The Research of Low Earth Orbit Prediction of Satellite Based on Deep Neural Network

Hai-tao YANG¹, Jun-peng ZHU²*, and Jian ZHANG³

¹Science and Technology on Complex Electronic System Simulation Laboratory, Space Engineering University, Beijing, China
²Department of Graduate Management, Space Engineering University, Beijing, China
³Department of Graduate Management, Space Engineering University, Beijing, China

*Corresponding author

Keywords: Satellite orbit prediction, Deep neural network.

Abstract. Satellite orbit prediction is a basic requirement in satellite applications. The current orbit prediction mainly depends on the dynamic model. Because of the limitations of the detection equipment and the satellite orbit data cannot be updated in time, which cause the dynamical model long-term orbit divergence to be serious. Using deep neural network as a method of orbit prediction which can predict the future data by training the satellite orbit data and grasp the implicit relationship between the data. The neural network model is optimized and the prediction data is compared with the actual data. The error of 20 days forecast is reduced to 2km, which improves the accuracy of neural network forecasting satellite orbit.

Introduction

In area of Space application, the analytical orbit model is widely used. Orbit prediction were usually achieved through the two lines elements (TLE) provided by North American Aerospace Defense Command (NORAD). By specific method, the average orbit element of cycle disturbance was removed from TLE. While algorithms of TLE have not been public, and part of the satellites TLE data are not released in time, the precision of satellites orbit forecast remain low [1].

With the development of Deep Neural Networks in recent years, the neural network can deal with the nonlinear and high dimensional problem without knowing the structure and relevant parameters between its input and output functions. To find out inner connection of input and output functions through the training of the neural network learning, and training set data were input, the correct mapping relation of neural network will be acquired.

In recent years, the neural network show the enormous effect in the area of image and video identification, stock analysis. The successful application of deep neural network in pattern recognition also shows its effectiveness in uncertain system of mathematical models. Deep long-short memory neural network model is introduced into the study of satellite orbit prediction by this paper, to get rid of the dynamic model of the neural network and to carry on the forecast by learning to get the data of the implicit rule, so as to improve the lack of dynamic prediction model.

Basic Structure

In this paper, the long and short-term memory neural networks are mainly used for orbit prediction research because of their time-series structure. The depth is mainly between the inputs and outputs of the sequential logic, and each sequence can be understood as a single training unit [2]. The training results at the next moment will be affected by the weights of the parameters at the previous moment. As shown in Figure 1.
When the hidden layer receives input $X_t$ at time $t$, the value of the hidden layer is $S_t$, and the output is $O_t$, the $S_t$ which depends not only on $X_t$, but also on $S_{t-1}$, which is calculated as shown in the equation 1 and 2.

$$O_t = \sum_{j=1}^{n} W_{ij} g( VS_t )$$  \hspace{1cm} (1)$$

$$S_t = f( UX_t + WS_{t-1} )$$  \hspace{1cm} (2)$$

Equation 1 is the calculation formula of all output layers. The output layer is a fully connected layer, that is, each of its time nodes is connected with each time node of the hidden layer. $V$ is the output layer weight matrix, $g(\cdot)$ is the activation function. Equation 2 is the calculation formula of hidden layer, which is the cyclic layer, $U$ is the weight matrix of input $X$, $W$ is the weight matrix input $S_{t-1}$ at the previous moment and $f(\cdot)$ is the activation function.

The long short-term memory neural network hidden layer for the length of the short-term memory module, which consists of input layer, hidden layer, output layer [3]. It differs from recursive neural networks in that the hidden layer contains a storage module that stores information, and the storage unit in the storage module has self-connection, self-adaptation and logic gates to control the flow of information. There are two logic gates at the input and output layers that control the activation and recall of memory modules at the input and output layers, as shown in Figure 2.

The main purpose of long short-term memory neural network is to predict the next time-series data under the condition of a priori information rather than how many steps to forecast the data. In order to achieve this goal, the input timing data will be iteratively calculated by the following formula. The equation is as follows [4].

$$i_t = \sigma(W_{ix_t} x_t + W_{im_{t-1}} m_{t-1} + W_{ib} c_{t-1} + b_i)$$  \hspace{1cm} (3)$$

$$f_t = \sigma(W_{fx_t} x_t + W_{fm_{t-1}} m_{t-1} + W_{fc} c_{t-1} + b_f)$$  \hspace{1cm} (4)$$
\[ o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \]  \hspace{1cm} (5) \\
\[ c_t = f_t \odot c_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1} + b_c) \]  \hspace{1cm} (6) \\
\[ m_t = o_t \odot h(c_t) \]  \hspace{1cm} (7) 

In long short-term neural network, we first construct three gates to control the flow of information, \(i_t, f_t\) and \(o_t\), respectively, input gate, forgotten gate and output gate, different gates have different effects on information flow control, details as follows \cite{5}:

- **Input Gate**: Mainly controls the amount of information flowing into a memory cell.
- **Forgotten gate**: Mainly control the amount of memory cells accumulated in the previous moment to the current moment.
- **Output Gate**: The main control at some point in the memory cell flow of information can flow into the current hidden state \(m_t\).

\(c_t\) is the memory cell activation function, \(x_t\) is the input eigenvector, \(m_t\) is the hidden output vector. Equation (6) shows the accumulation of historical information in the long short-term memory neural network structure, mainly through the self-connection of memory cells. In the accumulation process, forgotten gate can limit the last moment of the storage unit information, and rely on the input gate to limit the new information. This also reflects that the self-connection of memory cells is linearly cumulative.

Equation (7) shows that the current hidden state \(m_t\) is calculated from \(c_t\) which is self-updating in a linear way. It is added to \(h(\cdot)\) with non-linearity first, and then filtered by output \(o_t\) to get the current hidden state \(m_t\).

**Analysis and Optimization**

**Analysis**

In order to verify the performance of deep neural network in satellite orbit prediction, the long short-term memory neural network prediction is carried out based on the actual data of X, Y and Z coordinates of the satellite's geocentric inertial coordinate system. The experimental object was the TacSat2 US scientific imaging satellite, launched in December 2006 with orbital height of 413KM × 424KM near-circular orbit, and the experiment was performed at UTC time from 12:00:00 on July 1, 2007 to October 8, 2017 12:00:00 total of 100 days of data as a learning sample. The 100-day satellite orbital X, Y, Z values are calculated by STK (satellite Tool Kit) High Precision Orbital Predictor (HPOP) and the learning samples are sampled at 1-minute intervals for the next 20 days. The results of X-axis prediction are partially shown in Figure.3. The error analysis mainly records the maximum absolute error and training time according to the training times. As shown in Table 1.

![Figure 3. The results of X-axis prediction.](image)

<table>
<thead>
<tr>
<th>Epoch (20</th>
<th>50</th>
<th>100</th>
<th>150</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum absolute error</td>
<td>650km</td>
<td>608km</td>
<td>486km</td>
<td>524km</td>
</tr>
<tr>
<td>Training time</td>
<td>600s</td>
<td>1500s</td>
<td>3000s</td>
<td>4500s</td>
</tr>
</tbody>
</table>

It can be seen from the graph analysis that the prediction data changes cyclically with the real data, which shows that the learning ability of long short-term memory neural network can learn the
inherent relationship between the cyclical changes of the original data through learning. However, the
effect of prediction data fitting with real data is poor, and the accuracy of forecasting data is not
enough for satellite orbit prediction. Table 1 shows that the training error is better than 20,50 times in
100 times, but when the number of training levels exceeds 100 times, the accuracy of training results
rebounds again at 150 times and 200 times, causing the training error to rise, and the increase of the
layers also makes the training time also increases. Neither timeliness nor accuracy can meet the
demand for satellite orbit prediction.

**Optimization**

For small-scale sample data, forgotten gate to open the storage of implicit relations and parameters
can improve the training accuracy, and the impact on the timeliness is not obvious. However, for
large-scale sample data, the hidden layer output information should be stored at each moment. As the
training proceeds, the storage unit will store the implicit relationship and parameters at each moment.
Due to the limited storage space of the storage unit, eventually resulting in new hidden relationships
and parameters can not be stored.

The improved long short-term memory neural network is an extension of the long short-term
memory model structure. As the improved idea is to close the forgotten gate completely, which can
result in the storage unit unable to store the related parameters. Therefore, this section through the
long short-term memory module to add a time gate K to achieve a flexible closure of the storage unit
or open function. The model structure is shown in Figure 4.

![Image](image.jpg)

Figure 4. The improved structure of the long and short-term memory model.

The opening and closing of this gate is controlled by an independent rhythmic oscillation specified
by three parameters; The first parameter $\tau$ controls the period of the real-time pulse; the second
parameter $r_{on}$ controls the proportion of the time gate that continues to open during the entire period;
and the third parameter $s$ controls the phase shift of each memory cell pulse in the modified. These
parameters should be learned during the training process. Although the changes may occur, the time
gate still has a mature linear formula to describe, as shown in Equation 2-20 [6].

$$\varphi_i = \frac{(t - s) \mod \tau}{\tau}$$

(8)

$$K_i = \begin{cases} 
\frac{2\varphi_i}{r_{on}}, & \text{if } \varphi_i < \frac{1}{2} r_{on} \\
2 - \frac{2\varphi_i}{r_{on}}, & \text{if } \frac{1}{2} r_{on} < \varphi_i < r_{on} \\
\alpha \varphi_i, & \text{otherwise}
\end{cases}$$

(9)
\( \varphi \) is a transitional variable, representing the various stages in the transition period. There are three main phases of the time gate \( K_t \), the first phase of the time gate from 0 to 1; the second phase from 1 to 0; in the third phase, the time gate closed to maintain the current state of the storage element.

The loss rate \( \alpha \) plays an active role in the closing phase of time gate \( K_t \). When parameters are lost in linear cells that convey important gradient information, the time gate will be closed, thus completely avoiding the phenomenon of gradient disappearance. Because the time gate is closed, the information in the current storage unit is also better preserved, ensuring sample data training accuracy.

**Experiment and Evaluation**

**Experiment**

During the experiment, the experimental samples are the same as the previous ones. The data of X, Y and Z axes in the next 5 days, 10 days and 20 days are predicted, and the results of X-axis prediction are still displayed.

![Figure 5. The results of X-axis prediction.](image)

**Evaluation**

The orbital prediction assessment mainly focuses on the quantitative analysis of the pros and cons of orbital forecasting methods. In this paper, satellite orbit prediction is the key to the analysis of the advantages and disadvantages of the improved long short-term memory neural network in orbit prediction. As the error analysis is generally used as the main method to evaluate the prediction results in orbit prediction analysis, error analysis can be used as the main basis for constructing the evaluation framework.

Combined with the existing orbit error analysis method, it is classified into two types of analysis: contrast and confidence. These two types of assessment can reflect the performance of the improved long short-term memory neural network on the orbit prediction of key nodes from different angles.
Comparative evaluation involves the comparison of methods, so the data error, the mean square error and error improvement rate as the main indicators of evaluation. Confidence assessment mainly based on the confidence of the error data and confidence interval as an assessment indicator.

\[
e = \hat{x} - x
\]

(10)

\(e\) is the error value, \(\hat{x}\) is the prediction value, \(x\) is the sample value.

In the orbital prediction assessment, the criteria for assessing the prediction error mainly include the mean value, the mean square value and the maximum value. The mean value refers to the average value of the absolute value of the prediction error in a certain period of time. The mean square value refers to the average value of the square of the error in a certain period of time, and the maximum value is the maximum value of the error in the prediction period. Considering that the mean square value is usually used to characterize the error distribution and has good persuasiveness, the mean square error of the satellite orbit prediction error is used in the analysis of the satellite orbit prediction error, and the equation is as shown in (11).

\[
MSE = \frac{1}{M} \sum_{i=1}^{M} (x - \hat{x})^2
\]

(11)

\(M\) is the number of points involved in the calculation in a specific time period, \(x\) is the true value, \(\hat{x}\) is the predicted value.

This article focuses on improved long short-term memory neural networks and mathematical models to improve the rate of comparison. In the paper, the expression of prediction error improvement rate is as shown in equation (12).

\[
imp \_rate = \frac{Max_{Dyn} - Max_{NN}}{Max_{Dyn}}
\]

(12)

Where \(imp \_rate\) is the improvement rate of prediction error; \(Max_{Dyn}\) is the maximum error value predicted by mathematical model, \(Max_{NN}\) is the maximum error value of the improved neural network. According to Equation (13), the mean square error of each prediction period of mathematical model, long short-term memory neural network and improved long short-term memory neural network is calculated. The calculation results are shown in Table 2-4.

**Table 2. Mean square error of long short-term memory neural network**

<table>
<thead>
<tr>
<th></th>
<th>X/KM</th>
<th>Y/KM</th>
<th>Z/KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>101.21</td>
<td>94.89</td>
<td>89.56</td>
</tr>
<tr>
<td>Medium-term prediction</td>
<td>132.23</td>
<td>157.78</td>
<td>127.32</td>
</tr>
<tr>
<td>Long-term prediction</td>
<td>189.36</td>
<td>192.45</td>
<td>162.78</td>
</tr>
</tbody>
</table>

**Table 3. Mean square error of mathematical model**

<table>
<thead>
<tr>
<th></th>
<th>X/KM</th>
<th>Y/KM</th>
<th>Z/KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>0.43</td>
<td>0.58</td>
<td>0.41</td>
</tr>
</tbody>
</table>
Table 4. Improved mean square error of long short-term memory neural network

<table>
<thead>
<tr>
<th></th>
<th>X/KM</th>
<th>Y/KM</th>
<th>Z/KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>1.04</td>
<td>1.22</td>
<td>0.84</td>
</tr>
<tr>
<td>Medium-term prediction</td>
<td>1.13</td>
<td>1.42</td>
<td>0.86</td>
</tr>
<tr>
<td>Long-term prediction</td>
<td>1.56</td>
<td>2.01</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 5 shows the improved accuracy of the long short-term memory neural network with improved satellite prediction.

Table 5. Neural Network Improvement of orbit accuracy

<table>
<thead>
<tr>
<th></th>
<th>X/%</th>
<th>Y/%</th>
<th>Z/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>-58.65%</td>
<td>-52.46%</td>
<td>-51.19%</td>
</tr>
<tr>
<td>Medium-term prediction</td>
<td>5.31%</td>
<td>4.23%</td>
<td>-4.65%</td>
</tr>
<tr>
<td>Long-term prediction</td>
<td>63.59%</td>
<td>61.42%</td>
<td>62.55%</td>
</tr>
</tbody>
</table>

Analysis shows that the improved long short-term memory neural network has been greatly improved compared with the former forecast accuracy. The prediction accuracy of long short-term memory neural network improved in short-term prediction is lower than that of the mathematical model. The medium-term prediction accuracy is equal to that of the two methods. In long-term forecasting, the accuracy of long short-term prediction is higher than the mathematical model. Compared with the mathematical model, the accuracy of the improvement is above 60%, and the improvement effect is better, reaching the requirement of satellite operation monitoring.

Satellite orbit operation, due to the instability of the space environment, making the orbit fluctuations, which result in lower prediction accuracy. Therefore, it is more practical to give the confidence interval of the prediction error than to give the prediction data.

The expected error of the above calculation is set as $\mu$, and the confidence interval $[\mu_1, \mu_2]$ is the confidence $(1 - \alpha)$, the distribution range of the error value$^{[7]}$. In error analysis, this is a method of parameter estimation, with two main analysis methods. The first one is the confidence $(1 - \alpha)$, the probability of calculating the parameters in the random interval $[\mu_1, \mu_2]$; the second is the expected length of the random interval, the smaller the expected length, the more accurate. These two methods complement each other. Combined with the sample data of error and standard deviation, we can calculate the confidence interval of the data with reliable confidence.

The confidence interval of the expected value is set on the premise of total $X \sim N(\mu, \sigma^2)$. By the central limit theorem know, when $n$ is large, no matter what the distribution of $X$, are approximate:

$$Z = \frac{\bar{X} - EX}{DX/n} \sim N(0,1)$$

(13)

According to $X \sim N(\mu, \sigma^2)$, $\bar{X} \sim N(\mu, \sigma^2/n)$, satisfy the following equation:

$$E\bar{X} = \mu$$

(14)

$$DX = \sigma^2/n$$

(15)

$$Z = \frac{\bar{X} - \mu}{\sigma^2/n} \sim (0,1)$$

(16)
Confidence is \((1 - \alpha)\), order:

\[
P\left\{ \frac{\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha}}{z_{\alpha} \sigma} \leq \mu \leq \frac{\bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha}}{z_{\alpha} \sigma} \right\} = 1 - \alpha \tag{17}
\]

The confidence interval is

\[
\left[ \frac{\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha}}{z_{\alpha} \sigma}, \frac{\bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha}}{z_{\alpha} \sigma} \right] \tag{18}
\]

The greater the confidence level \((1 - \alpha)\), the longer the confidence interval\([8]\). In order to make the confidence intervals for better forecasting results more convenient for engineering reference, this section respectively calculates the confidence intervals of forecast errors under the confidence levels of 90\%, 95\% and 99\%. In the calculation process, in order to facilitate the actual reference, the error value is converted to a positive number for all calculations, the results shown in Table 6-8.

Table 6. Confidence interval confidence interval at 90\% confidence level

<table>
<thead>
<tr>
<th></th>
<th>X/KM</th>
<th>Y/KM</th>
<th>Z/KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>[0.12, 1.10]</td>
<td>[0.14, 1.48]</td>
<td>[0.02, 1.02]</td>
</tr>
<tr>
<td>Medium-term prediction</td>
<td>[0.06, 1.37]</td>
<td>[0.38, 1.72]</td>
<td>[0.19, 1.17]</td>
</tr>
<tr>
<td>Long-term prediction</td>
<td>[1.25, 2.90]</td>
<td>[1.12, 3.09]</td>
<td>[0.21, 1.22]</td>
</tr>
</tbody>
</table>

Table 7. Confidence interval confidence interval at 95\% confidence level

<table>
<thead>
<tr>
<th></th>
<th>X/KM</th>
<th>Y/KM</th>
<th>Z/KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>[0.02, 1.20]</td>
<td>[0.03, 1.59]</td>
<td>[0.00, 1.10]</td>
</tr>
<tr>
<td>Medium-term prediction</td>
<td>[0.00, 1.50]</td>
<td>[0.31, 1.84]</td>
<td>[0.10, 1.27]</td>
</tr>
<tr>
<td>Long-term prediction</td>
<td>[1.10, 3.06]</td>
<td>[0.93, 3.28]</td>
<td>[0.13, 1.32]</td>
</tr>
</tbody>
</table>

Table 8. Confidence interval confidence interval at 99\% confidence level

<table>
<thead>
<tr>
<th></th>
<th>X/KM</th>
<th>Y/KM</th>
<th>Z/KM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-term prediction</td>
<td>[0.00, 1.38]</td>
<td>[0.00, 1.84]</td>
<td>[0.00, 1.28]</td>
</tr>
<tr>
<td>Medium-term prediction</td>
<td>[0.00, 1.75]</td>
<td>[0.05, 2.11]</td>
<td>[0.00, 1.46]</td>
</tr>
<tr>
<td>Long-term prediction</td>
<td>[0.79, 3.37]</td>
<td>[0.56, 3.66]</td>
<td>[0.00, 1.49]</td>
</tr>
</tbody>
</table>

The table above gives the confidence intervals for the mean of forecast errors for each period under different confidence levels. Confidence interval of 90\% confidence interval is smaller, the actual reference range is too small, the reference is not strong; The confidence interval of 99\% confidence is large, the actual reference range is too large, and the accuracy is not high. In engineering applications, the confidence interval of 95\% confidence level is generally used as a reference to ensure the reliability and accuracy of the confidence interval.

Summary

In this paper, the idea and the improved model of track prediction using long short-term memory neural networks are given. A long period of historical orbit data of satellite orbit data is used as a training sample to predict the later orbit data. The prediction accuracy basically reaches the people's using precision for satellite transit and measurement and control, which has some guiding significance in this application. As the prediction gets rid of the kinetic model, it makes the satellite fixed orbit, rendezvous and docking, the error is still a large deviation, which in the next study also need to neural network model, training parameters, sample data and other aspects which affect the accuracy of training for further study.

298
Reference


