

# 유전자 알고리즘을 이용한 퍼지 시계열예측 방법에 관한 연구

## A Study on Fuzzy Time Series Prediction Method using the Genetic Algorithm

지현민\*, 장우석\*, 이성목\*, 강환일\*\*

Hyun Min Jee, Woo Seok Chang, Sung Mok Lee, Hwan Il Kang

**Abstract** - This paper proposes a time series prediction method for the nonlinear system using the fuzzy system and its genetic algorithm. At first, we obtain the optimal fuzzy membership function using the genetic algorithm. With the optimal fuzzy rules and its input differences, a better time prediction series system may be obtained. We obtain a good result for the time prediction of the electric load.

**Key Words** : evolutionary algorithm, genetic algorithm, time series prediction

### 1. 서론

시계열(Time series)은 정해진 순간마다 측정된 값  $x_1, x_2, \dots, x_n$ 들의 모임이다. 시계열예측의 응용은 사업계획 수립, 일기예보, 주가예측과 다양한 신호처리분야등에서 활발히 응용발전하고 있다. 일반적으로 시계열은 혼돈된 신호이거나 불규칙적인 특성을 가지고 있어 이들을 함수로 모델링하기에는 부족함이 있다. 따라서 근래에는 시계열에서 퍼지 규칙과 적당한 소속함수를 선택하여 시계열을 모델링하려는 경향이 있다.[1] 시계열중 특정주기기를 가지는 값 사이에는 얼마간의 종속성이 발견되고, 이런 성격에 근거하여 시계열의 발생원을 분석할 수 있다. 즉 시계열의 발생원은 어느 정도의 결정론적인 동력학 시스템(Deterministic Dynamic System)에 의해 지배를 받는다고 볼 수 있기 때문에 우리가 얻는 시계열은 예측이나 모델링을 통해 분석이 가능하다[1]. 신경망과 퍼지 논리를 적용한 모델링과 예측은 과거부터 많은 연구가 수행되었던 문제이다.[2,3,4,5]

Wang과 Mendel[6]은 학습을 통하여 퍼지규칙을 발생시켰고 Jang과 Sun[7]은 퍼지 IF-Then규칙을 혼돈신호예측에 이용하였다. Ye와 Gu[8]는 상하이 주가시장예측을 위한 퍼지 시스템을 개발하였다. 또한 smart trading을 위한 퍼지규칙이 Benachenhou[9]에 의해 개발되었다.

이 논문에서는 변형된 입력 방법을 도입함으로써 더 나은 시계열 예측 결과[10]를 얻고 또한 소속함수의 개수가 고정되어 있을 때 소속함수의 변형이 보다 더 좋은 예측결과를 나타냄을 보여주기 위함이다. 성능을 비교하는 척도로는 평균자승오차의 제곱근(RMSE:Root Mean Squared Error)를 사용하여 사용된 표 데이터는 Mackey-Glass time series를 사

용하였다. 또한 소속함수의 변형을 형성할때 유전자알고리즘을 이용하여 최적의 결과를 얻고자 한다.

### 2. 퍼지추론방법과 유전자알고리즘

#### 2.1 퍼지추론방법:

퍼지추론시스템은 아래(1)식을 만족하는 퍼지규칙  $R$ 를 찾아내는 작업이다.

$$\widehat{x}_{n+k} = R(x_1, x_2, \dots, x_n) \quad (1)$$

위의 식은 단기간의 예측에 이용된다. 보다 더 개선된 방법으로 식(1)대신에 변형된 입력  $(x_1 - x_2), (x_2 - x_3), \dots, (x_{n-1} - x_n)$ 을 사용하면 보다 더 좋은 결과를 얻을 수 있다.[10] 즉 다음 식(2)을 이용하여 예측할 수 있다.

$$\widehat{x}_n - \widehat{x}_{n+k} = R((x_1 - x_2), \dots, (x_{n-1} - x_n)) \quad (2)$$

이 논문에서는 식(2)을 이용하고 퍼지규칙은 우선 변형된 입력과 변형된 출력을 이용하여 퍼지소속함수중 어느곳에 속하는가를 결정한다. 예를들면  $(x_1 - x_2)$ 의 경우 주어진 퍼지소속함수중 어느 함수에 가장 많이 속하는 퍼지소속함수를 기억하고 그 소속함수값도 기억한다. 이때 동일한 전건부에 대하여 상이한 후건부가 발생하면 모순된 규칙이 발생하게 된다. 이 경우 Max-Min의 방법을 이용하면 해결할 수 있다.[10] Max-Min의 방법은 모순된 규칙의 입력과 출력의 모든 소속함수의 최소값이 큰 것이 있는 것을 후건부로 택하는 방식이다. 이런 방법으로 퍼지규칙을 구성한다. 이 퍼지규칙을 이용하여 예측치를 구하기 위해 Middle of Area(MOA), Centroid(중앙 평균법), Middle of Maxima(MOM)등의 퍼지회화기법을 사용하고 출력을 얻기위해서는 Mandani의 Min-Max방법과 Larson product 방법을 사용할 수 있다.[12]

저자 소개

\*明知大學 情輔工學科 學士課程

\*\*明知大學 情輔工學科學科 教授 · 工博

### 3. 유전자알고리즘법의 도입

퍼지함수의 모양을 동일한 모양의 삼각형으로 구성할 수 있으나 입력과 출력의 상황에 따라 퍼지함수가 각각 다른 삼각형의 모습으로 변형시키기 위해 유전자알고리즘을 도입한다. 우선 교배가 주연산자인 유전자알고리즘을 이용하여 최적의 퍼지함수모양을 정하여 예측결과를 구하였다.

#### 3.1 유전자알고리즘

유전자알고리즘은 생물학적 진화와 적자생존의 원리를 이용한 반복적인 탐색최적화기법으로 경사하강법의 단점인 지역극소점에 빠질 가능성을 제거하는데 있다. 유전자알고리즘의 특징은 주어진 변수에 해당되는 이진수를 만들고 적합도 함수를 구하여 적합도함수의 값에 따라 다음세대에 선택되어 지는데 2개의 기본연산자로서 교배와 돌연변이가 있으며 선택이라는 방법을 통해 반복수행된다.[11] 선택의 방법은 비례선택, Roulette Wheel, Tournament, Tournament with Eliticism 등이 있다.

### 4. 시계열예측에의 응용

본 실험에서는 1997년의 평일의 매시간 데이터를 이용하여 한 시간 후를 예측하는 실험을 하였다. 그림1에서는 평일의 전기부하를 나타낸다.

#### 4.1 소속함수의 정의

기존의 방법은 소속함수를 입력변수의 최소값과 최대값 사이를 임의의 개수로 등분하여 일률적으로 정의하였다. 그러나 시계열은 어떤 특정한 값근처에 많이 분포할 가능성이 있으므로 그 특성에 따라 소속함수를 정의하는 것이 보다 합리적인 방법이다. 본 논문에서는 유전자알고리즘을 이용하여 소속함수를 정의한다. 소속함수 정의 알고리즘은 다음과 같다.

단계1: 랜덤수를 50개를 생성시킨다. 이 50개를 염색체로 이용한다.그중에서 가장 작은 4개를 선정하고 그것이 있는 자리는 1로 다른자리는 0으로 하여 50개의 비트를 가진 이진수를 만든다.

단계2: 6개의 소속함수를 다음과 같이 정의한다.

$$\mu_1 = \frac{(x - p_{\min} + \epsilon)}{\epsilon} \quad x < p_{\min} = b(0)$$

$$\frac{(b(1) - x)}{b(1) - p_{\min}} \quad p_{\min} < x < b(1)$$

$$\mu_i = \frac{(x - b(i-2))}{b(i-1) - b(i-2)} \quad b(i-2) < x < b(i-1)$$

$$\frac{(b(i) - x)}{b(i) - b(i-1)} \quad b(i-1) < x < b(i)$$

$$(i = 2, \dots, 5)$$

$$\mu_6 = \frac{(x - b(4))}{p_{\max} - b(4)} \quad b(4) < x < p_{\max} = b(5)$$

$$\frac{(p_{\max} - x)}{\epsilon} \quad p_{\max} < x < p_{\max} + \epsilon$$

여기서  $b(i) = p_{\min} + \frac{j(i)(p_{\max} - p_{\min})}{51}$  이며

$i = 1, 2, 3, 4$ 이고  $j(i)$ 는 50개의 비트를 가진 이진수에서 자릿수를 왼쪽에서 오른쪽으로 세어  $i$ 번째 1이 나오는 자릿수가  $j$ 라는 것을 뜻한다. 예를들면

이진수  
000000000100000000010000000010000000000000000000에서  
 $j(2) = 20$ 이 된다.

단계3. 단계2의 소속함수를 이용하여  $x(10)$ 에서부터  $x(100)$ 까지 이용하여 퍼지규칙을 찾아내고 이 최적의 소속함수를 구하기 위해  $x(10)$ 에서  $x(100)$ 까지의 데이터를 이용하여 MAPE(mean absolute percentage error)의 역수를 적합도 함수로 정하였다. 유전자 알고리즘을 이용하여 최적의 소속함수의 모양을 찾는다. 유전자알고리즘을 구현하기 위한 각 변수들의 정의와 값은 다음과 같다. 세대수는 40세대이며 문자열의 최대길이는 50개 교배율은 0.3, 돌연변이율은 0.01, 교배의 위치는 한번만 허락하며 한 세대의 개수는 50개, 선택은 룰렛선택방식과 엘리트선택방식을 반반씩 이용하였다. 그리고 교배는 50개의 랜덤수를 가지는 두 염색체를 이용하여 행하며 돌연변이는 돌연변이가 일어난 곳은 1에서 원래 실수를 뺀 값으로 한다.

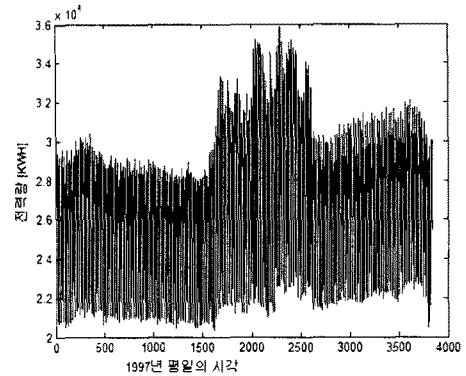


Fig. 1 Electric load for the normal days per each hour

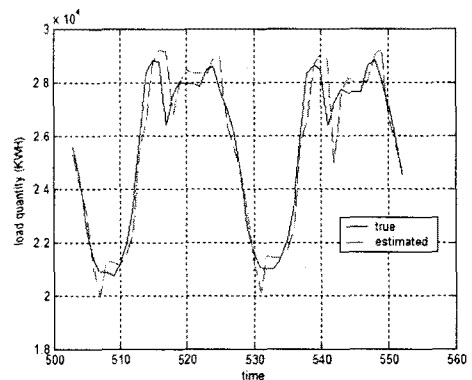


Fig. 2. The predicted value vs the true value

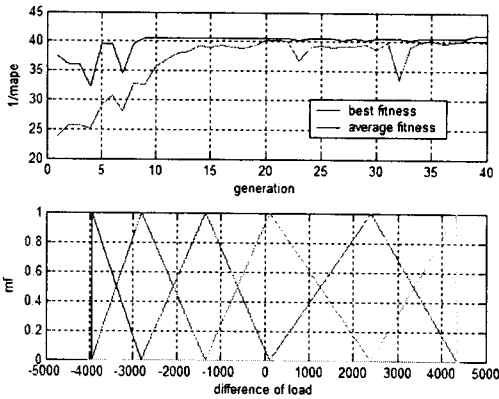


Fig. 3. Above: The best fitness and average fitness. Bottom: The optimal membership function

### 5. 실험

본 논문에서는 1997년의 평일의 매시간 데이터를 이용하여 한시간 후를 예측하는 실험을 하였다. 그림1에서는 평일의 전기부하를 나타낸다. 그림2에서는 6개의 퍼지 함수와 퍼지 규칙 데이터베이스를 이용하여 예측치와 실제값을 비교하였다. 그 결과 MAPE(mean absolute percentage error)는 2.43%로 나타난다. 그림3의 윗그림은 적합도 함수의 최댓값과 평균치의 변화를 나타낸다. 그림3의 아래그림은 최적의 소속함수의 모양을 나타내었다.

### 6. 결론

본 논문에서는 소속함수의 갯수가 고정되어 있을 때 균등한 퍼지소속함수에서 최적의 퍼지소속함수를 구함으로서 시계열을 정확히 예측할 수 있음을 보였다. 소속함수의 변형을 만들 때 유전자알고리즘을 이용하여 조사해 본 결과 유전자 알고리즘이 우수한 특성을 나타내었다.

### 감사의 글

본 연구는 전력산업지원사업(R-2004-B-201)의 도움을 받아 이루어진 연구로서, 관계부처에 감사 드립니다.

### 참 고 문 헌

[1] 김인택, 공창욱, "시계열예측을 위한 퍼지학습알고리즘," 퍼지 및 지능시스템학회 논문지, 제7권 제3호, pp.34-42. 제7권 제3호, 1997.

[2] R. M. Tong, "The evaluation of Fuzzy Models derived from Experimental Data," Fuzzy sets and Systems, vol. 4, pp. 1-12, 1980.

[3] W. Pedrycz, "Fuzzy Control & Fuzzy Systems," John Wiley & Sons, 1989.

[4] K. Hornik, M. Stinchcombe and H. White, "Multilayer feedback networks are universal approximators," Neural Network, vol. 2, pp. 359-366, 1989.

[5] Wang, L. E., "Fuzzy systems are universal approximators," Proc. IEEE International Conf. on Fuzzy Systems, San Diego, pp. 1163-1170, 1992.

[6] L.X. Wang & J. M. Mendel, "Generating fuzzy rules by learning from examples," IEEE Trans Syst., Man, Cybern., vol. 22, pp. 1414-1427, 1992.

[7] R. Jang & C. Sun, "Prediction Chaotic Time series with Fuzzy If-Then rules," Second IEEE International Conference on Fuzzy Systems, San Francisco, pp.1079-1084, 1993.

[8] Z. Ye & L. Gu, "A fuzzy system for trading the shanghai stock market," Trading on the Edge, Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets, G. J. Deboeck, Ed. New York: Wiley, 1994. pp. 207-214.

[9] D. Benachenhou, "Smart trading with (FRET)," in Trading on the Edge, Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets, G. J. Deboeck, Ed., New York: Wiley, 1994, pp.215-242

[10] 이성록, 김인택, "A study on new fuzzy time series prediction method," pp.565-569, 명지대학교 산업기술연구소 논문집 vol. 19, 1999.

[11] Zbigniew Michalewicz, "Genetic Algorithms+ Data Structures=Evolution Programs, Springer Verlag, 1994.

[12] 임영도, 이상도, 퍼지, 신경망, 유전진화, pp. 215-216, 1996, 도서출판 영과 일.