

Design of HCBKA-Based IT2TSK Fuzzy Prediction System

방 영 근* · 이 철 희†
(Young-Keun Bang · Chul-Heui Lee)

Abstract - It is not easy to analyze the strong nonlinear time series and effectively design a good prediction system especially due to the difficulties in handling the potential uncertainty included in data and prediction method. To solve this problem, a new design method for fuzzy prediction system is suggested in this paper. The proposed method contains the followings as major parts ; the first-order difference detection to extract the stable information from the nonlinear characteristics of time series, the fuzzy rule generation based on the hierarchically classifying clustering technique to reduce incorrectness of the system parameter identification, and the IT2TSK fuzzy logic system to reasonably handle the potential uncertainty of the series. In addition, the design of the multiple predictors is considered to reflect sufficiently the diverse characteristics concealed in the series. Finally, computer simulations are performed to verify the performance and the effectiveness of the proposed prediction system.

Key Words : First-order difference, Hierarchically classifying clustering, IT2TSK fuzzy logic, Multiple predictors

1. 서 론

현대 사회는 기술의 다양화와 구조의 복잡화로 인해, 다루어져야 할 데이터들의 특성들 또한 매우 복잡하고 정교한 형태를 보이고 있다. 특히, 신호처리 분야나 예측 제어 분야, 경제 계획, 지구학, 기상학 등과 같은 분야들에서는 그들에게서 나타나는 일종의 신호와 같은 개념의 데이터들의 속성 분석을 위해 많은 노력을 기울이고 있는 중이다. 이는 보다 정교하고 세밀한 출력 특성을 가지는 시스템 및 그들의 복잡한 구조적 특성에 능동적으로 대처할 수 있는 시스템을 효과적으로 구현하기 위해 필요한 중요 과정으로 이해될 수 있다 [1-3]. 어떤 시스템이 다양한 형태로 수집된 신호 또는 데이터들이 가지는 속성들을 합리적 과정을 통해 분석한 후, 이러한 분석결과를 토대로 목적에 맞게 능동적으로 동작할 수 있다면, 여러 분야에서 높은 효율성을 보일 수 있을 것이다. 이러한 시스템 설계를 목적으로 하는 분야로 시계열 예측 분야 또한 중요한 하나의 분야로 간주 될 수 있다. 이는 인간의 변화되는 행동패턴에 따른, 또는 자연현상의 변화 패턴에 따른 분석을 토대로 미래와 관련된 정보를 제공하는 것을 주요 목적으로 하므로 인간의 생활과 밀접한 관계를 보일 수 있기 때문이다. 하지만 예측 분야에서 취급되는 데이터들의 속성들을 분석하는 것이 쉽지만은 않다. 이는 사회의 고도화에 따른 여러 분야의 취급 데이터들이 보다 복잡한 구조로 연관되어 지기 때문이다.

따라서 고전적 기법으로 간주 될 수 있는 선형 모델들의 경우 데이터들의 복잡한 구조에 기인한 강한 비선형적 특성이나 데이터가 담고 있는 정보의 불확실성을 취급하기에는 어려움이 존재하여 왔다 [4]. 또한 최근에 활발히 연구되어지고 있는 퍼지나 신경망, 또는 유전알고리즘과 같은 소프트웨어 기법들이 비교적 고전 기법들에 비해 데이터의 불확실한 정보의 취급에 용이한 특성을 담고 있었지만, 능동적 시스템의 구현을 위해선 그들 사이의 결합이나 구조의 복잡성을 취할 수밖에 없었으며 항상 좋은 성능을 유지하는 것은 아니었다[5-7]. 이러한 맥락에서 우리는 논문 [8-9]를 통해 고전기법이 가지는 장점들과 퍼지모델이 가지는 불확실성의 취급능력을 동시에 활용할 수 있는 예측시스템을 제안한 바 있다. 또한, 논문 [9]에서는 보다 구조적으로 덜 복잡한 형태, 다시 말하면 시스템의 행동 특성을 결정짓는 퍼지 규칙기반을 보다 효과적으로 생성할 수 있는 계층적 분류 구조를 갖는 퍼지 시스템의 설계 법을 제시하였다. 이 논문들에서는 원형 데이터의 특성을 충분히 반영할 수 있고, 또한 1차 차분 데이터가 가지는 안정된 정보 특성을 이용하기 위해 데이터의 전처리 과정과 다수의 TSK 퍼지 예측기를 이용하는 방법을 사용하였으며, 또한 효율적인 퍼지 규칙기반의 생성을 위해 계층구조 클러스터링 기법(HCBKA : Hierarchical Correlation-Based K-means clustering Algorithm)을 제안하여, 설계된 예측기의 효율성을 높일 수 있도록 하였다. 또한 강한 비선형 특성의 시계열 예측 시뮬레이션을 통해 실제 좋은 예측 기대치를 만족하는 결과를 보여줌으로써 사용된 방법의 효용성을 입증하였다.

하지만 Mendel은 일반적으로 Type-1 퍼지 논리만으로는 데이터 이면에 내재된 불확실성을 충분히 다룰 수 없으며, 결국 초기 Zadeh에 의해 고려되었던 퍼지 논리의 확장 개념

* 정 회 원 : 강원대학교 삼척캠퍼스 전기공학과 강사

† 교신저자, 정회원 : 강원대학교 전기전자공학과 교수

E-mail : chlee@kangwon.ac.kr

접수일자 : 2011년 3월 8일

최종완료 : 2011년 5월 14일

을 일반화 시킨 Type-2 퍼지논리를 체계화함으로써 이를 극복할 수 있음을 그의 논문과 저서를 통해 증명하였다 [10-11]. 또한, Mendel은 Type-2 퍼지논리의 추론과정의 복잡하고 난해한 수학적 접근을 개선하고자 Interval Type-2 퍼지논리를 제안하여, 구조적으로 보다 쉽게 접근 가능한 Type-2 퍼지논리 개념을 일반화 시켰다 [12].

따라서 본 논문에서는 기존 논문[9]의 1차 차분과 연계된 Type-1 퍼지 예측기의 구조를 1차 차분 특성을 그대로 이용하면서도 차분 데이터 내에 내재된 데이터의 불확실성을 보다 효과적으로 취급할 수 있는 1차 차분과 연계된 Interval Type-2 퍼지 예측 시스템의 설계 법을 다룬다. 시스템을 구성하고 있는 각각의 퍼지 예측기의 설계 시에는 TSK 퍼지모델과 계층구조 클러스터링 기법을 적용함으로써 고전 선형모델이 가지는 근사추론의 이점과 퍼지 규칙의 데이터에 대한 기술 능력을 개선함으로써 보다 단순한 구조와 적은 수의 규칙기반으로도 다양한 특성으로 표현되는 시계열들의 예측 문제에서도 충분히 좋은 결과를 나타낼 수 있도록 하였다. 결국, 이러한 구조로 설계되어지는 제안된 예측시스템은 1차 차분데이터가 가지는 안정된 정보 제공 능력과, TSK 퍼지 모델을 통한 선형모델이 가지는 강점, 그리고 Interval Type-2 퍼지 논리를 통한 불확실성의 취급능력을 모두 수용할 수 있도록 설계되었음을 의미하며, 마지막으로 비정상, 비선형 시계열 데이터의 예측 시뮬레이션을 통해 제안된 시스템의 성능과 적용된 방법들의 효용성을 입증하였다.

2. Interval Type-2 TSK 퍼지 예측 시스템의 구조

그림 1은 제안된 Interval Type-2 TSK(이하 IT2TSK) 퍼지 예측 시스템의 기본 구조를 보여준다.

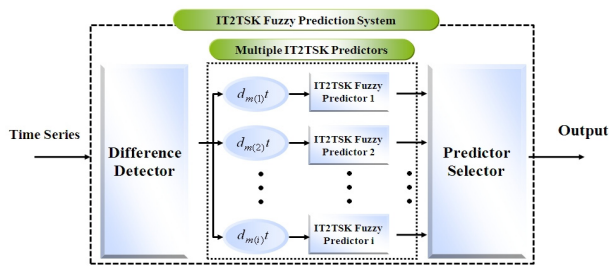


그림 1 제안된 IT2TSK 퍼지 예측 시스템의 전체 구조
Fig. 1 Total structure of the proposed IT2TSK fuzzy prediction system

먼저, 그림 1에서 차분 검출기는 앞서 서론에 언급된 1차 차분데이터를 생성하는 과정을 의미한다. 이 검출기의 역할은 훈련데이터로부터 생성가능한 모든 차분데이터들 중에서 원형 데이터의 특성을 잘 드러낼 수 있는 최적의 1차 차분 데이터를 생성하는 부분이다. 차분 검출기에서 원형 데이터의 특성표현에 적합한 최적의 차분 간격 판별을 위해 자기상관함수[8-9] 분석 기법이 사용되며, 3장에서 설명되어진다. 다음으로 선택된 차분간격으로부터 생성되는 각각의 차분데이터들의 특성 분석을 통해 상응하는 각각의 퍼지 예측기들이 설계되어진다. 즉, 설계 되어야 할 예측기의 개수는

선택된 차분 간격 수와 동일함을 의미하며, 각각의 퍼지 예측기의 설계를 위해, IT2TSK 퍼지 모델, HCBKA, LSM(least squares method), 그리고 상응하는 차분데이터가 사용된다. 이러한 IT2TSK 퍼지 예측기의 설계과정은 4장에서 소개된다. 마지막으로, 5장에서는 제안된 예측시스템의 마지막 과정인 예측기 선택자를 설계하는 과정으로 다수의 예측기들의 병렬구조 속에서 성능 평가 지수를 이용하여 최적 예측기를 선택할 수 있는 방법을 설명하게 되며, 이를 통해 제안된 예측 시스템의 출력을 얻는 방법을 설명하게 된다.

3. 차분 검출기(데이터의 전처리)

차분 검출기는 원형 데이터로부터 생성될 수 있는 모든 차분데이터들 중에서 원형데이터의 특성, 즉 패턴이나 법칙성, 또는 계절성과 같이 유용한 정보를 충분히 드러낼 수 있는 차분데이터를 생성하는 역할을 하게 된다. 이러한 차분데이터의 생성을 위해 우리는 논문 [9]에서 타당한 방법을 제안한 바 있다. 제안된 방법을 간단히 요약하면, N 길이의 주어진 훈련데이터로부터 생성가능한 모든 차분데이터의 개수는 $N-1$ 개가 되고, 이들 중 적절한 개수만을 추출하여 시스템 설계에 이용하는 것으로 다음과 같은 방법으로 진행된다. 먼저 식 (1)과 같이 차분 간격에 따른 원형 데이터와의 자기상관함수를 분석한다.

$$r_j = \frac{\sum_{i=1}^{N-j} (y(i) - \bar{y})(y(i+j) - \bar{y})}{\sum_{i=1}^N (y(i) - \bar{y})^2} \quad (1)$$

여기서, $y(i)$ 는 i 번째 훈련데이터를 의미하고, \bar{y} 는 훈련데이터의 평균, j 는 차분 간격, 그리고 r_j 는 자기상관함수를 의미한다. 다음으로 전체 분석된 $N-1$ 개의 자기상관계수를 크기순으로 정렬한 후 처음부터 다섯 번째까지의 자기상관계수에 해당하는 차분간격 값들을 1차 선택하게 된다. 이러한 이유는 계절이 가진 특성을 최소로 고려하기 위한 것이다. 첫 번째 선택이 끝나면, 나머지 차분간격 값들 사이에서 차 연산을 실행하게 되며, 이때 차 연산 값이 가장 큰 폭을 보이는 두 상관함수 사이에서 상관성이 급격히 떨어진다고 볼 수 있으므로 이를 임계점으로 지정하고 그 이상의 상관계수에 해당하는 차분간격 값을 선택하게 된다. 이러한 방법은 전문가의 지식이나 반복적 실험을 필요로 하지 않고 단지 계절의 자기 상관함수 분석을 통해서만 이루어지므로 매우 효율적이라 할 수 있다. 따라서 첫 번째 선택과정 이후의 $N-6$ 개의 나머지 상관계수 값들에서 차 연산의 결과 값들은 $N-7$ 개의 값들을 나타낼 것이며, 연산의 효율성을 위해 상관계수 값이 0이하의 상관 계수 값에 해당하는 차분 간격 값들을 고려하지 않는다면, 다시 말해 0이하의 자기상관계수가 \emptyset 개라면 결국 차 연산 결과 값은 $N-7-\emptyset$ 개가 된다. 이 연산 결과들에서 α 번째 연산 결과가 가장 크게 나타났다면, 이는 α 번째 상관계수와 $\alpha+1$ 번째 계수 사이의 가장 큰 변화를 의미하는 것으로 α 번째 상관계수 이상에 상응하는 차분 간격 값들이 이차 최적 차분 간격 후보군으로 선택됨을 의미한다. 따라서 시스템의 설계를 위해 생성되는 차

분데이터의 수는 $5+\alpha$ 개이며, 제안된 시스템을 구성하는 예측기들의 수 또한 이와 동일함을 의미한다. 이렇게 선정된 $5+\alpha$ 개의 차분 간격 후보군은 아래와 같은 방법으로 각각의 차분간격에 상응하는 차분 데이터들을 생성하게 된다.

$$\begin{aligned} d_{m(i)}t_1 &= y(N) - y(N-m(i)) \\ d_{m(i)}t_2 &= y(N-1) - y(N-m(i)-1) \\ &\vdots \\ d_{m(i)}t_n &= y(N-n-1) - y(N-m(i)-n-1) \\ &\vdots \\ d_{m(i)}t_{N-m(i)} &= y(m(i)+1) - y(1) \end{aligned} \quad (2)$$

여기서, $m(i)$ 는 선택된 차분 간격 후보군 중에서 i 번째 차분 간격을 의미하며, $d_{m(i)}t_n$ 은 상응하여 생성된 차분데이터 중에서 n 번째 차분데이터를 의미한다. 이렇게 생성된 차분데이터들은 그들 각각의 특성을 반영할 수 있는 독립된 예측기의 설계를 위한 입력 데이터로 사용된다. 따라서 설계되는 전체 시스템은 각각의 차분데이터가 표현하는 원형 데이터의 특성과 각각의 차분데이터의 특성을 반영하는 각각의 예측기들을 통해 원형 데이터에 내재된 불확실성과 비선형적 특성을 충분히 수용할 수 있는 구조로 설계되어 진다.

4. 차분 검출기(데이터의 전처리)

TSK 퍼지 모델의 경우 일반적으로 전반부의 퍼지규칙 표현을 위한 입력 데이터 쌍과 후반부 선형식의 파라미터 추정을 위한 출력 값을 요구하게 된다. 이를 위해 생성된 차분데이터들로부터 입력과 파라미터 추정을 위한 출력 값을 포함하는 데이터 쌍을 정의하면 다음과 표현될 수 있다.

$$\begin{aligned} \text{Sets} &= [d_{m(i)}t_k, d_{m(i)}t_{k+1}, d_{m(i)}t_{k+2}, \dots, d_{m(i)}t_{k+p}] \quad (3) \\ \text{where } k &= [1, N-m(i)-p] \end{aligned}$$

여기서, $d_{m(i)}t_k$ 는 파라미터 추정을 위한 출력 값을 나타내며, $k+1$ 번째 데이터부터 $k+p$ 번째 데이터까지가 입력데이터로 사용된다. 따라서 위의 입력 데이터 쌍을 표현할 수 있는 IT2TSK 퍼지 모델의 j 번째 퍼지 규칙 R^j 는 다음과 같이 정의 될 수 있다.

$$\begin{aligned} R^j: \text{ If } & d_{m(i)}t_{k+1} \text{ is } \tilde{F}_1^j \text{ and } d_{m(i)}t_{k+2} \text{ is } \tilde{F}_2^j \text{ and} \\ & \dots \text{ and } d_{m(i)}t_{k+p} \text{ is } \tilde{F}_p^j \\ \text{Then } Y^j &= \sum_{q=0}^p C_q^j d_{m(i)}^q t_{k+p} \end{aligned} \quad (4)$$

$$\text{where } d_{m(i)}t_{k+0} = 1 \text{ and } C_q^j = [c_q^j - s_q^j, c_q^j + s_q^j]$$

여기서, C_q^j 는 추정되어야 할 파라미터를 의미하고, c_q^j 는 C_q^j 의 각각의 중심 또는 평균을 의미하며, s_q^j 는 C_q^j 의 폭을 의미한다. 또한, \tilde{F} 는 interval Type-2 퍼지 집합을 의미하며, Y^j 는 $d_{m(i)}t_k$ 에 대한 Interval Set으로 $[d_{m(i)}^L t_k^L, d_{m(i)}^R t_k^R]$ 와 같이 정의된다. 결국, Interval Type-2 퍼지 추론을 위해선 Interval Set으로 주어지는 출력 값 Y^j 에서 최종 출력 값을

탐색하는 과정을 거치게 되며, Interval Set으로 주어지는 각각의 출력 집합을 얻기 위해선 Type-1 퍼지논리의 점화강도(firing strength)에 해당하는 Type-2 퍼지 집합의 점화 집합(firing set)을 먼저 추론하여야 한다. Interval Type-2 퍼지 집합의 경우 Secondary 소속함수의 모든 값이 1로 정의되므로 점화 집합은 다음과 같이 상위 소속함수(UMF: upper membership function)와 하위 소속함수(LMF: lower membership function)에 대한 점화 강도의 쌍으로 정확히 기술 되어질 수 있게 된다.

$$F^j(\mathbf{x}) = [f^j(\mathbf{x}), \bar{f}^j(\mathbf{x})] \quad (5)$$

$$f^j(\mathbf{x}) = \underline{\mu}_{\tilde{F}_1^j}(x_1) \star \underline{\mu}_{\tilde{F}_2^j}(x_2) \star \dots \star \underline{\mu}_{\tilde{F}_p^j}(x_p) \quad (6)$$

$$\bar{f}^j(\mathbf{x}) = \bar{\mu}_{\tilde{F}_1^j}(x_1) \star \bar{\mu}_{\tilde{F}_2^j}(x_2) \star \dots \star \bar{\mu}_{\tilde{F}_p^j}(x_p) \quad (7)$$

여기서, f 는 하위 소속함수에 대한 점화 강도, \bar{f} 는 상위 소속함수에 대한 점화 강도를 의미하며, F 는 이들 점화 강도로 구성되는 점화집합을 의미하게 된다. 또한, μ 는 각 입력 데이터의 퍼지집합에 대한 소속 정도를 나타내며, \star 은 T-norm 연산자로, 본 논문에서는 minimum T-norm 연산자를 사용한다. 이렇게 구해진 점화집합은 Type-2 퍼지 집합의 최종 출력을 얻기 위해 사용되어 진다.

4.1 HCBKA-based IT2TSK 퍼지 규칙 생성

그림 2는 본 논문에서 사용된 HCBKA에 기반된 입력공간의 퍼지 분할 예를 보여준다.

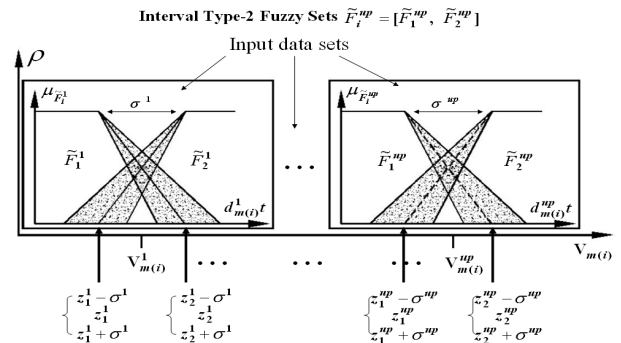


그림 2 HCBKA에 기반된 입력공간 퍼지 분할 예
Fig. 2 Example of HCBKA-based fuzzy partition

그림 2에서 $V_{m(i)}$ 는 상위 클러스터의 중심을 의미하며, z 는 클러스터 내에 분할된 퍼지 집합의 중심이다. 또한, σ 는 Interval Type-2 퍼지 집합의 불확실의 정도를 표현하는 FOU(footprint of uncertainty)의 크기를 의미한다. 그림처럼 본 논문은 HCBKA를 이용한 퍼지분할을 실행 하였다. 먼저 입력 데이터들은 데이터들의 상관성을 기반으로 하여 각각의 상위 클러스터(upper cluster)로 일차 분류되며, 분류된 데이터들을 이용하여 각각의 상위 클러스터 내에서 퍼지 집합으로 제 분할되는 과정을 거치게 된다. 상위 클러스터의 생성을 위한 방법은 다음과 같다. 먼저 up 번째 상위 클러스터의 임의의 중심 $V_{m(i)}^{up}$ 이 $[v(1)_{m(i)}^{up}, v(2)_{m(i)}^{up}, \dots, v(p)_{m(i)}^{up}]$ 이라

면, 입력데이터 쌍들은 다음의 상관성 판별식을 통해 판별 값이 더욱 큰 상위 클러스터로 일차 분류된다.

$$\rho_{DV}^{up} = \frac{\sum_{q=1}^p (d_{m(i)t_{k+q}} - \bar{D})(v(q)_{m(i)}^{up} - \bar{V}_{m(i)}^{up})}{\sqrt{\sum_{q=1}^p (d_{m(i)t_{k+q}} - \bar{D})^2} \sqrt{\sum_{q=1}^p (v(q)_{m(i)}^{up} - \bar{V}_{m(i)}^{up})^2}} \quad (8)$$

여기서, ρ 는 입력데이터 쌍들과 임의의 상위 클러스터 중심과의 교차 상관계수를 의미하며, \bar{D} 는 입력 데이터쌍의 평균을 의미하고 \bar{V} 는 임의의 상위 클러스터 중심의 평균을 의미한다. 따라서 식 (8)을 이용하여 식 (3)과 같이 생성된 모든 입력쌍들이 상위 클러스터로 분류가 되면, up 번째 상위 클러스터의 중심 $V_{m(i)}^{up}$ 은 다음과 같은 방법으로 갱신된다.

$$V_{m(i)}^{up} = \frac{1}{J^{up}} \sum_{n=1}^{J^{up}} D_n^{up} \quad (9)$$

여기서, D^{up} 은 up 번째 상위 클러스터에 분류된 입력 데이터쌍을 의미하고 J^{up} 은 up 번째 상위 클러스터에 분류된 데이터쌍의 개수를 의미한다. 이러한 방법으로 상위 클러스터의 중심이 갱신되면 식(8)과 식(9)를 이용하여 상위 클러스터의 중심을 지속적으로 반복 탐색하게 되며, 아래의 정지 조건을 만족하게 되면 탐색을 중지하게 된다.

$$Stop = \left| \frac{V_{m(i)}^{up}(pr) - V_{m(i)}^{up}(cu)}{V_{m(i)}^{up}(pr)} \right| \leq 10^{-4} \quad (10)$$

여기서, $V_{m(i)}^{up}(pr)$ 는 갱신되기 바로 전단계의 up 번째 상위 클러스터의 중심을 의미하고, $V_{m(i)}^{up}(cu)$ 는 갱신된 상위 클러스터의 중심을 의미한다.

이렇게 상위 클러스터의 생성이 끝나고, 모든 입력 데이터쌍들이 상위 클러스터 내로 분류가 되면, 각각의 상위 클러스터 내에 분류된 입력 데이터쌍을 이용하여 퍼지집합을 생성하게 된다. 상위 클러스터내의 퍼지집합의 생성을 위해 각 클러스터에 분류된 입력 데이터쌍들 중 최대값과 최소값 사이를 전체 입력공간으로 하여 그림 2와 같이 각 상위 클러스터마다 K-means 클러스터링 알고리즘을 이용하여 2개씩의 퍼지 집합을 생성하며, 이때에 Interval Type-2 퍼지집합의 FOU를 위한 σ 는 다음과 같이 정의되어진다.

$$\sigma^{up} = \sqrt{\sum_{h=1}^d (d_{m(i)t_k}^{up(c)h} - \bar{d}_{m(i)t_k}^{up(c)}) / (d-1)} \quad (11)$$

여기서, σ^{up} 는 up 번째 상위 클러스터의 Interval Type-2 퍼지집합의 FOU를 위한 표준편차를 의미하고, d 는 up 번째 상위 클러스터 내에 2개로 정의된 Interval Type-2 퍼지집합의 각 중심 사이에 포함된 데이터의 개수를 의미한다. 또한, $d_{m(i)t_k}^{up(c)h}$ 는 up 번째 상위 클러스터 내의 두 개의 퍼지집합 중심 사이에 위치하는 입력 차분 데이터들을 의미한다. 따라서 이러한 방법으로 Interval Type-2 퍼지집합의 생성이

끝나면, 입력데이터가 각각의 퍼지집합을 만족하는 소속 정도는 다음과 같이 정의 된다.

$$\begin{aligned} &\text{if } d_{m(i)t}^{up} \in \tilde{F}_1^{up} \text{ and } d_{m(i)t}^{up} \leq z_1^{up} \\ &\quad \underline{\mu}_{\tilde{F}_1}^L(d_{m(i)t}^{up}) = 1, \quad \underline{\mu}_{\tilde{F}_1}^R(d_{m(i)t}^{up}) = 1 \\ &\text{elseif } d_{m(i)t}^{up} \in \tilde{F}_2^{up} \text{ and } d_{m(i)t}^{up} \geq z_2^{up} \\ &\quad \underline{\mu}_{\tilde{F}_2}^L(d_{m(i)t}^{up}) = 1, \quad \underline{\mu}_{\tilde{F}_2}^R(d_{m(i)t}^{up}) = 1 \\ &\text{elseif } z_1^{up} < d_{m(i)t}^{up} < z_2^{up} \end{aligned} \quad (12)$$

$$\begin{aligned} \underline{\mu}_{\tilde{F}_1}^L(d_{m(i)t}^{up}) &= \frac{d_{m(i)t}^{up} - (z_1^{up} - \sigma^{up})}{z_2^{up} - (z_1^{up} - \sigma^{up})}, \\ \underline{\mu}_{\tilde{F}_1}^R(d_{m(i)t}^{up}) &= \frac{d_{m(i)t}^{up} - (z_1^{up} + \sigma^{up})}{z_2^{up} - (z_1^{up} + \sigma^{up})}, \\ \underline{\mu}_{\tilde{F}_2}^L(d_{m(i)t}^{up}) &= \frac{(z_1^{up} + \sigma^{up}) - d_{m(i)t}^{up}}{(z_2^{up} + \sigma^{up}) - z_1^{up}}, \\ \underline{\mu}_{\tilde{F}_2}^R(d_{m(i)t}^{up}) &= \frac{(z_1^{up} - \sigma^{up}) - d_{m(i)t}^{up}}{(z_2^{up} - \sigma^{up}) - z_1^{up}} \end{aligned}$$

여기서, μ^L 은 입력데이터가 만족한 퍼지집합의 중심으로 부터 왼쪽 경사면에 대한 소속 정도를 의미하고, μ^R 은 오른쪽 경사면의 소속 정도를 의미한다. 또한, 각각의 입력 데이터쌍이 각각의 상위 클러스터 내에서 생성하는 퍼지 규칙기반의 전반부를 표현하면 다음과 같이 표현될 수 있다.

$$\begin{aligned} R^{up(1)} : & d_{m(i)t_{k+1}}^{up} \text{ is } \tilde{F}_1^{up} \text{ and } d_{m(i)t_{k+2}}^{up} \text{ is } \tilde{F}_1^{up} \text{ and} \\ & \dots \text{ and } d_{m(i)t_{k+p}}^{up} \text{ is } \tilde{F}_1^{up} \\ R^{up(2)} : & d_{m(i)t_{k+1}}^{up} \text{ is } \tilde{F}_1^{up} \text{ and } d_{m(i)t_{k+2}}^{up} \text{ is } \tilde{F}_1^{up} \text{ and} \\ & \dots \text{ and } d_{m(i)t_{k+p}}^{up} \text{ is } \tilde{F}_2^{up} \\ & \vdots \\ R^{up(j)} : & d_{m(i)t_{k+1}}^{up} \text{ is } \tilde{F}_2^{up} \text{ and } d_{m(i)t_{k+2}}^{up} \text{ is } \tilde{F}_2^{up} \text{ and} \\ & \dots \text{ and } d_{m(i)t_{k+p}}^{up} \text{ is } \tilde{F}_1^{up} \\ & \vdots \\ R^{up(M)} : & d_{m(i)t_{k+1}}^{up} \text{ is } \tilde{F}_2^{up} \text{ and } d_{m(i)t_{k+2}}^{up} \text{ is } \tilde{F}_2^{up} \text{ and} \\ & \dots \text{ and } d_{m(i)t_{k+p}}^{up} \text{ is } \tilde{F}_2^{up} \end{aligned} \quad (13)$$

만약, 하나의 입력데이터쌍이 3개의 데이터로 구성되어 있다면, 이 입력은 총 8개의 퍼지 규칙을 생성할 수 있으며, 전체 퍼지규칙은 생성된 모든 입력 데이터쌍에 대하여 중복되는 규칙은 삭제하면서 생성할 수 있게 된다. 퍼지 규칙기반의 전반부 규칙의 생성이 끝나면, 전반부의 규칙을 만족하는 입력데이터쌍을 이용하여 TSK 퍼지 규칙의 후반부 파라미터를 식별하게 된다.

4.2 HCBKA-based IT2TSK 퍼지 규칙 파라미터 식별

식 (13)과 같이 생성된 퍼지규칙에서 up 번째 상위 클러스터에서 생성된 j 번째 퍼지규칙 $R^{up(j)}$ 를 만족하는 입력 데이터 쌍의 수가 n 개라면, 규칙 $R^{up(j)}$ 의 후반부 파라미터들은 n 개의 입력 데이터쌍을 이용하여 추정되어 되며, 비록 TSK 퍼지 규칙의 전반부 규칙이 Type-2 형태를 나타내고 있더라도 Type-1 TSK 퍼지모델의 파라미터 추정기법에 사용되는 최소자승법(LSM : least squares method)으로 간략하게

추정할 수 있다. 이러한 접근은 Type-2 퍼지 집합의 FOU가 0일 경우 정확히 Type-1 퍼지논리로 사상될 수 있기 때문이다. 따라서 FOU를 0로 가정하여 파라미터를 식별한 후, 전반부에서 주어지는 점화집합과 이들로부터 연산되어지는 최대 무게중심 출력과 최소 무게 중심 출력을 얻게 된다면 Type-2 퍼지 논리의 출력을 얻을 수 있게 된다. 먼저, n 개의 입력데이터 쌍으로부터 퍼지규칙 $R^{up(j)}$ 의 후반부는 다음과 같이 n 개의 연립방정식으로 표현 가능해 진다.

$$\begin{aligned} d_{m(i)}^{up(j)} t_k^1 &= c_0^{up(j)} + c_1^{up(j)} d_{m(i)}^{up(j)} t_{k+1}^1 + c_2^{up(j)} d_{m(i)}^{up(j)} t_{k+2}^1 + \dots + c_p^{up(j)} d_{m(i)}^{up(j)} t_{k+p}^1 \\ &\vdots \\ d_{m(i)}^{up(j)} t_k^l &= c_0^{up(j)} + c_1^{up(j)} d_{m(i)}^{up(j)} t_{k+1}^l + c_2^{up(j)} d_{m(i)}^{up(j)} t_{k+2}^l + \dots + c_p^{up(j)} d_{m(i)}^{up(j)} t_{k+p}^l \\ &\vdots \\ d_{m(i)}^{up(j)} t_k^n &= c_0^{up(j)} + c_1^{up(j)} d_{m(i)}^{up(j)} t_{k+1}^n + c_2^{up(j)} d_{m(i)}^{up(j)} t_{k+2}^n + \dots + c_p^{up(j)} d_{m(i)}^{up(j)} t_{k+p}^n \end{aligned} \quad (14)$$

또한, 식 (14)를 행렬식과 벡터로 표현하면 다음과 같이 표현될 수 있다.

$$\begin{bmatrix} d_{m(i)}^{up(j)} t_k^1 \\ \vdots \\ d_{m(i)}^{up(j)} t_k^l \\ \vdots \\ d_{m(i)}^{up(j)} t_k^n \end{bmatrix} = \begin{bmatrix} 1 & d_{m(i)}^{up(j)} t_{k+1}^1 & d_{m(i)}^{up(j)} t_{k+2}^1 & \dots & d_{m(i)}^{up(j)} t_{k+p}^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & d_{m(i)}^{up(j)} t_{k+1}^l & d_{m(i)}^{up(j)} t_{k+2}^l & \dots & d_{m(i)}^{up(j)} t_{k+p}^l \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & d_{m(i)}^{up(j)} t_{k+1}^n & d_{m(i)}^{up(j)} t_{k+2}^n & \dots & d_{m(i)}^{up(j)} t_{k+p}^n \end{bmatrix} \begin{bmatrix} c_0^{up(j)} \\ c_1^{up(j)} \\ c_2^{up(j)} \\ \vdots \\ c_p^{up(j)} \end{bmatrix} \quad (15)$$

$$Y^{up(j)} = X^{up(j)} C^{up(j)} \quad (16)$$

여기서, Y 는 식 (15) 왼쪽 항의 출력 벡터, X 는 오른쪽 항의 입력데이터 쌍 벡터, 그리고 C 는 파라미터 벡터를 의미한다. 따라서 식 (16)의 파라미터 벡터는 오차과위의 합을 최소로 만드는 최소자승법을 통해 다음과 같이 추정될 수 있다.

$$\hat{C}^{up(j)} = ((X^{up(j)})^T X^{up(j)})^{-1} (X^{up(j)})^T Y^{up(j)} \quad (17)$$

4.3 HCBKA-based IT2TSK 퍼지 예측기의 출력

식 (13)과 같이 일반적으로 하나의 입력 데이터쌍은 여러 개의 퍼지 규칙을 만족할 수 있다. 따라서 하나의 입력데이터 쌍이 M 개의 퍼지 규칙을 만족하였다면, 이 입력쌍을 위한 출력은 식 (17)에 의해 추정된 파라미터와 식 (5)의 점화집합에 의해 다음과 같이 Type-1 퍼지 논리로 주어지는 하한 출력과 상한 출력의 2개의 출력을 얻을 수 있게 된다.

$$\begin{aligned} y_{TSK,1}^{l(up)}(\mathbf{x}) &= \frac{\sum_{j=1}^M f^{up(j)}(\mathbf{x}) y^{up(j)}(\mathbf{x})}{\sum_{j=1}^M f^{up(j)}(\mathbf{x})} \\ &= \frac{\sum_{j=1}^M f^{up(j)}(\mathbf{x}) (\hat{c}_0^{up(j)} + \hat{c}_1^{up(j)} d_{m(i)}^{up(j)} x_{k+1} + \hat{c}_2^{up(j)} d_{m(i)}^{up(j)} x_{k+2} + \dots + \hat{c}_p^{up(j)} d_{m(i)}^{up(j)} x_{k+p})}{\sum_{j=1}^M f^{up(j)}(\mathbf{x})} \end{aligned} \quad (18)$$

$$\begin{aligned} y_{TSK,1}^{r(up)}(\mathbf{x}) &= \frac{\sum_{j=1}^M \bar{f}^{up(j)}(\mathbf{x}) y^{up(j)}(\mathbf{x})}{\sum_{j=1}^M \bar{f}^{up(j)}(\mathbf{x})} \\ &= \frac{\sum_{j=1}^M \bar{f}^{up(j)}(\mathbf{x}) (\hat{c}_0^{up(j)} + \hat{c}_1^{up(j)} d_{m(i)}^{up(j)} x_{k+1} + \hat{c}_2^{up(j)} d_{m(i)}^{up(j)} x_{k+2} + \dots + \hat{c}_p^{up(j)} d_{m(i)}^{up(j)} x_{k+p})}{\sum_{j=1}^M \bar{f}^{up(j)}(\mathbf{x})} \end{aligned} \quad (19)$$

여기서, y^l 은 하위 소속함수에 의해 연산되는 하한 출력을 의미하고, y^r 은 상위 소속함수에 의해 연산되는 상한 출력을 의미하게 된다. 또한, \mathbf{x} 는 예측하기 위한 입력 데이터 쌍으로 $[d_{m(i)} x_{k+1}, d_{m(i)} x_{k+2}, \dots, d_{m(i)} x_{k+p}]$ 을 의미한다. 이렇게 주어지는 상한 출력과 하한 출력에서 Type-2 퍼지논리의 최적의 출력을 얻는 방법 중 하나는 KM(Karnik-Mendel) 알고리즘[13] 통해 최소 가중 무게중심 출력 y^L 과 최대 가중 무게 중심 출력 y^R 을 얻어 이들의 평균을 취하는 방법으로, 제안된 시스템의 각 예측기 출력 y^L, y^R 은 다음과 같이 정의될 수 있다.

$$y_{TSK,1}^{L(up)}(\mathbf{x}) = \frac{\sum_{j=1}^L (y_{TSK,1}^{r(up)}(\mathbf{x}))^j + \sum_{j=L+1}^M (y_{TSK,1}^{l(up)}(\mathbf{x}))^j}{\sum_{j=1}^L \bar{f}^{up(j)}(\mathbf{x}) + \sum_{j=L+1}^M f^{up(j)}(\mathbf{x})} \quad (20)$$

$$= \frac{\sum_{j=1}^L \bar{f}^{up(j)} (y_{TSK,1}^{r(up)}(\mathbf{x}))^j + \sum_{j=L+1}^M f^{up(j)} (y_{TSK,1}^{l(up)}(\mathbf{x}))^j}{\sum_{j=1}^L \bar{f}^{up(j)}(\mathbf{x}) + \sum_{j=L+1}^M f^{up(j)}(\mathbf{x})}$$

$$y_{TSK,1}^{R(up)}(\mathbf{x}) = \frac{\sum_{j=1}^R (y_{TSK,1}^{l(up)}(\mathbf{x}))^j + \sum_{j=R+1}^M (y_{TSK,1}^{r(up)}(\mathbf{x}))^j}{\sum_{j=1}^R f^{up(j)}(\mathbf{x}) + \sum_{j=R+1}^M \bar{f}^{up(j)}(\mathbf{x})} \quad (21)$$

$$= \frac{\sum_{j=1}^R f^{up(j)} (y_{TSK,1}^{l(up)}(\mathbf{x}))^j + \sum_{j=R+1}^M \bar{f}^{up(j)} (y_{TSK,1}^{r(up)}(\mathbf{x}))^j}{\sum_{j=1}^R f^{up(j)}(\mathbf{x}) + \sum_{j=R+1}^M \bar{f}^{up(j)}(\mathbf{x})}$$

여기서, L 과 R 은 무게 중심이 변동하는 switching point를 의미하며, KM알고리즘에 의해 쉽게 찾을 수 있게 된다.

따라서 결국 제안된 Type-2 퍼지 예측기의 출력은 다음과 같이 얻을 수 있게 된다.

$$\hat{y}(t + \Delta) = x(t) + Y_{TSK,2}(\mathbf{x}) = x(t) + \frac{y_{TSK,1}^L(\mathbf{x}) + y_{TSK,1}^R(\mathbf{x})}{2} \quad (22)$$

여기서, \hat{y} 은 예측기 출력을 의미하고, $x(t)$ 는 현재의 데이터 값을 의미한다. 또한, $Y_{TSK,2}(\mathbf{x})$ 는 차분 입력에 대한 차분 예측값을 의미한다. 이는 본 논문이 입력데이터로 차분데이터를 사용하기 때문이다. 따라서 식 (22)와 같이, 예측 값은 현재값에 대한 차분 예측값의 증가분으로 주어질 수 있게 된다. 마지막으로 Δ 는 전망 예측 스텝을 의미한다.

이러한 과정을 통해 $5 + \alpha$ 개의 예측기의 설계가 끝나면 그들은 병렬구조에서 제안된 하나의 예측 시스템을 구성하게 된다.

5. 제안된 시스템의 출력

제안된 예측 시스템은 $5+\alpha$ 개의 독립된 예측기들의 병렬 구조로 이루어져 있다. 이러한 병렬구조의 시스템의 출력을 얻기 위해서 다양한 방법이 고려될 수 있지만, 본 논문에서는 훈련과정에서 보여준 각각의 예측기들의 성능을 평가한 후 이들 중 가장 우수한 예측기가 최종 예측을 수행할 수 있도록 하였으며, 예측 선택자에서는 성능지표를 이용하여 시스템 전체의 출력을 담당할 최종 예측기를 선택하는 역할을 수행하게 된다. 이는 예측시스템의 구현이 수집된 과거의 데이터들의 패턴이나 법칙성과 같은 속성들이 미래에도 이와 유사한 형태를 보일 것이라는 전제로 구현되기 때문에, 훈련과정에서 우수한 성능을 보여준 예측기는 실제 예측구간에서도 계열의 특성에 의해 그 성능을 유지할 수 있기 때문이다. 예측기 선택과정에서의 성능지표로 다음과 같은 MSE(mean squared error)를 사용하였으며, 이를 최소화하는 예측기가 최종 예측기로 선택되게 된다.

$$MSE = \frac{1}{N-m(i)-p-1} \sum_{n=m(i)+p+1}^N (y(n) - \hat{y}(n))^2 \quad (23)$$

여기서, $y(n)$ 은 실제 데이터 값을 의미하고 $\hat{y}(n)$ 은 예측기의 출력을 의미한다. 또한, $N-m(i)-p-1$ 은 각각의 차분 간격과 입력데이터의 개수에 따라 각각의 예측기에서 예측이 수행된 데이터의 길이를 의미한다. 이러한 방법은 예측이 이루어진 데이터들에 대해서만 정확히 성능검증을 함으로써 예측기 선택 시 정밀성을 높일 수 있게 된다. 결국 이러한 예측기 선택구조는 실제 예측 시 모든 예측기의 구동으로 초래될 수 있는 연산량의 부담을 경감할 수 있게 된다.

6. 시뮬레이션

본 논문에서는 제안된 예측시스템의 성능 검증을 위해 비정상 시계열과 비선형 시계열 데이터를 이용하였으며, 각각의 시뮬레이션의 수행을 위해 분할되는 훈련데이터의 길이와 예측 데이터의 길이는 성능비교를 위해 인용된 논문들에서 사용한 길이와 동일한 길이로 사용하여 효과적인 비교가 가능하도록 하였다. 표 1은 시뮬레이션을 위해 사용된 제안된 시스템의 변수들과 데이터의 변수들을 보여준다.

표 1 데이터와 시스템 변수
Table 1 Parameters of the system

지표 데이터	전체 길이	훈련 길이	예측 길이	시스템 변수(공통)	
				변수 명	변수 값
호주 전력생산량	155	70	85	예측 스텝(Δ)	1
				입력 수(p)	3
				상위 클러스터 수(up)	3
Mackey Glass 데이터	1000	500	500	퍼지 집합 수(\hat{F}_i)	2

6.1 호주 전력생산량 데이터 예측

제안된 시스템의 성능 검증을 위해 첫 번째 사용된 시뮬

레이션 데이터는 호주의 분기별 전력생산량 데이터[14]이며 성능평가를 위한 성능 지수로는 다음과 같은 MRE(mean relative error)를 사용한다.

$$MRE = \frac{1}{N} \sum_{i=1}^N \left(\frac{|y(i) - \hat{y}(i)|}{|y(i)|} \right) \times 100\% \quad (24)$$

여기서, N 은 예측데이터의 길이를 의미하며, $y(i)$ 는 현재 값, $\hat{y}(i)$ 는 예측 값을 의미한다.

아래의 표 2는 제안된 시스템의 예측기 구현 시 TITSK 퍼지모델을 사용한 경우와 IT2TSK 퍼지 모델을 사용한 경우에 대한 성능 비교를 보여준다.

표 2 예측시스템 사이의 성능 비교
Table 2 Comparison results between the systems

시스템	지표		생성 규칙 수
	훈련데이터	예측데이터	
TITSK 예측시스템	MSE 2.1027E+4	MSE 3.1165E+4	MRE 1.4883
IT2TSK 예측시스템	MSE 2.0671E+4	MSE 3.0929E+4	MRE 1.4828

표 2를 살펴보면, TITSK 퍼지모델을 적용한 예측시스템의 성능보다 IT2TSK 퍼지모델을 적용한 예측시스템의 성능이 훈련구간이나 예측구간 모두에서 그 성능이 향상되었음을 알 수 있다. 이는 Interval Type-2 퍼지 집합이 데이터의 이면에 내재된 불확실성을 잘 반영한 결과로 볼 수 있다. 또한, 시스템의 규칙기반을 형성하는 퍼지 규칙의 수가 14개로 상당히 적은 수의 퍼지규칙으로도 비교적 좋은 예측을 수행하였음을 알 수 있다. 이것은 계층구조 클러스터링 기법의 상관성과 통계적 특성을 동시에 고려하는 분류 구조가 생성되는 규칙의 적합성을 높이고 그에 따른 파라미터 추정 부정확성을 줄임으로써 생성된 퍼지규칙이 데이터들의 특성을 잘 기술한 결과로 볼 수 있다. 아래의 그림 3은 IT2TSK 예측시스템의 최종 예측결과를 보여준다.

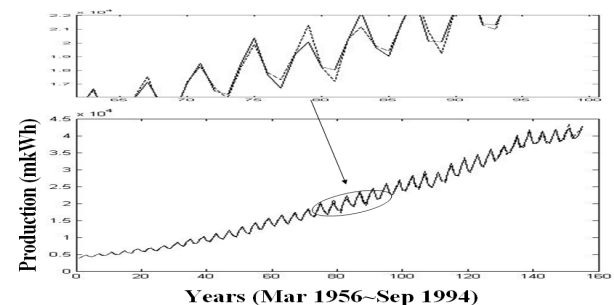


그림 3 호주 전력생산량 데이터의 예측 결과
Fig. 3 Prediction result of the Electricity Production

그림 3에서 실선은 실제 데이터 값을 의미하고 점선은 시스템의 예측 값을 의미한다. 전체 예측된 결과를 살펴보면 제안된 시스템의 예측 값들이 실제 데이터에 대하여 전 구간에서 거의 같은 모양으로 예측되고 있으며, 확대된 그림상에서도 시스템이 상당히 근사적으로 실제 값을 예측함을 알

수 있다. 이러한 결과는 제안된 예측시스템의 성능이 우수함을 보여주는 것이다. 아래의 표 3은 제안된 시스템의 예측성능과 다른 시스템들[9,15]의 예측성능을 비교한 것이다.

표 3 제안된 시스템과 다른 시스템들 사이의 성능비교
Table 3 Comparison results between our system and other systems

system index	Mamdani 모델	Fuzzy AR	다중 퍼지모델	GA-RS	our system	
					T1TSK	IT2TSK
MRE	7.8123	3.1254	2.7125	1.8100	1.4833	1.4828

표 3을 살펴보면 제안된 시스템의 예측성능이 비교된 다른 시스템들의 예측성능보다 우수함을 알 수 있다. 또한 제안된 시스템의 설계 기법에 있어 Interval Type-2 퍼지논리를 적용한 시스템의 예측성능이 데이터의 불확실성을 보다 잘 취급할 수 있어 Type-1 퍼지논리를 적용한 시스템 보다 우수하게 나타났음을 알 수 있다.

6.2 Mackey-Glass Time series 예측

제안된 시스템의 성능 검증을 위한 두 번째 시뮬레이션 데이터는 Mackey-Glass 시계열 데이터로 비선형 시스템의 모델링이나 검증을 위해 매우 빈번히 사용되며 다음과 같은 식에 의해 발생된다.

$$\frac{dx(t)}{dt} = \frac{\alpha x(t-\tau)}{1+x^{10}(t-\tau)} - \beta x(t) \quad (25)$$

여기서, α, β, τ 는 인용된 논문에서 사용한 동일하게 각각 0.2, 0.1, 17로 정의 되며, 식에 의해 생성되는 데이터들 중 $x(124)$ 에서부터 $x(1123)$ 까지 1000개의 데이터를 추출하여 시뮬레이션에 사용한다. 또한 시스템의 성능 검증을 위한 성능지표로는 다음과 같이 정의되는 RMSE(root mean square error)를 사용한다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y(i) - \hat{y}(i))^2} \quad (26)$$

아래의 표 4는 T1TSK 퍼지논리와 IT2TSK 퍼지논리를 각각 제안된 시스템에 적용하였을 경우의 성능 비교를 보여준다.

표 4 예측시스템 사이의 성능 비교
Table 4 Comparison results between the systems

시스템	지표	예측데이터			생성 규칙 수
		MSE	MSE	RMSE	
T1TSK 예측시스템		6.7149E-7	6.6812E-7	8.1738E-4	24
IT2TSK 예측시스템		6.3428E-7	6.1818E-7	7.8624E-4	

표 4의 예측성능을 살펴보면 제안된 예측시스템에 T1TSK 퍼지논리나 IT2TSK 퍼지논리를 적용하였을 경우, 24개의 시스템 규칙이 생성되었으며 이를 통해 두 시스템 모두 비교적 좋은 성능을 나타냄을 알 수 있다. 또한, IT2TSK 퍼지논리를 적용한 제안된 예측시스템의 성능이 Type-2 퍼지논리의 불확실성 취급 능력에 의해 Type-1 퍼지논리를 사용한 경우 보다 그 예측성능이 우수함을 알 수 있다. 그림 4는 제안된 예측시스템의 최종 예측 결과를 보여준다.

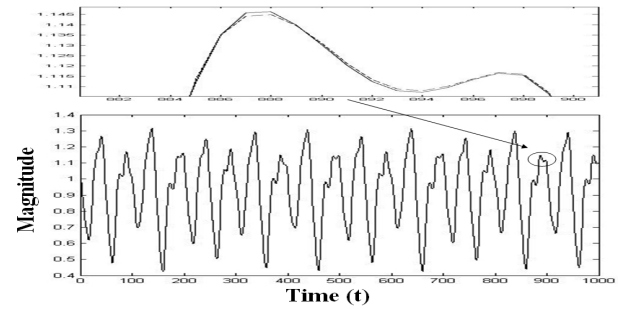


그림 4 Mackey-Glass 시계열 데이터의 예측 결과
Fig. 4 Prediction result of the Mackey-Glass time series

그림 4도 역시 실선은 실제 데이터 값을 의미하고 점선은 시스템의 예측 값을 의미한다. 그림을 살펴보면 예측 값이 실제 데이터 값을 거의 모든 구간에서 덮고 있음을 알 수 있으며, 이는 제안된 시스템의 성능이 매우 우수함을 보여주는 것이라 할 수 있다. 아래의 표는 제안된 예측시스템과 다른 시스템들[6-7] 사이의 성능을 비교한 것이다.

표 5 제안된 시스템과 다른 시스템들 사이의 성능비교
Table 5 Comparison results between our system and other systems

system index	ANNs	ARIMA-ANN	ANFIS	ANFIS2	our system	
					T1TSK	IT2TSK
RMSE	0.0122	0.0027	2.7125	0.0012	8.1738E-4	7.8624E-4

표 5를 살펴보면 제안된 예측시스템의 성능이 T1TSK 퍼지논리를 적용한 경우나 IT2TSK 퍼지논리를 적용한 경우 모두에서 다른 시스템들 보다 우수한 성능을 보여준다. 결국, 시뮬레이션에서 나타난 결과들은 제안된 시스템의 설계에 사용된 차분 검출방법이나 다중 예측기들로 구성된 시스템의 구조, 각각의 예측기 설계를 위해 사용된 기법들이 매우 효과적임을 보여주는 것으로 간주할 수 있다.

7. 결 론

본 논문에서는 시계열 데이터의 예측을 위한 퍼지 예측시스템의 설계를 다루었다. 자연현상에 기인하는 강한 비선형적 특성을 나타내는 시계열 데이터를 효과적으로 분석하고, 분석된 정보를 효율적으로 이용함으로써 보다 강인한 예측시스템이 구현될 수 있도록 하였다. 이를 위해 1차 차분 검출과정을 설계함으로써 안정된 정보를 기반으로 각각의

예측기들이 설계될 수 있도록 하였으며, 다수의 예측기들로 하나의 시스템을 구현함으로써 계열 내에 숨겨진 다양한 특성들을 고려할 수 있도록 하였다. 또한, 계층구조 클러스터링과 IT2TSK 퍼지논리에 기반하여 예측기를 설계함으로써, 적은 수의 규칙으로도 시스템의 기술능력을 개선함과 동시에 불확실한 정보의 취급에 효과적으로 적용할 수 있도록 하였다. 마지막으로 다수의 예측기들 중 모델 선택자의 설계를 통해 하나의 예측기만을 최종적으로 구동함으로써, 연산의 부담이나 시스템 운영의 복잡함을 완화하도록 하여 효과적인 예측이 수행될 수 있도록 하였다. 비정상 시계열 데이터와 비선형 데이터를 이용한 시뮬레이션 결과는 제안된 시스템의 성능이 우수함을 보여 주었으며, 이는 제안된 시스템의 설계과정에 사용된 기법들이 비정상 또는 비선형 데이터를 취급하는 시스템의 설계를 위해 효과적일 수 있음을 보여주는 것으로 간주 할 수 있다. 따라서 제안된 시스템의 설계 기법은 예측 시스템의 설계뿐만 아니라 비선형 제어 시스템의 설계 등 복잡한 구조의 데이터들을 취급해야 하는 분야에서 효과적으로 사용될 수 있을 것으로 기대된다. 향후에는 데이터들의 특성 반영뿐만 아니라 예측기들의 예측 특성들도 충분히 반영할 수 있는 효과적인 시스템 운영방법에 관한 연구가 필요할 것으로 사료된다.

참 고 문 헌

- [1] G. Box, G. M. Jenkins, G. C. Reinsel, Time Series Analysis : Forecasting and Control, third edition, Prentice-Hall, 1994. Design Automation Conf., pp. 253-259, 1992.
- [2] 김해경, 김태수, 시계열 분석과 예측 이론, 교우사, 2003. Systems, pp. 446-455, Dec. 1994.
- [3] D. M. Kim, I. S. Bae, J. M. Cho, J. O. Kim, K. C. Chang, "Prediction of Dynamic Line Rating Based on Thermal Risk Probability by Time Series Weather Models", Trans. KIEE, vol. 55A, no. 7, pp. 273-280, 2006.
- [4] G. Janazcek, L. Swift, Time Series Forecasting, Simulation, Application, Ellis Horwood, 1993.
- [5] T. Senjyu, H. Takara, K. Uezato, T. Funabashi, "One-Hour-Ahead Load Forecasting Using Neural Network", IEEE Trans Power Syst, vol. 17, pp. 113-118, 2002.
- [6] O. Valenzuela, I. Rojas, F. Rojas, H. Pomares, L. J. Herrera, A. Guillen, L. Marquez, M. Pasadas, "Hybridization of intelligents and ARIMA models for time series prediction", Fuzzy Sets and Systems, vol. 159, pp. 821-845, 2008.
- [7] M. EI-Koujok, R. Gouriveau, N. Zerhouni, "Towards a Neuro-Fuzzy System for Time Series Forecasting in Maintenance Applications", 17th Triennial Word Congress of the International Federation of Automatic Control, hal-00298361, version 1, 2008.
- [8] Y. K. Bang, C. H. Lee, "Design of Multiple Model Fuzzy Predictors using Data Preprocessing and its Application", Trans. KIEE, vol. 58, pp. 173-180, 2009.
- [9] Y. K. Bang, C. H. Lee, "Design of Fuzzy System with Hierarchical Classifying Structure and its Application to Time Series Prediction", Journal of Korean Institute of Intelligent Systems, vol. 19, pp. 595-602, 2009.
- [10] L. A. Zadeh, "A computational approach to fuzzy quantifiers in natural languages", Comput. Math, vol. 9, pp. 149-184, 1983.
- [11] J. M. Mendel, R. I. John, "Type-2 Fuzzy Sets Made Simple", IEEE Trans. on Fuzzy Systems, vol. 10, pp. 117-127, 2002.
- [12] J. M. Mendel, R. I. John, F. Lui, "Interval Type-2 Fuzzy Logic Systems Made Simple", IEEE Trans. on Fuzzy Systems, vol. 14, pp. 808-821, 2006.
- [13] N. N. Karnik, J. M. Mendel, "Centroid of a Type-2 Fuzzy Set", Information Sciences, vol. 132, pp. 195-200, 2001.
- [14] <http://www.personal.buseco.monash.edu.au/hyndman>
- [15] Y. S. Joo, Fuzzy System Modeling Using Genetic Algorithm and Rough Set Theory, M. S. Thesis, Dept. of Electrical and Electronic Eng, Kangwon Univ, Korea, 2003.

저 자 소 개



방 영 근 (方寧根)

1974년 4월 7일

2000년 삼척대학교 전자공학과(학사)

2003년 삼척대학교 전기공학과(석사)

2010년 강원대학교 전기공학과(박사)

E-mail : b2y2c1@hanmail.net



이 철 희 (李哲熙)

1959년 7월 4일

1983년 서울대학교 전기공학과(학사)

1985년 서울대학교 전기공학과(석사)

1989년 서울대학교 전기공학과(박사)

1990년~현재 : 강원대 전기전자공학부 교수

E-mail : chlee@kangwon.ac.kr