A Mobile Traffic Prediction Model Based on Area Spatio-Temporal Feature

Jian Ma, Chaoran Zhou*, Xin Zhang, Ming Yue

Changchun University of Science and Technology, Changchun, China

Abstract

Accurate prediction of mobile traffic can help operators plan network resources in advance and strengthen the management of network resources. Accuracy of mobile traffic prediction is affected by such spatio-temporal factors as the change of pedestrian flow and historical flow in surrounding areas. In this paper, we propose a Prediction Model for Spatio-Temporal Feature of mobile traffic (STFP) based on Residual Network (ResNet) and Long Short-Term Memory Network (LSTM). By analyzing the Pearson Correlation Coefficient (PCC) of mobile traffic data, we determined that proximity data and periodic data were selected as inputs to the STFP model. To avoid gradient explosion, the STFP model uses ResNet as spatial feature extraction network. We assign different weights to the outputs of the two branches for dynamic fusion according to the degree of influence of different input, and then use LSTM to extract the temporal features of the two inputs, and finally realize the prediction of mobile traffic. We select four deep learning models as baselines. The experimental results show that compared with the baseline models, the STFP model has better prediction accuracy and better indicators of RMSE, MAPE and R^2.

Keywords

mobile traffic prediction; spatio-temporal features; residual network; Long Short-Term Memory Network; Pearson Correlation Coefficient

1. Introduction

With the vigorous development and popularization of mobile smart phones, the scale of the mobile Internet market continues to expand, which is closely related to learning, work, entertainment, and other aspects of daily life. The arrival of the 5G era provides a foundation for big data, artificial intelligence, and the Internet of Things [1-3]. The growing network demand makes the resource allocation and reasonable allocation of mobile traffic particularly important. Accurate prediction of mobile traffic is the basis for intelligent management of mobile network resource allocation [4].

To achieve accurate prediction of mobile traffic, it is necessary to analyze the temporal and spatial correlation of mobile traffic and build a mobile traffic prediction model with the ability to extract spatio-temporal feature and high prediction accuracy. The research work mainly includes:

We propose mobile traffic prediction model STFP (Spatio-temporal Feature Prediction) based on ResNet and LSTM. The model uses ResNet to extract the spatial feature information of the input data while avoiding gradient explosion, and then uses LSTM to extract the temporal features of the data.

To capture proximity and periodicity of mobile traffic data, we analyze the PCC of mobile traffic data, select proximity data and periodic data as the input of the model.

To dynamically fuse the two inputs, the model uses different parameter matrices and learns weight values from historical data to fuse the two outputs in a weighted manner without losing the important information of the two characteristics.

ICCEIC2022@3rd International Conference on Computer Engineering and Intelligent Control
EMAIL: *2022800008@cust.edu.cn (Chaoran Zhou)

2. Related work

In the field of mobile communication, many scholars had carried out related research on the prediction of mobile traffic, which proved that mobile traffic has time correlation, and adopted the time series analysis method. Literature [5] used Autoregressive Integrated Moving Average model (ARIMA) to predict mobile traffic, but ARIMA model requires time series data to be stable, which may lead to model prediction accuracy decline. Other typical machine learning mobile traffic prediction methods also has limitations, such as: Support Vector Machine (SVM) [6], K-Nearest Neighbor (KNN) [7], etc. Compared with time series analysis methods, these methods can learn data regularity. However, the prediction accuracy will decrease when dealing with complex high-dimensional data.

Deep learning methods that can process high-dimensional data and extract nonlinear data features have a better ability to capture the characteristics of mobile data than typical machine learning methods. However, the models proposed in literature [8-10] such as LSTM and Deep Belief Network (DBN) only focus on the temporal feature of mobile traffic data, without considering the spatial feature of the data. Literature [11] analyzed the mobile base station data and found that adding adjacent mobile base station mobile traffic data for prediction can improve the prediction accuracy. Literature [12] proposed a combined model that used Convolutional Neural Network (CNN) to extract spatial features and LSTM to extract temporal features to predict mobile traffic data, which improved the prediction accuracy. In literature [13], Generative Adversary Networks (GAN) and transfer learning strategies were used to address data scarcity and improve prediction performance. Literature [14] proposed a model based on Graph Convolutional Network (GCN) and Gated Linear Units (GLU) to predict mobile traffic consumption at different time horizons by simulating the mobility of mobile networks and crowds.

3. Prediction model

3.1. Problem definition

The prediction of mobile traffic is to predict the mobile traffic usage of this area at the next moment through the historical mobile traffic usage of an area. In this paper, by dividing the prediction area into M×N squares, each square represents the mobile traffic generated in this area at time t, and then constructs the mobile traffic spatio-temporal matrix X_t of the entire prediction area at time t. The problem is to predict the mobile traffic spatio-temporal matrix X_{t+1} at the next moment through the historical mobile traffic spatio-temporal matrix X_t.

3.2. STFP model structure

To capture the temporal and spatial correlation of mobile traffic and achieve accurate mobile traffic prediction, we propose a mobile traffic prediction model STFP based on ResNet and LSTM. The model structure is shown in Figure 1. We determine proximity sequences and periodic sequences according to the characteristics of mobile traffic. The STFP model extracts the spatial characteristics of mobile traffic of different sequences through two-dimensional convolution and residual blocks, and then weights the two outputs and fuses them. The fusion matrix is input to the LSTM to extract temporal features. Finally, a fully connected layer is used to fuse the spatial and temporal features to obtain the result.

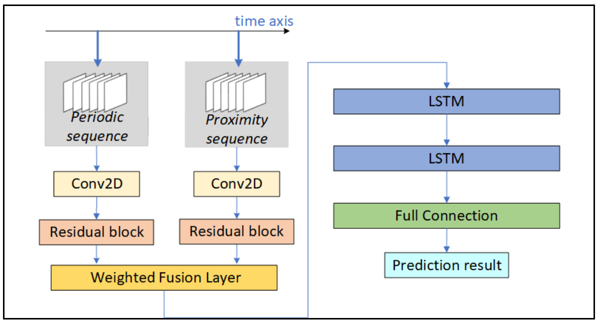

Figure 1 STFP model

3.3. Input sequence

Through correlation analysis of the PCC of mobile traffic, the proximity sequence and the periodic sequence are determined.

(1) proximity sequence X_c: It is composed of the mobile traffic spatio-temporal matrix at the previous time t adjacent to the predicted time t+1 of the predicted day T.

$$X_c = \{X_t^T, X_{t-1}^T ... X_{t-c}^T\}$$

(2) periodic sequence X_p: It is composed of the mobile traffic spatio-temporal matrix at time t+1 of historical day T-1.

$$X_p = \{X_{t+1}^{T-1}, X_t^{T-1} ... X_{t-q}^{T-1}\}$$

3.4. Convolution module

The predicted area in the mobile traffic space-time matrix has spatial correlation with adjacent areas. Using two-dimensional convolutional layers to process images can extract image features without losing spatial information and can extract high-dimensional complex features from simple features. The two-dimensional convolution is shown in Figure 2. In the STFP model, the spatio-temporal matrix of mobile traffic at time t in X_c and X_p is processed by two-dimensional convolution. The matrix is regarded as a single-channel image data and sent to the convolutional layer to extract the spatial features.

3.5. Residual module

Adding residual units to convolutional layers can effectively solve the gradient explosion problem. The residual unit directly adds the unit's input to the unit's output through residual connection before activating. Batch normalization is added to the residual unit to alleviate the gradient vanishing of the deep network and make the training of the deep network model more stable. To avoid network degradation caused by stacking convolutional layers, residual units are stacked after the convolutional layers, and the spatio-temporal matrices of mobile traffic of different sequences processed by two-dimensional convolution are input into the residual module.

3.6. Weighted fusion module

The outputs of the two branches are X_c^{out} and X_p^{out} respectively and have different weights. To dynamically fuse the outputs of the two branches, two trainable weight matrices are used to learn weight values from historical data, and the outputs are weighted and fused. Finally, the fusion output is obtained through the activation function. The weighted fusion is defined as:

$$X^F = f(W^C X_c^{out} + W^P X_p^{out}) \tag{1}$$

among them, W^c and W^p are the trainable weight matrices of different branches, f is the activation function, and X^F is the output after weighted fusion.

3.7. LSTM module

To capture long-term temporal features of mobile traffic data, the STFP model uses LSTM to process the data. LSTM processes data through three gating units. Flatten the spatio-temporal matrix obtained after weighting fusion into a one-dimensional vector and input it into LSTM. The input gate extracts the input on demand, retains important information, the forget gate selectively discards the information of the previous time step, and obtains the output value through the output gate.

4. Experiment

4.1. Dataset

4.1.1. Mobile traffic dataset source

The experimental data in this paper comes from the open mobile traffic data set of Milan in the "Telecom Italia Big Data Challenge"[15].

4.1.2. Data visualization and analysis

Figure 2 shows the change curve of mobile traffic in the area (44, 59) from 0:00 on November 4, 2013, to 24:00 on November 10, 2013. As shown in Figure 2, mobile traffic demand on weekdays is significantly higher than that on weekends, and the mobile traffic demand during the day is greater than that at night. Some sudden activities will lead to a surge in mobile traffic, such as the 135th to 140th hours in Figure 2.

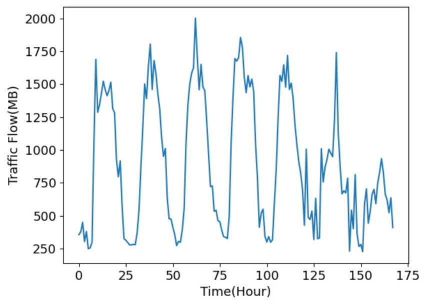

Figure 2 The mobile traffic curve of the area (44, 59)

Figure 3 shows the mobile traffic usage in Milan at 10:20 am on November 21, 2013. The demand for mobile traffic in the urban center is significantly higher than that in the surrounding areas because the urban population is mainly concentrated in the urban center and its adjacent areas, and the mobile traffic demand in the central area is higher than that in the less populated urban areas. The demand for mobile traffic in residential areas is lower than that of commercial areas in the city center, but still higher than that in the surrounding areas.

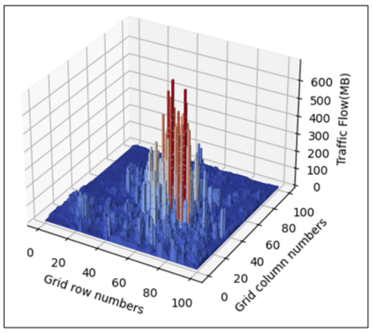

Figure 3 3D schematic diagram of mobile traffic

4.2. Evaluation metric

We choose to use Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2) as model evaluation metrics to evaluate the performance of the model.

4.3. Experimental setup

The model training method is as follows: The optimization function is the Adam optimizer. The learning rate is 0.001, the batch size is 64, and the number of iterations is 500. The loss function is mean squared error (MSE). To eliminate the influence of dimension, the input data are min-max normalized before training.

4.4. Experimental results and analysis

To verify the performance of the STFP model, LSTM [8], GRU [16], 3DCNN [17] and CNN-LSTM [12] are selected for comparison. The results are shown in Table 1.

Table 1 Comparison of evaluation metrics of each model

Model	RMSE	MAE	R^2
LSTM	9.867	5.278	0.792
GRU	9.235	5.013	0.814
3DCNN	8.265	4.597	0.857
CNN-LSTM	7.052	3.861	0.881
STFP	**6.436**	**3.494**	**0.905**

The experimental results in Table 1 shows that LSTM and GRU are models for predicting mobile traffic based on time series, and the ability to extract spatial characteristics of mobile traffic is insufficient. 3DCNN is a prediction model with the ability to extract spatiotemporal features. The prediction effect is better than that of LSTM and GRU, but it is insufficient for long-term temporal feature extraction. CNN-RNN is a combined prediction model, which extracts spatial features and long-term temporal features through CNN and LSTM respectively, so the prediction error is low, but the periodicity of the data is not considered. The STFP model considers the proximity and periodicity of mobile traffic in the time dimension, and extracts the feature of the spatial dimension, and the prediction performance is better.

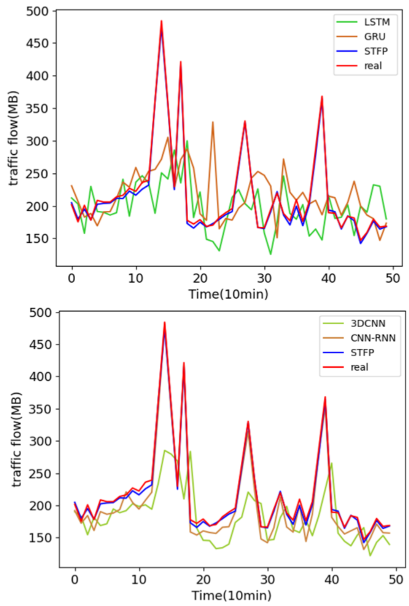

Figure 4 Comparison of model prediction results

As shown in Fig.4, LSTM and GRU learn the changing laws of mobile traffic, but they are less effective in predicting mobile traffic during peak periods. 3DCNN learned change law of mobile traffic peak period, but the error between the predicted value and the actual value is large. The curve fitting effect of CNN-RNN is stronger than that of LSTM and GRU, but the periodicity of mobile traffic is not considered, so the prediction accuracy is not as good as that of STFP model. The STFP model considers the proximity and periodicity of mobile traffic, so the prediction accuracy is better.

5. Conclusion and future work

Aiming at the spatio-temporal features of mobile traffic, we propose a mobile traffic prediction model STFP based on ResNet and LSTM.

To capture proximity and periodicity of mobile traffic data, the model determined the proximity sequence and periodic sequence by analyzing the PCC of the data. Convolution module and residual unit module are used to extract spatial features while avoiding gradient explosion. Learn two branch weights from historical data and dynamically fuse the inputs of the two branches. The model extracts long-term temporal features of mobile traffic data by using LSTM. Experiments show that the prediction effect of the STFP model is better than the baseline models.

However, since the STFP model does not introduce external factors (such as weather, holidays, large-scale events, etc.), this will affect the accuracy of predicting mobile traffic. Based on the STFP model, we will design an external factor module to introduce the feature information of external features, to improve the robustness of the model and make the model have better predictive ability in complex situations.

6.Acknowledgment

This work is supported by Jilin Provincial Department of Science and Technology—Jilin Provincial Natural Science Foundation (NO. 20200201182JC)

7. References

[1] Liu X, Wu S, Guo Y, et al. The demand and development of Internet of Things for 5G: A survey[C]//2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW). IEEE, 2018: 1-2

[2] Anand R, Sindhwani N, Juneja S. Cognitive Internet of Things, Its Applications, and Its Challenges: A Survey[M]//Harnessing the Internet of Things (IoT) for a Hyper-Connected Smart World. Apple Academic Press, 2022: 91-113

[3] ALRikabi H T H, Hazim H T. Secure Chaos of 5G Wireless Communication System Based on IOT Applications[J]. International Journal of Online & Biomedical Engineering, 2022, 18(12)

[4] Xu Y, Gui G, Gacanin H, et al. A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges[J]. IEEE Communications Surveys & Tutorials, 2021, 23(2): 668-695

[5] Zhou B, He D, Sun Z. Traffic predictability based on ARIMA/GARCH model[C]//2006 2nd Conference on Next Generation Internet Design and Engineering, 2006. NGI'06. IEEE, 2006: 8 pp.-207

[6] Liu X, Fang X, Qin Z, et al. A short-term forecasting algorithm for network traffic based on chaos theory and SVM[J]. Journal of network and systems management, 2011, 19(4): 427-447

[7] Chang H, Lee Y, Yoon B, et al. Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences[J]. IET intelligent transport systems, 2012, 6(3): 292-305

[8] Abbasi M, Shahraki A, Taherkordi A. Deep learning for network traffic monitoring and analysis (NTMA): A survey[J]. Computer Communications, 2021, 170: 19-41

[9] Alawe I, Ksentini A, Hadjadj-Aoul Y, et al. Improving traffic forecasting for 5G core network scalability: A machine learning approach[J]. IEEE Network, 2018, 32(6): 42-49

[10] Narejo S, Pasero E. An application of internet traffic prediction with deep neural network[M]//Multidisciplinary Approaches to Neural Computing. Springer, Cham, 2018: 139-149
[11] Zhou X, Zhao Z, Li R, et al. The predictability of cellular networks traffic[C]//2012 international symposium on communications and information technologies (ISCIT). IEEE, 2012: 973-978
[12] Huang C W, Chiang C T, Li Q. A study of deep learning networks on mobile traffic forecasting[C]//2017 IEEE 28th annual international symposium on personal, indoor, and mobile radio communications (PIMRC). IEEE, 2017: 1-6
[13] Wu Q, He K, Chen X, et al. Deep transfer learning across cities for mobile traffic prediction[J]. IEEE/ACM Transactions on Networking, 2021
[14] Fang Y, Ergüt S, Patras P. SDGNet: A Handover-Aware Spatiotemporal Graph Neural Network for Mobile Traffic Forecasting[J]. IEEE Communications Letters, 2022, 26(3): 582-586
[15] Barlacchi G, De Nadai M, Larcher R, et al. A multi-source dataset of urban life in the city of Milan and the Province of Trentino[J]. Scientific data, 2015, 2(1): 1-15
[16] Fu R, Zhang Z, Li L. Using LSTM and GRU neural network methods for traffic flow prediction[C]//2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). IEEE, 2016: 324-328
[17] Ji S, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(1): 221-231