Spatio-Temporal Traffic Prediction of Wireless Communication Network Based on Multi-source Data

Yu Wang[1(✉)], Yangyang Sun[1], Yanlin Fan[1], Tao Jiang[1], Jiansheng Xiong[1], Ying Zhou[1], and Zhibo Han[2]

[1] Intelligent Network Innovation Center, China United Network Communications Corporation Limited, Beijing 100048, China
{wangyu216,sunyy204,fanyl51,jiangt30,xiongjs,zhouy36}@chinaunicom.cn

[2] Beijing Key Laboratory of Work Safety Intelligent Monitoring, Beijing University of Posts and Telecommunications, Beijing, China
Hzb123159@bupt.edu.cn

Abstract. Accurate prediction of wireless communication network traffic can assist operators in precise operation, improve communication network management, and reduce energy consumption. However, due to the highly complicated spatio-temporal dependence and the influence of multi-source cross domain data, the accurate prediction of cellular traffic is facing great challenges. In this work, we propose a Dense-convolutional-neural-network-based traffic prediction model for fusion of Multi-Source Data(MS-DCN). The model includes spatio-temporal module and external feature module. We leverage DenseUnit architecture to capture temporal characteristics with different degree of dependence and study spatial characteristics. In external feature module, the same DenseUnit architecture is employed to capture multi-soure factors. Spatiotemporal features and external features are effectively integrated to achieve accurate prediction of large-scale wireless communication traffic. In the experimental part, MS-DCN is proved to have higher prediction accuracy than the existing models on the actual cellular data set.

Keywords: Wireless Traffic Prediction · Spatio-Temporal Data · Multi-source Data

1 Introduction

With the rapid development of mobile internet and the continuous growth in the number of mobile devices, along with the widespread adoption of services like high-definition videos and file downloads, mobile data traffic has undergone explosive growth. In 4G and 5G networks, the spatial and temporal fluctuations within each cell have significantly increased. If traditional static design methods

Supported by organization x.

X. Chen et al. (Eds.): IoTaaS 2023, LNICST 585, pp. 255–267, 2025.
https://doi.org/10.1007/978-3-031-70507-6_19

continue to be employed, many base stations will remain in a low-load operational state for prolonged periods. Therefore, accurate modeling and prediction of wireless traffic will assist operators in optimizing their operations.

However, there are still many difficulties in effective prediction of cellular traffic. The essence of community traffic is the reflection of crowd activity. First, the spatio-temporal characteristics of crowd activities lead to the spatio-temporal dependence of community traffic. Nowadays, individuals in real world have dual identities–individuals in the social network and users with intelligent terminals in the communication network. Users' behavior (daytime activities, night rest) in the temporal dimension combined with their activities (commuting, shopping, entertainment, work) in the spatial dimension will infiltrate into the communication network through intelligent terminal carried by them, and further form the cell business flow in a specific time and space frame. In addition, with the popularity of modern means of transportation (subway, car), the range of users' activities in the spatial dimension is increasing, which leads to the global dependence of traffic between geographically distant communities. Therefore, how to capture the highly complex spatio-temporal dependence of cell traffic in a large area is an arduous task. Second, the crowd activity is affected by many external factors, which brings out traffic's multi-source cross-domain property. For instance, the deployment density of cellular base stations determines the upper limit of traffic in the area; User data on social network reflects the demand of mobile users for online services, and the weather and external environment will affect the activity of the crowd. As a result, how to efficiently integrate cross domain heterogeneous data (base station distribution, social network, weather, etc.) with spatio-temporal traffic data is also a problem to be solved. Moreover, once the above difficulties are solved and accurate traffic prediction is realized, how to realize the function of traffic prediction in a wider scope simultaneously (different regions in the same city, and even different countries) is a problem that needs to be considered.

Motivated by the aforementioned problems, this paper aims to solve. propose a deep learning based cross domain spatio-temporal data traffic prediction model MFMS-DenseNet. The contributions of this work can be summerized as follows:

- Firstly, this paper proposes the architecture of MS-DCN, which jointly learns the spatial-temporal correlation features and interdependence of multimodality traffic data by multimodal deep learning architecture.
- Secondly, Extensive experiments are conducted using a realistic dataset of Beijing city, which demonstrates the efficiency of our proposed MS-DCN.

The remainder of the paper is organized as follows.

2 Related Work

With the development of mobile communication technology and Internet industry, the demand for wireless traffic prediction technology is increasing. Now the rise of artificial intelligence and machine learning has injected infinite vitality into the field of wireless traffic prediction. From the traditional time series model to the complex spatio-temporal traffic prediction model integrated with deep learning technology, many scholars at home and abroad have made large amounts of related research, which brings traffic prediction a more complete foundation.

In the traditional traffic prediction, the prediction objects are mostly single base station traffic, which can be regarded as time series without spatial characteristics. For instance, Sadek et al. [1] proposed a k-factor Geigenbauer Autoregressive Moving Average Model to predict high-speed network traffic time series with long-term dependence.

With the introduction of Convolutional Neural Network (CNN), an increasing number of researchers have taken advantage of its ability to capture spatial characteristics quickly and efficiently and applied it to traffic prediction models. The Residual Network(ResNet) can train a deep convolutional neural network model, which has been applied to a large number of traffic prediction research since it was first proposed in 2016. For instance, work [2] designed a simplified deep residual network model, which used the combination of two convolution layers and one residual units to learn the spatio-temporal characteristics of traffic.

In order to obtain more accurate prediction results, more researchers prefer to combine convolutional neural network and recurrent neural network to build traffic prediction model, which capture the spatial and temporal characteristics of traffic respectively. Gu et al.[3] established a traffic flow prediction model based on Spatio-temporal Convolution Recurrent Neural Network (STGCRNN), which uses graph convolution to obtain the spatial dependence of traffic flow data, and uses gated recurrent neural network to obtain the time dependence of traffic flow data. For instance, Zhang et al.[4] designed a Multi-Channel Sparse LSTM model, which can capture multi-source network traffic information while considering long-term and short-term dependence. Similarly, Wu et al. [5] proposed a cellular network traffic prediction model STA-LSTM based on improved Spatio-Temporal Multi-Level Attention Mechanism and LSTM, which tends to retain long-term dependent input data[6].

The above research has made some achievements in traffic prediction. Based on the previous research, this work establishes a cellular traffic prediction model for multi-source spatio-temporal data. The DenseNet convolution neural network is used to capture the spatio-temporal characteristics more completely. At the same time, the model combines the impact of multi-source cross domain data on cell traffic to extract other attribute characteristics of traffic. It can accurately predict the flow of large-scale cells.

3 Traffic Prediction Model

In this section, we first introduce our wireless communication network prediction model MFMS-DenseNet, including the design principle of the spatio-temporal module and the external feature module.

3.1 Model Framework

Figure 1 shows the architecture of MS-DCN, which is based on multi-source cross-domain data and dense connection network. The model is composed of two main components: spacio-temporal module and external feature module. The spatio-temporal module includes closeness layer, period layer, trend layer and parameter matrix fusion part. The external feature module learns the social attributes (weekdays and holidays) and weather attributes of traffic.

Based on the analysis of the closeness, period and trend of traffic in time dimension, we first extracts the data set by sliding window. Since the selected data set can be seen as the change of traffic distribution in two-dimensional space with time and has the characteristics of video stream, for the historical observations which possess three kinds of temporal properties, we collect data with different timestamps, and then feed collected data into three same sub modules for training, so as to capture three kinds of time dependence: closeness, period and trend, respectively. The input data is normalized by Min-Max to improve the convergence speed of the model. After data entering the sub module, the spatial dependence is captured by convolutions in the DenseUnit which is mainly composed of stacked DenseLayers, and the learned features of sub modules are combined by parameter matrix fusion.

The influence of closeness, period and trend characteristics of wireless traffic on the final predicted output is different. We use the method of parameter matrix fusion to reflect the strength of different time characteristics with different weights, and define the weights as W_{p1}, W_{p2}, W_c, and W_t respectively. During the training process, the learning parameters are transformed into tensors with the same form as the input features, so that the Hadamard product can be carried out, and then we get the output of matrix fusion, i.e. the output of spatio-temporal module. The output can be written as

$$X_T = W_{p1} \circ X_{p1}^7 + W_{p2} \circ X_{p2}^7 + W_c \circ X_c^7 + W_t \circ X_t^7 \tag{1}$$

where $\circ$ represents Hadamard product, the learning parameter matrix and the corresponding feature output matrix are multiplied respectively, and the period, closeness and trend are added to obtain the predicted feature output. Through the iterative training of neural network, the model calculate the gradient and update weight continuously, so as to accurately determine the degree of determination of period, closeness and trend characteristics on wireless traffic prediction.

The external feature module uses the same DenseUnit as the spatio-temporal module to learn from the quantified data of working days, holidays and weather factors. The input is regarded as a three channel feature, and the two channel prediction results are output after convolution of the module, which is defined as X_{Ext}. Afterwards, the output of the spatio-temporal module is spliced with the

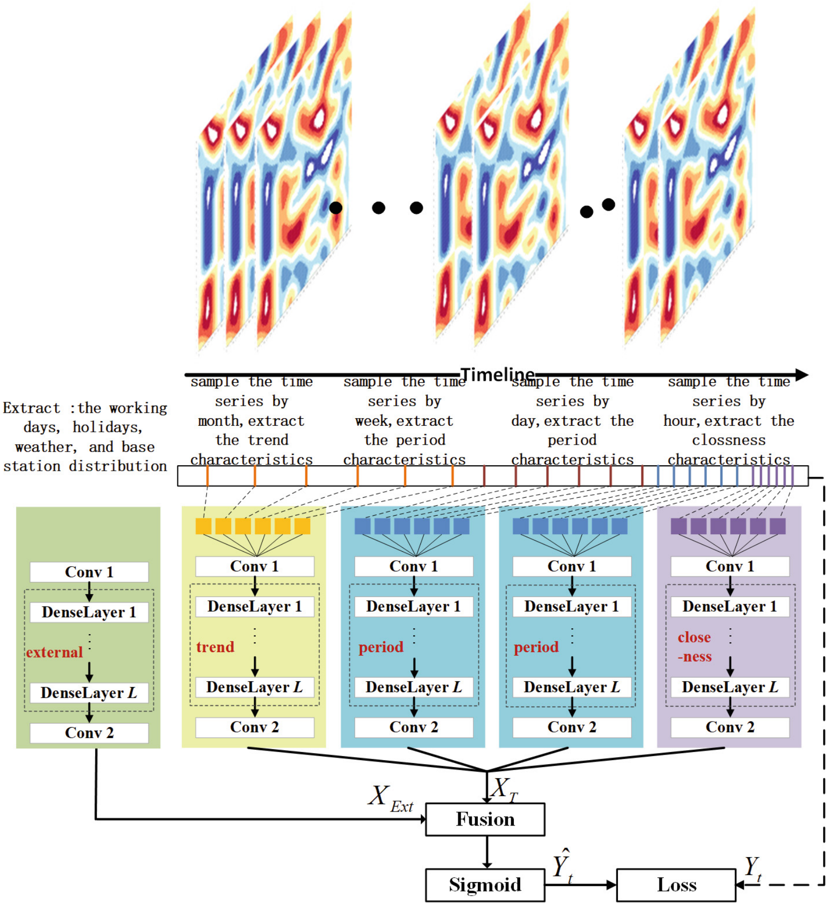


Fig. 1. MS-DCN Architecture

output matrix of the external feature before activation. We achieve the prediction output of model MFMS-DenseNet at time t, which can be written as:

$$\hat{Y}_t = \sigma(X_T \oplus X_{Ext}) \tag{2}$$

where $\oplus$ represents splicing operation, σ represents Sigmoid activation function.

Although we have used Min-Max normalization method to scall all input data into the range [0,1], the range of the output prediction data has changed after the complex operation of the model. However, the range of the real traffic label is [0,1]. In order to facilitate the error calculation between the prediction data and the real data, it is necessary to select the sigmoid activation function to map the prediction results to [0,1] one by one.

According to the predicted results and the real results, we use L2 norm (Mean Squared Error, MSE) to represent the loss function and update the parameters of the whole model during the training process:

$$L(\boldsymbol{W}) = ||Y_t - \hat{Y}_t||_2^2 \tag{3}$$

where $\boldsymbol{W}$ refers to the set of parameters trained by the model to minimize the loss function, Y_t refers to true value. When the loss function is the smallest, the weights W_{p1}, W_{p2}, W_c, W_t and W_{Ext} also get the optimal value.

3.2 Spatio-Temporal Module

In the previous chapter, we proved that the cell traffic has temporal and spatial attributes, thus wen should set the spatio-temporal module first when building the traffic prediction model. This module includes the period layer (short period and long period layer), the closeness layer, the trend layer and the parameter matrix. Among them, the period layer, the closeness layer and the trend layer adopt independent dense connection network units. We sample the historical traffic data with a certain dependence, and then map sampled data into spatio-temporal data, put them into period layer, closeness layer and trend layer in the form of "video stream" respectively for training and feature learning, so as to mine the spatiotemporal dependence of traffic. In parameter matrix fusion part, we fuse the features learned in each time layer with appropriate weights to obtain the final prediction output of the spatio-temporal module.

Period Layer In order to get the final prediction output of the spatio-temporal module more accurately, we sample the time series in days and weeks respectively in the period layer to express the short period and long period.

Short Period Layer If $p1$ is set as the size of short periodic dependence, the historical traffic data of length l_{p1} is selected as the input of period layer. For instance, in order to predict the wireless traffic of a cell at 15:00 on January 5th, the $p1$ value is 3, then we need to take the traffic of the cell at 15:00 on January 2th, January 3th and January 4th as the input which is similar to a three-channel video. The input traffic is trained by DenseUnit, and the period layer model is shown in Figure 9 (a).

- **DenseUnit**

Denseunit contains a 3 × 3 convolution layer, DenseBlock framework and the last layer 1 × 1 convolution layer. We set the input traffic size to 20 × 20, $p1$ is 3. Because of the need to consider the input and output traffic, the initial input of DenseUnit in period layer should be regarded as a 6-channel input. The multi-channel input first passes through the 3 × 3 convolution layer. The depth of 3 × 3 convolution kernel is also 6. We set the number of output features of the first layer is 32, thus we need to use 32 3 × 3 × 6 convolution kernel to turn the 6-channel input of this layer into 32 feature outputs, which are defined as tensor X_{p1}^0. In addition, the zero filling operation is used to keep the output feature size unchanged. The output of the first convolution layer passes through DenseBlock architecture, and then passes through the last convolution layer. The last convolution layer uses 1×1 kernel function to ensure output two flow feature graphs, i.e. two 20×20 3D prediction graph, which represents the uplink traffic and downlink traffic of the periodic layer respectively.

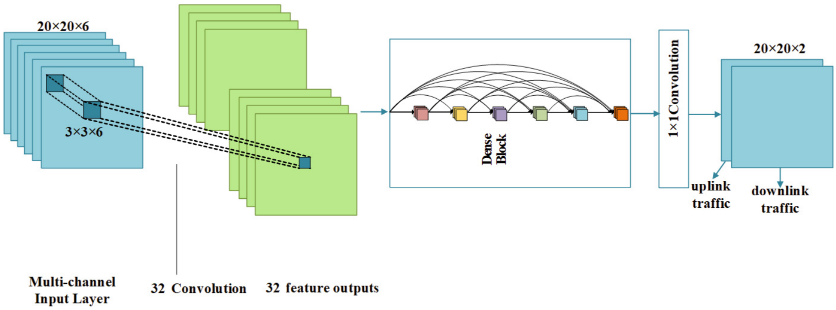


Fig. 2. DenseUnit

- **DenseBlock**

The 32 output features of the first convolution layer is taken as the initial input of DenseBlock. In this work, we define that DenseBlock contains six DenseLayers, and the dense full connection is used between each layer, namely the input of each layer is channel merging for the learned features of all previous layers. Assuming that the number of layers is i, and we define the number of output channels of each DenseLayer is 12, then the input of layer i is 32+12i and the output of the whole DenseBlock is 104.

- **DenseLayer**

The structure of one DenseaLayer in DenseBlock is shown in Fig. 10, including Batch Normalization(BN) + ReLU activation function + 1×1 convolution kernel, BN + ReLU + 3×3 convolution kernel. The combination of these six steps is defined as a composite nonlinear function $f_l(.)$, $l = 1, 2, \ldots, 6$. BN normalizes the input to avoid the input data offset to affect the training speed; ReLU function activates one part of neurons and makes the output of the other part of neurons to 0, which can alleviate the phenomenon of over-fitting; convolution kernel is used to extract features. The input features and convolution kernels are point multiplied and superimposed to get the corresponding features. The above nonlinear function BN-ReLU-Conv(1×1)-BN-ReLU-Conv(3×3) is the bottleneck structure, of which 1×1 convolution can be used to reduce the dimension of the picture, thus solving the problem of large number of channels caused by DenseNet full connection, greatly reducing the calculation amount.

The vector obtained after DenseBlock is denoted as:

$$X_{p1}^6 = f_l(X_{p1}^0 \oplus X_{p1}^1 \oplus X_{p1}^2 \oplus X_{p1}^3 \oplus X_{p1}^4 \oplus X_{p1}^5) \tag{4}$$

where $\oplus$ represents splicing operation, X_{p1}^0 represents initial input and X_{p1}^1 refers to the output vector of the first DenseLayer. The output of DenseBlock goes through the last 1×1 convolution kernel and a dual-channel output is obtained, which respectively represents the final prediction result of short period layer: incoming flow and outgoing flow. The result is written as X_{p1}^7.

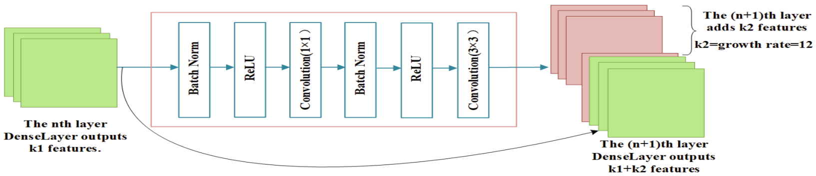


Fig. 3. One DenseLayer Architecture

Long Period Layer If $p2$ is set as the size of long period dependence, the historical traffic data of length l_{p2} is selected as the input of long period layer. For example, in order to predict the wireless traffic of a cell at 15:00 on January 5th, the $p2$ value is 2, then we need to take the traffic of the cell at 15:00 on December 22nd and December 29nd of the previous year as a dual-channel input. We still use DenseUnit to obtain the prediction output vector of the long period layer, which is written as X_{p2}^7.

Closeness Layer If c is set as the size of closeness dependence, the historical traffic data of length l_c is selected as the input of closeness layer. Assume we still predict the wireless traffic of a cell at 15:00 on January 5th, the c value is 3, then we need to take the traffic of the cell at 12:00, 13:00 and 14:00 on January 5th of the cell as a three-channel input. Closeness layer still use DenseUnit to obtain the prediction output vector, which is written as X_c^7.

Trend Layer If t is set as the size of trend dependence, the historical traffic data of length l_t is selected as the input of trend layer. Assume we still predict the wireless traffic of a cell at 15:00 on January 5th, the t value is 2, then we need to take the traffic of the cell at 15:00 on November 5th and December 5th of the previous year as a dual-channel input. Trend layer still use DenseUnit to obtain the prediction output vector, which is written as X_t^7.

3.3 External Feature Module

In addition to the most basic spatio-temporal attributes, the community traffic also has other attributes. We focuse on the social attributes and weather attributes that have great impacts on traffic prediction. Social attributes include: working days and holidays; weather attributes include: weather conditions.

In the model MS-DCN, we propose an external feature module, and use the same DenseUnit to extract the working days, holidays and weather features of traffic. We first add labels on the data sets of working days, holidays and weather: add label 1 on working days/holidays, define non-working days/non-holidays as 0; add labels 1, 2 and 3 on sunny days, cloudy days and rainy days respectively. We extract their feature vectors respectively in hours, and then reshape each feature into 20×20 size. These three features are defined as X_D, X_H and X_W, which is regarded as a three-channel input. After going through DenseUnit, the dual-channel output is obtained, which is namely the prediction output of the whole external feature module X_{Ext}.

4 Experiment Results

In this section, we present the experiment results. Subsection 2 examines the performance of the proposed model from three aspects: traffic prediction performance, comparing methods, and comparing results between different functional regions.

4.1 Data Discription and Hyperparameter Setting

In this section, to evaluate the proposed model, we conduct experiments on the Beijing dataset. This paper utilizes the air interface total traffic data for August 2021 in Haidian, Beijing, and conducts analysis and processing on an hourly basis. The dataset is dividied into two segments: the training set encompasses hourly measurements spanning from August 1st, 2021, to August 25th, 2021, totaling 600 records, while the testing set comprises hourly measurements captured from August 26th, 2021, to August 31st, 2021, amounting to 144 records.The simulation parameters are set as Table 1.

The hardware configuration of the experimental environment includes an Intel(R) Core(TM) i9-10940X CPU @ 3.30GHz, an RTX 3090 GPU with 24GB of VRAM, and 128GB of CPU memory. The software configuration consists of Python 3.8, PyTorch 2.0.1 with CUDA 11.7 and cuDNN 8.2.4 support.

Table 1. SIMULATION PARAMETERS

Parameters	Values
Length of data	3
Learning rate	0.001
Batch size	32
Epoch size	300
close size	3
period size	3
Size with periodic dependence	3

4.2 Model Performance

Figure 4 represents traffic curves depicting actual and predicted values for a randomly selected region. The MS-DCN model proposed in this paper demonstrates strong performance on the multi-source dataset, with a high degree of curve fitting and accurate control over traffic peak occurrences.

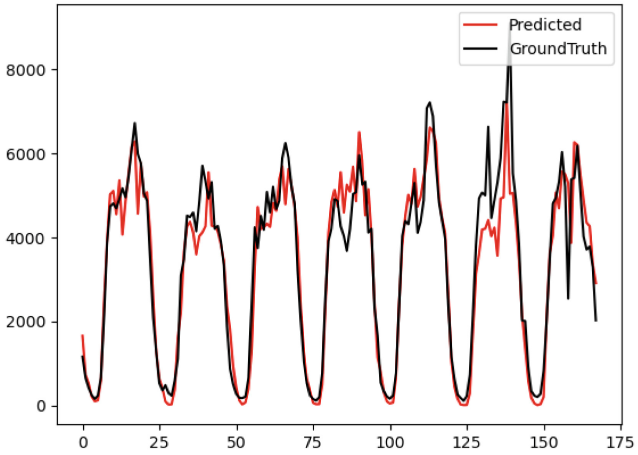


Fig. 4. Traffic volume predicted by MS-DCN

To furtherly verify the predictive performance of the proposed model in this paper, this paper compares the predictive errors of several models, including ARIMA[7], LSTM [8], XGBoost [9], DenseNet[10], and Holt-Winters [11], with the prediction error of the MS-DCN model proposed in this paper, using the same dataset as shown in Figure 5. In this experiment, the evaluation criterion

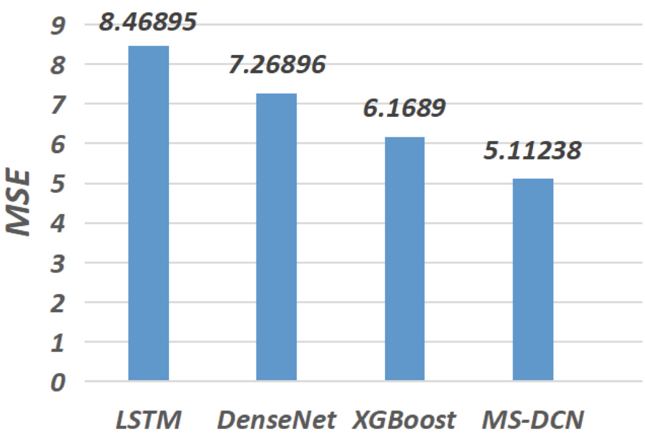


Fig. 5. Different Models' Traffic Prediction Errors

utilized was the Mean Squared Error (MSE). The LSTM and DenseNet models performed poorest, achieving MSE values of 8.46895 and 7.26896, respectively. On the other hand, the XGBoost and DenseNet models exhibited the most favorable results among the four considered models. Particularly, the MS-DCN excelled with the most optimal performance, achieving an MSE of 5.11238.

To validate the predictive accuracy of the proposed model across multiple source datasets, this paper selects two typical geographic functional areas in Beijing: University Region(UR) and Scenic Region(SR) for model prediction. In Figure 6 of the map of Beijing, the red region represents UR, and the black region SR. The predicted results by MS-DCN are shown in Figure 7.

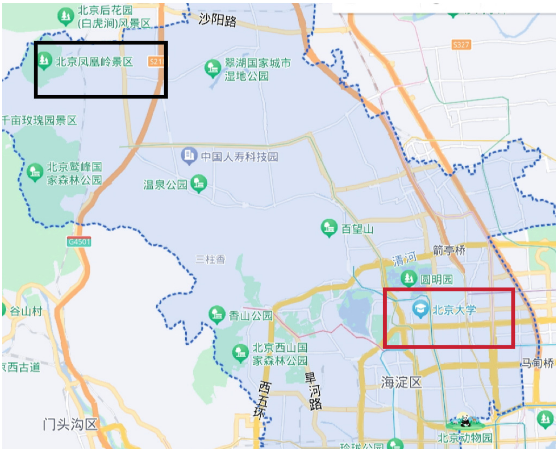
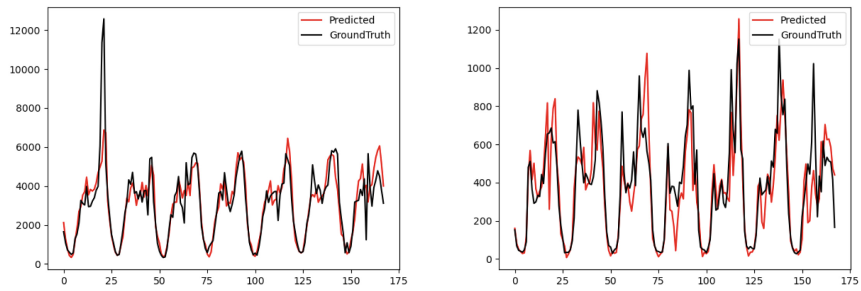


Fig. 6. Different functional regions in Haidian District, Beijing

From the Figure 7 (a) and (b), it is evident that MS-DCN performs well in terms of overall predictive performance. However, during periods of high network traffic, the model's predictions show some slight inconsistencies with the actual values. This may be attributed to certain challenges the model faces in handling data characteristics, complexity, or nonlinear dynamic changes during peak traffic hours, resulting in a slight fluctuation in prediction accuracy during these time periods.



(a) Traffic volume predicted by MS-DCN in UR (b) Traffic volume predicted by MS-DCN in SR

Fig. 7. Traffic Prediction for Different Functional regions by MS-DCN

5 Conclusion

In this work, we propose the architecture of MS-DCN, which is based on multi-source cross-domain data and dense connection network. The model is composed of two main components: spacio-temporal module and external feature module. The spatio-temporal module includes closeness layer, period layer, trend layer and parameter matrix fusion part. The external feature module learns the social attributes (weekdays and holidays) and weather attributes of traffic.In conclusion, the experiments demonstrate the effectiveness of the proposed model's performance.

References

1. Sadek, N., Khotanzad, A.: Multi-scale high-speed network traffic prediction using k-factor gegenbauer arma model. In: 2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577), vol. 4, pp. 2148–2152 (2004)
2. Hu, X., Dai, G., Ge, Y., Ning, Z., Liu, Y.: A simplified deep residual network for citywide crowd flows prediction. In: 2018 14th International Conference on Semantics, Knowledge and Grids (SKG), pp. 60–67 (2018)
3. Cao, Y., Chen, Y.: Qoe-based node selection strategy for edge computing enabled internet-of-vehicles (ec-iov). In: IEEE Visual Communications and Image Processing (VCIP) **2017**, 1–4 (2017)
4. Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H.: T-gcn: a temporal graph convolutional network for traffic prediction. IEEE Trans. Intell. Transp. Syst. **21**(9), 3848–3858 (2020)
5. Xiaolong, L.I., Pan, G., Zhaohui, W.U., Guande, Q.I., Shijian, L.I., Zhang, D., Zhang, W., Wang, Z.: Prediction of urban human mobility using large-scale taxi traces and its applications. Front. Comp. Sci. **6**(001), 111–121 (2012)
6. Tian, Z.: Network traffic prediction method based on wavelet transform and multiple models fusion. Int. J. Commun. Syst. (2) (2020)

7. Wang, K., Ma, C., Qiao, Y., Lu, X., Dong, S.: A hybrid deep learning model with 1dcnn-lstm-attention networks for short-term traffic flow prediction. Physica A: Statistical Mechanics and its Applications. (12), 126293 (2021)
8. Wang, J., Tang, J., Xu, Z., Wang, Y., Xue, G., Zhang, X., Yang, D.: Spatiotemporal modeling and prediction in cellular networks: a big data enabled deep learning approach. In: IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, pp. 1–9 (2017)
9. Zhang, C., Zhang, H., Qiao, J., Yuan, D., Zhang, M.: Deep transfer learning for intelligent cellular traffic prediction based on cross-domain big data. IEEE J. Sel. Areas Commun. **37**(6), 1389–1401 (2019)
10. Zhang, C., Zhang, H., Yuan, D., Zhang, M.: Citywide cellular traffic prediction based on densely connected convolutional neural networks. IEEE Commun. Lett. **22**(8), 1656–1659 (2018)
11. Zhao, Z., Chen, W., Wu, X., Chen, P.C.Y., Liu, J.: Lstm network: a deep learning approach for short-term traffic forecast. IET Intel. Transport Syst. **11**(2), 68–75 (2017)