Video Transmission Control Method Based on Recurrent Neural Network with Long Short-Term Memory Patterns (LSTM-RNN)

Peng Zhou[a], Zhaohua Long[b]

School of Chongqing University of Posts and Telecommunications, Chongqing 400065, China.

[a]1281153289@qq.com, [b]longzh@cqupt.edu.cn

Keywords: Loop neural network; Real-time transmission control, Network bandwidth.

Abstract: A video transmission control method based on cyclic neural network is proposed for video communication. The scheme users LSTM-RNN to estimate the network bandwidth and adjusts the video transmission rate according to the predicted results. The experimental results show that the model can accurately predict the network bandwidth, which can describe the trend of changes better in network bandwidth and ensure the real-time and stability of the video.

1. Introduction

With the development of Internet technology, network congestion control of video transmission has become the focus of people's attention. Network congestion affects the user experience and data transfer to a certain extent. However, just improving the hardware will not prevent network congestion. Therefore, how to control the network congestion of video transmission well has become the focus and difficulty of the research.

The emergence of neural networks with high nonlinearity and strong adaptive learning ability brings new ideas to the prediction technology [1]. In recent years neural networks have been widely used in the work and achieved great results. Common neural network includes BP (back propagation) neural network model, RBF (radial basis function) neural network, RNN (recurrent neural network), CNN (convolutional neural networks), etc. Among them, BP neural network is a multi-layer feedforward neural network trained according to the error back propagation algorithm, which is one of the most widely used neural network models. However, its existence is sensitive to the initial weight threshold, easy to fall into the local minimum, and slow convergence [2].

RNN is a deep neural network with recurrent feedback [3]. Considering the temporal correlation of time series, it has a stronger practicality in learning sequence data with long-term dependence [4]. LSTM is a special RNN network model, which can learn the long-term dependence on sequential data. LSTM effectively solves the problem of gradient disappearance and gradient explosion during the conventional RNN training process, and is widely used in the field of time series prediction [5].

This article uses the LSTM estimates to predict the network bandwidth utilization, according to the predicted results to adjust the sending rate of video data, make whole video communication network utilization in a certain range. Improving the utilization of the network without changing the quality of the video, thereby improving the overall quality of video communication.

2. BP neural network model

BP neural network is a multi-layer feedforward neural network trained according to error back-propagation algorithm. Its basic idea is gradient descent method, which makes use of gradient search technology to minimize the error mean square error between the actual output value and the expected output value of the network. Fig. 1 shows BP network model structure [6].

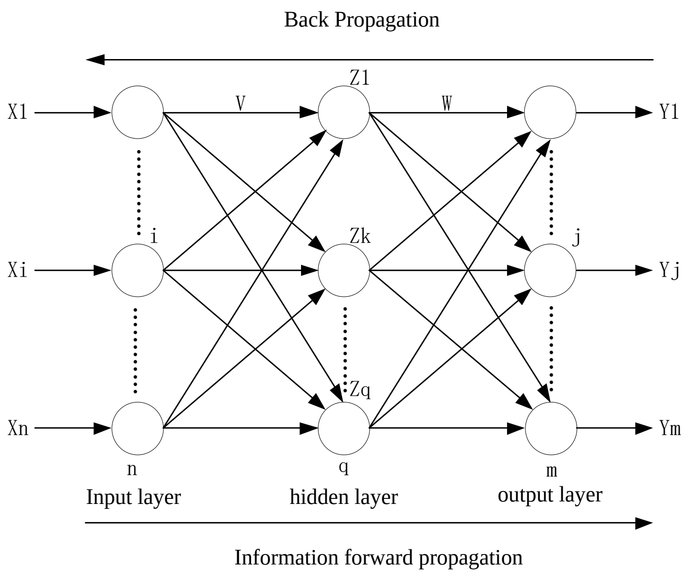

Figure 1. BP neural network structure model

The structure of BP neural network is composed of an input layer, an output layer and multiple hidden layers. Layers are fully connected with each other. Each neuron in the two layers is associated with a weight. Neurons in the same layer are connectionless.

There is a BP neural network that the input layer has d neurons, the output layer has l neurons and q neurons with hidden layer. For the training sample (x_k, y_k), it is assumed that the output of the neural network is y_j' shown in Equation. 1.

$$y_j' = f(\beta_j - \theta_j) \tag{1}$$

$$\mathrm{E}_k = \frac{1}{2}\sum_{j=1}^{l}(y_j - y_j')^2 \tag{2}$$

Then the mean squared error E_k of the network at (x_k, y_k) is as shown in Equation. 2.

The update estimate for parameter v is shown in Equation. 3.

$$\mathrm{v} \leftarrow \mathrm{v} + \Delta\mathrm{v} \tag{3}$$

$$\Delta\omega_{hj} = -\eta\frac{\partial E_k}{\partial \omega_{hj}} \tag{4}$$

BP algorithm is based on gradient descent strategy, and the modified values of weights between layers are shown in Equation. 4:

ω_{hj} Represents the connection weight from neuron h to neuron j, η is the learning rate, and the target's negative gradient direction adjusts the parameters, which is the steepest descent algorithm of BP network.

3. Recurrent neural network model

RNN is a feedforward neural network that adds memory cells to traditional neural networks. The nodes between hidden layers change from connectionless to connected. Fig. 2 shows a typical RNN model structure, including input layer x, hidden layer h, and output layer o. One of the unidirectional information flow from the input layer to the hidden layer, and then from the hidden layer to the output layer, constitutes the basic neural network structure. and the other information flow from the hidden layer to the hidden layer forms a closed loop, forming a self-connection hidden layer. RNN is a kind of neural network specially used for processing time series data samples. Each layer of RNN not only outputs to the next layer, but also outputs a hidden state for the current layer to use in processing the next sample.

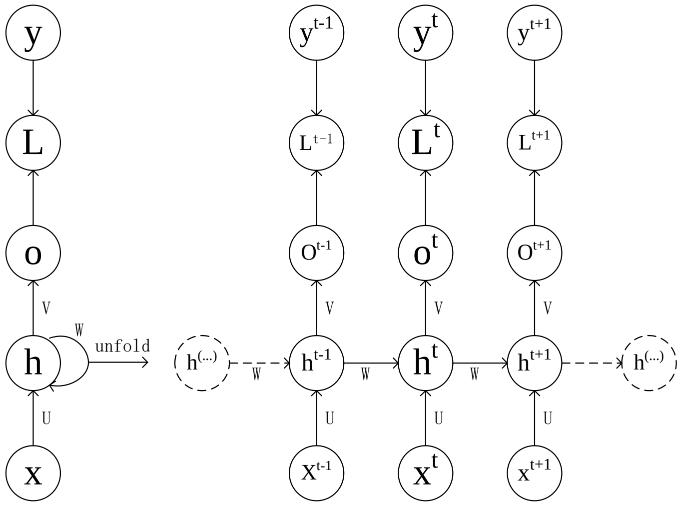

Figure 2. RNN structure model

Where x represents the input, h is the hidden unit, o is the output, L is the loss function, and y is the label on the training set. The t in the upper right band of these elements represents the state at time t. In addition to the input layer information, the hidden layer at the current time t also contains the information h^{t-1} of the previous time $\mathrm{t}-1$, and h^t will also affect the hidden layer at the next time $\mathrm{t}+1$.

For t moments, as shown in Equation. 5. the φ as activation function, generally using tanh function, as shown in Equation. 6, b is bias.

$$h^t = \varphi(Ux^t + Wh^{t-1} + b) \tag{5}$$

$$h^t = \tanh(Ux^t + Wh^{t-1} + b) \tag{6}$$

The output o^t at time t is as shown in Equation. 7.

$$o^t = Vh^t + c \tag{7}$$

The predicted output of the final model is shown in Equation. 8.

$$y_t' = \sigma(o^t) \tag{8}$$

Where σ is the activation function, usually used for RNN classification, so the softmax function is generally used, as shown in Equation. 9.

$$y_t' = softmax(o^t) \tag{9}$$

U, W and V are the corresponding weight matrices.

The structural characteristics of RNN hidden layer enable the circulatory neural network to 'memorize' the previous information and use it for the output at the moment, so it can effectively solve the problem of long-term dependence.

However, as the time interval increases, RNN will lose the ability to learn information from the past and the gradient will disappear.

4. LSTM Neural network model

LSTM neural network is a special model of RNN, which can learn long-term dependent information. Its principle is similar to RNN in essence. The difference is that LSTM model introduces the structure of "memory cell" in the hidden layer and uses different functions to calculate the state of the hidden layer. [7] In a "memory cell," the ability to remove or add information to the cellular state through a carefully designed 'gate' structure. The cell structure of LSTM is shown in Fig. 2

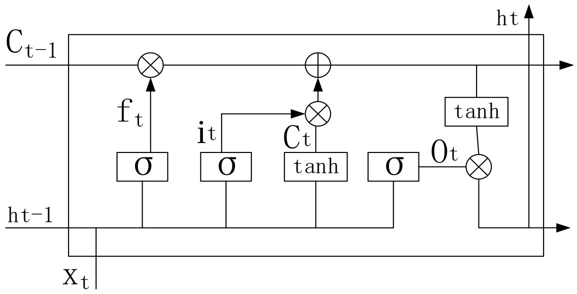

Figure 3. LSTM cell structure

In Fig. 2, i_t is the input threshold layer used to control information input; f_t is the forgetting threshold layer used to control the preservation of the historical state of cells. o_t is the output threshold layer used to control the output of information. The activation function σ is the sigmoid function corresponding to the three doors. The value generated is between [0, 1]. When the output value of the forgotten door is 0, it means that all the information of the previous state is discarded. LSTM Equation is calculated as follows (10)~(16).

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{10}$$

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{11}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{12}$$

$$C_t' = \tanh(W_c[h_{t-1}, x_t] + b_c) \tag{13}$$

$$C_t = f_t * C_{t-1} + i_t * C_t' \tag{14}$$

$$h_t = o_t * \tanh(C_t) \tag{15}$$

$$y_t = \sigma\left(W_{yh}h_t\right) \tag{16}$$

Where x_t is the input of the input layer, i_t is the output of the input gate, f_t is the output of the forgetting gate, o_t is the output of the output gate, C_t is used to update the cell state, h_t is the output of the hidden layer, and y_t is the output of the output layer. Where W is the weight matrix of the corresponding layer, b is the deviation vector of each output layer output, and σ and tanh are activation functions, respectively using the sigmoid function and the hyperbolic tangent function.

5. Network bandwidth prediction model based on LSTM neural network

5.1. Acquisition and processing of training samples and test samples

The prediction of network bandwidth is to organize the collected historical traffic data into the training set of neural network, determine the network model through the training set, and use this model to predict the future bandwidth utilization. The data of this experiment were collected by professional software wireshark. A total of 23,000 samples were obtained through 5 days of time-sharing data collection. The first 80% were selected as training samples, and the remaining 20% were used as test samples to analyze the performance of the model.

Due to the sudden and non-stationary change of bandwidth usage in the network, the changes of peak value and low value of data collected in different time periods are relatively large, which will reduce the prediction performance of the network. In order to avoid this situation, we need to normalize the collected data, which is between 0 and 1. The normalization Equation is shown in 17.

$$x_{(i)}' = \frac{x_{(i)} - x_{min}}{x_{max} - x_{min}} \tag{17}$$

$x_{(i)}'$ is the input value of the normalized neural network; $x_{(i)}$ is the original value collected; x_{min} is the minimum value of the original input value; x_{max} is the maximum value in the original input value.

5.2. Establishment of LSTM neural network prediction model

The experiment in this paper is based on python3.5 programming in ubuntu16.04 system and uses Keras deep learning computing framework to build LSTM cyclic neural network model. A four-layer network structure model is adopted, consisting of a layer of input layer, 2 LSTM layers, a layer of full connection layer and a layer of output layer. The first hidden layer has 128 LSTM units, the second hidden layer has 64 LSTM units, and the output layer is a single-valued output prediction result. The input layer is a single neuron. The LSTM storage unit uses the default activation function sigmoid. and finally the network training model converges. Root mean square error (RMSE) and mean relative error (MRE) was used to verify the prediction accuracy of the evaluation model. The specific calculation Equation is shown in 18-19.

$$\mathrm{RMSE}(y_t, y_t') = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - y_t')^2} \tag{18}$$

$$\mathrm{MRE}(y_t, y_t') = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y_t' - y_t}{y_t}\right| * 100\% \tag{19}$$

Where y_t is the actual value of the sequence sample, y_t' is the predicted value of the network, and N is the total number of samples.

5.3. Simulation results and analysis

In order to verify the effect of the LSTM model proposed in this paper, the training samples collected in the early stage were used as test objects. Fig. 4 shows part of the raw data.

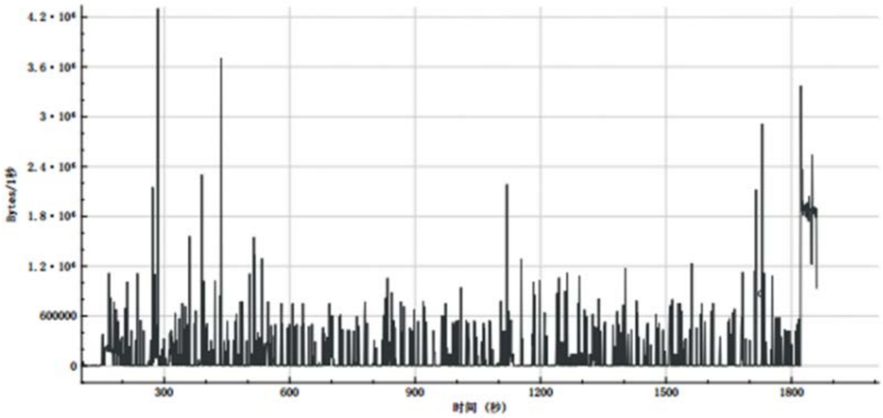

Figure 4. Partial raw data statistics

Fig. 5 shows the predicted results using the LSTM model. In this picture, the blue part is the actual value of the data and the red part is the predicted value. The predicted curve of LSTM model can well reflect network bandwidth utilization. Help us to better observe the trend of the network. According to the current network bandwidth utilization to adjust the video transmission rate. Achieve the purpose of controlling congestion.

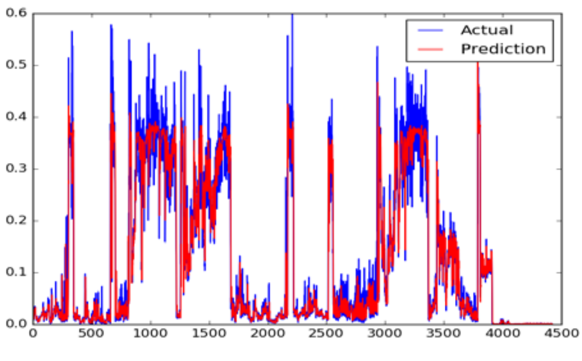

Figure 5. Network bandwidth utilization prediction results

As can be seen from Table 1, the prediction error of LSTM is small, and the prediction effect is better than that of BP network. In terms of MRE, the LSTM network is 8.52%, while the result of the BP network is 20.13%. The accuracy of the LSTM network prediction is higher than that of the BP network.

Table.1. Error comparison of different models

Model	RMSE	MRE
BP	81.45	20.13%
LSTM	53.85	8.52%

6. Conclusion

In this paper, artificial intelligence is combined with traditional communication to propose a congestion control algorithm based on LSTM neural network. Through experimental comparison, this method can accurately predict the bandwidth utilization of the network and play a very important role in the congestion control of video transmission.

Acknowledgments

I would like to express my gratitude to my supervisor professor Long Zhaohua for his guidance on the theory, and for his advice on modification in the process of algorithm research and implementation. This research is supported by the extension of huixun wireless projection cloud management system (KJZH17118), which is the excellent achievement transformation project of colleges and universities of Chongqing education commission.

References

[1] Li Yuling, Wang Wei. Remote Video Transmission Control Method Based on BP [J]. Computer Engineering, 37 (22): 276-278. (2011).

[2] Zhang Lifang. Application of Genetic BP Neural Network in Network Traffic Prediction [J]. Computer and Digital Engineering, 42 (04): 660-663. (2014).

[3] Huang Bin, Lu Jinjin, Wang Jianhua, et al. Object recognition algorithm based on deep convolutional neural network [J]. Journal of Computer Applications, 36 (12): 3333-3340. (2016).

[4] Yang Weiyue, Fu Qian, Wan Dingsheng. Time Series Prediction Model Based on Deep Cyclic Neural Network [J]. Computer Technology and Development, 27 (3): 35-38. (2017).

[5] Kong Weicong, Dong Zhaoyang, Jia Youwei, et al. Short-term residential load forecasting based on LSTM recurrent neural network [J]. IEEE Trans on Smart Grid, PP (99): 1. (2017).

[6] Zhang Lifang. Application of Genetic BP Neural Network in Network Traffic Prediction[J]. Computer and Digital Engineering, 42 (04): 660-663. (2014).

[7] Li Xiaolin, Wu Teng. Efficient network traffic prediction method based on PF-LSTM network [J/OL]. Computer Application Research: 1-6. http://kns.cnki.net/kcms/detail/51.1196.TP.20181129.1137.015.html. (2019-01-23)