A LSTM Based Campus Network Traffic Prediction System

Yue Geng
School of Computer Science
Shaanxi Normal University
Xi'an 710119, China
gengyue@snnu.edu.cn

Shuyu Li
School of Computer Science
Shaanxi Normal University
Xi'an 710119, China
lishuyu@snnu.edu.cn

***Abstract*- Accurate and real-time campus network traffic prediction is very important in network management. Aiming at security analysis of network traffic and prediction problems caused by the nonlinearity and multi-dimensional dynamics of campus network traffic. A network traffic prediction system based on long-term/short-term memory (LSTM) model is presented for analysis of campus users' network behaviors in the paper. The prediction system uses immortal log analysis tool of Xijia Education to gather and pre-process multi-source heterogeneous log data from various network applications, and adopts an improved LSTM model to analyze and predict network traffic of campus users. Experimental results validate the effectiveness of the proposed system by using real campus network data.**

Keywords*- *Network traffic Prediction; Time Series Model; Deep Learning; LSTM

I. INTRODUCTION

With the expansion of campus network, the use of network bandwidth increases significantly, which makes the quality of network service not guaranteed. If the network traffic information is obtained by certain specific methods and corresponding optimization measures are taken, the output bandwidth pressure of the campus network can be effectively reduced and the overall performance of the campus network can be improved.

In this paper, we propose a model based on LSTM and apply it into the campus network traffic prediction system. Specifically, the proposed model uses historical network traffic data to predict traffic flow in the coming hours. The main contributions of the paper are as follows: a) We propose the NTLSTM model, and test its performance with real campus network traffic, and compare it with other models. b) In order to improve computational efficiency, we use the RMSprop optimizer to replace the standard SGD(stochastic gradient descent). c) We implement the campus network traffic prediction system based on NTLSTM model.

II. THEORETICAL BASICS AND RELATED WORKS

A. The LSTM Model

Hochreiter and Schmidhuber [1] first proposed the LSTM model. The success of LSTM relies on the ability to learn time series features and automatically determine several hyper parameters from the data [2-3]. The LSTM network consists of an input layer, one or more hidden layers and an output layer. The hidden layer is treated as a memory block with memory cells to remember the long-time characteristics.

B. Related Works

Aldhyani and Joshi [4] proposed using an integrated model of different time series models with soft clustering techniques to predict network usage. They propose that linear time series models provide better results in network traffic predicting when integrated with a clustering approach. Eterovic et al. [5] compared the performance of ARIMA and artificial neural network (ANN) models. They showed that ARIMA worked for short term predictions but failed for long term predictions. ANN gives better prediction results in this study when taking long term forecasts into consideration. Nie et al. [6] proposed a network traffic prediction method based on a deep belief network and a Gaussian model. A proposed hybrid model of two-dimensional correction and single exponential smoothing (SES) prediction method presented in [7]. In [8] a prediction model is established by using multiple kernel support vector regression for network traffic. Wei [9] proposed a prediction model based on RBF neural network optimized by improved gravitation search algorithm, and apply to Lorenz chaotic time series and network traffic to test the validation of the algorithm. Sahrani et al. [10] proposed two Nonlinear Auto-Regressive Moving Average (NARMA) models for prediction of network traffic: Multi-Layer Perceptron (MLP) and polynomial. Katris et al. [11] construct a model selection scheme based on the White's Neural Network test for non-linearity.

III. DESIGN OF NETWORK TRAFFIC PREDICTING SYSTEM

A. Requirement Analysis of System

The system uses the improved circular neural network model to analyze and predict the big data of network traffic log. It uses the Immortal log tools to decompose heterogeneous multi-type log data and establish the sequential characteristics of network traffic content data. Based on the content feature representation, the NTLSTM model is introduced. The system uses the characteristics of network traffic metadata to train the neural network and detects the anomalies of network traffic through the training model.

Therefore, the main functions of the campus network traffic prediction system designed in this paper are as follows:

- Network traffic prediction: By using the historical network traffic data uploaded by users, the traffic usage in the future period of time can be seen.
- Prediction model visualization: Users can customize the parameters of the model and view the effects of the models under different parameters, so that users can choose the parameter model that suits their needs.
- User behavior analysis: Through the statistical algorithm, the system displays the comparison of the number of line users within two weeks in the form of histogram, and shows the active users.
- Scrutiny user's log: The system displays the historical online records of all users in a chronological order.

B. System Design

1) Architecture of the system

The schematic diagram of the campus network traffic prediction system is presented in Figure 1.

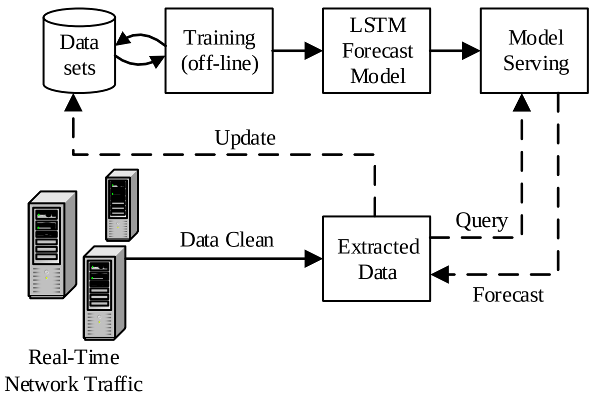

Figure 1. Network traffic prediction system framework

As can be seen from Figure 1, the process of the system to implement the network prediction function is as follows: First, the campus network traffic log data is collected from the network information center management, and the system calls the Immortal log processing tool for data cleaning. Then, the system updates the data set with the extracted data, and trains model offline with the updated data to obtain the NTLSTM prediction model, and finally implement the model prediction service function. At the same time, the trained model can also directly predict and analyze the cleaned data.

2) Functions design of the systems

The system mainly analyzes the campus network traffic log data of students and teachers from two sides: network traffic analysis and user behavior analysis. Network traffic analysis includes the functions of model display and network traffic prediction; user behavior analysis includes the functions of active users' analysis and user scrutiny.

a) Network traffic analysis

The model display function displays the loss value curve and model effect display graph of the network traffic prediction model used by the current system. The system provides the user with an interface to change the model parameters on the front page. Then the system uses the new parameters to train the model in the background and obtain the training results. Finally, the loss model curve and model effect display diagram of the new model are displayed on the home page. In this way, the user can select the appropriate model parameters based on the obtained model performance diagram.

The network traffic prediction is the main function of the system. This function visually displays the predicted results of the model by comparing the actual and predicted values. First, users upload historical network traffic log data, and then the system calls the Immortal log analysis tool of Xijia Education to clean up the data. The system uses the processed data as the input data of the training model to obtain the prediction results. Users can choose to view network traffic values for the next 24 hours (or 48 hours). As shown in Figure 2, the black line represents the historical network traffic value, the blue line represents the real network traffic value, and the red line represents the predicted network traffic value obtained by the model.

b) User behavior analysis

The user behavior analysis function mainly analyzes and statistics the online log data of students and teachers through statistical algorithms. The active user analysis function mainly analyzes users' online time, address and other characteristics, and displays the total number of people online in the past week from the histogram of the front page.

The user review function mainly displays the information of all visited users on the page according to the uploaded network traffic log data, including user name, IP, MAC address, offline reason and recent visit time.

C. The NTLSTM Based Prediction Framework

The proposed NTLSTM model consists of several stages which are as follows:

Data Preprocessing: We use Immortal tools to deal with big data of multi-source heterogeneous network traffic logs. After the data set is converted to a clean data set, the data set is divided into training sets and test sets for evaluation.

Training NTLSTM Neural Network: In this stage, the data is fed to the neural network and trained for prediction assigning random biases and weights. Our NTLSTM model is composed of a sequential input layer followed by 2 LSTM layers and 2 dropout layers followed respectively, and then finally a fully connected layer with linear activation function.

Output Generation: In this stage, the output value generated by the output layer of the NTLSTM is compared with the target value. The error or the difference between the target and the obtained output value is minimized by using BPTT which adjusts the weights and the biases of the network.

The proposed network traffic prediction algorithm is shown in Algorithm 1.

The proposed model is implemented based on the LSMT model. Similarly, each memory cell of the model has three gates for maintaining and adjusting its cell state: forgotten gates f_t, input gates i_t and output gates o_t. The formulations of all nodes in an LSTM structure are given by formula (1):

$$
\begin{aligned}
f_t &= sigmoid\left(W_{f,x}x_t + W_{f,h}h_{t-1} + b_f\right)\\
\tilde{C}_t &= \tanh(W_{\tilde{c},x}x_t + W_{\tilde{c},h}h_{t-1} + b_{\tilde{c}})\\
i_t &= sigmoid\left(W_{i,x}x_t + W_{i,h}h_{t-1} + b_i\right)\\
C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t\\
o_t &= sigmoid\left(W_{o,x}x_t + W_{o,h}h_{t-1} + b_o\right)\\
h_t &= o_t \odot \tanh\left(C_t\right)
\end{aligned} \tag{1}
$$

where $W_{f,x}$, $W_{f,h}$, $W_{\tilde{c},x}$, $W_{\tilde{c},h}$, $W_{i,x}$, $W_{i,h}$, $W_{o,x}$ and $W_{o,h}$ are weight matrices for the corresponding inputs of the network activation functions. ⊙ denotes element-wise product. $\tilde{C}_t$ defines input node. For objective function we use square loss function given by the following formula (2):

$$
\sum_{t-1}^{n}\left(y_t - p_t\right)^2 \tag{2}
$$

Where y_t represents the real output and p_t represents the predicted traffic value at t time, n represents the length of history traffic sequence. RMSProp optimizer, a modification of stochastic gradient descent (SGD) optimizer with adaptive learning rates, is applied for back propagation through time (BPTT). Compared with SGD, RMSProp algorithm solves the problem that the loss function oscillates during the update process, speeding up the convergence of the function. Specifically, at t iteration the formula (3):

$$
\begin{aligned}
s_{dw} &= \beta s_{dw} + \left(1-\beta\right)dW^2\\
s_{db} &= \beta s_{db} + \left(1-\beta\right)db^2\\
W &= W - \alpha\frac{dW}{\sqrt{s_{dw}}+\varepsilon}\\
b &= b - \alpha\frac{db}{\sqrt{s_{db}}+\varepsilon}
\end{aligned} \tag{3}
$$

s_{dw} and s_{db} are the gradient momentum accumulated by the loss function during the prior $t-1$ iteration, and β is an exponent of the gradient accumulation. ε is used for smoothing to prevent the denominator from being zero, which is a small value and generally be 10^{-8}.

Algorithm 1: NTLSTM

Input: network traffic log *DataSource*
Output: Model structure and related parameters W

// Stage 1: Data preprocessing
1. $D_M \leftarrow Transfer\left(DataSource\right)$ // Use the Immortal tool to extract the feature from the multi-source heterogeneous network traffic log and output the multidimensional feature data set D_M.
2. $D \leftarrow ConvertSeriesToMatrix\left(D_M\right)$ // Normalize the data set D_M and convert it to a matrix form.

// Stage 2: model training
3. **for** iter < MaxIter **do**//Iterate training.
4. Take a random sample D_B from training dataset D.
5. Calculate the $L\left(y,p\right)$ via Equ. (2).
6. Update parameters with RMSProp optimizer.
7. **end for**
8. Output NTLSTM model structure and related parameters W.

IV. EXPERIMENTS AND EVALUATION

A. Experimental Datasets

The network traffic records are collected from Shaanxi Normal University Network Information Management Center. Our dataset contains six series data, where S1-S4 represent the network traffic data aggregated by hour, which distribution are relatively stable with over 2,000 records. S5-S6 represent network traffic data aggregated by minute, which distribution fluctuated greatly with over 5,000 records.

B. Experimental results and analysis

In our experiment, ARIMA, SVR and NTLSTM models are compared. To test our model prediction accuracy better, we use both mean square error (MSE) and mean absolute error (MAE) defined as follows formula (4):

$$
\begin{aligned}
MSE &= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i\right)^2\\
MAE &= \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i\right|^2
\end{aligned} \tag{4}
$$

Table 1 demonstrate the prediction performance of different algorithms with six time-series datasets. Additionally, Figure 3 shows the comparison of actual and predicted values of NTLSTM.

TABLE I. THE PREDICTION PERFORMANCE OF DIFFERENT ALGORITHMS

Methods	Evaluation	Datasets					
		S1	**S2**	**S3**	**S4**	**S5**	**S6**
ARIMA	MSE	3.76	4.22	2.38	4.11	10.2	12.3
	MAE	1.06	1.04	0.86	1.07	7.46	8.04
SVR	MSE	8.90	8.75	9.38	10.1	1.29	0.08
	MAE	2.98	2.96	3.06	3.18	1.13	0.29
NTLSTM	MSE	0.47	0.41	0.30	0.46	4.17	3.62
	MAE	0.55	0.51	0.41	0.49	2.02	1.48

Experimental results show that for S1-S4, NTLSTM is superior to ARIMA and SVR models in most cases. It is proved that NTLSTM model has good performance for stable time series data sets. In the actual campus network environment, although there are several peak periods of network traffic in a day, its overall distribution is stable. But for S5-S6, NTLSTM has large MSE values. That means SVR can better capture the sudden change of network traffic, while NTLSTM has a tendency of predicting steady future network traffic.

V. CONCLUSION

This paper introduces the proposed NTLSTM model to predict network traffic in the future hours. We compared the predictive performance of ARIMA, SVR and NTLSTM models based on the real campus network traffic records of Shaanxi Normal University. According to our experimental results, we found that NTLSTM model is superior to ARIMA and SVR model in hourly network traffic prediction. In addition, we designed and developed a network traffic prediction system to achieve better network management and optimize network performance. In future work, we will test NTLSTM with more hidden states and variable length time series input data.

Figure 2. 24(48)-hour network traffic prediction results

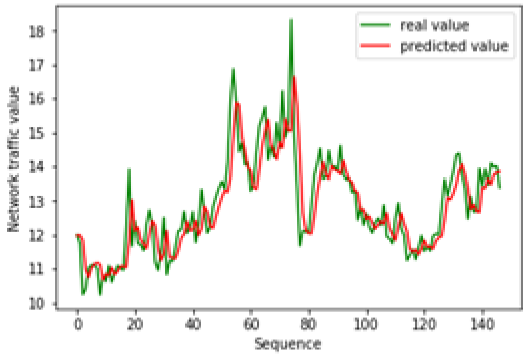

Figure 3. Predicted traffic flow vs. Real traffic flow

ACKNOWLEDGMENT

This research is supported by the National Key R&D Program of China (No. 2017YFB1402102) and the CERNET Innovation Project (No.NGII20170703).

REFERENCES

[1] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.

[2] 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.

[3] 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.

[4] Aldhyani T H H, Joshi M R. Integration of time series models with soft clustering to enhance network traffic forecasting[C]//2016 Second International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). IEEE, 2016: 212-214.

[5] Eterovic T, Mrdovic S, Donko D, et al. Data mining meets network analysis: Traffic prediction models[C]//2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2014: 1479-1484.

[6] Nie L, Jiang D, Yu S, et al. Network traffic prediction based on deep belief network in wireless mesh backbone networks[C]//2017 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2017: 1-5.

[7] Aldhyani T H H, Joshi M R. An integrated model for prediction of loading packets in network traffic[C]//Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies. ACM, 2016: 30.

[8] Xiang C, Qu P, Qu X. Network traffic prediction based on MK-SVR[J]. JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE, 2015, 12(8): 3185-3197.

[9] Wei D. Network traffic prediction based on RBF neural network optimized by improved gravitation search algorithm[J]. Neural Computing and Applications, 2017, 28(8): 2303-2312.

[10] Sahrani M N, Zan M M M, Yassin I M, et al. Artificial Neural Network Non-linear Auto Regressive Moving Average (NARMA) Model for Internet Traffic Prediction[J]. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 2017, 9(1-3): 145-149.

[11] Katris C, Daskalaki S. Comparing forecasting approaches for Internet traffic[J]. Expert Systems with Applications, 2015, 42(21): 8172-8183.