A New Approach for Short-term Time Series Forecasting

Henghai Fan¹, Shuai Wu²*, Ning Chen¹, Bo Gao¹ and Yiming Xu²
¹Xi’an Satellite Control Centre, China
²School of Computer Science, Fudan University, Shanghai, China

* wushuaixl@sina.com

Abstract. Short-term series forecasting is one of the essential issues in a variety of tasks, such as traffic flow prediction, stock market tendency analysis, etc. Most current methods based on stable or abundant historical data. In this paper, we proposed a novel model called LA-NN. It takes advantages of both long short-term memory (LSTM) network and autoregressive integrated moving average (ARIMA) by a relation integration of them. So as to deal with the situation of insufficient historical data and sudden abnormal changes in data. A comparison with other representative forecast models validates that the proposed LA-NN network can achieve a better performance.

1. Introduction
A time series is a set of observations $x_t$, each one being recorded at a specific time $t$ [1]. Time series forecasting is a task of collecting past observations, capturing the underlying pattern and then to predict the future. While the objective of short-term time series forecasting is to develop a predictor which could extrapolate the time series into the near future, which ranges from several minutes to dozens of minutes.

Time series data widely exist in various fields, such as transportation, medical, economic, energy, and so on. Time series forecast can promote the development of various fields. At present, many researchers have applied time series forecast technologies to various fields, such as economic forecasting [2], traffic speed forecasting [3], patient flow forecasting [4], electricity-related forecasting [5], and so on.

Although there are many correlational researches, many challenges still exist. At present, methods either need to be improve accuracy or base on massive historical data. But due to some realistic factors, historical data may be insufficient or irregular variations (or unstable pattern), which we simply called mutation period, will emerge in the time series. This characters lead to difficulties in predicting such period of time series. Different from conventional time series forecasting, the accuracy of short-term forecasting in mutation period is extremely important. Therefore, establish a forecasting model that can answer insufficient history data and mutation period is the main task of this work.

To address above issues, we propose a more effective model, called LA-NN. The main contributions of this paper can be summarized as two aspects. Firstly, LA-NN integrate LSTM and ARIMA in a rational way to make use of their advantages. Secondly, we evaluate our model on real traffic data.

The remainder of this paper is organized as follows. Section 2, introduces a general overview of existing literatures on time series forecasting. The methodology is introduced in Section 3. Experiments are shown in Section 4. Section 5, concludes the paper and suggests future directions.
2. Related work
Over the past few decades, a lot of researches and effort has been devoted to develop and improve short-term time series models. ARIMA family of models are quite flexible, they need less data and can accurately predict linear correlation sequence data. But no nonlinear patterns can be captured by the ARIMA models. ARIMAs to complex real-world problem is not always satisfactory. ARIMA and its variants are widely used in time series data. In [6], a new approach called ARIMA-GARCH was used to forecast ridership time series data.

Time series forecast base on deep learning approaches has become a new trend. Compared with conventional learning architectures, deep neural network can model complex non-linear relationship by using distributed and hierarchical feature representation [7]. So far, many NNs variants have been proposed to assist time series forecasting. In recurrent neural network (RNN) has been used in [8] to predict traffic speed time series data. Among all NNs, RNN is widely regarded as a suit model to capture underlying evolution of time series. But many researchers proved that RNN fail to capture long term evolution. To address this problem, long short-term memory (LSTM) are used to predict time series data. In term of time series date, LSTM is excellent. Many LSTM variants or hybrid methods are applied in time series forecast. In [9], a LSTM based deep learning approach is applied in traffic forecast. In [10], a hybrid methods called is proposed to predict traffic speed. The basic idea is to feed the features of surrounding areas extracted by CNN to RNN for learning time-series patterns.

Distinct from other prediction approaches, this paper takes into account the shortage of historical data and sudden changes in data and constructs a new NN model called LA-NN with multi layers based on ARIMA and LSTM. New approach has a better performance on accuracy even in the case of insufficient historical data or sudden changes in data.

3. Methodology
Short-term forecast is a task of modelling underlying patterns of time series, which includes both regular and irregular variations. The major challenge is to predict these irregular variations or unstable patterns. This paper deals with the problem with a combing model, which models ether linear or non-linear relationship in order to make a reliable forecast result. The proposed short-term forecast model is based on the available technologies, which include ARIMA algorithm, RNNs, LSTM network. The detail of the methodology will be explained in this section.

ARIMA can use very little historical data to predict, but it requires stable data. LSTM can mine time series patterns, but it needs magnanimous historical data. Time series data exist widely, so there are two issues. First, when we need to predict time series, we don't necessarily have enough historical data. Second, when the time series data changes abnormally, the prediction results of the prediction model trained based on historical data are poor. In order to address this situation, a hybrid neural network algorithm called LA-NN based on ARIMA and LSTM is proposed in this paper. The structure is as figure 1.

Figure 1. The structure of LA-NN

Figure 1 presents the architecture of LA-NN, which is comprised of one ARIMA layer, one LSTM layer and a fully connected layer, predicting base on small quantities data, modeling long temporal dependency and adjusting weights of $y_{ARIMA}$ and $y_{LSTM}$ respectively. For improve the efficiency and accuracy, the length of the input series can different, because LSTM may need long series to model
long temporal dependency, ARIMA may need short series to address sudden changes in data. Finally, a fully connected layer consisting of multiple activation functions is constructed to adjust weights of $y_{ARIMA}$ and $y_{LSTM}$.

4. Experiment

4.1 Data preparation
The proposed time-series data forecast model was applied to the traffic speed time-series data collected by Shanghai Traffic Information Center as a numerical example. We get the traffic speed of days from 01 May 2017 to 30 August 2017 as original dataset, including 598 viaduct sections with a frequency of 2 min.

In addition, in order to verify the ability of the new model to cope with sudden changes in data, some modifications were made so that some data did not conform to the normal pattern. The selected dataset were divided into four subsets: a small training dataset containing several days of data, a bigger training dataset containing several weeks of data, a test dataset containing several days of data, a test dataset containing modification data.

4.2 Benchmarks
We compare our model with the following baseline methods.

ARIMA: This model [11] is a generalization of autoregressive moving average model with an initial differencing step applied to remove the non-stationarity of the data.

LSTM: LSTM can effectively handle long temporal dependence and reduce gradient vanishing which is appropriate for speed prediction[12].

4.3 Evaluation for forecast result
Three criteria are commonly used to evaluate the performance of traffic forecast model. They are mean absolute error (MAE), root mean square error (RMSE) and mean relative error (MRE).

4.4 Experiment result
The experimental data we selected are periodic. Most of the time, the data has certain regularity, some time period data is larger, some time period data is smaller. Figure 2 show our data intuitively. As mentioned in Section 4.1, part of the test data in this experiment is used to test the prediction accuracy of the algorithm for abnormal mode data, so some of the data are modified. Figure 3 is the modified data of figure 2, and the blue part is the modified data.

![Figure 2. Eight cycles of data](image1)

![Figure 3. One modified data](image2)

First, determine the parameters of ARIMA. The parameters of ARIMA model are determined as follows: $p=2$, $d=1$, $q=0$. Second, Train LA-NN Model and LSTM Model. Each model is trained with small training set. Third, Test ARIMA, LA-NN Model and LSTM Model on normal data and modified data respectively. Evaluate prediction results with MRE, RMSE, MAE. The values given in this paper are the average values of multiple time series data. In order to display the experimental results more intuitively, this paper presents the experimental results of one of the time series data with pictures. The red line is the real value and the blue line is the prediction result.
4.4.1 Experiments on modified data. One of the issues to be solved by the new algorithm is when the law of time series data changes greatly for some reason, the accuracy of time series data is still high. In this stage, the experiment is to predict the blue part data of figure 3. Table 1 show that the accuracy of LAN is higher than that of the other two algorithms. Intuitive prediction results can be seen in figure 4, figure 5, figure 6.

| Algorithm | MAE | RMSE | MRE |
|-----------|-----|------|-----|
| ARIMA     | 6.70| 9.37 | 0.14|
| LSTM      | 8.89| 10.43| 0.18|
| LA-NN     | 7.47| 8.92 | 0.14|

4.4.2 Experiments on normal data when training data is insufficient. Another issue we want to address is that models are not well trained when historical data is insufficient. In this stage, the model is trained by using small training sets, and the predicted results are shown in table 2. Because ARIMA model needs less historical data, its accuracy has not been reduced. When training LAN and LSTM models, the same small data sets are used. The results show that the prediction accuracy of LAN is better than that of LSTM, because ARIMA is integrated in LAN, and ARIMA still has higher accuracy when historical data is insufficient. Intuitive prediction results can be seen in figure 7, figure 8.
Table 2. Forecast performances of different algorithms for normal data when training data is insufficient

| Algorithm | Evaluation Criteria |
|-----------|---------------------|
|           | MAE    | RMSE   | MRE    |
| ARIMA     | 7.02   | 9.88   | 0.13   |
| LSTM      | 9.77   | 13.12  | 0.19   |
| LA-NN     | 6.53   | 8.32   | 0.10   |

4.4.3 Experiments on modified data when training data is insufficient. Finally, the modified data is predicted using the LAN model and LSTM model trained by ARIMA and small data sets. As shown in table 3, the accuracy of LANN and LSTM decreases in varying degrees, but LANN can still maintain high accuracy. Intuitive prediction results can be seen in figure 9, figure 10.

Figure 9. LSTM experimental results on modified data when training data is insufficient  
Figure 10. LANN experimental results on modified data when training data is insufficient

5. Conclusion
In order to deal with the problem that the accuracy of prediction model decreases due to insufficient training data and sudden change of time series data, a new neural network model, LANN, is proposed in this paper. Experiments are carried out in four cases. The results show that LANN has higher prediction accuracy and stronger adaptability, and can maintain higher accuracy in various situations.

Table 3. Forecast performances of different algorithms for modified data when training data is insufficient

| Algorithm | Evaluation Criteria |
|-----------|---------------------|
|           | MAE    | RMSE   | MRE    |
| ARIMA     | 6.70   | 9.37   | 0.14   |
| LSTM      | 13.98  | 16.98  | 0.35   |
| LA-NN     | 9.39   | 12.07  | 0.25   |

Reference
[1] Shanmugam R, Brockwell P J, Davis R A. Introduction to Time Series and Forecasting[J]. Technometrics, 1997, 39(4):426.
[2] Koceska N, Koceski S. Financial-Economic Time Series Modeling and Prediction Techniques – Review[J]. Journal of Applied Economics & Business, 2014.
[3] Lin L, Li J, Chen F, et al. Road Traffic Speed Prediction: A Probabilistic Model Fusing Multi-Source Data[J]. 2018.
[4] Kadri F, Harrou F, Chaabane, Sondès, et al. Time Series Modelling and Forecasting of Emergency Department Overcrowding[J]. Journal of Medical Systems, 2014, 38(9):107-1235.
[5] Hong T, Pinson P, Fan S, et al. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond [J]. International Journal of Forecasting, 2016: S0169207016000133.

[6] Ding C, Duan J, Zhang Y, et al. Using an ARIMA-GARCH Modeling Approach to Improve Subway Short-Term Ridership Forecasting Accounting for Dynamic Volatility [J]. IEEE Transactions on Intelligent Transportation Systems, 2017: 1-11.

[7] Huang W, Song G, Hong H, et al. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5): 2191-2201.

[8] Tang J, Liu F, Zou Y, et al. An Improved Fuzzy Neural Network for Traffic Speed Prediction Considering Periodic Characteristic [J]. IEEE Transactions on Intelligent Transportation Systems, 2017: 1-11.

[9] Zhao Z, Chen W, Wu X, et al. LSTM network: a deep learning approach for short-term traffic forecast [J]. IET Intelligent Transport Systems, 2017, 11(2): 68-75.

[10] LC-RNN: A Deep Learning Model for Traffic Speed Prediction

[11] Bartholomew D, Box G E P, Jenkins G M. Time Series Analysis: Forecasting and Control [J]. Journal of the Operational Research Society, 1971, 22(2): 199-201.

[12] Ma X, Tao Z, Wang Y, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data [J]. Transportation Research Part C: Emerging Technologies, 2015, 54: 187-197.