

# Time PLOT과 이동평균 융합 시계열 데이터 예측

이준연

동명대학교 미디어공학과

## Forecasting the Time-Series Data Converged on Time PLOT and Moving Average

Jun-Yeon Lee

Department of Media Engineering, Tongmyong University

**요약** 시계열 데이터를 예측하는 것은 매우 어려운 작업이다. 비선형적인 특성을 갖는 신호에서 얻어지는 데이터들이 불확실성을 가지고 있기 때문이다. 본 논문은 특정 시계열 데이터의 정확한 예측을 위하여 시계열 자료가 어떤 패턴에 따라 변화한다는 전제하에서 과거 자료들을 평균하여 미분으로써, 시계열 변화 패턴의 찾았다. 또한 미분 데이터의 반영 비율에 따라 특이성을 갖는 시계열데이터를 일반화하기 위하여 확률변수를 적용하였다. 순환변동과 계절변동을 소거하고, 불규칙 변동만을 추출하여 경향의 추세를 더한 예측값을 계산하게 된다. 이렇게 예측된 값은 이동평균과 단순이동평균에 의하여 가장 좋은 결과값을 갖는 알고리즘과 비교를 통하여 제안 알고리즘의 우수성을 입증하였다.

• **Key Words** : 이동평균, Time PLOT, 시계열 예측, 순환변동, 계절변동

**Abstract** It is very difficult to predict time-series data. This is because data obtained from the signal having a non-linear characteristic has an uncertainty. In this paper, By differentiating time-series data is the average of the past data under the premise that change depending on what pattern, and find the soft look of time-series change pattern. This paper also apply the probability variables to generalize time-series data having a specific data according to the reflection ratio of the differentiation. The predicted value is estimated by removing cyclic movement and seasonal fluctuation, and reflect the trend by extracting the irregular fluctuation. Predicted value has demonstrated the superiority of the proposed algorithm and compared with the best results by a simple moving average and the moving average.

• **Key Words** : Moving Average; Time PLOT; Forecasting of Time-series, Cyclical Movement, Seasonal Fluctuation

### 1. 서론

일반적으로 얻어지는 데이터들은 일정한 시간을 주기로 하여 얻어지는 시계열 데이터가 대부분이고, 많은 분야에서 이러한 시계열 데이터의 분석과 예측을 필요로 한다. 시계열 생체신호를 분석하는 것은 그 속에 내포된

다양한 정보를 통해 그 신호를 발생시키는 기관의 상태를 정확하게 모델링하기 위함이다. 모델링은 이미지 형태이거나 수치로 제시되어 임상 현장에서의 자동진단 시스템이나 병을 진단하는데 보조 자료로 활용될 수 있다. 현재 생체신호를 처리하는 과정으로는 수학적인 모델을

\*교신저자 : 이준연(jylee@ac.kr)

접수일 2015년 6월 16일

수정일 2015년 7월 31일

게재확정일 2015년 8월 20일

통해 임시로 이미지 혹은 시각적 데이터를 그려내는 것이 전부이며, 스펙트럼이나 어트랙터 등이 이에 속할 수 있다. 수치로 표현되는 경우 각 증상별로 정확하면서도 서로 확실하게 구별될 수 있는 값을 도출해내는 것이 패턴 분류 목적중의 하나가 될 것이다. 어떤 생체신호가 인간이 감지하기에 증상별로 유사하여 잘 구별할 수 없어도 만약 특정한 알고리즘에 의해 계산된 값이 같은 증상끼리는 유사하거나 오차 범주 안에 들어가면서도 다른 증상들끼리는 확연하게 차이가 있는 값을 제시한다면 패턴 분류 목적에 아주 부합되는 일이 될 것이다[1]. 이러한 과정이 가능하도록 하는 수학적 모델이 카오스이론이다.

뇌파와 심박동 등 생체신호는 이미 카오스 현상이 내재된 동력학과 마찬가지로 정형화할 수 있는 정보를 내포하고 있다고 알려져 있으므로 비선형 분석법(nonlinear dynamical analysis)에 기초하여 새로운 분석 방법을 개발하고 있다. 비선형 분석이란 신호의 동역학 특성을 측정함으로써 계의 상태가 시간의 변수에 의해 달라지는 것을 정량화하고 계 모델링을 위해 필요한 모든 정보를 추출해 내는데 있다[2].

시계열 데이터를 다차원 공간에서 운동하는 새로운 기하학적 형태로 제시해주는 어트랙터는 이미 여러 시계열 데이터를 분석하는 방법에 이용되기 시작하였다. 어트랙터를 구성하기 위해서는 지연시간과 매립차원이라는 변수가 도입되어야 한다. 어트랙터는 궁극적으로 시계열 데이터가 서로 어떤 연관성이 있는지를 나타내어 주는데, 지연시간 값에 따라 그려지는 어트랙터 모양이 크게 좌우된다. 결국 최종적으로 계산되는 동역학 측정값이 달라져서 적절한 시간 지연 값과 매립차원을 결정하는 것은 매우 중요한 일이 된다. 특히 생체신호의 경우, 패턴매칭이나 분류에 상당한 영향을 미치게 된다[3].

시계열 자료를 분석한다는 것은 어떤 자료가 과거 관찰값들의 변화패턴과 비슷하게 현재 시점 이후에도 변화될 것이라는 전제하에서 향후 시점의 값을 예측하는 기법이다. 정확한 예측을 위하여 보다 중요한 것은 시계열 자료의 성격과 특징을 파악하고, 그 자료와 관련된 주변 환경 변수들은 어떤 것들이 있는지, 그 자료 자체를 분석할 것인지 변환 또는 조정이 필요한지 등을 고려해 보아야 할 것이다. 본 논문에서는 시계열 자료가 어떤 패턴에 따라 변화한다는 전제하에서 과거 자료들을 평균하여 차분으로써, 시계열 변화 패턴의 부드러운 모습을 찾고자 한다. 제안된 방법에서는 데이터 발생 시스템에 대한

사전정보가 부족하거나 시변 특성을 갖는 경우에도 잘 동작할 수 있도록 원데이터로부터 차분값을 사용하였다. 그래서 데이터의 특성과 유사성이 잘 드러나게 하였고, 데이터에 대한 예측의 정확도를 향상시킬 수 있도록 하였다.

## 2. 예측의 정확도 측정

시계열자료를 표현하는 방법은  $\{Y_t, t=1, 2, \dots\}$ 로서  $t$ 는 시점을 나타내고  $Y_t$ 는 그 시계열자료를 나타내는 변수이다. 즉,  $\{Y_1, Y_2, \dots\}$  등으로 시계열자료를 표현한다면  $Y_1$ 은 시점 1에서의  $Y$ 값,  $Y_2$ 는 시점 2에서의  $Y$ 값을 나타내는 것이다.

예측은 과거 시점들에서 얻어진 시계열자료들을 이용하여 장래시점에서의 그 시계열 변수 값들을 알아내는 것인데, 그 예측값들이 얼마나 정확하다고 할 수 있는가를 측정해야만 할 것이다. 그리고 특정 시계열자료에 대해 예측을 하는 방법으로 여러 가지 방법들이 가능한데 어느 방법으로 예측한 결과가 가장 적당한가를 알기 위해서는 예측값들의 정확도를 측정하여 비교해 볼 수 있다. 먼저, 현재 시점( $t$ )에서  $n$ 개의 시계열자료를 갖고 있다면,  $\{Y_{t-(n-1)}, \dots, Y_{t-1}, Y_t\}$ 의 시계열자료들을 가지고 예측을 수행하게 된다는 것이다. 그리고  $n$ 개의 시계열자료들으로써 어떤 예측모형을 사용하여 향후  $m$ 개 시점들에 대한 예측값들을 얻는다면, 그 예측값들은  $\{F_{t+1}, F_{t+2}, \dots, F_{t+m}\}$ 으로 표현될 것이다. 또한 이 예측모형에 의하여 현시점부터 과거  $n$ 개의 자료들이 계산될 수 있는데 그 값들은  $\{F_{t-(n-1)}, \dots, F_{t-1}, F_t\}$ 로 표현한다. 따라서, 현재 얻어진 시계열자료와 예측모형을 통하여 계산된 시계열 값들, 간의 차이는 (예측모형에 의하여 얻어지는) 오차들이 되는데, 이 오차들으로써 그 예측모형이 적절한 예측 방법인가를 판단하는 기준을 얻어야 할 것이다. 즉,

$$\begin{aligned} e_t &= Y_t - F_t \\ e_{t-1} &= Y_{t-1} - F_{t-1} \\ &\vdots \\ e_{t-(n-1)} &= Y_{t-(n-1)} - F_{t-(n-1)} \end{aligned}$$

들을 오차라고 부르며, 이들 오차들을 가지고 예측 방법의 정확도를 측정해야 할 것이다.

### 3. 시계열 데이터 분석

시계열 데이터의 분석이란 관측된 과거의 자료를 분석하여 일련의 법칙을 발견하고, 이를 모델링하여 추정하는 것을 뜻한다. 이러한 시계열 분석의 가장 중요한 목적은 예측이다. 미래에 대한 예측이 가능해지면 그 시스템에 어떤 인위적인 조작을 가함으로써 시계열 실현값이 우리가 원하는 목표에 부응하도록 유도할 수 있다. 또 하나의 목적은 자료의 기술이다. 자료를 시간의 흐름에 따라 그림으로 표현하고, 자료의 특성을 파악하여 법칙성을 발견하는 것이다. 대체적으로 추세, 순환, 계절성, 특이점, 변환점 등 가지고 있는 시계열 데이터의 특징을 찾아 분석할 수 있다는 것이다[4,5].

시계열 데이터의 분석과정은 먼저 데이터의 전처리 단계를 거쳐 선형분석과 비선형분석을 거쳐 실제 모델을 구성하고 응용에 적용하는 것이다. 데이터의 전처리단계는 체계적이고 일관된 분석을 수행하기 위해서 수집된 시계열을 살펴 보면서 측정 잡음이 과도하게 들어간 부분을 제거하며, 전체가 잡음에 노출된 경우 해당 데이터를 폐기하고 다시 실험한다.

선형 분석은 생체 시계열 데이터 자체에 대해 수행되며, 시계열에 관여하는 에너지 성분 및 시계열간 관련성을 조사한다. 대부분의 생체신호에 대해 생체계에 작용하는 에너지 성분과 크기, 분포 등을 조사하는 푸리에 분석을 기본적으로 수행한다. 그리고, 자기상관함수, 복합 선형회기 등을 구하여 시계열내의 관련성을 찾고 예측모형을 만든다. 그러나 선형분석은 시계열 자체의 성질과 계에 관여하는 에너지 성분의 종류 및 크기를 알려줄 뿐계의 실제 동력학에 관해서는 정보를 주지 못한다[6].

비선형 분석은 복잡하게 변하는 생체시계열 데이터를 만들어내는 생체 동력학의 구조와 성질을 정량적 또는 정성적으로 표현하고 조사한다. 그런데 생체 시계열을 만드는 동력학은 현재 정확히 알려진 것이 거의 없는 블랙박스과 같으나, 카오스 이론에서 발전된 위상공간매립법을 이용하면 스칼라 생체시계열데이터로부터 계의 동력학에 관한 많은 정보를 알아낼 수 있다. 스칼라 시계열로부터 복잡한 동력학의 독립변수의 수를 계산하여 실제하는 차원을 찾아낸 후 그 공간 내에서 동력학을 어트랙터로서 재현한다. 그리고 어트랙터의 기하학적 구조와 시간에 따라 변화하는 성질을 나타내는 정량적인 양을 정의하고 계산하여 그 상관관계를 구한다[7]. 모델링 및 응용단계에서는 선형 및 비선형 분석을 근거로 생체에

관여하는 에너지 성분 및 독립변수의 수를 최소로 하는 모델을 구성하여 실제 생체현상을 발생시키는 생체동력학을 이해할 수 있다.

#### 3.1 데이터 전처리

많은 시계열자료들은 균일하지 않고 편차가 큰 값들을 갖는다. 특정 시점에서는 평균에 비해 매우 큰 값을 가지거나 또는 작은 값을 갖게 되고, 또 일정 기간동안 시계열 자료는 상승하거나 하강하는 형태를 갖기도 한다. 평활법(Smoothing Method)이란 시계열의 들쭉날쭉함을 어떻게 부드럽게 조정하여 예측할 수 있느냐 하는 문제이다. 다시 말하면, 어떤 부드러운 곡선을 기준으로 임의로(Random) 이탈하여 시계열자료 값들이 얻어진 것이라는 전제하에서 그 부드러운 곡선의 패턴을 찾아내자는 방법이다[8].

따라서, Smoothing 방법은 과거 시계열자료들에 어떤 방법으로 가중치를 부여하여 예측하느냐에 따라 크게 이동평균(Moving Average)방법과 지수평활법(Exponential Smoothing) 방법의 두 가지로 나누어 생각할 수 있다.

##### 3.1.1 이동평균방법

이동평균방법이란 어느 시계열 자료가 일정한 주기를 갖고 비슷한 패턴으로 움직이고 있을 경우에 적용시킬 수 있는 방법이다. 주기가 길어질수록 직선에 가까운 부드러운 선이 얻어질 것임을 쉽게 알 수 있다[9].

평균을 사용하여 예측을 한다는 것은 매우 직관적인 방법이다. 왜냐하면, 평균이란 누구나 알고 있는 개념이어서 쉽고 간단하게 사용할 수 있기 때문이다.

시계열 자료,  $\{Y_1, Y_2, \dots, Y_T\}$ 가 있다고 하자.

그러면 (T+1) 시점에서의 예측값  $F_{T+1}$ 을

$$F_{T+1} = \overline{Y_{(T)}} = \sum_{i=1}^T Y_i / T$$

로 하자는 것이다. 만일 (T+1) 시점의 실제 시계열 값이 얻어진다면 예측오차는

$$e_{T+1} = Y_{T+1} - F_{T+1} = Y_{T+1} - \overline{Y_{(T)}}$$

로 계산되어 진다. 마찬가지로 시점 (T+2)에서의 예측값,  $F_{T+2}$ 는

$$F_{T+2} = \overline{Y_{(T+1)}} = \sum_{i=1}^{T+1} Y_i / (T+1) = \frac{T \times F_{T+1} + Y_{T+1}}{(T+1)}$$

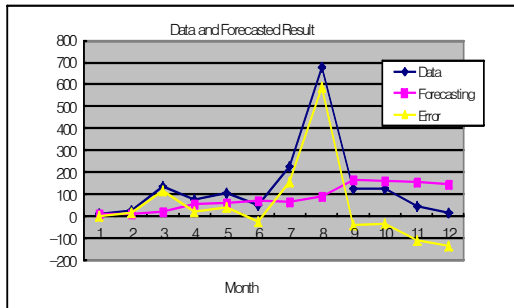
로 계산될 것이다.

평균을 예측의 방법으로 적용시킬 수 있는 시계열자료는 추세(Trend)도 없고 계절성(Seasonality)도 없는 시계열자료이다. 이러한 시계열자료에 대해서라면 평균에 사용되는 자료의 수가 많으면 많을수록 안정적(Stable)인 평균값이 얻어지기 때문에 평균을 이용할 경우 좋은 예측값이 될 것이다[10].

그러나, 시계열자료가 추세나 계절성이 있을 경우, 또는 어느 시점에서 급격히 변화하는 모습을 나타낸다면 평균의 방법으로 좋은 예측값을 기대할 수는 없을 것이다.

<Table 1> The method of average prediction

월	1	2	3	4	5	6	7	8	9	10	11	12
자료	12	28	135	79	108	51	229	677	126	127	47	15
예측값	10	12	20	58.3	63.5	72.4	68.8	91.6	164.8	160.4	157.1	147.1
eps=Y-F	2	16	115	20.7	44.5	-21.7	159.2	585.4	-38.8	-33.4	-110.1	-132.1



[Fig. 1] Data and Forecasting Result

### 3.1.2 단순이동평균

시계열자료가 충분히 확보되어 있을 경우, 평균의 방법을 개선하여 일정한 개수의 자료들을 평균하여 예측값으로 사용하는 방법이다. 즉, 예측시점을 기준으로 과거 일정 개수의 자료들 평균한 다음 시점의 예측값으로 사용하는 방법으로서 모든 과거자료를 동원하여 평균을 얻는 평균 방법과는 다르다[11,12].

예측하고자 하는 시점 (T)를 기준으로 과거 몇 개의 자료로써 평균을 얻을 것인가 하는 주기 (N)가 결정되면, 다음과 같은 방법으로 이동평균(moving average)을 얻을 수 있다. 시계열 자료,  $\{Y_1, Y_2, \dots, Y_T\}$ 가 있다고 하

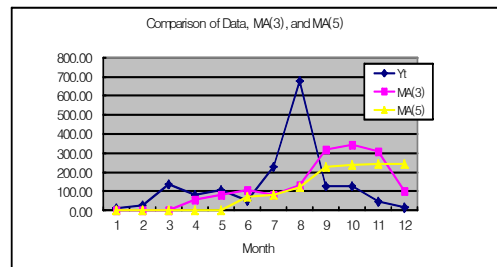
자. 그리고 주기를 N이라고 하면, 시점 (T+1, T+2, T+3)에서의 예측값,  $(F_{T+1}, F_{T+2}, F_{T+3})$ 을 얻는다.

그러나, 단순 이동평균방법은 시계열자료가 추세를 갖고 있을 경우에는 부적절하다. 이동평균방법을 사용함에 있어서 주기 (N)를 얼마로 해야 하는가 하는 문제가 있는데, 첫 번째는 MA(1)은 현재 시점 (T)의 시계열 값으로 다음 시점 (T+1)의 시계열자료를 예측한다는 것이다. 그리고 두 번째 문제는 MA(4)는 분기별 자료에 사용할 수 있는 데, 4개의 자료들 평균을 얻으면 계절 효과를 제거하는 데 도움이 되지만, 그 평균값을 다음 시점의 예측값으로 한다는 것은 추세나 계절효과를 나타내기에 바람직하지 않다[13,14].

4개의 자료들 평균을 4개 자료들의 가운데 위치시킨다면 시계열자료들의 요소들을 파악하는데 도움이 될 것이다. 세 번째 문제는 이동평균방법에서 주기를 크게 하면 할수록 시계열자료를 부드럽게 하는 효과는 있다. 왜냐하면 많은 자료들의 평균을 내어서 예측값으로 사용한다는 것은 그 기간내의 큰 편차를 고려하지 않는다는 것이 때문이다[15].

<Table 2> The Comparison of Moving Average value in 3 Cycle and 5 Cycle.

월	1	2	3	4	5	6	7	8	9	10	11	12
Yt	12	28	135	79	108	51	229	677	126	127	47	15
MA(3)	x	x	x	58.3	80.7	107.3	79.3	129.3	319	344	310	100
MA(5)	x	x	x	x	x	72.4	80.2	120.4	228.8	238.2	242	241.2



[Fig. 2] Comparison of MA(3) and MA(5)

<Table 2>와 [Fig. 2]에서 알 수 있듯이 시계열자료와 MA(3), MA(5)의 이동평균값들, 그리고 각각의 방법으로 얻어진 각 월의 예측값을 나타내고 있다. MA(3)가 MA(5)보다는 훨씬 시계열자료를 부드러운 패턴으로 나타내고 있음을 알 수 있다.

그리고 위의 자료에 대한 MA(3), MA(5)에 대한 예측의 정확도 값들은 다음과 같이 정리될 수 있다. 여러 가지 예측의 정확도를 나타내는 척도 값들을 고려할 때, MA(3)가 MA(5)보다는 좋은 예측의 방법임을 알 수 있다.

(Table 3) Error comparison of cycle 3 and 5 in Moving Average

	MA(3) t=4~t=12	MA(5) t=6~t=12
term		
Mean Error	-7.7	6.9
Mean Absolute Error	172.4	197.5
Mean Absolute Percentage Error	151.4	283.8
Standard Deviation of Error	77.5	74.1
Durbin-Watson Statistics	0.79	0.39

### 3.1.3 Time PLOT

앞에서 설명한 이동평균기법은 시계열 자료가 어떤 패턴에 따라 변화한다는 전제하에서 과거 자료들을 평균 함으로써, 시계열 변화 패턴의 부드러운 모습을 찾아보자는 것이었다. 따라서 시계열자료의 부드러운 패턴이 얻어지면 그 패턴의 연장선상에서 예측값을 찾는 방법이다. 이에 반해 시도표에 의한 분석은 시계열자료는 여러 변동의 결합으로 이루어지는 것이므로 시계열자료를 형성하고 있는 변동요소를 찾아내고 시계열자료를 그 요소들로 표현하여 예측하는 것이다. 본 논문에서는 스플라인 피트(spline fit)에 의한 시도표 분석을 통하여 시계열 자료의 구조를 파악하고자 한다.

#### 1) 경향(Trend)

경향이란 자료가 장기적으로 변화해 가는 큰 흐름을 나타내는 것이다. 자료가 장기적으로(Long-run) 커지는지, 작아지는지, 변화가 없는지를 나타내 주는 요소를 추세라고 한다.

#### 2) 순환변동 (Cyclical Movement)

순환변동은 경제전반의 또는 특정 산업의 부침(Ups and Downs)를 나타내 주는 것을 말한다.

#### 3) 계절변동 (Seasonal Fluctuation)

계절변동은 일정한 주기를 갖고 반복적으로 같은 패턴으로 변화하는 것을 말한다. 순환변동과 계절변동의 차이점은 계절변동은 주기가 일정하게 정해진 반면 순환

변동은 정해지지 않은 기간 동안에 사이클(Cycle)을 보여주는 것이라 할 수 있다.

이상과 같이 시계열자료는 세 가지 변동을 내포하면서 불규칙하게 변화하는 것이므로 시계열자료( $Y_t$ )는

$$Y_t = f(\text{경향, 순환, 계절}) + \text{오차}$$

로 표현할 수 있다. 여기서 오차란 불규칙 변동으로 설명되는 항인데 실제 시계열자료 값과 추세, 순환, 계절 변동의 조합에 따른 패턴,  $f(\text{경향, 순환, 계절})$ 간의 차이라고 보면 된다. 그러므로 시계열자료는 실제로 위에서 설명한 네 가지 요소가 뒤섞여 있는 것으로 간주하여

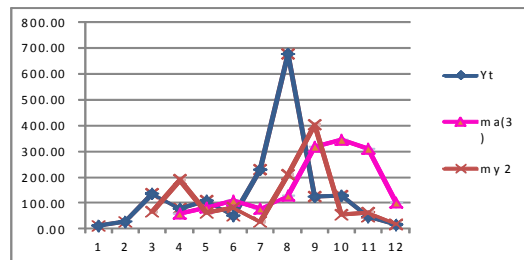
$$Y_t = f(T_t, C_t, S_t, I_t)$$

로 표현한다. 여기서,  $Y_t$ 는  $t$ 시점에서의 시계열 값,  $T_t$ 는  $t$ 시점에서의 추세요소,  $C_t$ 는 순환요소,  $S_t$ 는 계절요소,  $I_t$ 는 불규칙요소를 나타낸다.

이들 요소의 분해 과정은 다음과 같다.

- (1) 계절변동 주기( $M$ )를 파악한 후  $N$ 기의 이동평균을 구하여 계절성과 불규칙성(임의성)을 제거시킬 수 있다.
- (2) 시계열자료( $Y_t$ )로 부터 (1)에서 구한 이동평균 값들은 분리시키게 되면 추세변동과 순환변동만 남게 된다.
- (3) 시계열자료의 추세가 어떤 형태로 이루어져 있는지를 알 수 있게 되고, 그 추세의 형태( $T_t$ )를 구할 수 있다.
- (4) (2)에서 얻어진 추세변동과 순환변동만 남아 있는 시계열에서 (3)에서 구한 추세변동을 제거하면 순환변동( $C_t$ )이 얻어지고
- (5) 위의 과정에서 얻어진  $S_t, T_t, C_t$ 를 모두 제거할 때 남는 것이 불규칙변동( $I_t$ )이다.

$$I_t = (1-x) * \text{Deviation} + x * |Y_{t-1} - Y_{t-2}|$$



[Fig. 3] Comparison of MA(3) and proposed algorithm

여기서 Deviation은 일반적으로 이전 시계열자료의 기울기에서 얼마나 멀리 떨어져 있는가를 나타내게 된다. x는 적용확률변수로 이전의 평균값대비 직전 기울기값 반영 비율이다. 이는 급격한 변화요인을 가진 시계열자료에 따라 예측값이 큰 폭으로 변화하는 현상을 막고, 경향을 선형적으로 반영하기 위한 매개변수이다.

[Fig. 3]은 적용확률변수의 값을 0.2, 즉 20%의 비율만큼 직전 변화요인 값을 반영한 결과와 앞 절의 이동평균에 의한 예측에서 가장 우수한 결과를 보인 MA(3)와의 비교 결과이다.

#### 4. 결론

일반적으로 시계열 데이터를 정확히 예측하는 것은 매우 어려운 작업이다. 비선형적인 특성을 갖는 신호에서 얻어지는 데이터들이 불확실성을 가지고 있기 때문이다. 이를 해결하기 위하여 본 논문에서는 시계열 데이터의 예측에 사용될 적절한 입력값을 여러 차분 간격을 통하여 선택적으로 사용하는 예측알고리즘을 제안하였다. 제안된 알고리즘은 동일한 데이터에서 가장 예측오차가 적은 3차분 이동평균값(MA(3))과 비교를 통하여 제안된 알고리즘의 우수성을 입증하였다. 예측값과 실제값의 평균 오차는 MA(3)의 경우 173.3인데 비해 제안된 알고리즘은 135.5로 약 22%가량의 오차 개선을 보였다.

#### ACKNOWLEDGMENTS

이 논문은 2013년도 동명대학교 교내학술연구비지원에 의하여 연구되었음(2013A015).

#### REFERENCES

- [1] Alexandar D. Shin, M.D., "Effect of Attractor Construction Methods to Fractal Dimension for Heart Sound Analysis", <http://blog.naver.com/mathic/120009708196>.
- [2] Taesik Kim, "Speech Signal Processing for Analysis of Chaos Pattern", Voice Science Vol. 8, No. 3, pp. 149-157, 2001.
- [3] Taesik Kim, Hong Jung, "Nonlinear Data Presentation Method for Chaotic Analysis of Heart Rate", Healthcare Informatics Research, Vol. 9 No. 2, pp. 93-100, 2003.
- [4] George E. P. Box, Gwilym M. Jenkins, "Time Series Analysis : forecasting and control", Revised Ed., Prentice Hall, 1976.
- [5] G. Janacek, L. Swift, "Time Series Forecasting, Simulation, Applications", Ellis Horwood, 1993.
- [6] Seungwhan Park, "The Recognition Technique of Biological Signal", Information Processing, Vol. 6, No. 4 pp. 44-52, 1999.
- [7] Richard L. Burden, J. Douglas Faires, "Numerical Analysis", Sixth Ed., International Thomson Publishing Asia, 1997.
- [8] Li-Xin Wang, "A Course in Fuzzy Systems and Control", Prentice Hall, 1997.
- [9] IS Choi, DS Kang, JH Lee, MW Kang, "Prediction of the industrial stock price index using domestic and foreign economic indices", Journal of the Korean Data & Information Science Society, Vol 23, No. 2 pp. 271-283, 2012
- [10] KH Hwang, WP Kim, C Jeog, "Demand Forecasting of Dok-do Tourism using Comparison of Univariate Time Series", Korea Academic Society of Tourism and Leisure, Vol. 27, No. 2, pp. 59-77, 2015.
- [11] Jae-whan Son, "A probabilistic model of school violence based on the markov chain process", Studies on Korean Youth, Vol. 26, No. 2 pp. 35-57, 2015.
- [12] JC Kim, JH Lee, GJ Kim, SS Park, DS Jang, "Data Engineering" Time Series Analysis of Patent Keywords for Forecasting Emerging Technology", Journal of Information Processing, Vol 3, No. 9, pp. 355-360, 2014.
- [13] Jinhyun Kim, Changhyung Lee, Kyuseok Shim, "Time Series Prediction using Clustering Algorithm", Journal of Information Science, Vol. 20, No. 3, pp. 191-195, 2014.
- [14] Seunggho Mun, Jeonghyung Lee, "A Study on the Disaggregation Method of Time series Data",

Journal of Digital Policy, No. 12, No. 6, pp. 155-160, 2014

- [15] Sunho Min, Changho Seo, "The Study of Performance Improvement of Dejitter Algorithm applying Time Series Model for VoicePlatform Security Data", Journal of the Korea Institute of Information Security & Cryptology, Vol. 23, No. 5, pp. 963-968, 2013.

### 저자소개

이 준 연(Jun-Yeon Lee)

[정회원]



- 1992년 8월 : 중앙대학교 대학원  
컴퓨터공학과 (공학석사)
- 1992년 8월 ~ 1995년 8월 :  
Microsoft Ltd. Developer
- 2000년 2월 : 중앙대학교 대학원  
컴퓨터공학과 (공학박사)

· 2000년 3월 ~ 현재 : 동명대학교 미디어공학과 교수  
<관심분야> : 오피니언 마이닝, 큐레이션 시스템