A Novel Time Series Forecasting Approach with Multi-Level Data Decomposing and Modeling*

Xuemei Han, Congfu Xu, HuiFeng Shen and Yunhe Pan

College of Computer Science,
Zhejiang University,
Hangzhou, 310027, P.R.China

zjuwendy@yahoo.com.cn, xucongfu@cs.zju.edu.cn
useraddcn@yahoo.com.cn, panyh@zju.edu.cn

Abstract—Time series produced in complex systems are always controlled by multi-level laws, including macroscopic and microscopic laws. These multi-level laws bring on the combination of long-memory effects and short-term irregular fluctuations in the same series. Traditional analysis and forecasting methods do not distinguish these multi-level influences and always make a single model for prediction, which has to introduce a lot of parameters to describe the characteristics of complex systems and results in the loss of efficiency or accuracy. This paper goes deep into the structure of series data, decomposes time series into several simpler ones with different smoothness, and then samples them with multi-scale sizes. After that, each time series is modeled and predicted respectively, and their results are integrated finally. The experimental results on the stock forecasting show that the method is effective and satisfying, even for the time series with large fluctuations.

Index Terms—time series forecasting, complex system, data decomposing, multi-scale sampling

I. INTRODUCTION

Time series is one of the most frequently encountered forms of data. The forecasting of time series is an important topic in data mining since it plays great role in the economic decision making, prevention of natural calamity, and so on.

This paper pays more attention to the time series in complex systems, such as financial data, meteorologic data, etc. Since these data are controlled by multi-level laws, including macroscopic and microscopic laws, they are more stochastic than other types of data, such as engineering data and physical data. Statisticians, economists and mathematicians have paid much attention to the analysis of time series structure of complex systems. Most of economists consider that time series are the components of trends, cycles, seasonal variations and irregular fluctuations [1]. In 1951, Hurst [2] investigated the long-term storage capacity of reservoirs and found the long-memory trait of the hydrology series firstly. After 1980s, researchers found this trait was very common in time series of different areas [3].

On the other hand, Rosenblatt [4] presents the concept, strong mixing condition, which reflects the short-term process in time series.

Many methods are proposed to cope with time series forecasting, such as Box-Jenkins [5][6], Neural Networks [7][8], Genetic Algorithms [9][10], Kalman filter method [11], etc. These methods generally construct a single model with complicated parameters on the raw data when describing the complex system, while ignore preprocessing before modeling. However, since both long term trends and short term fluctuations coexist in the same sequence, it is a dilemma to balance the accuracy and the efficiency: discarding mass historic data which may be useful for analysis and forecasting will decrease the accuracy, while giving the same weight to the historic data will increase the processing time inevitably.

This paper proposes a new forecasting approach with multi-level decomposing and multi-scale sampling. In our method, the time series are decomposed into several simpler ones which are called as separated-series, then we sample every new separated-series with diverse scales. The new separated-series can be described with different models or the same model with different parameters. Then the forecasting of every new separated-series is conducted. The final forecasting results of the original series can be obtained by integrating those of separated-series.

The rest of this paper is organized as follows. In Section II, we give several definitions and formulations of time series preprocessing. Section III explains the modeling and forecasting processes. Experimental results of the approach are shown in Section IV. The last section offers our conclusion.

II. TIME SERIES PREPROCESSING

The data produced in complex systems are often influenced by multi-level factors, and the influence periods of these factors are diverse. Some factors may result in a long-term evolution, for example, the inherent value mechanism of stock, which is decisive in the long-term trends of stock price. While some factors’ durations are much shorter, such as people’s psychological factors in the stock market, which

---

* This paper is supported by National Natural Science Foundation of China (No. 60402010), Advanced Research Project of China Defense Ministry (No. 413150804), and partially supported by the Aerospace Research Foundation (No. 2003-HT-ZJDX-13).

Congfu Xu is the correspondent author.
may always cause short-term fluctuations. This requests multi-scale models which can deal with not only long-term trends forecasting but also short-term fluctuations prediction.

In this paper, original time series are decomposed into several ones with different cycles. The traits of the new time series are more visible and the new time series are easier to modeled. After decomposing, each new time series is sampled with different cycle which can bring the gain of efficiency without decreasing the accuracy. This section gives the formulation of the approach.

A. Definitions

We start by giving some definitions related to time series which may be convenient for describing our approach.

Definition 1 (standard sampling cycle): Standard sampling cycle, which is denoted as \( \delta \), is the original interval of recordings or statistics. This paper assumes that \( \delta \) is invariable for a given system, and the other sampling cycles are all integral multiple of \( \delta \).

Definition 2 (operator \( \oplus \) and separated-series): \( T \) is an ordered set \( \{t_1, t_2, \ldots , t_n, \ldots \} \) with \( t_{i+1} - t_i = \delta, i \in \mathbb{N} \), and \( 1 \leq i \leq n \). A time series \( X = \{x_t, t \in T\} \) can be decomposed to several time series and this operation is denoted as \( \oplus \)

\[
X = X(1) \oplus X(2) \oplus \ldots \oplus X(m),
\]

and the \( i \)-th position of the sequence \( X(j) \) satisfies

\[
\sum_{j=1}^{m} x(j)_{t_i} = x_{t_i}, \quad 1 \leq j \leq m.
\]

\( X(j) \) is named as separated-series which is equinumerous to \( X \), and the sampling time of the same order as \( X \) is the same.

B. Time Series Decomposing

In this subsection we present a method to decompose time series into several separated-series satisfying equation (1). This method, which we name as multi-smoothness-factor decomposing, is an inductive process as below:

a) Set smoothness-factor array: Given an array

\[
\mathcal{L} = \{l_1, l_2, \ldots , l_{m-1}\}
\]

where \( l_i \in \mathbb{N}, 1 \leq i \leq m - 1 \), \( \mathcal{L} \) is a monotonically decreasing sequence, for instance, a binary-base exponent array like \( \{2^6, 2^4, 2^2, 2^1\} \). \( \mathcal{L} \) can be determined by experiments or experiences.

b) Evaluation of \( X(1) \): By using smoothness-factor \( l_1 \), the \( p \)-th element in \( X(1) \) can be calculated by

\[
x(1)_{t_p} = \begin{cases} 
\frac{1}{p} \sum_{q=1}^{p} x_{t_q}, & p \leq l_1 + 1 \\
\frac{1}{l_1} \sum_{q=p-l_1}^{p} x_{t_q}, & p > l_1 + 1 
\end{cases}
\]

c) Evaluation of \( X(j), 1 < j < m \): Elements of time series \( X(j) \) can be gotten by subtracting elements of \( X(1), X(2), \ldots , X(j-1) \) from \( X \) with the same orders. Using smoothness-factor \( l_j \), the \( p \)-th element in \( X(j) \) can be calculated by

\[
x(j)_{t_p} = \begin{cases} 
\frac{1}{p} \sum_{q=1}^{p} x_{t_q}, & p \leq l_j + 1 \\
\frac{1}{l_j} \left( \sum_{q=p-l_j}^{p} x_{t_q} - \sum_{k=1}^{j-1} x(k)_{t_q} \right), & p > l_j + 1 
\end{cases}
\]

d) Evaluation of \( X(m) \): Elements of time series \( X(m) \) can be gotten by subtracting elements of \( X(1), X(2), \ldots , X(m-1) \) from \( X \) with the same orders.

Algorithm 1 shows the process of evaluating \( x(j)_{t_p} \).

\[
\text{Algorithm 1: Evaluate-Elements-of-X(j)}
\]

\[
\text{Data: } X, \mathcal{L} \\
\text{Result: } x(j)_{t_p}, 1 \leq j \leq m \\
\begin{align*}
\text{begin} & \\
& x(1)_{t_p} \leftarrow x_{t_p} \\
& \text{for } j = 1 \text{ to } m-1 \text{ do} \\
& \quad \text{if } p \leq l_j + 1 \text{ then} \\
& \quad \quad x(j)_{t_p} \leftarrow (x(j)_{t_p-1} \ast (p-1) + x_{t_p})/p \\
& \quad \quad \text{else} \\
& \quad \quad \quad x(j)_{t_p} \leftarrow (x(j)_{t_p-1} \ast l_j - x(j)_{t_p-l_j-1} + x_{t_p})/l_j \\
& \quad \quad \quad x_{t_p} \leftarrow x_{t_p} - x(j)_{t_p} \\
& \quad \quad x(m)_{t_p} \leftarrow x_{t_p} \\
\text{end}
\end{align*}
\]

Therefore, the original time series is decomposed into several ones with different traits. The number of \( X(j) \) can be determined by experiments, in most cases, 3 to 5 may be appropriate. The separated-series will turn from smooth to coarse with the increment of \( j \).

C. Multi-Scale Sampling

A typical decomposing result is shown in fig. 1. The original time series is decomposed into three separated-series by two smoothness factors. We can find that the separated series turn from smooth to coarse and the periods of fluctuations become shorter with the increment of \( j \). For the smoother separated-series, since they reflect the extending of long-term trends, their long-term history must be considered in modeling. However, each point in the smoother separated-series carries less information than that in the coarse ones. It is a waste of time if the smoother separated-series are sampled with standard sampling cycle \( \delta \). Multi-scale sampling has been discussed in recent years [12]. The essence is that the new data are sampled with high frequency and the older data with lower frequency in the same time series. This seems simple, but the variable measurements always make an extra trouble for the later processes.
This paper proposes a new multi-scale sampling method. We sample the same separated-series with the constant frequency, while the sampling frequencies are variable when handling the different separated-series. Let the ordered set \( \mathcal{S} = \{s_1, s_2, \ldots, s_m\} \) be the sampling cycle array, where \( s_j \) (1 \( \leq j \leq m \)) is the sampling cycle of \( X(j) \) and the elements of \( \mathcal{S} \) are monotonically decreasing. To be convenient for calculation, we choose \( s_j \) to be the sampling cycle array, where \( s_j \) (1 \( \leq j \leq m \)) is the sampling cycle of \( X(j) \) and the elements of \( \mathcal{S} \) are monotonically decreasing. To be convenient for calculation, we choose \( s_j \) to be the sampling cycle array, where \( s_j \) (1 \( \leq j \leq m \)) is the sampling cycle of \( X(j) \) and the elements of \( \mathcal{S} \) are monotonically decreasing.

**III. MODELING AND FORECASTING**

In the following subsections, we show the process of modeling and forecasting. The process of new data entering and other details related to modeling and forecasting are also discussed.

**A. New Data Entering**

The time series analysis and forecasting is always an online process. New data will enter continuously in the process of forecasting, and they should be processed at once. Algorithm 2 shows the update procedure of sampled separated-series when new data entering.

**Algorithm 2: New-Data-In**

**Data:** Sampled separated-series \( \{\hat{X}(j)\} \), \( x_t \), \( L \), \( \mathcal{S} \), \( \{\text{Count}_i[j]\} \) is set of sampling counters.

**Result:** \( \{X'(j)\} \) with new data, 1 \( \leq j \leq m \)

begin
  \hspace{1em} Evaluate-Elements-of-\( \hat{X}(j) \)
  \hspace{1em} for \( j = 1 \) to \( m \) do
    \hspace{2em} \text{Count}_i[j] \leftarrow \text{Count}_i[j] + 1
    \hspace{2em} \text{if} \ (\text{Count}_i[j] \mod \mathcal{S}[j] = 0 \text{ then} \\
    \hspace{3em} \hat{X}(j) \leftarrow \hat{X}(j) + \{x(j)_t\}
  \hspace{1em} \end{algorithm}

**B. Construct Forecasting Models**

Neural Networks, Genetic Algorithms, etc., can be employed for the separated-series modeling. Different models or one model with different parameters can be adopted for different separated-series. Here we exemplify the modeling and forecasting process by using Box-Jenkins method.

A stationary time series \( \{y_k\} \) of mean zero can be taken for responses of linear time-invariant random system with the input of white noise. Then \( y_k \) satisfies the difference equation

\[
y_k = \sum_{i=1}^{p} \phi_i y_{k-i} + \varepsilon_k - \sum_{j=1}^{q} \theta_j \varepsilon_{k-j}.
\]

which is denoted as ARMA\((p, q)\), where \( \sum_{i=1}^{p} \phi_i y_{k-i} \) is the weighted sum of the most recent \( p \) responses, and \( \sum_{j=1}^{q} \theta_j \varepsilon_{k-j} \) is the weighted sum of the recent \( q \) white noise. The estimation of \( \phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q \) and \( \sigma^2 \) can be obtained by the least squares method or maximum likelihood estimating method, whereafter, \( y_k \) can be predicated by equation 2. Details of constructing ARMA model can be found in [3].

Assume that there are \( H \) sampling points in \( \hat{X}(j) \), then the processes of modeling are:

1) Select the most recent \( h \) sampling points, \( h \leq H \). If \( \hat{X}(j) \) is not a stationary sequence, it must be transformed to be a stationary one by first or multistage difference method.

2) Construct an ARMA model for the sampling points with appropriate parameters. The discussion of determining the order of ARMA model can be seen in [13]. The reasonability of the model should be tested through the Box Pierce test [3].

3) Predict the value of next moment by using equation 2.

4) Add new data and compare the difference of new data with the prediction of the step 3. The residuals should be calculated.

The model should be verified periodically. For the time-variant complex system, the parameters may not be appropriate for prediction any longer with the passage of time. When the residuals can not satisfy presumable assumptions, the most recent \( h \) sampling points are selected and the model is updated through step 1 and step 2.

**C. Result Integration**

Now the prediction of each separated-series is gotten. The sampling cycle of each \( \hat{X}(j) \) is different, the prediction cycle turns shorter as the sampling frequency goes higher. They are very useful, for example, from the prediction of the longer sampling cycle, the potential trend of the time series can be seen. But for the short-term accurate prediction, the results of all the separated-series must be integrated to obtain the final forecasting result.

As shown in fig. 2, the final forecasting results should be integrated from those of all the frequency sampling sequences. Assuming to predict the result of the moment \( \tau \), the prediction
of $\tau$ of each separated-series should be obtained according to equation 1. Since the sampling cycles are different, not all the predictions of $\tau$ can be obtained directly, but the nearest observing and forecasting are available. Since these cases most happen on the smoother separated-series, the absent predictions can be calculated by polynomial interpolation simply. Therefore, we get the final result.

IV. EXPERIMENTS AND RESULTS

Stock data are chosen to validate our approach in this paper. Some forecasting results are demonstrated in Figures 3, 4 and 5. We choose 530 days close prices of Minsheng Bank for modeling and forecasting. The smoothness array $L = \{60, 15\}$, which decompose the original time series into three separated-series. As fig. 3 shows, the top left subplot is the fitting curve of the original time series, and the other subplots are the fitting curves of separated-series.

The sampling cycle array $S$ is chosen as $\{10, 5, 1\}$ according to the smoothness of each separated-series. The final fitting result is shown in fig. 4 (Close price as the vertical axis and time as the horizontal axis).

The curve of relative errors is shown in fig. 5. The mean of relative errors is 0.0105, and 92.83% relative errors are smaller than 3%. Compared with the approaches which are used for modeling with fixed parameters such as in [14][15], our approach gets better forecasting results.

V. CONCLUSION

In this paper, we propose a new approach for time series analysis and forecasting. Through multi-level decomposing, the complex time series turns to be a series of simpler ones, then the separated-series are sampled according to their smoothness and modeled respectively. Treating the long-term trends and short-term fluctuations with different weight bring on gains of accuracy and efficiency.

REFERENCES


