Missing Data Prediction Using LSTM for Radio Frequency Transfer System

Jiahui Cheng[1], Hao Gao[1], Chenxia Liu[2], Zhuoze Zhao[1], Yaojun Qiao[1,*], Bin Luo[1,*], Song Yu[1] and Hong Guo[3]

[1]State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

[2]Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071066, China

[3]State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics, and Center for Quantum Information Technology, Peking University, Beijing 100871, China

E-mail: qiao@bupt.edu.cn, luobin@bupt.edu.cn

***Abstract*—Missing data affects stability analysis of Radio Frequency (RF) transfer system. In this paper, a missing data prediction method using Long Short-Term Memory (LSTM) network is proposed for the short-term failure in RF transfer system. The performance is verified by experimenting on the measurement data of 2.4 GHz RF transfer system using 3000 km single optical fiber bi-directionally. The results demonstrate that our proposed method can achieve higher prediction accuracy compared to back propagation neural network (BPNN) and Support Vector Regression (SVR). When the prediction time is 240 s, the Root Mean Squared Error (RMSE) between the predicted value and the actual value is less than 0.001. This provides an effective and feasible method to impute missing data for RF transfer system.**

Keywords—frequency transfer; missing data prediction; Long Short-Term Memory network

I. Introduction

High stable radio frequency (RF) transfer via optical fiber has important applications in many fields, such as atomic clock comparison [1], navigation and positioning [2] and fundamental physics research [3]. During the long-term measurement of the RF transfer system, the frequency signal can't be continuously transmitted to the receiver due to the reasons such as system power down and phase locked loop (PLL) loses lock. This leads to the measurement data missing and affects the stability analysis of RF transfer system. A number of methods have been devised for dealing with missing data. Mean-imputation method (MIM) is the most commonly method, which is simple but not very effective [4]. Some machine learning algorithms such as Support Vector Regression (SVR) [5] and back propagation neural network (BPNN) [6] are used to predict missing data, which can get higher processing accuracy than MIM. Although they can learn nonlinear relationships in the data, they cannot model the temporal dependency explicitly. Considering that the measurement data of RF transfer system is the nonlinear data with time dependence, it is important to find a method that can learn the time dependence of the nonlinear data to impute the missing data of RF transfer system.

We propose a missing measurement data prediction method for RF transfer system using Long Short-Term Memory (LSTM) network. In the training phase, the measurement data during the normal operation of the system is used to train the LSTM network. In the prediction phase, the constructed LSTM model is used to predict missing data for a period of time in the future. By experimenting on the measurement data of 2.4 GHz RF transfer system using 3000 km single optical fiber bi-directionally, the result show that the proposed method can achieve higher prediction accuracy compared to the BPNN and SVR. When the prediction time is 240 s, the Root Mean Squared Error (RMSE) between the predicted value and the actual value is less than 0.001.

II. Methods and Results

LSTM network is an extension of recurrent neural network (RNN), which effectively overcomes the problem of vanishing gradients that exists in RNN [7]. LSTM network can not only learn the series correlation in nonlinear series prediction problems, but also memorize information for a long time. Therefore, we propose a missing measurement data prediction method for RF transmission system using LSTM network and verify by experimental data. We obtain measurement data by transmitting the 2.4 GHz RF signal over 3000 km single optical fiber bi-directionally. During the training phase, the obtained data is used to construct the LSTM network. During the prediction phase, the constructed LSTM model is used to predict the missing data.

The simplified schematic diagram of the RF transfer system is shown in Fig. 1 [8]. The system can synchronize the 2.4 GHz RF signal at the remote site (RS) with the 2.4 GHz RF signal at the local site (LS), which is locked to a rubidium clock (Quartzlock, A1000). The phase noise cancellation process mainly includes the following steps. At first, the 2.4 GHz RF signal at RS is intensity modulated on the optical carrier (1550.12nm, 34#) and transmitted to LS via single-mode fiber (SMF). The optical/electrical (O/E) converter at the LS can detect the 2.4GHz RF signal carrying the phase noise caused by fiber-optic link. Then, the PLL and two auxiliary RF signal generators (2.3 GHz, 2.5 GHz) complete the phase conjugation at the LS. Finally, the phase conjugate signal is modulated on the optical carrier (1550.92nm, 33#) and transmitted to the RS, which is used to cancel the phase noise of the fiber link. The PLL at the RS acts as signal regenerator, it improves the signal to noise ratio (SNR) of the transmitted signal. The phase difference of the 2.4 GHz RF signal between the LS and the RS can be measured by the voltmeter. The measured voltage data is used to analyze the stability of the RF transfer system.

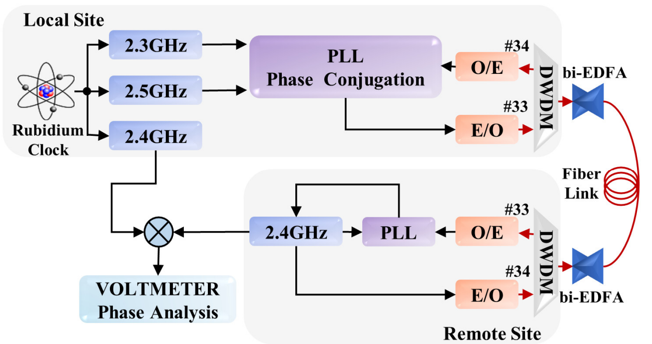

Fig. 1. Simplified schematic diagram of the RF transfer system. PLL: phase locked loop. DWDM: dense wave-length division multiplexing. bi-EDFA: bi-directional erbium-doped fiber amplifier. O/E: optic/electro converter. E/O: electro/optic converter.

1200 s of measurement data are used to train, verify and test the LSTM network, and the ratio of training set, validation set and test set is 6:2:2. The prediction result is shown in Fig. 2. The gray and orange curves represent the actual and predicted values of the missing data, respectively. The results indicate that our proposed method can achieve high prediction accuracy. Within the time range of 240 s, the predicted result well matches the actual value. The bule sphere represents the RMSE of the LSTM network at different prediction times. The smaller the RMSE between the predicted value and the actual value, the higher the prediction accuracy of the model. It can be seen that as the prediction time increases, the prediction accuracy of the model gradually decreases.

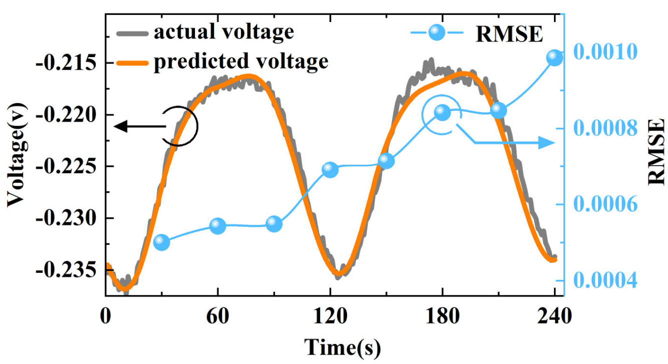

Fig. 2. The prediction result of LSTM network.

We compared the prediction performance of the LSTM prediction model with BPNN and SVR in Fig. 3. The blue and orange curves represent the RMSE and Decision Coefficient (R Squared, R^2) of different models, respectively. When RMSE is close to 0, the prediction accuracy of the model increases. When R^2 is close to 1, the correlation between the predicted value and the actual value becomes stronger. The metrics in Fig. 3 show that the LSTM network can achieve higher prediction accuracy compared to other models. When the prediction time is 240 s, the RMSE between the predicted value and the actual value is less than 0.001, the R^2 between the predicted value and the actual value is higher than 0.98.

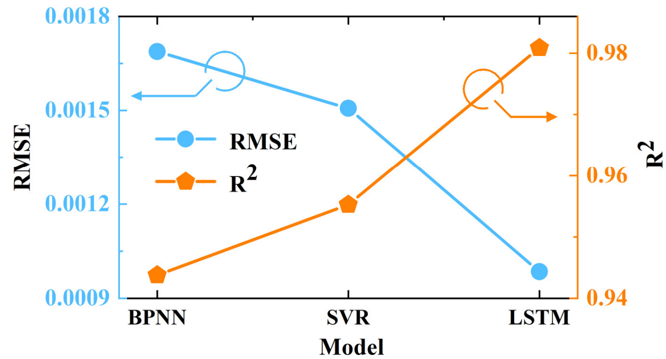

Fig. 3. The performance metrics for different models.

III. Conclusions

In this paper, we propose a missing data prediction method using LSTM network for the short-term failure in RF transfer system. By experimenting the measurement data of 2.4 GHz frequency signal transmitted via 3000 km single optical fiber bi-directionally, the results show that our proposed method can achieve higher prediction accuracy compared to BPNN and SVR. When the prediction time is 240 s, the RMSE between the predicted value and the actual value is less than 0.001. Our proposed model can improve the utilization rate of measurement data in RF transfer system.

Acknowledgment

This paper was supported in part by the National Natural Science Foundation of China under Grant 62271080; in part by Fund of State Key Laboratory of IPOC (BUPT) under Grant IPOC2022ZT06; and by the Beijing University of Posts and Telecommunications (BUPT) Excellent Ph.D. Students Foundation (CX2022112).

References

[1] M. Cizek, L. Pravdova, T. M. Pham, et al., “Coherent fibre link for synchronization of delocalized atomic clocks,” Opt. Express, vol. 30, pp. 5450-5464, February 2022.

[2] C. Clivati, G. Cappellini, L. F. Livi, et al., “Measuring absolute frequencies beyond the GPS limit via long-haul optical frequency dissemination,” Opt. Express, vol. 24, pp. 11865-11875, May 2016.

[3] R. Santagata , D. B. A. Tran, B. Argence, et al., “High-precision methanol spectroscopy with a widely tunable SI-traceable frequency-comb-based mid-infrared QCL,” Optica, vol. 6, pp. 411-423, April 2019.

[4] M.N. Noor, A.S. Yahaya, N.A. Ramli, et al., “Mean imputation techniques for filling the missing observations in air pollution dataset,” Trans Tech Publications Ltd, December 2013.

[5] Q. Shang, Z. Yang, S. Gao and D. Tan, “An imputation method for missing traffic data based on FCM optimized by PSO-SVR,” Journal of Advanced Transportation, vol. 2018, pp. 1-21, January 2018.

[6] J. Wang, M. Anisetti, G. Jeon, “Reconstruction of missing color-channel data using a three-step back propagation neural network,” International Journal of Machine Learning and Cybernetics, vol. 10, pp. 2631-2642, July 2019.

[7] B. Su and E. M. Gutierrez-Farewik, “Gait Trajectory and Gait Phase Prediction Based on an LSTM Network,” Sensors, vol. 20, pp. 7127-7143, December 2020.

[8] C. Liu, S. Zhou, Z. Zhao, et al., “Ultrastable long-haul fibre-optic radio frequency transfer based on PLL frequency mixing,” IFCS-ISAF, pp. 1-2, 2020.