대 공대 전자정보공학부

* 연세대 공대 전기 및 컴퓨터공학과

Jong-Il Choi, Yeun-Woo Lee, Young-Hoon Joo, Jin Bae Park

(Tel:82-63-469-4706; Fax:82-63-466-2086; E-mail:yhjoo@ks.kunsan.ac.kr)

Abstract - This paper deals with the fuzzy modeling for the complex and uncertain systems, in which conventional and mathematical models may fail to give satisfactory results. mGA(messy Genetic Algorithm) has more effective and adaptive structure than sGA with respect to using changeable-length string and VEGA(Virus Evolution Genetic Algorithm) can search the global and local optimal solution simultaneously with reverse transcription operator and transduction operator. Therefore in this paper, the optimal fuzzy model is obtained using Virus-messy Genetic Algorithm(Virus-mGA). In this method local information is exchanged in population so that population may sustain genetic divergence. To prove the superiority of the proposed approach, we provide the numerical example.

1. 서론

복잡하고 불확실한 비선형 시스템을 수학적으로 모델링하는 경우에는 비선형성의 고차 모델을 이용하여거나 모델 근사화 방법을 이용한다. 하지만 만족할 만한 모델을 얻기는 어려우므로 퍼지 모델링 기법이 활발히 연구되어 왔다[1-3]. 퍼지 모델링 방법의 주관심사는 전문가적 경험을 이용하여 미지의 시스템을 퍼지 규칙으로 정확히 묘사하는데 있다. 이에 많은 유전 알고리즘과 클러스터링 기법 [4-5]을 이용해 퍼지 모델의 파라미터와 구조를 동정하는 자동 동정 기법이 연구되어 왔으며, D.Goldberg[6]와 Deb, Kargupta[7]등을 중심으로 메시-유전 알고리즘이 연구 되어왔다. sGA의 비례적인 선택은 효과적인 스키마타를 증가 시키기도 하지만 스키마타를 직접적으로 다루지 않으며, 초기 지역 수렴과 비효과적인 스키마타의 증가의 문제점이 있다. 이에 바이러스-진화 유전 알고리즘[8-9]은 이러한 지역 수렴에 대한 방안이 될 수 있다. 따라서 본 논문에서는 호스트의 바이러스 개체 사이의 수평적 진화와 세대간 유전적 정보의 수직적 진화 모두를 기반으로 하여 공동 진화를 의미하며, 가변길이 스트링을 사용함으로써 동정의 유연성을 증가 시킬 수 있는 새로운 암호화 기법인 바이러스-메시 유전 알고리즘(Virus-messy genetic algorithm)을 사용한 퍼지 모델링 기법을 제안한다. 또한, 바이러스 메시-유전 알고리즘은 퍼지모델의 파라미터와 구조 동정에 적용되고 효율성과 가능성을 입증하기 위해 비선형 시스템의 모델링에 적용한다.

2. 본론

2.1 퍼지모델

본 논문에서는 동정될 모델로 TS 퍼지 모델로 사용하였다. TS 퍼지 모델의 i 번째 규칙은 식 1과 같이 표현된다.

$$\begin{aligned} \text{Plant Rule } i : \\ \text{IF } x_1(t) \text{ is } \Lambda_1^i \text{ and } x_n(t) \text{ is } \Lambda_n^i, \text{ THEN } y \text{ is } \omega_i \\ (i=1,2,\dots,r, j=1,2,\dots,n) \end{aligned} \quad (1)$$

식 (1)은 Rule i 의 i 번째 규칙을 나타낸다. 여기에서 Λ_j^i 는 퍼지 집합, $x_j(t)$ 와 y 는 각각 입력변수, 출력변수 그리고 ω_i 는 퍼지 싱글톤으로 실수이다. 또한 r 은 규칙들의 총 개수를 나타낸다. 본 논문에서는 삼각형 소속함수가 사용되었으며, 전체부의 소속함수 Λ_j^i 는 식(2)처럼 표현할 수 있다.

$$\Lambda_j^i(x_j; a_{ij}, b_{ij}, c_{ij}) = \begin{cases} \frac{x_j - a_{ij}}{b_{ij} - a_{ij}} & a_{ij} \leq x_j \leq b_{ij} \\ \frac{b_{ij} - x_j}{b_{ij} - a_{ij}} & b_{ij} \leq x_j \leq c_{ij} \\ 0 & x_j > c_{ij} \text{ or } x_j < a_{ij} \end{cases} \quad (2)$$

a_{ij}, b_{ij}, c_{ij} 는 각각 왼쪽 꼭지점, 중심, 오른쪽 꼭지점을 의미한다. 본 논문에서는 출력값을 구하는

비퍼지화 방법으로 무게 중심법을 사용하였으며 식 (4)에 나타내었다. 식 3은 입력에 대한 i 번째 규칙의 적합도이며 $\Lambda_i^j(x_j(t))$ 는 Λ_i^j 에 대한 $x_j(t)$

의 소속 정도이다.

$$\mu_i(x(t)) = \prod_{j=1}^n \Lambda_i^j(x_j(t)) \quad (3)$$

$$y^* = \frac{\sum_{i=1}^r \mu_i(x(t)) \omega_i}{\sum_{i=1}^r \mu_i(x(t))} \quad (4)$$

2.2 Virus-mGA의 구조

sGA는 풀고자 하는 해의 집합인 염색체에 있어서 세대가 바뀌어도 고정된 길이를 갖는 단점을 갖는다. 이에 Virus-mGA는 가변스트링을 사용함으로써 더 순응적인 스트링 구조를 갖을 수 있다. 또한 그림 1과 같이 진화시에 원시상과 병치상으로 불리우는 두 개의 상으로 이루어진다.

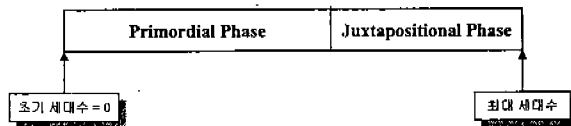


그림 1. 개체군 변형 표

원시상(Primordial phase)에서는 선택 연산자를 사용하여 호스트 개체군의 많은 후보 스트링 중에서 진화시킬 개체를 선별하기만 하고 병치상(Juxtapositional phase)에서 보통의 유전 알고리즘에서 사용하는 돌연변이 연산자와 교차 연산자 대신 다루기 쉬운 자르기와 붙이기 연산자를 사용하여 개체들을 진화 시키며 부모 개체에서 무작위로 선택된 위치에서 스트링을 자르기 때문에 후손 스트링은 이전의 스트링의 길이와 다르게 변경되어 동정의 유연성을 증가 시킨다. 그림 2는 자르기와 붙이기 연산과정을 나타낸 것이며 자르기와 붙이기는 각각 자르기 확률 p_c 와 붙이기 확률 p_s 에 따라 연산이 수행된다. 또한 병치상에서는 바이러스 감염이 일어난다.

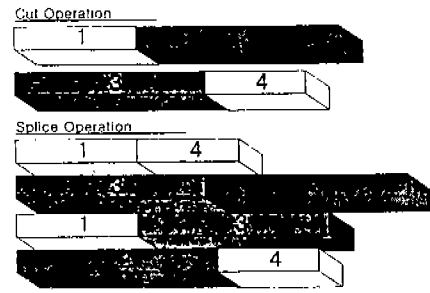


그림 2 자르기와 붙이기 연산 sGA에서 비례적인 선택은 조기 지역수렴의 원인이 될 수도 있을 뿐만 아니라 효과적인 스키마타를 증진 시키면서 동시에 비효과적인 스키마타도 함께 증가 시킨다는 문제점이 있다. 이를 막기 위해 바이러스 연산자를 사용하여 호스트와 바이러스 개체 사이의 수평적 진화와 세대간 유전적 정보의 수직적 진화 모두를 기반으로 하여 공동 진화를 모색한다. 따라서, 바이러스 감염에 의해 후보해의 현재 상태에 따라 지역과 전역 사이의 탐색을 적응적으로 변화시켜 지역과 전역의 탐색공간을 동시에 고려한다. Virus-mGA는 그림 3과 같이 두 개의 개체군 즉, 호스트 개체군과 바이러스 개체군을 갖으며, 각각은 후보 해의 집합과 호스트 개체의 부분자열의 집합으로 정의 할 수 있다.

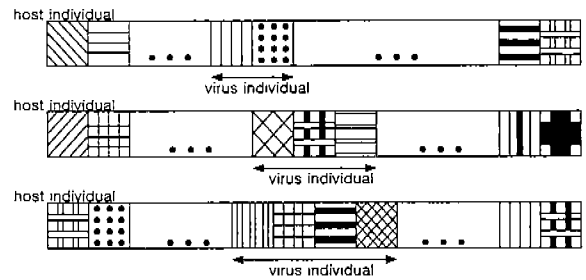
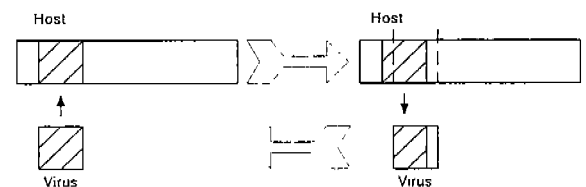


그림 3. 감염된 호스트 개체

그림 4는 역전사 연산자와 형질도입 연산자이다. 이것은 자르기와 붙이기 연산자에 추가된 연산자이다



(1) 역전사 연산자

(2) 형질도입 연산자

그림 4. 역전사 연산자와 형질도입 연산자.

역전사 연산자는 바이러스 개체가 호스트 개체에 바이러스 개체의 유전 형질을 전해주는 연산자로서 정해진 비율의 임의로 선택된 호스트 개체를

감염시키게 된다. 또한 형질도입 연산자는 호스트 개체로부터 부분자열의 길이를 증가 하거나 감소 하여 형질을 도입하게 되고 바이러스의 적합도에 따라 연산을 수행 하게 된다. 그러므로 감염을, 바이러스 적합도, 생명력의 파라미터들을 갖게 되며, 식 (5)는

$$\begin{aligned} fitvirus_{ij} &= fithost_j - fithost_i \\ fitvirus_{ij} &= \sum_{j \in S} fitvirus_{ij} \end{aligned} \quad (5)$$

바이러스의 적합도를 나타내는 식이다. 여기서 $fithost_j$ 와 $fithost_i$ 는 각각 j 번째 호스트 개체에 감염하기 전과 후의 적합도, $fitvirus_{ij}$ 는 i 번째 바

이러스가 j 번째 호스트를 감염 시키기 전과 감염시킨 후의 호스티 개체의 적합도의 차이, 그리고 $fitvirus_{ij}$ 는 i 번째 바이러스의 적합도이다. 감염을

은 바이러스 개체가 호스트 개체를 감염시키는 것을 정의 하는 계수로서 초기에 정해진 값으로 시작 해서 각 세대마다 바이러스 개체의 적합도가 양수면 증가 시키고 음수이면 감소 시킨다. 생명도 계수는 식 (7)과 같이 나타낼 수 있다.

$$life_{i,t+1} = r \times life_{i,t} + a \times fitvirus_i \quad (7)$$

여기에서, $life_{i,t+1}$ 는 $t+1$ 세대에서의 생존율이고, $life_{i,t}$ 는 t 세대 에서의 생존율, r 은 감소율, a 는 생명도 계수를 나타낸다. 만약 생명도가 양의 값을 가지면 바이러스 개체는 다음 세대에 남아 있을 것이고 그렇지 않다면 새로운 부분자열을 형질 도입한다.

3. 모의 실험 및 결과

본 논문 에서는 제안된 퍼지 모델링 기법의 우수성을 평가 하기 위해 식 (8)과 같은 비선형 시스템을 이용한다.

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2, \quad 1 \leq x_1, x_2 \leq 5 \quad (8)$$

이 시스템은 Sugeno[10]가 사용한 50개의 입출력 데이터 로서 0과 5 사이에서 무작위로 추출한 데이터이며, 입력이 두 개, 출력이 하나인 비선형 시스템이다. virus-mGA를 사용한 퍼지 모델의 동정을 위한 초기 파라미터로써 초기 개체군의 크기 100, 원시상과 병치상의 세대수 200, 자르기 확률 0.2, 불이기 확률 1.0, 돌연변이율 0.2, 는 0.9, 감소율 0.1, 생명을 0.9, 미세동조시 반복되는 세대수는 1000이다. 그림 5는 본 논문에서 제안한 퍼지 모델링 알고리즘 적용시 평균 제곱 오차의 변화를 나타낸 그림이다.

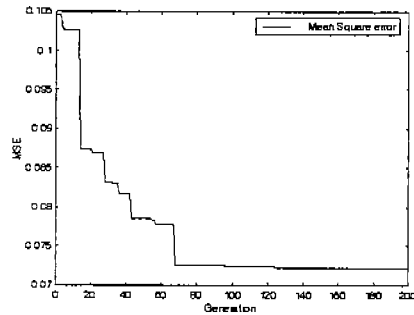


그림 5 동정 단계에서의 MSE 변화

표 1은 제안된 퍼지 모델링 방법과 기존의 퍼지 모델링 방법의 성능을 비교한 표이다. 규칙수는 4개이며, 그 때의 평균 제곱 오차는 0.0721로 기존의 방법보다 적은 수의 규칙으로 우수한 모델링 성능을 나타냄을 알 수 있다.

표 1 기존의 방법과의 비교

Model	No. of rules	MSE
Sugeno's	6	0.079
Virus-mGA	4	0.072

4. 결론

본 논문에서는 virus-mGA를 이용한 퍼지 모델링 방법을 제안하였다. 제안된 방법은 sGA에서 무시되었던 지역적 탐색 공간을 고려함으로써 지역과 전역 탐색 공간을 동시에 고려할 수 있는 동시에 가변스트링을 사용함으로써 동정의 유연성도 증가시킬 수 있다는 장점을 가진다. 또한 기존의 퍼지 모델링 방법과 비교를 통해 제안된 방법이 규칙 수나 평균 제곱 오차 면에서 우수함을 증명하였다.

[참고 문헌]

- [1] R. M. Tong, "The Evaluation of Fuzzy Models Derived form Experimental Data", Fuzzy Sets and Systems, Vol. 4, pp. 1-12, 1980.
- [2] 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.
- [3] H. Takagi and 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.
- [4] Y. H. Joo, H. S. Hwang, K. B. Kim and K. B. Woo, "Linguistic Model Identification for Fuzzy System", Electronics, Vol. 31, No. 4, pp. 330-331, 1995.
- [5] 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.
- [6] D. E. Goldberg, B. Korb, and K. Deb, "Messy Genetic Algorithms: Motivation, Analysis, and

- First Results", Complex Systems, Vol. 3, No. 5, pp. 493-530, 1989.
- [7] H. Kargupta, "The Gene Expression Messy Genetic Algorithm", Proc. of IEEE Int. Conf. on Evolutionary Computation, Nagoya, Japan, 1996
 - [8] 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.
 - [9] 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
 - [10] M. Sugeno and T. Yasukawa . " A Fuzzy logic -based Approach to Qualitative Modeling". IEEE Trans. Fuzzy Sys. Vo 1. No 1. pp. 7-31, Feb. 1993.