

시계열 예측을 위한 구간 연산 신경망 모델

(An Interval Arithmetic Neural Network Model for Time Series Prediction)

김 호 준 [†] 김 우 성 ^{**}
(Ho-joon Kim) (Woo-sung Kim)

요 약 본 연구에서는 효과적인 시계열 예측을 위하여 구간연산 기능을 갖는 신경망모델을 제안한다. 이는 기존의 FIR 네트워크의 동작특성을 구간연산으로 일반화함으로써 시계열 신호의 동적 특성을 효과적으로 반영할 뿐만 아니라 학습데이터 표현의 유연성을 증대시킨다. 이는 또한 실제 응용에서 신경망 입력 데이터에 내재할 수 있는 측정치 오류의 영향을 보완할 수 있게 하며 데이터 그룹화를 효과적으로 이룰 수 있게 함으로써 입력데이터의 양을 감축시키고 이로부터 학습의 효율을 개선한다. 본 논문에서는 구간연산을 이용하는 네트워크의 동작특성, 학습알고리즘을 제시하고 실제 응용시스템에 이를 적용함으로써 제안된 이론의 유용성을 평가한다.

Abstract We propose an interval arithmetic neural network model for effective prediction of time series data in this paper. The model effectively reflects the dynamic characteristics of time series signals by generalizing the activation of the FIR network for interval arithmetic operations. The proposed model exhibits several advantageous characteristics. From the interval arithmetic functions, more flexible representations can be provided for the learning data. The model also can reduce the effect of noise and increase the learning efficiency in practical applications. In this paper, we first describe the network structure and behavior, and then introduce a learning algorithm using interval arithmetic operations. To evaluate the effectiveness of the proposed model, we applied the model to a weather forecasting system. From the experimental results of the long term forecast of Korean regional precipitation and temperature, the validity of the proposed model is discussed empirically.

1. 서 론

시계열패턴 예측문제는 기상정보 예측이나 교통량 예측시스템 등 실생활의 다양한 분야에 응용될 수 있는 연구이다. 그런데 대부분의 실제 문제에서 적용되는 자료의 형태는 그 종류가 매우 다양할 뿐만 아니라 양적으로도 방대하여 효율적인 시스템 구축을 어렵게 한다. 최근의 연구에서 이러한 시계열 예측문제를 위한 신경망 구조 및 학습방법에 대한 연구가 다양하게 시도되고 있다[1-12]. 기존의 패턴인식이나 분류 등의 문제에 사

용되는 일반적인 신경망 모델의 구조적 특성은 시간에 따라 변하는 각종 패턴의 예측문제에 직접적으로 적용하기에 적합하지가 않다. 이에 입력신호의 동적 특성을 효과적으로 반영하도록 하기 위하여 네트워크 구조상의 피드백 연결이나 시간지연 효과를 반영하는 별도의 노드구조를 두는 연구가 시도되었는데 대표적인 예로 TDNN(Time Delayed Neural Network), 순환구조 신경망(recurrent neural networks) 및 FIR(Finite Impulse Response) 네트워크 등을 들 수 있다[1-6].

또 다른 연구로 Kehagias 등은 하위계층에 선형 및 비선형 예측모듈을 구성하고 이들이 상위계층의 의사결정 모듈에서 상호 결합되는 형태의 이른바 모듈구조의 시계열 패턴분류 신경망 구조를 제안하였으며[7] 유사한 연구로 Weigend 등은 모듈간의 예측결과를 융합하는 과정에서 별도의 학습을 거치는 형태의 고유한 네트워크 구조를 제시한 바 있다[8].

· 본 연구는 첨단정보기술연구센터를 통하여 과학계단의 지원을 받았음

[†] 중신회원 : 한동대학교 전산전자공학부 교수
hjkim@han.ac.kr

^{**} 비 회 원 : 한국과학기술원 전산학과
rigel@paradise.kaist.ac.kr

논문접수 : 2000년 2월 17일

심사완료 : 2000년 7월 25일

이처럼 여러 신경망 모델과 예측방법론이 제시되었음에도 불구하고, 실제 문제에서 계측장비특성으로 인한 관측데이터의 오류, 데이터 양의 방대함으로부터 오는 계산량의 증대, 노이즈 및 누락데이터의 영향 등이 성능저하의 요인으로 지적된다.

본 논문에서는 기존의 전형적인 시계열예측 신경망인 FIR 네트워크[4]의 특성을 확장하여 상기의 문제들을 해결 및 보완할 수 있는 가능성을 제시한다. FIR 네트워크는 신경망의 시냅스를 FIR 선형 필터로 모델링한 신경망으로 시계열 예측문제에 전형적으로 사용되는 모델이다. 본 연구에서는 기존의 FIR 네트워크가 갖는 시간 지연효과의 구조와 더불어 노드의 활성화 특성에 구간 연산기능을 수용하는 모델을 제안하며 이에 따른 학습알고리즘을 제시한다.

일반적인 시계열 예측 신경망에서 개선된 예측성능을 얻기 위해서는 학습 및 인식에 사용되는 자료의 종류와 범위를 확대할 필요가 있다. 그런데 그러한 경우 데이터의 양이 방대해짐으로써 신경망의 규모와 계산량의 증대, 불필요한 노이즈의 영향 등으로 실용적 측면에서 오히려 성능을 저하시킬 우려와 과도학습(over-training)의 가능성을 내재한다. 본 연구에서는 방대한 양의 학습데이터에 대하여 예측에 필요한 정보를 최대한 유지하는 전제에서 데이터의 양을 감축시키는 방법론으로서 구간형태의 데이터 표현 및 연산기능을 고려한다. 즉 데이터를 일정 시간 범위로 그룹화하고 각 그룹에 대한 대표값을 구간형태로 표현하는 방안을 채택하면 신경망에 적용되는 특징의 수를 줄일 수 있다. 이는 또한 일부 응용에서 미세한 시간단위로 변화하는 신호의 형태에 지나치게 민감하게 학습되는 현상을 보완 할 수 있게 할 뿐 만 아니라 예측 결과값도 구간형태로 표현할 수 있게 함으로써 시스템 구현 시 사용자에게 제공되는 정보표현의 유연성을 증대시키는 등 부수적인 효과를 기대할 수 있다.

이러한 배경에서 본 논문에서는 구간형태의 자료에 대한 연산기능을 지원하는 시계열예측 신경망 모델을 제안하고 그 유용성을 고찰한다. 제안된 신경망모델은 기존의 FIR네트워크나 TDNN이 갖는 시간지연 효과의 이점을 그대로 반영할 뿐만 아니라 신경망의 동작특성 및 학습과정을 구간연산으로 일반화하여 상기의 유용한 특성을 지원하게 한다.

본 논문에서는 이러한 신경망 모델의 구조와 학습알고리즘을 제시하고, 이를 실제 시계열 예측 응용문제에 적용한 실험결과를 통하여 제시된 이론의 타당성을 고찰한다.

2. 시계열 예측을 위한 신경망 기법

기존의 많은 예측이론과 비교할 때, 신경망을 사용한 예측기의 특성은 예제패턴에 의한 학습기능을 통하여 스스로 지식을 획득한다는 점과 변화하는 환경에 대한 적응능력 등이 장점으로 지적될 수 있으며, 이는 실제세계의 각종 응용시스템 구축을 위한 기반기술로써 효과적으로 활용될 수 있을 것으로 기대된다. 신경망이론에서 시간에 따라 변화하는 사상을 실현하는 문제는 일반적으로 순차적 학습문제로 다루어 진다. 이는 전형적인 신경망 응용 문제인 패턴인식이나 분류 문제와는 달리 시계열 데이터를 처리하기 위한 방법론이 부수적으로 고려되어야 한다.

FIR 네트워크는 시계열데이터 예측에 쓰이는 전형적인 신경망모델이다. 이러한 시계열 데이터 예측의 특성은 어떤 유한기간의 입력을 통하여 유한기간의 출력을 결정하게 되고, 시간의 흐름에 따라 모든 입력 데이터들이 주어진 일정 시간주기로 이동되어지는 특성이 있다. 가장 일반적인 신경망모델인 다층퍼셉트론을 시계열예측에 직접 적용한다면 입력신호의 집합에서 출력신호로의 복잡한 형태의 매핑을 찾아내는 과정으로 구현될 수 있지만, 이는 신호의 시간차원을 고려하지 않은 뉴런들의 정적인 매핑이다. 반면에 FIR 네트워크는 뉴런 내부에 동적인 특성을 추가하여 시계열 데이터의 고유특성을 효과적으로 반영할 수 있도록 한다.

FIR네트워크의 구조는 기존의 다층퍼셉트론에서 각각의 뉴런들에 대하여 상위층으로의 연결형태를 그림1과 같이 FIR 필터로 확장한 구조이다. 출력층을 제외한 임의의 뉴런 X 는 $x(k)$ 에서부터 $x(k-T)$ 까지 일정 개수의 연속시점을 반영하는 노드들과 그에 해당하는 가중치의 집합 W 로 구성되어 있고 이들 연산에 의해 뉴런의 출력값을 생성하게 된다. 이 출력값은 같은 층의 다른 뉴런의 출력값 및 바이어스(bias)노드 값과 결합된 형태로 활성화함수를 거쳐 다음 층의 해당시점 입력값으로 전달되게 된다. 동일한 방법으로 전방향 활성화과

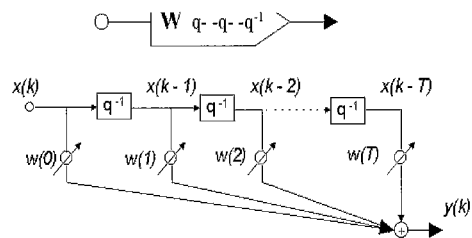


그림 1 FIR 네트워크의 필터 구조

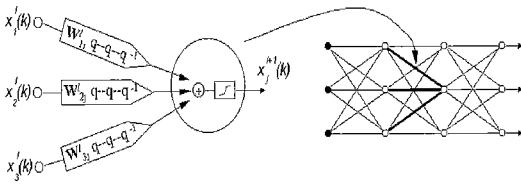


그림 2 FIR 네트워크의 구조

정을 거쳐 출력층 노드값이 생성되게 되는데 이러한 과정을 그림 2에 나타내었다. 다시 말해서 신경망의 각 노드는 주어진 시점의 출력값을 생성한 후에 다음 시점의 연산을 위해 모든 노드의 값이 한 단계씩 이동하여 일정기간동안의 시간지연을 반영하도록 한다. 이것은 입력의 시계열 데이터가 시간이 지남에 따라 일정주기로 한 단계씩 이동되는 것 뿐만 아니라, 그림 2와 같이 중간층의 모든 뉴런의 값도 같은 방식으로 이동됨을 의미한다. 이러한 네트워크 내부구조에 동적인 특성을 갖는 것이 다층퍼셉트론에 비교한 FIR 네트워크 모델의 고유한 특성이다. 또한 FIR 네트워크의 각 입력뉴런에는 서로 다른 종류의 시계열 데이터를 적용시킬 수 있음으로써 상호 연관된 다양한 데이터를 사용하는 예측이 가능하다. 그러나 이 경우 학습데이터 증가로 인한 계산량증대가 문제로 지적된다. 따라서 다음절에서 이를 완화시킬 수 있는 신경망 모델을 제시한다.

3. 시계열 예측 문제에서 구간 연산의 유용성

본 연구의 핵심은 시계열 예측신경망의 동작을 구간 연산으로 일반화하는데 있다. 기존의 다층퍼셉트론이나 FIR 네트워크의 동작특성은 덧셈, 곱셈 및 지수함수 등의 조합으로 이루어지는데 이러한 기본 산술연산은 구간연산으로 일반화가 가능하다. 이러한 경우 신경망의 특성은 다음과 같이 몇 가지 이점을 제공한다.

우선 신경망의 입력 및 출력값에 구간형태의 데이터 표현을 직접적으로 적용하는 학습과 예측이 가능하다는 점이다. 신경망예측기를 구현할 때 적용하는 특징의 종류와 값의 표현 형태는 예측성능에 지대한 영향을 미친다. 물론 일부 문제에서 신경망이론의 학습능력에 의하여 임의로 구성된 대량의 특징 집합으로부터 문제에 필요한 특징의 종류와 그들과 출력 값과의 관계를 스스로 찾아내기를 기대할 수는 있다. 그러나 아무런 사전 지식 없이 적용하는 문제에서, 데이터의 종류가 매우 다양하고 그들 중 상당수가 노이즈를 포함하고 있다면 이로부터 유용한 지식을 신경망 스스로 획득하기가 쉽지 않다.

반면에 해당 분야 전문가의 이론적 지식 또는 경험적 지식의 표현에서 사용되는 특징들은 일반적으로 예측결과를 좌우하는 중요한 요인이 된다. 따라서 이들을 신경망학습을 위한 특징집합에 포함시킬 수 있다면 보다 향상된 성능을 얻을 수 있을 것이다. 그런데 많은 응용문제에서 전문가의 지식표현에 사용되는 특징의 형태가 구간형태의 표현을 사용한다는 것이다. 이에 구간형태의 데이터 표현을 직접 수용하여 연산할 수 있는 신경망이 구현된다면 보다 효과적인 성능을 기대할 수 있을 것이다.

또 다른 유용성으로 특징수와 계산량의 감축효과를 들 수 있다. 실제 문제에서 성능을 개선하기 위하여 학습에 사용되는 데이터의 종류와 시계열 표현에서 적용하는 시간범위를 확대하면 상대적으로 데이터의 양이 방대해지고 신경망의 규모와 계산량이 증가하는 문제가 발생한다. 이는 실용적 측면에서 시스템 구현상의 단점으로 지적될 수 있다. 이에 대한 대안으로 입력데이터를 공간적 및 시간적으로 분할, 그룹화하고, 그룹내에서 상한값과 하한값의 범위를 갖는 구간형태의 데이터로 표현함으로써 특징의 수를 감소시키면 신경망의 노드 수를 줄일 수 있고 더불어 학습의 효율을 개선할 수 있다.

이 외에 부수적인 유용성으로 시계열 예측기법과 연관된 실세계의 응용에서 데이터의 표현이 구간형태로 나타나는 경우가 빈번하다는 점을 들 수 있다. 예로서 측정장비의 특성상 상당량의 측정오차를 포함할 가능성이 있다면 일정기간 동안 반복 관측된 값의 범위를 구간형태로 표현 할 수 있다. 또한 예측 결과값의 표현도 구간형태의 범위로 표현할 수 있다는 점은 사용자에게 제공되는 정보표현의 융통성이라 할 수 있다.

4. 시계열 예측을 위한 구간 연산 신경망의 구조

구간 연산기능의 네트워크 구조 및 동작특성을 정의하기 위하여 우선 기본이 되는 구간 데이터에 대한 산술연산식을 보인다. 네트워크 구조 및 학습알고리즘에 기본이 되는 기본 산술연산은 식 (1), (2) 및 (3)에 보인 구간값의 덧셈, 곱셈 및 지수함수를 사용한다. 이 때 신경망에서 사용되는 가중치 적용 연산을 위하여 구간 값과 상수와의 곱셈을 고려하였으며, 이는 그 후호에 따라 식(2)와 같이 값이 달라진다. 각 식에서 첨자 L과 U는 각각 하한 및 상한을 의미하며 구간벡터는 대문자를 사용하고 각각의 값들은 소문자로 나타내었다.

$$X + Y = [x^L, x^U] + [y^L, y^U] = [x^L + y^L, x^U + y^U] \quad (1)$$

$$c \cdot X = c \cdot [x^L, x^U] = \begin{cases} [c \cdot x^L, c \cdot x^U], & \text{if } c \geq 0 \\ [c \cdot x^U, c \cdot x^L], & \text{if } c < 0 \end{cases} \quad (2)$$

$$\exp(X) = \exp([x^L, x^U]) = [\exp(x^L), \exp(x^U)] \quad (3)$$

또한 신경망에서 가장 일반적으로 쓰이는 활성화함수인 시그모이드함수(식 (4))는 단조 증가의 특성을 가지므로 식 (5)에 보인 것과 같이 구간연산을 적용할 수 있다.

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (4)$$

$$f(\text{Net}) = f([\text{net}^L, \text{net}^U]) = [f(\text{net}^L), f(\text{net}^U)] \quad (5)$$

본 논문에서 제안하는 신경망모델은 앞서 정의한 구간연산을 기본으로 하여 FIR 네트워크에서 뉴런의 입력과 출력을 구간의 벡터로 확장하고, 그 내부의 모든 연산을 구간연산으로 일반화한 것이다. 이를 이루는 필터의 구조는 그림 3과 같다. 우선 필터의 입력 뉴런 X는 FIR 네트워크와 마찬가지로 식 (6)과 같이 X(k)에서부터 X(k-T)까지의 노드들과 그에 해당하는 가중치벡터 W로 구성된다(식 (7)). 그러나 FIR 네트워크와 달리 각각의 입력노드는 식 (8)과 같이 상한과 하한값을 갖는 구간의 벡터로 이루어 진다. 이때 쓰인 첨자의 의미는 다음과 같다. 식(7)에서 가중치벡터 W_{ij}^l 는 l층에서 하위 계층의 뉴런 i와 상위계층의 뉴런 j에 연결된 필터의 연결강도를 가리키며 T는 과거 T개만큼의 데이터를 유지함을 의미한다.

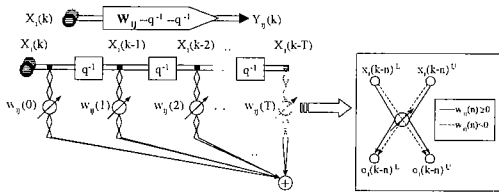


그림 3 제안된 신경망 모델의 필터 구조

FIR 네트워크와는 달리 필터의 입력은 $X_i(k)$ 구간의 벡터로 주어지며 그림 3의 왼쪽과 같이 주어진 시점의 값이 시간의 지연으로 q^{-1} 만큼씩 움직이게 되어 식 (6)과 같이 입력벡터 X(k)를 구성하게 된다. 입력구간 X(k)에 대한 필터의 출력값인 Y(k) 역시 구간의 벡터이며, 그것의 하한과 상한은 그림 3의 오른쪽에 표현한 바와 같이 해당 가중치값이 음이면 하한과 상한을 바꾸어 결과값을 얻는 구간연산을 적용하여 산출된다. 이는 수식의 (9)와 (10)에 해당하는 내용을 의미한다.

$$X_i^l = [X_i^l(k), X_i^l(k-1), \dots, X_i^l(k-T)] \quad (6)$$

$$W_{ij}^l = [w_{ij}^l(0), w_{ij}^l(1), \dots, w_{ij}^l(T)] \quad (7)$$

$$X_i^l(k) = [x_i^l(k)^L, x_i^l(k)^U] \quad (8)$$

$$y(k)^L = \sum_{i=0, w(i) \geq 0} w(i)x(k-i)^L + \sum_{i=0, w(i) < 0} w(i)x(k-i)^U \quad (9)$$

$$y(k)^U = \sum_{i=0, w(i) \geq 0} w(i)x(k-i)^U + \sum_{i=0, w(i) < 0} w(i)x(k-i)^L \quad (10)$$

위에서 나타낸 필터의 전과정을 간단히 식 (11)과 같이 입력벡터 X와 연결강도벡터 W의 연산으로 나타내었고, 이것은 그림 3의 전체 연산과정을 의미한다. 여기서 t는 벡터의 연산을 위한 전치(transpose)를 의미한다. 한편 다음 층 뉴런으로의 입력은 그림 4에 보인 것처럼 그 뉴런에 연결된 하위층에 있는 모든 필터출력들의 합을 구하고(식 (12)), 활성화함수인 시그모이드함수를 거친 후 해당 뉴런의 입력으로 들어가게 된다(식 (13)). 이 때의 연산 역시 모두 구간연산으로 식 (11), (12) 및 (13)은 하위계층에서 산출된 값이 상위계층에 반영되는 내용을 보인 것이다.

$$Y_j^{l+1}(k) = W_{ij}^l \circ (X_i^l)^t \quad (11)$$

$$Y_j^l(k) = \sum_{i=1}^{N_l-1} Y_{ij}^l(k) = \sum_{i=1}^{N_l-1} W_{ij}^{l-1} \circ (X_i^{l-1})^t \quad (12)$$

$$X_j^l(k) = f(Y_j^l(k)) \quad (13)$$

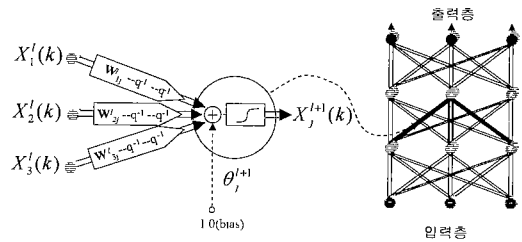


그림 4 뉴런의 연결구조 및 동작 특성

실제로 네트워크상의 각 뉴런은 그림 4에 보인 것처럼 바이어스(bias)노드를 갖고 있으나 이의 작용은 본 논문의 핵심내용을 전개하는 과정과 무관하므로 수식표현에서는 편의상 생략하였다. 이와 같이 고안된 구간연산 네트워크 모델의 이점은 앞서 언급한 바와 같이 구간연산으로의 일반화를 통한 이점과 FIR 네트워크가 갖는 동적 특성의 이점을 동시에 갖게 되는 것이다.

5. 학습 알고리즘

본 연구에서 제안한 시계열 예측을 위한 구간연산 신

경망의 학습알고리즘은 FIR 네트워크의 Temporal Backpropagation 학습알고리즘을 구간연산으로 일반화하여 개발하였다. 학습의 전체구조는 그림 5와 같이 오차값 δ 의 하한과 상한값을 갖는 구간으로 이루어진 오차벡터 Δ 가 역방향 필터를 통하여 역전파되는 형태로 이루어진다.

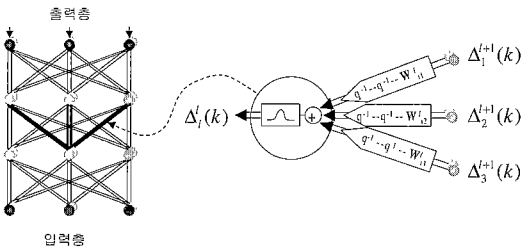


그림 5 제안된 신경망 모델의 학습과정

제안된 모델에서 임의시점의 연속된 입력 구간데이터 $X(k)$ 에 대한 기대출력 $T(k)$ 역시 구간으로 주어진다. 즉, 식 (14)에 나타낸 것과 같이 구간의 벡터이다. 또한 식 (15) 및 (16)과 같이 해당 시점에 대한 전체오차의 제곱을 $e^2(k)$ 로 나타내었고, 모든 시점(T 개의 시점)에 대한 전체오차제곱은 e^2 으로 나타내었다.

$$T_i(k) = [t_i(k)^L, t_i(k)^U] \tag{14}$$

$$e^2(k) = \frac{1}{4} \sum_{i=1}^{N_i} \{t_i(k)^L - x_i^L(k)^L\}^2 + \{t_i(k)^U - x_i^U(k)^U\}^2 \tag{15}$$

$$e^2 = \sum_{k=0}^T e^2(k) \tag{16}$$

이와 같이 학습을 위하여 학습패턴에 대한 오차식을 정의했을 때, 임의의 연결에 적용될 가중치 변화분은 식 (17)과 같이 최종시점의 전체오차제곱이 아닌 모든 시점에 대한 전체오차제곱에 대하여 해당 연결강도를 미분함으로써 얻을 수 있다. 그리고 이것은 다시 식(18)에 나타난 바와 같이 해당 가중치 값의 부호에 따라 오차값 δ 와 입력값에 비례한 식으로 나타낼 수 있다. 이때 μ 는 양의 상수로 학습률(learning rate)을 의미한다.

$$\Delta w_{ij}^t(n) = \mu \frac{\partial e^2}{\partial w_{ij}^t(n)} \tag{17}$$

$$\frac{\partial e^2}{\partial w_{ij}^t(n)} = \begin{cases} \delta_j^{t+1}(k)^L \cdot x_i^L(k-n)^L + \delta_j^{t+1}(k)^U \cdot x_i^U(k-n)^U, & \text{if } w_{ij}^t(n) \geq 0 \\ \delta_j^{t+1}(k)^L \cdot x_i^L(k-n)^U + \delta_j^{t+1}(k)^U \cdot x_i^U(k-n)^L, & \text{if } w_{ij}^t(n) < 0 \end{cases} \tag{18}$$

이때 오차값 δ 는 학습하고자 하는 가중치의 상위뉴런이 출력층인가 또는 중간층인가에 따라 내용이 달라진다. 학습하려는 연결강도의 상위뉴런이 출력층인 경우

오차값 δ 항의 하한과 상한은 각각 식(19)와 (20)으로 주어진다.

$$\delta_j^L(k)^L = \frac{1}{2} (t_j(k)^L - x_j^L(k)^L) f'(y_j^L(k)^L) \tag{19}$$

$$\delta_j^U(k)^U = \frac{1}{2} (t_j(k)^U - x_j^U(k)^U) f'(y_j^U(k)^U) \tag{20}$$

중간층에서의 오차값 항은 그림 5에 나타낸 바와 같이 구하려는 오차값의 상위층에서의 오차벡터인 Δ 가 앞 절에서 언급한 필터연산을 통하여 역전파된다. 식 (21)에 중간층의 오차값을 구하는 수식을 나타내었는데, 연결강도를 제외한 모든 값들이 구간의 벡터이다. 여기서 오차벡터 Δ 는 입력 벡터와 마찬가지로 식(22)와 같이 T 개의 시점을 유지하는 벡터로 구성되며, 이들 원소의 각각은 다시 구간의 벡터로 구성된다.

$$\delta_i^t(k) = \left\{ \sum_{j=1}^{N_{j+1}} W_{ij}^{t+1} \circ (\Delta_j^{t+1})^T \right\} \cdot f'(Y_i^t(k)) \tag{21}$$

$$\Delta_i^t = [\Delta_i^t(k), \Delta_i^t(k-1), \dots, \Delta_i^t(k-T)] \tag{22}$$

위와 같은 과정으로 학습이 이루어진다. 신경망의 계층구조와 노드수 등의 매개변수 값은 응용의 특성에 따라 선택되며, 학습률 및 초기가중치 등은 반복 학습과정 중에 최적의 값으로 조정될 수 있다.

6. 응용 시스템 구현

6.1 지역별 중장기 강수량 예측 문제

제안된 모델의 유용성을 평가하기 위하여 한반도 지역별 중장기 강수량 예측문제에 적용하고 실험결과를 분석하였다. 강수량예측을 위한 신경망예측기의 입력자료로 사용되는 기상데이터의 종류는 매우 다양한 형태로 나타난다. 또한 관측된 자료는 종류별로 각각 지역별, 시간대별로 분류될 수 있다. 이와 같이 형태와 범위가 다양한 기상데이터를 신경망예측기의 입력으로 사용하기 위해서는 적절한 전처리과정이 필요하게 된다. 일차 실험으로 한반도 각 지역의 월별, 계절별 강수량 예측문제를 대상으로 하였다. 신경망예측기의 학습과정은 우선 지역별로 관측된 과거 강수량데이터를 고려하였다. 방대한 양의 전체데이터로부터 기준 지역과 인접지역을 중첩하여 포함하는 분할의 집합을 생성하였고 반복 실험을 통하여 최적의 분할구조를 설정하였다. 생성된 각 분할에 대하여 최대값 및 최소값을 정규화하여 구간형태의 데이터를 생성하였다. 강수량데이터와 더불어 월평균 기온데이터를 학습자료로 사용하였다. 이는 평균기온이 강수량에 영향을 미칠 수 있다고 알려져 있는 점에서 그 유용성이 인정된다. 이 외에 태평양의 해수표면은

도(SST: Sea Surface Temperature)를 추가로 고려하였는데, 이는 그 양이 방대함으로 인해 시간적, 공간적 그룹화를 통하여 특정 수를 감소시켰으며 이도 역시 구간형태의 데이터로 생성하였다.

학습자료로 선정된 원시기상데이터는 대상 신경망모델의 특성에 적합한 형태로 정규화 되어야 한다. 본 연구에서는 0과 1사이의 값을 수용하는 신경망을 고려하였으며, 각 유형의 기상데이터에 대하여 일정기간 관측값의 최대 및 최소치의 범위 내에서 선형매핑하는 방식으로 정규화 하였다. 이들은 대상기간이 달라질 때마다 값이 재적용되며 이들 값은 각 학습과정 및 예측단계에서 동적으로 변화하여 적용된다. 또한 신경망의 출력값으로부터 예측강수량 값의 산출은 이러한 과정을 역으로 적용하여 산출하였다.

학습과정에 사용되는 데이터의 수는 네트워크의 구조와 밀접한 관계를 갖는다. 예를 들면 출력노드의 값이 다음 단계의 연산에 활용되기 위해 저장되는 버퍼의 크기나 입력방식의 설정 및 과거 데이터의 범위를 결정하는 요소는 시스템의 매개변수로서 설정되게 된다. 이들에 대한 최적의 값은 여러 회의 튜닝을 통하여 실험적으로 찾아내게 된다.

6.2 학습데이터의 생성

원시데이터로부터 시계열 예측신경망을 위한 학습패턴의 생성은 신경망의 구조설정과 연관된 일련의 전처리 작업을 통하여 이루어진다. 예측을 위한 학습데이터 집합은 자료의 종류별로 일정 시간범위 창(window)의 크기를 매개변수로 설정하여 이를 예측규모 단위로 중첩 이동시키므로써 학습을 위한 입력패턴과 기대출력값을 결정한다. 예를 들어 학습자료로 선택된 기상자료가 식(23)과 같이 n 개의 값으로 이루어진 시계열패턴의 형태로 주어졌을 때, 예측을 위한 학습데이터 집합의 크기가 결정되고, 이로부터 학습패턴 설정을 위한 시간범위 창의 크기 W 가 결정되면 학습패턴 P 는 식(24)와 같이 $(n-W)$ 개의 학습패턴이 생성될 수 있다.

$$S = (s_1, s_2, s_3, \dots, s_n) \quad (23)$$

$$P = (p_1, p_2, p_3, \dots, p_{n-W}) \quad (24)$$

다시 말해서 i 번 째 학습패턴 p_i 의 내용은 식(25)와 같이 결정되고, 이에 대한 기대출력값은 s_{i+W} 가 된다.

$$P_i = (s_i, s_{i+1}, s_{i+2}, \dots, s_{i+W}) \quad (25)$$

FIR 네트워크 모델의 경우는 필터에 과거의 데이터를 저장하기 때문에 학습패턴은 입력력 패턴이 각각 1개 씩 존재하게 된다. 이것은 앞서 설명한 W 의 크기를

1로하여 학습패턴을 생성하는 것과 같다. 이때 다층퍼셉트론에서의 창의 개념은 필터에 적용이 되며, 필터의 크기는 창의 크기와 같게 된다. 다시 말해서 창의 크기를 W 라 할 때, 이는 필터의 크기와 일치하며, 과거 W 개의 시점을 고려함을 의미한다. FIR 네트워크의 경우 여러 데이터를 입력으로 취할 경우 단순히 뉴런을 확장하면 된다. 그리고 각각의 데이터에 대하여 입력력 패턴 한 쌍씩을 취해나간다면 학습패턴이 생성된다.

제안된 신경망에서는 이러한 학습패턴이 구간데이터로 확장된다. 그러나 만일 제안된 신경망에서 단일값을 입력으로 취하고, 싶은 경우에는 하한과 상한값에 같은 값을 넣어 주면 되므로 구간 표현을 수용하지 않는 기존의 모델에 비해 보다 일반화된 능력을 갖게 된다.

6.3 신경망 예측기의 구현 및 실험결과 고찰

본 논문에서 제시된 내용의 유용성을 평가하기 위하여 월별 및 계절별 강수량 예측실험을 수행하였다. 실험은 월별 강수량 데이터를 사용한 월별강수량 예측실험, 계절별 강수량 데이터를 사용한 계절별 강수량 예측 실험, 월별강수량과 평균기는 데이터를 동시에 사용한 월별 강수량 예측실험 등 세 부분으로 나누어 수행하였다. 각각의 실험에 대하여 제안된 모델과의 비교를 위하여 FIR 네트워크를 이용한 예측 실험을 병행하고 상호 성능을 비교하였다. 가능한 객관적으로 평가하기 위하여 두 모델 모두 동일 데이터를 사용하였으며 가급적 환경과 구조가 일치하도록 설정하였다.

실험에 있어 공통적인 부분은 다음과 같다. 우선 신경망은 3개의 계층구조를 가지며 입력노드 값은 앞장에서 언급한 바와 같이 관측된 강수량 및 평균기는 데이터에서 해당 월의 최대값과 최소값을 기준으로 정규화하여 0과 1사이의 실수값으로 산출하였다. 또한 출력노드 값으로부터의 예측강수량 산출은 같은 방법을 적용하여 역환산 되도록 하였다. 또한 실험에서 은닉층의 뉴런의 수는 모두 15개로 설정하였는데 이는 여러 회의 실험을 통하여 선정된 값이다. 여기서 각 뉴런에 대하여 시간이 흐름에 따라 신호값이 일정주기로 이동(shift)되는 7개의 가상노드를 두어 과거 데이터를 유지하였다. 또한 입력층과 은닉층의 모든 뉴런에 바이어스 노드를 추가하여 학습효과를 개선할 수 있도록 하였다. 예측 대상은 서울지역으로 하였으며, 학습자료로는 서울지역과 서울 인근지역의 데이터를 고려하여 생성하였다. 신경망의 학습은 제시된 학습알고리즘을 사용하여 학습자료 생성에 설정된 기간에 대하여 전체를 학습하는 것을 한번의 학습으로 정의하여 모든 출력에 대하여 정해진 오차범위 5%내로 수렴할 때까지 반복하도록 하였다.

실험1. 월별 강수량 예측 실험

이 실험은 과거 x 년 기간의 특정 월 강수량데이터를 가지고, 예측하려는 t 년 해당 월의 강수량을 예측하는 구조이다. 이것은 $t-1$ 년까지의 입력데이터를 가지고 다음 해 t 년의 해당 월, 즉 1년 후의 월간 강수량을 예측하는 실험이다. 이 실험에서는 1977년 이후의 서울 및 서울 인근지역의 관측자료를 입력데이터로 하여 서울지역의 월별강수량을 예측하도록 하였다. 이 때 FIR 네트워크는 여러 지역의 강수량데이터의 평균을 취함으로써 데이터를 감축하였으며, 제안된 모델은 지역의 상한과 하한을 구간표현으로 채택함으로써 데이터를 감축하였다.

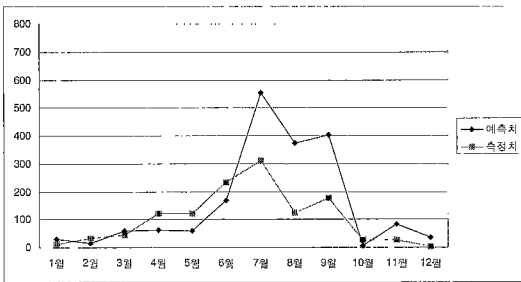


그림 6 FIR 네트워크를 이용한 월별 강수량 예측(서울 1998년)

그림 6에는 FIR 네트워크 모델을 사용하여 1998년 서울지역의 월별 강수량 예측실험을 수행한 결과를 예측치와 측정치를 비교한 그래프로 나타내었다. 그림 6을 보면 7, 8, 9월의 오차가 비교적 심한 편임을 알 수 있고, 1, 2, 3, 10월에는 비교적 정확한 예측을 하고 있음을 알 수 있다. 전체적 분포로 볼 때는 비교적 근사한 예측을 하고 있으나 허용오차의 한계를 벗어난 일부 결과를 발견할 수 있었다.

그림 7에는 제안된 모델을 사용하여 1998년 서울지역의 월별 강수량 예측실험을 수행한 결과를 예측구간과 측정치를 비교한 그래프로 나타내었다. 모델의 특성상 예측값은 구간으로 나타나며, 실제값인 측정치가 예측구간 내에 존재하며 예측구간의 크기가 작을수록 바람직한 예측을 하고 있다고 할 수 있다. 실험의 결과 4월, 5월 및 9월과 같이 예측구간을 다소 벗어나거나, 8월과 같이 예측구간이 지나치게 크게 나타나는 결점이 일부 발견되었지만 전체적으로 예측구간 내에 근접되는 결과를 보였다.

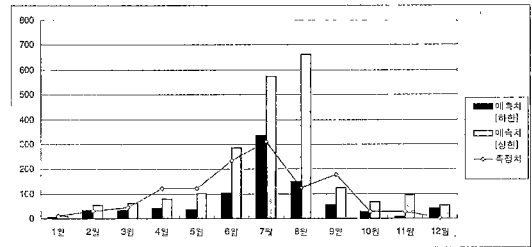


그림 7 제안된 모델을 이용한 월별 강수량 예측(서울 1998년)

실험2. 계절별 강수량 예측 실험

이 실험은 과거 일정기간의 특정 계절에 대한 강수량 데이터를 가지고, 향후 일정 년수(t)의 해당계절의 강수량을 예측하는 구조이다. 이것은 $t-1$ 년까지의 입력데이터를 가지고 다음 해 t 년의 해당 계절, 즉 1년 후의 해당 계절 강수량을 예측하는 실험이다. 이때 12~2월을 겨울, 3~5월을 봄, 6~8월을 여름, 9~11월을 가을로 정의하였다. 이 실험에서는 1977년 이후의 서울 및 서울 인근지역의 계절별 강수량 데이터를 입력데이터로 하여 서울지역의 계절별 강수량을 예측하도록 하였다. 이 때 실험1과 마찬가지로 FIR 네트워크는 여러 지역의 강수량 데이터의 평균을 취하였으며, 제안된 모델은 구간표현을 채택하였다.

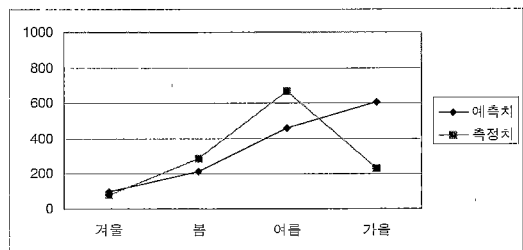


그림 8 FIR 네트워크를 이용한 계절별 강수량 예측(서울 1998년)

그림 8에는 FIR 네트워크 모델을 사용하여 1998년 서울지역의 계절별 강수량 예측실험에 대하여 예측치와 측정치를 비교한 그래프로 나타내었다. 이때 겨울철에 해당하는 값은 1997년 12월에서 1998년 2월까지의 데이터이다. 그림 8을 보면 가을에 해당하는 오차가 비교적 크게 발견됨을 알 수 있다.

그림 9에는 제안된 모델을 사용하여 1998년 서울지역의 계절별 강수량 예측실험을 수행한 결과를 나타내었

다. 실험의 결과 여름철에 다소의 오차를 보였지만 측정치가 예측구간에 가깝고 예측 구간이 작게 나타나는 바람직한 예측을 하고 있음을 알 수 있다.

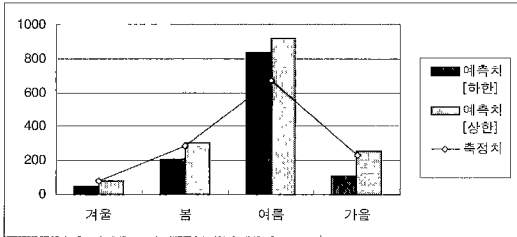


그림 9 제안된 모델을 사용한 계절별 강수량 예측(서울 1998년)

실험3. 기온자료를 포함한 학습 및 예측

실험1의 예측 성능을 개선하기 위하여 각각의 시점, 지역에 대하여 월별 평균기온 데이터를 추가하여 실험하였다. 실험 3은 과거 일정기간의 해당 월 강수량 및 평균기온 데이터를 가지고, 예측하려는 연도의 월별강수량을 예측하는 구조이다. 이 실험에서는 1977년 이후의 서울 및 서울 인근지역의 월별강수량에 평균기온 데이터를 포함하여 학습데이터를 구성하였으며, 이러한 평균기온 데이터는 예측과정에서도 사용되게 된다. 이 실험은 앞의 실험에서의 학습데이터 집합에 대하여 기온자료를 포함하는 형태로 확장함으로써 월별강수량 예측성능의 개선을 목표로 하는 것이므로 월별강수량의 예측결과만을 나타내었다. 우선 FIR 네트워크를 사용하여 1998년 서울지역의 월별강수량 예측실험 결과를 예측치와 측정치를 비교한 그래프로서 그림 10에 나타내었다. 이를 실험1과 비교할 때 전체적으로 오차가 감소하였으나 아직도 7월과 8월의 오차가 비교적 크게 나타나고 있음을 알 수 있다.

그림 11은 동일 환경에서 제안된 모델을 사용한 실험 결과이다. 실험1과 비교할 때 특히 심한 오차를 보였던 8월에서 매우 바람직한 예측을 하는 등, 전체적으로 성능이 개선되어 진 것을 알 수 있으며 FIR 네트워크와 비교하여서도 개선된 성능을 보였다. 그럼에도 불구하고 아직도 예측구간을 벗어나는 형태의 오차가 일부 발견되고 있다. 여러 회의 반복실험과 기존의 통계적 방법 등을 사용하는 예측실험의 성능 수준 등을 비교 분석한 결과, 이러한 오차의 원인은 제안된 모델이 갖는 원초적인 결함이라기 보다, 학습에 사용되는 데이터의 양적인 부족이나 특징 추출방법의 결함 및 특징에 포함되는 정보의 부족 등에 기인한다고 판단된다.

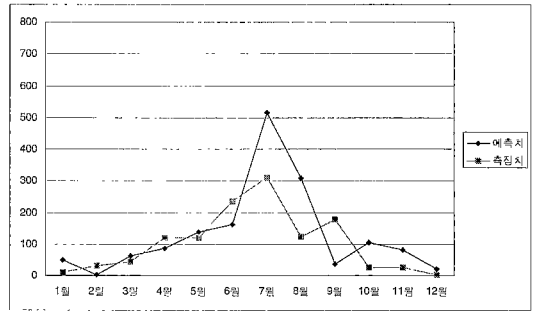


그림 10 확장된 학습데이터를 사용한 예측실험 결과 (FIR 네트워크)

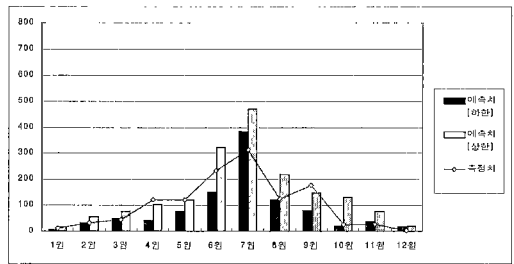


그림 11 확장된 학습데이터를 사용한 예측실험 결과(제안된 모델)

7. 결론

본 연구에서 제시한 신경망모델은 실제계의 시계열 예측문제에서 빈번하게 나타나는 구간형태의 데이터 표현을 수용한다. 이는 기존의 FIR네트워크가 갖는 동적 동작특성에 구간연산을 수행하는 활성화 특성을 갖도록 연산을 일반화함으로써 이루어 지는데, 신경망의 입력뿐만 아니라 내부적인 연산과정 및 출력에도 구간표현 및 연산을 사용함으로써 학습의 효율성과 유연성을 증대시킨다. 신경망예측기의 성능 개선을 위하여 학습에 사용되는 데이터의 종류와 범위를 확장하는 시도는 관측값에 내재하는 노이즈의 영향과 계산량 증대라는 문제점을 수반한다. 구간형태의 데이터 표현방식과 이에 대한 연산기능은 방대한 원시데이터에 대하여 시간적 및 공간적으로 분할하여 함축적인 특징을 추출할 수 있게 한다. 이는 신경망 예측기에 필요한 학습데이터의 양을 감축하고, 보다 유연한 자료의 표현 및 예측을 가능하게 한다.

제안된 이론은 한반도 지역별 강수량 예측 시스템을 대상으로 구현하여 실험에 적용하였다. 이 과정에서 학습자료는 지역별로 관측된 강수량 데이터와 평균기온

데이터를 사용하였는데, 신경망 예측기의 구조설정과 학습에 앞서, 제시된 데이터 처리기법을 적용함으로써 학습효과를 향상시킬 수 있었다. 또한 예측결과로부터 최대 및 최소값 형태의 예측이 가능하게 함으로써 보다 세분화된 정보를 얻을 수 있었다. 향후 연구로는 해수표면온도 등의 학습자료를 확대하는 방법과, 시계열 예측문제에서 해당 분야 전문가의 지식이나 관련분야의 이론적 지식을 신경망 학습에 반영하는 방법과 기존의 다른 시계열 예측기법과 상호 결합하는 방법 등이 연구되어야 하겠다.

참 고 문 헌

[1] Barak Cohen, Daid Saad and Emanuel Marom, Efficient Training of Recurrent Neural Network with Time Delays, Neural Networks, Vol. 10, No. 1, pp. 51-59, 1997.

[2] J. T. Connor, R. D. Martin and L. E. Atlas, "Recurrent Neural Networks and Robust Time Series Prediction," IEEE Transaction on Neural Networks Vol.5, No.2, 1994.

[3] Alex Aussem, Fionn Murtagh and Marc Sarazin, Dynamical Recurrent Neural Networks - Towards Environmental Time Series Prediction, International Journal of Neural Systems, Vol. 6, No. 2, p. 145-170, 1995.

[4] Eric A. Wan, 1996. "Time Series Prediction by Using a Connectionist Network with Internal Delay Lines," Time Series Prediction, pp. 195-217, 1996.

[5] C.A. Hernandez, J. Espf, K. Nakayama, M. Fernandez, "Interval Arithmetic Backpropagation," Proceedings of International Conference on Neural Information Processing, 1995.

[6] Cohen Barak, David Saad and Emanuel Maro, "Efficient Training of Recurrent Neural Network with Time Delays," Neural Networks Vol.10, No.1, pp. 51-59, 1997.

[7] Ath Kehagias and Vas. Petridis, Predictive Modular Neural Networks for Time Series, Neural Networks, Vol. 10, No. 1, pp. 31-49, 1997.

[8] Shanming Shi and Andreas S. Weigend, Markov Gated Experts for Time Series Analysis: Beyond Regression, Proceeding of IJCNN, Vol.1, 1997.

[9] Andreas S. Weigend, Morgan Mangeas, Ashok N. Srivastava, Nonlinear Gated Experts for Time Series: Discovering Regimes and Avoiding Overfitting, International Journal of Neural Systems, Vol. 6, No. 4, p.373-399, 1995.

[10] W. Hsu, L. S. Hsu and M. F. Tenorio, "The ClusNet Algorithm and Time Series Prediction,"

International Journal of Neural Systems Vol.4, No.3, pp.247-255, 1993.

[11] Paul J. Werbos: "Generalization of Backpropagation with Application to a Recurrent Gas Market Model," Neural Networks, Vol.1, No.4, 1988.

[12] Gustavo Deco, Ralph Neuneier and Bernd Schürmann, Non-parametric Data Selection for Neural Learning in Non-stationary Time Series, Neural Networks, Vol. 10, No. 3, pp. 401-407, 1997.



김 호 준
1987년 경북대학교 전자공학과 졸업. 1987년 ~ 1991년 한국원자력연구소 연구원. 1995년 한국과학기술원 전산학과 박사. 1995년 한국과학기술원 인공지능연구센터 연구원. 1996년 ~ 현재 한동대학교 전산전자공학부 조교수. 관심분야는 패턴인식, 신경회로망, 학습이론, 에이전트 등임.



김 우 성
1999년 2월 한동대학교 전산전자공학부 학사. 1999년 3월 ~ 2000년 2월 한동대학교 정보통신기술연구소 연구원. 2000년 ~ 현재 한국과학기술원 전산학과 석사 과정. 관심분야는 패턴인식, 신경회로망 등임.