Proceedings of CECNet 2022
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA220519

Traffic Flow Imputation Based on Multi-Perspective Spatiotemporal Generative Adversarial Networks

Guojiang SHEN, Nali LIU, Yinghui LIU, Wenfeng ZHOU and Xiangjie KONG [1]
College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China

Abstract. Traffic data occupies an important position in intelligent transportation systems (ITS). However, the collected traffic data is often incomplete. We propose a generative adversarial network (GAN) model based on multi-perspective spatiotemporal learning (MST-GAN) to repair data. To achieve the effect of interpolating data from three perspectives: temporal, spatial, and spatiotemporal, we utilize chained generator with independent parameters to progressively refine the learning of temporal and spatial features. In addition, we achieve high-level fusion of multi-perspective features by adversarial between multiple generators and one discriminator. We conduct experiments on two real datasets, showing that the imputation effect of the MST-GAN model is better than other baseline models under different missing patterns. For example, the root mean square error (RMSE) is less than 7.5% and the mean absolute error (MAE) is less than 5% in the random missing scenario, which is much lower than the best performance error of other models.

Keywords. Traffic flow imputation, Generative adversarial network, Multi-perspective spatiotemporal learning, Deep learning.

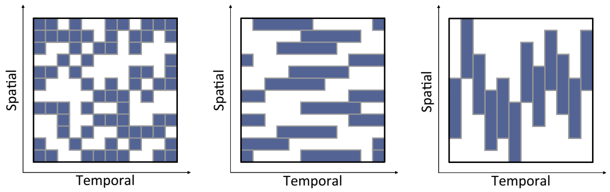
1. Introduction

With the rapid integration of big data analysis into ITS[1], the construction of intelligent cities has been promoted. Unfortunately, missing data can occur throughout a distributed urban network due to malfunctioning sensors or communication errors among the collection points[2]. Discarding records is the easiest way to deal with missing data. However, incomplete traffic information may seriously affect the prediction accuracy[3]. Therefore, it is crucial to propose an efficient traffic data imputation method. Most existing deep learning imputation models focus on a single road segment and only consider the temporal dependencies of the data. Most of these methods are based on recurrent neural network (RNN), like LSTM-M[4]. While these RNN-based methods ignore the spatial correlation between different roads. It is difficult to impute the traffic conditions of the entire road network. Recently, more and more works have applied graph-based deep learning techniques in various traffic tasks and have achieved excellent performances[5].

[1]Corresponding Author: Xiangjie Kong, College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China, E-mail:xjkong@ieee.org.

However, graph-based methods are mainly used to deal with traffic flow prediction and rarely for traffic data imputation. The few existing graph-based imputation methods do not consider the temporal correlation of traffic flow, so the imputation effect is not ideal. The current spatiotemporal data generation work integrates spatiotemporal features to a certain extent. But most of them just use the simple linear superposition of the temporal and spatial model from separate perspectives of temporal and spatial, they do not consider high-level spatiotemporal fusion from multiple perspectives[6].

Missing patterns can seriously affect the performance of the imputation method. We classify the missing patterns in ITS into three categories in this paper: 1) Random missing (RM) (Fig.1(a)), where missing values are completely independent of each other and displayed as randomly scattered points for each sensor (or road); 2) Temporal correlated missing (TCM) (Fig.1(b)), where missing values are dependent in the time dimension and appear as a consecutive time interval for each sensor (or road); 3) Spatially correlated missing (SCM) (Fig.1(c)), where missing values are dependent in the spatial dimension and appear at neighboring sensors or connected road links for each time slot[7]. We will verify the performance of our model on the above three missing patterns.



(a) RM (b) TCM (c) SCM

Figure 1. Patterns of missing data.

Our Contribution. To better estimate the lost traffic data, this paper draws on the idea of combining spatiotemporal interaction models and adversarial generative models. Compared with single-generator generative adversarial networks, we construct chained generators for multi-perspective fusion training based on iterated speech enhancement GAN (ISEGAN) and deep speech enhancement GAN (DSEGAN)[8]. Our contributions are as follows: (1) This is the first work to apply a chained generator to impute traffic data. We design a chained generator with independent parameters and one discriminator to capture spatiotemporal features in stages, achieving high-level fusion of spatiotemporal features during adversarial process. (2) We adopt a multi-perspective spatiotemporal learning strategy to impute traffic data by comprehensively analyzing spatiotemporal features from three perspectives: temporal, spatial, and spatiotemporal. (3)The model has been verified on two real large traffic datasets. The results show that the proposed MST-GAN model is much more efficient than other baseline models under different missing patterns.

Related Works. Missing data is inevitable in the Internet of Vehicles (IoV). At present, many methods have been proposed to impute data. Typical traditional imputation methods include support vector regression (SVR)[9], auto-regressive integrated moving average model (ARIMA)[10], etc. These methods ignore information with missing data and only use data before the missing data point, which can affect imputation accuracy because of not taking full advantage of the dataset.

Table 1. The limitations of these existing methods.

Category	Models	Limitations
Tradition	SVR	They ignore information with missing data and only use data before the missing data point.
	ARIMA	
RNN-Based	GRU-D	They only consider temporal information and ignore the global spatial dependencies in the traffic network.
	BRITS	
GAN-Based	GAIN	They focus on non-time-series datasets and do not take targeted measures to deal with spatiotemporal relationships.
	PC-GAIN	
Spatiotemporal	LSTM_AEs	They do not consider the effect of adversarial learning on the fusion of spatiotemporal features.
	GACN	

In recent years, the deep learning method for data imputation has gradually entered the public field. Che et al.[11] proposed the GRU-D model, GRU-D introduces two decay mechanisms so that the influence of the variable will gradually disappear over time when the variable is lost for a period of time. GRU-D uses two representations of the missing pattern, namely masking and time interval, and effectively integrates them into a deep model architecture. So it not only can capture the long-term time dependence in the time series but also use the missing mode to obtain better prediction results. In addition, Cao et al.[12] put forward BRITS model. The model can directly learn missing values in a bidirectional recursive dynamical system without the need for any specific assumptions. Both GRU-D and BRITS models are based on RNN, which only consider the temporal correlation of data, but do not fully consider the impact of spatial information on road network data.

Generative adversarial network is widely used in image data processing, it has been found to have good performance in data imputation in recent years. Yoon et al.[13] proposed the GAIN model based on GAN, they use the generator and discriminator adversarial learning to try to model the distribution of the original data and then achieve the effect of imputing the missing data. Wang et al.[14] proposed the PC-GAIN model, which adds a pre-training process based on GAIN. The PC-GAIN model is proposed to learn the potential category information contained in the subset of low-missing rate data, then it uses synthetic pseudo-labels to determine the auxiliary classifier, and improve the data imputation effect through the joint training of generator, discriminator and classifier. However, GAN-Based model does not consider the spatiotemporal correlation of data, so the imputation effect of traffic data is not very ideal.

Li et al.[15] proposed a deep spatiotemporal time-series missing data imputation model, called LSTM-AE$_s$, to enhance the imputation performance and handle multiple missing patterns. The proposed model combines deep auto-encoder (DAE) and long-short term memory (LSTM) for extracting spatiotemporal features to estimate missing values in multiple time series. Ye et al.[16] proposed a graph attention convolutional network model (GACN) for traffic-missing data imputation, which follows the encoder-decoder structure and introduces a graph attention mechanism to learn the traffic graph. The spatial correlation of traffic data is collected by adjacent sensors on the graph, and after the attention layer of the graph, a temporal convolutional layer is superimposed

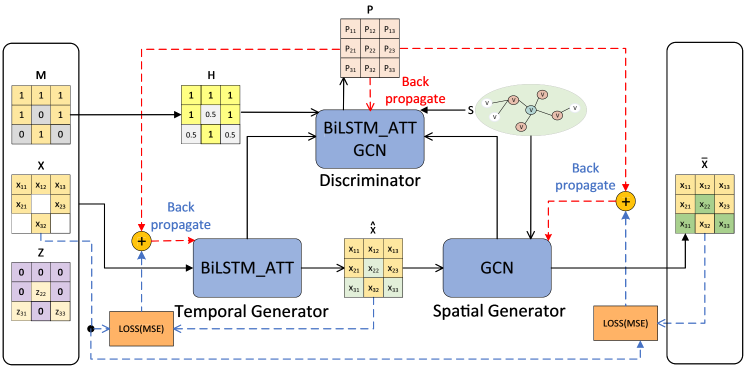


Figure 2. The MST-GAN framework.

to extract the relationship in the time series, so the traffic data can be imputed with higher quality by extracting the typical spatiotemporal features. However, time-domain convolution still has insufficient ability to obtain related information of different lengths in the extraction of time series information. Instead of a simple linear stack of spatial and temporal models, LSTM-AE$_s$ and GACN implement spatiotemporal feature fusion learning through encoder and decoder. At last, we list the summary and limitations of these existing methods in Table 1.

Organization. The remainder of this paper is organized as follows.Section 2 describes the architecture and details of the proposed MST-GAN framework. Section 3 discusses the results of the experiment. Finally, the paper is summarized in Section 4.

2. System Model

2.1. The Overall Framework Analysis

In this section, we will introduce the proposed MST-GAN framework, as shown in Fig.2. We adopt a learning strategy from three perspectives: temporal, spatial, and spatiotemporal. The temporal generator and spatial generator impute the missing data from the temporal and spatial perspective respectively. We embed temporal and spatial features into two different generators respectively, which can gradually refine the learning of temporal and spatial features in a staged manner. We design different generators with independent parameters so that the model can flexibly learn different spatiotemporal enhancement features at different stages[8]. In addition, the spatial generator inherits the enhanced features of the temporal generator. From the spatiotemporal perspective, the discriminator can combine the temporal and spatial generators to achieve a high-level fusion of spatiotemporal features in confrontation. Throughout the training process, the temporal and spatial generators strive to generate traffic flow data that is closest to the real distribution, and the discriminator judges whether the data is real or generated. The three networks achieve a dynamic balance in this confrontation. Specifically, we use bidirectional LSTM based on attention mechanisms (BiLSTM_ATT) to capture the time series features of traffic data and introduce graph convolutional network (GCN) to realize the spatial feature learning of the road network. Finally, the imputation of traffic data from three perspectives is achieved.

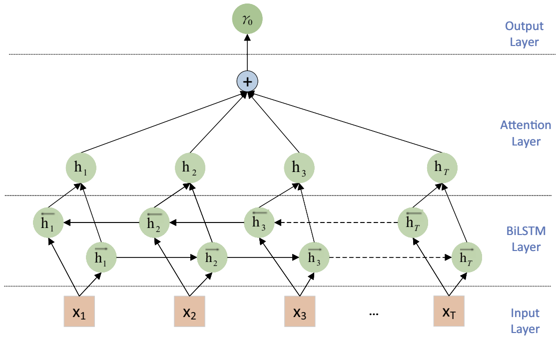


Figure 3. The network structure of BiLSTM_ATT.

2.2. Initial Input Data

The input to the model consists of five parts. Original observed traffic flow data matrix $X = (x_1, x_2, \cdots, x_N) \in R^{N*T}$ represents the traffic flow of N intersections. $x_i = (x_{i1}, x_{i2}, \cdots, x_{iT})$ represents the traffic flow of an intersection at T time gap, where x_{ij} represents the traffic flow situation of the i^{th} intersection in the road network at the j^{th} moment. Mask matrix $M \in R^{N*T}$ to characterize whether the road network data is missing, $M_{ij} = 0$ represents the lack of traffic flow data, otherwise $M_{ij} = 1$. The noise matrix Z is independent of other variables and makes the network produce a random distribution. To ensure the generator can generate samples based on the true underlying data distribution in the adversarial process, we introduce the indicator matrix H to provide additional information to the discriminator[13]. The road network matrix S maps the correlation information of different intersections in the road network.

2.3. Temporal Generator

The temporal generator (G_T) analyzes the characteristics of the road network from the temporal perspective. The G_T maps the initial random distribution to the real sample data. The core of G_T's network structure consists of BiLSTM_ATT, as shown in Fig.3. LSTM is suitable for processing time series data with long intervals and delays, but it only relies on the information of the previous moment to predict the next moment. We introduce BiLSTM_ATT whose main goal is to increase the available information to the network by taking contexts in both directions into account. However, the traffic conditions of road segments change over time, which makes the modeling challenging due to the high dynamics. To address this problem, we compute dynamic weights between different time points by applying a self-attention strategy, that can capture dynamic temporal correlations. Specifically, we first calculate the weight of each time series. Then, we weight and sum the vectors of the all-time series as the feature vector, which produces the output of the temporal generator. The Attention layer is calculated as follows:

$$
\begin{aligned}
V &= tanh(U) \\
\alpha &= softmax(W_\alpha^T V) \\
\gamma &= U\alpha^T
\end{aligned}
\tag{1}
$$

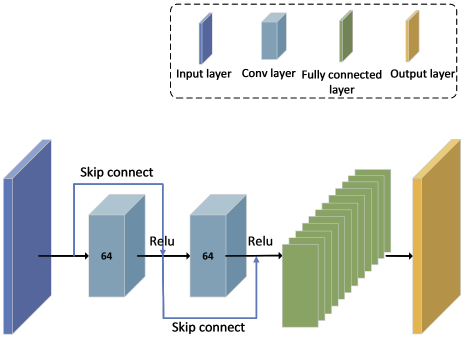


Figure 4. The network structure of GCN.

where the matrix $U = [u_1, u_2, \cdots, u_T] \in d^{N*T}$ (T is the length of time), $u_t = BiLSTM(x_t)$. $x_t = (x_{1t}, x_{2t}, x_{3t}, \cdots, x_{Nt})$ denotes the observation of each intersection in the t-time interval. $W_\alpha \in d^N$ is learned as a training parameter. $\gamma = U\alpha^T \in d^N$ is actually a weighted sum of each hidden state based on the distribution of attention.

In summary, we define the above as $\gamma_k = x_{T+k+1} = Attention(BiLSTM(x_{T+k}))$, $(k = 0, 1, \cdots, T-1, \gamma_k \in d^N)$, then there is $X_{BiLSTM\_ATT} = (\gamma_0, \gamma_1, \cdots, \gamma_{T-1})$, the final output value $\hat{X}$ of temporal generator is formulated as:

$$\begin{aligned} X_{BiLSTM\_ATT} &= Attention(BiLSTM(X)) \\ \hat{X} &= X \odot M + X_{BiLSTM\_ATT} \odot (1-M) \end{aligned} \tag{2}$$

2.4. Spatial Generator

The spatial generator (G_S) analyzes the characteristics of the road network from the spatial perspective. G_S strives to capture hidden spatial information to generate outputs that are closer to the real sample feature distribution. We know that the traffic flow between different intersections will affect each other. The relationship adjacency graph formed by different intersections is a non-euclidean structure. Graph neural networks (GNNs) can capture spatial dependencies of non-Euclidean graph structures[17]. Therefore, the core of G_S's network structure consists of GCN with two convolutional layers and a fully connected layer. The convolutional layer converts the input features into an internal hidden feature map and extracts local features. The fully connected layer reassembles these local features into complete features. Finally, we restore the feature maps to their original sizes. We know that as the depth of the network layer increases, it may bring about the problem of gradient dissipation. Therefore, we introduce skip connect to facilitate training deep networks in our model. In the G_S, we use ReLU as the activation function, which is relatively simple in gradient computation. As shown in Fig.4, both convolutional and fully connected layers are carefully designed to ensure that input and output spatial features are of the same size. At each time step of generation phase, G_S retains the position with the true value and populates the generated data into the missing position, thereby improving the guiding force of the observations to the model. After training with the spatial generator, we can impute missing values using the following equation:

$$\bar{X} = X \odot M + G_S(\hat{X}) \odot (1 - M) \tag{3}$$

2.5. Discriminator

The discriminator D achieves high-level fusion of spatiotemporal features from a spatiotemporal perspective. The generator strives to make the simulated data closer to the true value, and the discriminator strives to identify the observed and generated data, which is the process of two neural networks playing against each other. They eventually reach a dynamic equilibrium. In this paper, we achieve dynamic fusion of spatiotemporal features by confronting the discriminator with the spatiotemporal generator, abandoning the simple linear stacking scheme. We use a linear combination of the encoding loss and the adversarial loss to construct the generator loss according to a certain proportion, where the encoding loss function adopts the MSE function. We perform model optimization with two generators losses and a discriminator loss. The core structure of the discriminator network is composed of BiLSTM_ATT and GCN. The BiLSTM_ATT network structure is similar to the temporal generator. The GCN network contains two convolutional layers and a fully connected layer. Its structure is similar to the spatial generator. Besides, the indicator matrix H is introduced to provide additional information to the discriminator, where $H = X \odot M + 0.5 \odot (1 - M)$. The matrix H ensures that the generator generates samples based on the true underlying data distribution. The loss function of the discriminator can be expressed as:

$$L_D = -\frac{1}{n}\sum_{i=1}^{n}(M \cdot \log D(\bar{X}) + (1 - M) \cdot \log(1 - D(\bar{X})) \tag{4}$$

We define the minmax goals for MST-GAN as:

$$\underset{G_T, G_S}{min}\ \underset{D}{max} E_{\bar{X},M}[M \cdot \log D(\bar{X}) + (1 - M) \cdot \log(1 - D(\bar{X})] \tag{5}$$

3. Experiments Settings and Performance Evaluation

3.1. Data Description

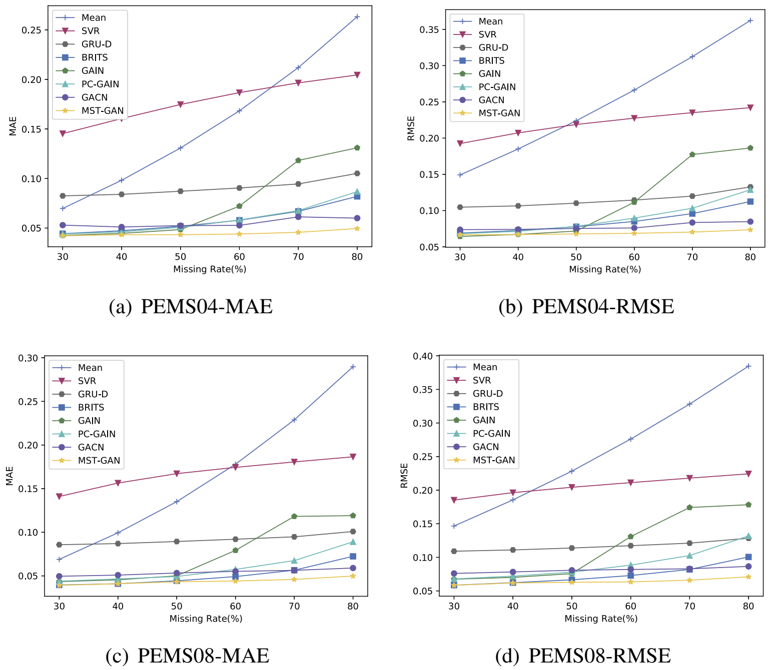
PEMS04 and PEMS08 datasets are from the California Department of Transportation Performance Measurement System. The data features include flow, occupy, speed, our paper only focuses on the flow data features for performance evaluation. PEMS04 dataset includes 307 detectors, the data range is 59 days, from January 1, 2018, to February 28, 2018. PEMS08 contains data collected by 170 detectors in a total of 62 days from July 1, 2016, to August 31, 2016. The time interval is 5 minutes, and 288 samples are generated at the same intersection every day.

3.2. Experiment Settings

In this paper, the first 50 days of the dataset are used as the training data, the remaining data is used as the test data. We choose T = 576 time steps (i.e., 5min * 576 = 48 hours) as the imputation window. During training, we use the sliding window method to perform imputation on [t, t + T), [t + T, t + 2T), [t + 2T, t + 3T), etc. The model was trained using the Adam optimizer with an initial learning rate of 0.01 and a batch size of 48. We normalize the attention coefficients of time series using Softmax. The performance of the model is compared with other baseline methods at different missing rates of 30%-80%. We choose MAE and RMSE as evaluation indicators. Here's how the evaluation metrics are calculated:

$$\begin{aligned} MAE &= \frac{1}{n}\sum_{i=1}^{n}\|\hat{y}_i - y_i\| \\ RMSE &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \end{aligned} \tag{6}$$

where n represents the number of missing data, $\hat{y}_i$ represents the prediction of missing value, and y_i represents the observed value.



(a) PEMS04-MAE (b) PEMS04-RMSE

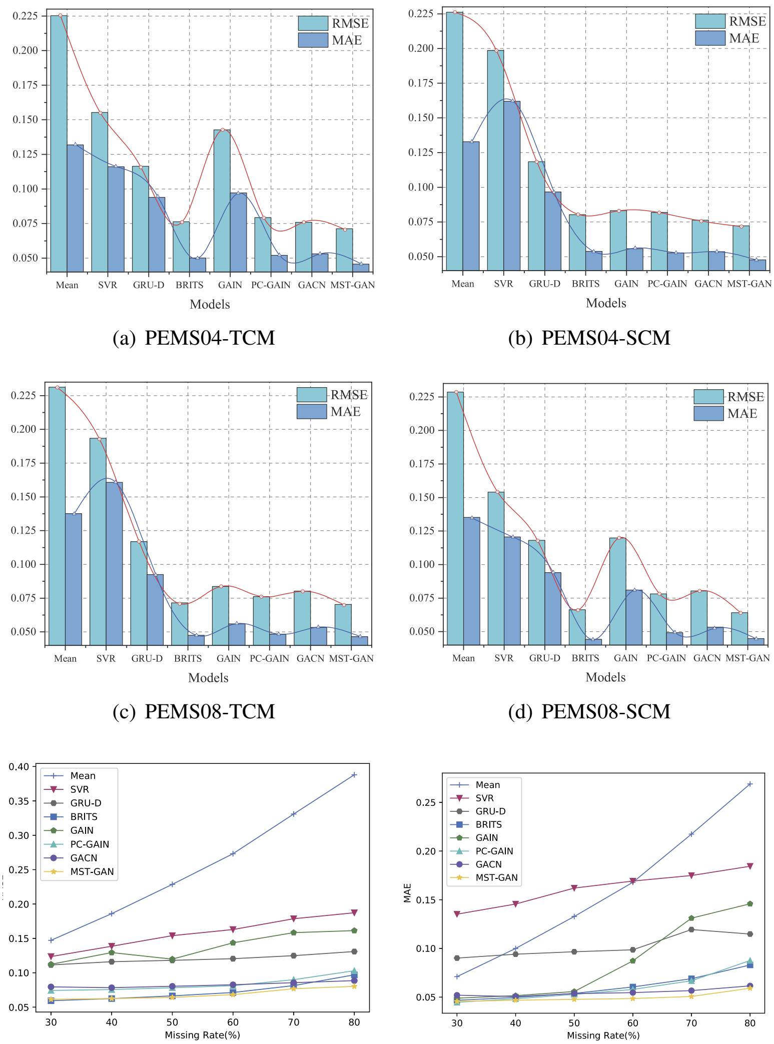
(c) PEMS08-MAE (d) PEMS08-RMSE

Figure 5. Performance comparison for imputation on PEMS04 and PEMS08 datasets in RM pattern.The abscissa represents different data missing rates, and the ordinate represents the loss of different models under different evaluation indicators.

3.3. Experiments Results and Evaluation

3.3.1. Baseline Analysis

Fig.5 shows the visualization results of data imputation errors under different evaluation metrics and different missing rates in RM pattern. We can draw the following conclusions: First, only the performance of MST-GAN and GACN based on spatiotemporal feature information is relatively stable when the missing rate is higher than 50%. But according to Fig.5, the imputation error of our model is lower than the GACN model. Second, in both PEMS04 and PEMS08 datasets, the RMSE of the MST-GAN model is less than 7.5%, and the MAE is less than 5%, which is lower than the error of the best performance of other models. Third, the performance of Mean is the worst in all datasets, and the impact of data missing rate is also the greatest. The performance of SVR is better than Mean, but the imputation performance of machine learning is significantly lower than that of deep learning. Fourth, GRU-D and BRITS are data imputation models based



(a) PEMS04-TCM

(b) PEMS04-SCM

(c) PEMS08-TCM

(d) PEMS08-SCM

(e) PEMS08-SCM-RMSE

(f) PEMS04-TCM-MAE

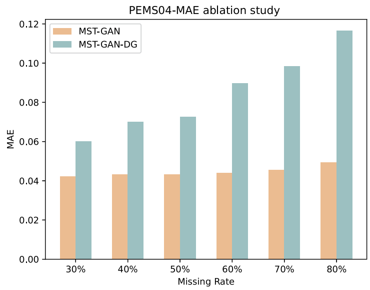
Figure 6. Performance comparison for imputation on PEMS04 and PEMS08 datasets in SCM and TCM patterns.

on time series features, and their performance is better than that of machine learning. Fifth, GAIN and PC-GAIN are data imputation models based on data distribution, they are difficult to apply to data imputation in the case of high missing rate, and their imputation error increases significantly when the missing rate is greater than 50%. This is because when the missing rate gradually increases, it is difficult for GANs based on learning data distribution to learn from historical data.

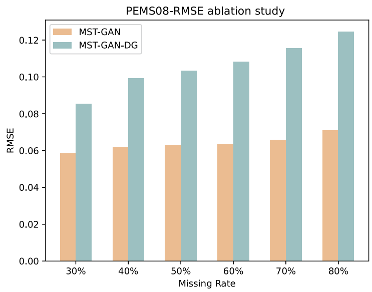
Fig.5 shows the performance of different models as the missing rate changes in the random missing mode. In Fig.6, we design the performance of each model in different missing patterns and datasets, where Fig.6 (a)-(d) are Typical of 50% missing cases. The abscissa represents different models, and the ordinate represents the evaluation indicators RMSE and MAE. We can find that our model has very stable traffic flow data imputation performance regardless of the evaluation metric. Fig.6 (e)-(f) are linear line charts of each baseline model under different datasets, missing patterns and evaluation metrics. It is evident that our model also outperforms other models in terms of TCM and SCM missing patterns. In conclusion, our method has better traffic data imputation performance than other baseline models in all missing cases.

3.3.2. Ablation Study

To verify that the multiple generators outperform the single spatiotemporal generator, we conduct ablation experiment MST-GAN-DG. Since RM is the most commonly studied missing type in literature, we use the RM scenario for experiments. MST-GAN-DG removes one generator and modifies both the generator and discriminator kernels to BiLSTM_ATT+GCN. The evaluation results are shown in Fig.7. According to the figure, we can know that the feature imputation method using temporal loss and spatial loss separation to train independent parameters is effective.



(a) PEMS04-MAE



(b) PEMS08-RMSE

Figure 7. Ablation Study. Fig.7 represents the data imputation loss of the MST-GAN and MST-GAN-DG models under different conditions.

4. Conclusion and Discussion

Missing values in traffic data are an unavoidable problem in ITS. Although many studies are addressing this problem, there are two important limitations: first, most of the existing spatiotemporal imputation methods are only simple linear superposition of temporal and spatial models, without considering high-level multi-perspective spatiotemporal fusion through chained generator adversarial learning; second, previous studies are mostly based on RM pattern and rarely consider other complex missing patterns. To fill these research gaps, we propose a deep learning model MST-GAN based on multi-perspective spatiotemporal learning to impute traffic data. MST-GAN adopts the strategy of analyzing and extracting road network features from three perspectives: temporal, spatial, and spatiotemporal, which makes our model learn the spatiotemporal information to the greatest extent. To achieve multi-perspective feature fusion, we employ adversarial between a chained generator and one discriminator to achieve high-level fusion of temporal and spatial information. The generator uses independent parameters to flexibly learn different enhanced features at different stages, making the overall model more flexible. Specifically, we capture the temporal and spatial correlations of traffic flow through bidirectional recurrent networks and graph convolutional networks, respectively. Furthermore, we introduce an attention layer to compute dynamic weights between different time points and focus on key temporal features. Finally, we conduct extensive experiments on real traffic datasets, comparing the performance of the MST-GAN model with other baseline models under three different missing patterns, showing that our results are all better than the baseline models. We validate the contribution of chained generators compared to single generator through ablation experiments, highlighting the superiority of the multi-generator module.

There are several directions for shortcomings and future works. First, the MST-GAN model does not consider the impact of other factors on traffic data imputation, such as weather factors, traffic accident information, etc. We plan to consider these external factors in our future work. Second, the MST-GAN model can be adapted to other spatiotemporal problems, such as route planning, traffic prediction, etc. Third, because deep learning methods are data-driven, models do not perform well without complete data. We consider combining tensor factorization method to improve the interpolation accuracy of the model. Fourth, the complex deep learning model has a slow convergence rate. We consider constructing a knowledge graph as a priori to improve the training efficiency and accuracy.

Acknowledgments

This work was supported in part by the ”Pioneer” and ”Leading Goose” R&D Program of Zhejiang under Grant 2022C01050, in part by the National Natural Science Foundation of China under Grant 62073295 and Grant 62072409, in part by the Zhejiang Provincial Natural Science Foundation under Grant LR21F020003, and in part by the Fundamental Research Funds for the Provincial Universities of Zhejiang under Grant RF-B2020001.

References

[1] Han L, Zheng K, Zhao L, Wang X, Wen H. Content-aware traffic data completion in ITS based on generative adversarial nets. IEEE Transactions on Vehicular Technology. 2020;69(10):11950–11962.

[2] Najafi B, Parsaeefard S, Leon-Garcia A. Estimation of missing data in intelligent transportation system. In: 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall). IEEE; 2020. p. 1–6.

[3] Liu A, Li C, Yue W, Zhou X. Real-time traffic prediction: A novel imputation optimization algorithm with missing data. In: 2018 IEEE Global Communications Conference (GLOBECOM). IEEE; 2018. p. 1–7.
[4] Tian Y, Zhang K, Li J, Lin X, Yang B. LSTM-based traffic flow prediction with missing data. Neurocomputing. 2018;318:297–305.
[5] Ye J, Zhao J, Ye K, Xu C. How to build a graph-based deep learning architecture in traffic domain: A survey. IEEE Transactions on Intelligent Transportation Systems. 2020;23(5):3904–3924.
[6] Li L, Zhang J, Wang Y, Ran B. Missing value imputation for traffic-related time series data based on a multi-view learning method. IEEE Transactions on Intelligent Transportation Systems. 2018;20(8):2933–2943.
[7] Liang Y, Zhao Z, Sun L. Dynamic spatiotemporal graph convolutional neural networks for traffic data imputation with complex missing patterns. ArXiv Preprint ArXiv:210908357. 2021;.
[8] Phan H, McLoughlin IV, Pham L, Chén OY, Koch P, De Vos M, et al. Improving GANs for speech enhancement. IEEE Signal Processing Letters. 2020;27:1700–1704.
[9] Wu CH, Ho JM, Lee DT. Travel-time prediction with support vector regression. IEEE Transactions on Intelligent Transportation Systems. 2004;5(4):276–281.
[10] Cetin M, Comert G. Short-term traffic flow prediction with regime switching models. Transportation Research Record. 2006;1965(1):23–31.
[11] Che Z, Purushotham S, Cho K, Sontag D, Liu Y. Recurrent neural networks for multivariate time series with missing values. Scientific Reports. 2018;8(1):1–12.
[12] Cao W, Wang D, Li J, Zhou H, Li L, Li Y. Brits: Bidirectional recurrent imputation for time series. Advances in Neural Information Processing Systems. 2018;31.
[13] Yoon J, Jordon J, Schaar M. Gain: Missing data imputation using generative adversarial nets. In: International Conference on Machine Learning. PMLR; 2018. p. 5689–5698.
[14] Wang Y, Li D, Li X, Yang M. PC-GAIN: Pseudo-label conditional generative adversarial imputation networks for incomplete data. Neural Networks. 2021;141:395–403.
[15] Li D, Li L, Li X, Ke Z, Hu Q. Smoothed lstm-ae: A spatio-temporal deep model for multiple time-series missing imputation. Neurocomputing. 2020;411:351–363.
[16] Ye Y, Zhang S, Yu JJ. Spatial-temporal traffic data imputation via graph attention convolutional network. In: International Conference on Artificial Neural Networks. Springer; 2021. p. 241–252.
[17] Jiang W, Luo J. Graph neural network for traffic forecasting: A survey. Expert Systems with Applications. 2022;207:117921.