




Research Article

Metro Traffic Flow Prediction via Knowledge Graph and Spatiotemporal Graph Neural Network

Shun Wang ¹, Yimei Lv,² Yuan Peng,³ Xinglin Piao ¹ and Yong Zhang ¹

¹Beijing Artificial Intelligence Institute, Faculty of Information Technology, Beijing University of Technology, 100124 Beijing, China

²Qingdao Engineering Vocational College, 266011 Qingdao, China

³China Electronics Technology Group, Taiji Co. Ltd., 100083 Beijing, China

Correspondence should be addressed to Xinglin Piao; piaoxl@bjut.edu.cn

Received 5 July 2022; Revised 6 August 2022; Accepted 29 August 2022; Published 22 September 2022

Academic Editor: Wen Liu

Copyright © 2022 Shun Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Existing traffic flow prediction methods generally only consider the spatiotemporal characteristics of traffic flow. However, in addition to the spatiotemporal characteristics, the interference of various external factors needs to be considered in traffic flow prediction, including severe weather, major events, traffic control, and metro failures. The current research still cannot fully use the information contained in these external factors. To address this issue, we propose a novel metro traffic flow prediction method (KGR-STGNN) based on knowledge graph representation learning. We construct a knowledge graph that stores factors related to metro traffic networks. Through the knowledge graph representation learning technology, we can learn the influence representation of external factors from the traffic knowledge graph, which can better incorporate the influence of external factors into the prediction model based on the spatiotemporal graph neural network. Experimental results demonstrate the effectiveness of our proposed model.

1. Introduction

Accurate traffic prediction has become a core component of intelligent transportation system (ITS) construction, providing a decision-making basis for traffic management and helping improve residents' travel efficiency. Many studies have shown that there are many external factors (such as weather, events, accidents, etc.) that have an impact on traffic flow. Different types of POIs around traffic nodes can also lead to different traffic flow patterns. With the continuous development of intelligent transportation systems, more and more factors affecting traffic flow can be collected. However, existing traffic flow prediction methods generally consider learning spatiotemporal features from traffic data to predict future traffic flow changes. These methods ignore the exploitation of the influence of external factors on traffic flow. Therefore, it is a challenging research problem to con-

sider these factors to obtain accurate traffic prediction results fully.

The development of deep learning methods brings new solutions to traffic forecasting. Generally, deep learning-based methods build deep networks by stacking different basic modules. Many architectures have been developed to handle complex spatiotemporal data. Graph convolutional networks can be used to model data in nonEuclidean space, which is more suitable for road network structure or metro structure in traffic flow prediction. Existing GCN-based methods usually use a purely data-driven approach to learn the spatiotemporal dependencies of traffic flow. Several studies attempt to exploit information from external factors in traffic flow forecasting. However, these methods usually directly encode external factors into representation vectors. The representation of the traffic flow is directly concatenated with the obtained representation vector in the prediction

model. In large-scale traffic data, there are complex correlations between external factors and the traffic flow of nodes. Existing representation vectors of external factors are still deficient in expressing their influence on traffic flow.

In order to deal with the limitations of data-driven forecasting models on the utilization of external factors, one solution is to take a traffic knowledge graph composed of various traffic factors into account when constructing a predictive model. Knowledge graphs store real-world concepts and knowledge in the form of graph structures. The intelligent transportation system has a wide range of data sources and rich semantic information. The construction of a traffic knowledge graph can help solve traffic problems by using the traffic information contained in the data. The external factor representation obtained by the existing methods cannot fully express the impact on the traffic flow in the prediction model. Knowledge graph representation learning provides a new research direction for obtaining more useful semantic representations. Knowledge graph representation can obtain low-dimensional dense vectors, which provide better entity representation and prior knowledge for machine learning or deep learning models for downstream tasks.

The metro knowledge graph contains many influencing factors related to the operation of the metro network, such as weather, events, and POI (points of interest). We propose a metro traffic flow prediction method based on knowledge graph representation learning and spatiotemporal graph neural network (KGR-STGNN). The knowledge graph representation learning technique can be used to obtain the low-dimensional vector representation of the knowledge graph, and the representation of the entities and relationships of the entire metro network can be obtained at regular intervals. These semantic representations can fully represent the correlation characteristics between external factors and metro networks. Besides, this method utilizes Graph Neural Network (GNN) to learn the spatial dependencies between traffic flow sequences. Temporal Convolutional Network (TCN) is used to learn the temporal features of traffic flow. The dynamic semantic representation learned in each time period is used as a part of the input, which is fused with the traffic flow representation obtained through the spatiotemporal learning module. Finally, a traffic flow representation including the influence of complex external factors is obtained, to achieve a more accurate metro traffic flow prediction.

The main goal of this research is to combine a knowledge graph representation containing metro operation information with a spatiotemporal traffic flow prediction model. The main contributions of this study are summarized as follows:

- (i) A metro traffic knowledge graph is constructed, and we propose a new metro traffic flow prediction model based on this knowledge graph, which effectively improves the modeling ability of the impact of traffic events on traffic prediction
- (ii) The knowledge graph representation model is used to obtain the traffic knowledge representation at dif-

ferent times, which is significantly different from the external factor utilization method in the previous prediction model

- (iii) A new metro traffic flow prediction model KGR-STGNN based on knowledge graph representation and spatiotemporal graph network is proposed. Experimental results demonstrate the effectiveness of using traffic knowledge in traffic forecasting

2. Related Work

In this section, we review related research on knowledge graph and traffic flow prediction.

2.1. Traffic Knowledge Graph. In the field of knowledge graph, early research mainly involves semantic networks that contain a lot of information, such as WordNet [1], Google Knowledge Graph [2], etc. The general knowledge graph is built on the basis of a large amount of real-world data and mainly focuses on the modeling and storage of knowledge in a wide range and multiple domains. Domain knowledge graph [3] is oriented to a single industry, which can realize knowledge modeling and reasoning and provide help for the analysis of industry problems and the formulation of strategies, such as geographic information graph [4], medical knowledge graph [5], and so on. In recent years, domain knowledge graphs for a single industry have emerged continuously. Grainger et al. [6] automatically and dynamically build domain semantic networks based on corpus data. Pan et al. [7] use the linked data of large organizations to build and deploy industry-oriented data. Li et al. [8] constructed an e-commerce knowledge graph based on e-commerce data to help understand user needs and improve recommendation effects.

The traffic field has a wide range of data sources and rich semantic information. Building a traffic knowledge graph can help solve traffic problems by using the traffic information contained in the data. Muppalla et al. [9] used traffic surveillance video data to build a video information graph to help understand traffic patterns. Zhang et al. [10] proposed an urban knowledge graph neural network model (Urban Knowledge Graph Neural Network, UKG-NN) to combine urban knowledge graph and neural network to solve the problem of traffic accident reasoning and optimizing cargo storage. Keller [11] builds an air traffic knowledge graph by fusing various aviation data, which provides help for air traffic management. Zhou and Chen [12] constructed an urban traffic knowledge graph containing various types of urban knowledge, and combined the deep spatiotemporal convolutional neural network (STCNN) to extract features from the urban knowledge graph to achieve urban congestion.

2.2. Precise Prediction of the Area. Zhu et al. [13] proposed a traffic flow prediction method based on graph convolution on the basis of knowledge representation driven. By constructing a traffic knowledge graph and using an entity-attribute relation-based knowledge graph representation model (KR-EAR) combines knowledge graph and

information in the transportation network, the accuracy of traffic flow prediction is effectively improved. Sun et al. [14] constructed a traffic knowledge graph and an affair graph based on open-source data and further used the traffic knowledge graph to identify relevant traffic events in social media data.

In general, the metro knowledge graph is of great significance to the construction of intelligent transportation systems, but there are few studies on the metro knowledge graph, and the application form of the knowledge graph in the transportation field is simple.

2.3. Knowledge Graph Representation. The purpose of knowledge graph representation learning techniques is to express semantic entities and relationships in knowledge graphs as low-dimensional dense vectors [1]. The vector representations obtained through knowledge graph representation learning can be used to compute complex semantic associations between different entities and relationships in the knowledge graph. On the basis of knowledge graph representation learning technology, researches such as knowledge graph completion, knowledge reasoning, and link prediction are developing continuously [15]. Through the obtained distribution representation, the heterogeneous information in the knowledge graph can be fused to obtain a more complex representation, which provides help for downstream tasks.

Inspired by word embedding techniques in the field of natural language processing such as word2vec [16], knowledge graph representation learning methods have been continuously proposed and practically applied. The most widely used knowledge graph representation learning direction is the TransE model proposed by Bordes et al. [17]. The idea of this model comes from word vector representation, and words with high semantic similarity also have a relatively close distance in a low-dimensional vector space. Bordes et al. project entities and relations into a low-dimensional vector space and predict new knowledge entities by summing entity vectors and relation vectors to obtain new vectors. Although the TransE model has the advantages of simplicity, effectiveness, and low computational complexity, it still has shortcomings in the representation of complex relationships such as one-to-many, many-to-one, and many-to-many. In order to improve the representation of complex semantic relations, many researchers have proposed a series of improved methods based on the TransE model. Wang et al. [18] proposed the TransH model to learn semantic representations of complex relations between entities and by projecting knowledge entities onto different planes related to the relations. Lin et al. [19] proposed the TransR model based on the TransH model. The core idea of this model is to learn different types of entity representations and relational representations in different types of vector spaces, and then map the entity vectors to the corresponding relational vector spaces for computation.

In conclusion, knowledge graph representation learning provides convenience for the application of knowledge graphs in downstream tasks. The use of metro traffic knowl-

edge graph also requires the help of knowledge graph representation learning.

2.4. Traffic Flow Prediction. Traffic flow prediction is an important research problem in the intelligent transportation system. Traditional flow traffic prediction methods are mainly based on statistical models. HA [20], ARIMA [21], and VAR [22] are classic time series analysis methods based on statistical models, widely used in traffic flow prediction. However, these methods are generally designed for small datasets and are not suitable for processing time-series data with complex dynamic characteristics in large-scale datasets. Machine learning methods such as SVR model [23], Bayesian model [24], and k-nearest neighbor model [25] can deal with modeling on complex data and can better reflect the nonlinear and uncertain characteristics of traffic data. However, traditional machine learning methods usually only consider the temporal information but ignore the spatial dependence in the traffic data. Moreover, these methods rely on the ability to design feature models manually and still have limitations in modeling complex and dynamic traffic flows.

Recently, deep learning methods have shown superiority in traffic flow prediction problems. RNN [26] and its variants are widely adopted as a component of traffic flow prediction models to capture the temporal dependence of traffic data. LSTM [27] and GRU [28] further enhance the RNN's ability to model long-range time-series traffic data. In addition, CNN [29] has also been proven to be effective in time modeling. STGCN [30] uses convolutional networks to extract temporal features from graph-structured traffic data. GSTNet [31] and Graph Wavenet [32] use dilated causal convolution to capture the temporal dependence of transportation network nodes. Dilated causal convolution [33] is a special CNN that effectively captures long-term dependence by changing the size of the receptive field. In terms of spatial dependency modeling, main studies use CNN to model spatial correlation, but CNN is mainly designed for modeling data in Euclidean space. Many researchers have adopted GCN [34] to effectively handle traffic data in nonEuclidean space to address this limitation. For example, STGCN [30] and TGCN [35] use GCN to learn the topological structure of the traffic road network.

In summary, traditional statistical models and deep learning models for traffic flow prediction are constructed based on the analysis of spatiotemporal features of traffic data. However, the semantic information contained in the traffic data also reflects the specific laws and patterns of traffic flow, which can be utilized to enhance traffic prediction results.

3. Traffic Knowledge Graph

The knowledge graph is a semantic structure that can express entities, concepts, and relationships in the real world. The knowledge graph can conveniently store data in a form that is more recognizable to humans, providing a better format for data management. The knowledge graph is developed on the basis of semantic network. In essence, a

knowledge graph is a more structured and abstract semantic network. The knowledge graph is mainly composed of entities and relationships. Entities represent individuals or concepts in the objective world, and relationships represent various semantic associations between entities.

In order to make more effective use of the semantic information contained in metro data, we construct a metro knowledge graph on the basis of processing and analyzing multisource heterogeneous traffic data. The data sources of the knowledge graph constructed in this paper are mainly Beijing metro card data, Beijing metro road network data, map data, weather data, and social media data. The entities of the knowledge graph mainly include track sites, routes, regions, map points of interest (POI), weather, traffic events, etc. The relationship mainly includes the attribution relationship between site and line, the spatial positioning relationship between site and latitude and longitude value, the attribution relationship between site and area, the attribution relationship between line and region, the spatial positioning relationship between POI (points of interest) and latitude and longitude value, and the influence of site and weather. Relationships and affected relationships of site-traffic events, etc.: in addition, the main entity has various attribute information, such as the site's passenger flow value, the traffic event's time, etc.

3.1. Knowledge Graph Construction. The construction of the metro knowledge graph mainly includes the following six steps:

- (1) *Traffic data acquisition.* The sources of traffic data related to metro network management and operation are very complex, including metro network data, metro card swiping data, points of interest data, weather data, and social media data. Some of the above data can be obtained through the cooperation channels with the Beijing Municipal Transportation Authority or the Metro Operation Department, and some of the data needs to be obtained through a given application programming interface
- (2) *Data processing.* The collected data has different formats and organizational structures, so further processing the acquired data into structured data is necessary
- (3) *Extracting traffic entities and relationships.* The key part of the construction of a metro knowledge graph requires artificially using professional knowledge to determine the pattern layer composition of the knowledge graph, and then extracting metro entities and relationships from the processed data
- (4) *Obtaining the attribute value of the traffic entity.* Consistent with the operation in the previous step, the required attribute value key-value pairs are extracted from the processed data and filled in as the attribute value of the knowledge graph entity
- (5) *Storage in the knowledge graph storage tool.* Install the corresponding knowledge graph storage soft-

ware, store the constructed entities, relationships, and attributes, and use the graph database as a graph storage tool

- (6) *Query usage of traffic knowledge graph.* Based on the storage in the previous step, using the query algorithm, the entities, attributes, and relationships in the graph can be retrieved

The primary task of metro knowledge graph construction is to define the mode layer, which is the knowledge graph's core part and key content. The first step is clarifying the metro traffic network's core elements and knowledge. The crucial elements in metro traffic system are stations, routes, and environments. Stations and lines are the metro network's basic elements and structural elements. Generally speaking, stations and lines do not change over a period of time and belong to the fixed and static part of the knowledge graph. The environment contains two parts, one part is static elements related to traffic, such as POI (points of interest), and the other part is dynamic elements, such as weather and traffic events. Figure 1 shows the overall process of building a knowledge graph. We use the "site-belonging-line" triplet as an example, and show the process of extracting entities, relationships, and attributes from raw data. The knowledge graph is finally stored in the graph database.

3.2. Main Components of Traffic Knowledge Graph. For the static part of the metro knowledge graph, its main entities include stations, routes, regions, and POI (points of interest). Station entities have attributes such as name, number, latitude and longitude, and transfer information. Line entities have attributes such as name and number. POI entities have attributes such as id, name, type, location, latitude and longitude, and the area to which they belong. The relationship between entities mainly includes: the adjacent relationship between stations in space, the attribution relationship between stations and lines, the adjacent relationship between station and POI. The upper part of Figure 1 shows the static entities and relationships of the traffic knowledge graph.

Unlike the static part, the dynamic part of the metro knowledge graph is mainly composed of elements that change over time, such as weather, traffic events, and metro station passenger flow. The entities in the dynamic part mainly include weather, events, and passenger flow status. The weather entity has attributes such as recording time, temperature, humidity, pollutant indicators, weather conditions, and wind conditions; the event entity has attributes such as event type, occurrence time, and event location; and the passenger flow status has attributes such as time and passenger flow. After realizing the design of the dynamic part entity and the static part entity, it is necessary to define the relationship between them further. As shown in the schematic diagram of the dynamic part of the metro knowledge graph in Figure 2, there is an "affected" relationship between stations and weather or events.

The traffic knowledge triples in the graph are defined as (h, r, t) , where h and t represent the head entity and tail

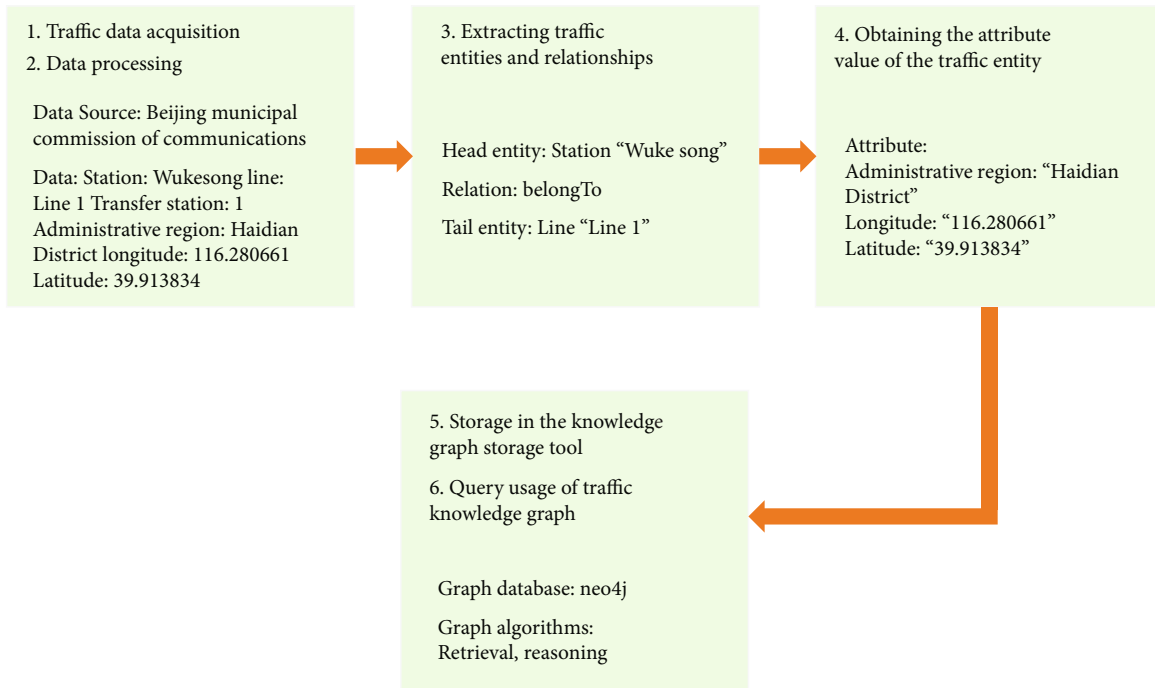


FIGURE 1: Knowledge graph construction process (take “station-belongTo-line” as an example).

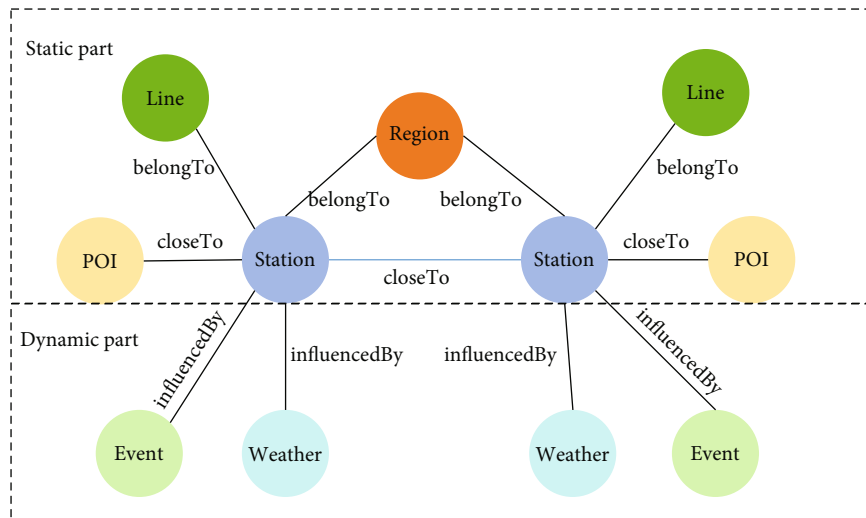


FIGURE 2: The main components of the metro traffic knowledge graph we constructed.

TABLE 1: Entities and attributes of metro knowledge graph.

Entity	Attribute	Number
Station	Station name, station id, latitude and longitude, transfer information	325
Line	Line name, line id	13
Region	Region name, region id	12
POI	Poi id, poi name, poi type, latitude and longitude, poi region	422,000
Event	Event type, event location, event time	6,199
Weather	Time, temperature, humidity, pollutants, weather conditions, wind conditions	52,992

TABLE 2: Edges of metro knowledge graph.

Head entity	Relation	Tail entity	Number
Station	closeTo	Station	528
Station	belongsTo	Line	381
Station	belongsTo	Area	12
POI	closeTo	Station	1.258 million
Station	influencedBy	Weather	1.435 million
Station	influencedBy	Event	14,000

entity, respectively, and r represents the relationship. The properties of the entity are defined as $(h, \text{hasproperty}, \text{property}_{value})$. The entire metro knowledge graph can be expressed as $KG = \{(h, r, t) | h, t \in EUP, r \in R\}$, E represents the set of knowledge entities, R represents the set of association relationships, and P represents the attribute collection. Table 1 shows the entities and attributes of the metro traffic knowledge graph, and Table 2 shows the edges in the metro traffic knowledge graph. Figure 3 shows the constructed traffic knowledge graph stored in the graph database.

4. Methodology

4.1. Knowledge Graph Representation. TransE is the most representative translation-based knowledge graph representation learning model. Inspired by the translation-invariant feature of word vectors in the word2vec model [16], the TransE model was proposed and used in the representation learning of knowledge graph entities and relations [17]. Specifically, for the triple (h, r, t) representing the knowledge ontology, h and t represent the head entity and the tail entity, respectively, and r represents the association between entities. The TransE model has the following settings: the embedding vector of the tail entity t should be equal to the sum of the embedding vector of the head entity h and the embedding vector of the relationship r , that is, $h + r \approx t$. When designing the knowledge graph representation learning model, the scoring function is used to measure the reliability of the triples, and it is an important tool to measure the rationality of the representation results. The scoring function of the TransE model is:

$$f(h, r, t) = \|h + r - t\|_2 \quad (1)$$

$\|\cdot\|_2$ indicates that the calculation method of vector distance is L_2 norm. TransE regards entities and relationships in the knowledge graph as two matrices. The structure of the entity matrix is $n \times d_e$, where n represents the number of entities, d_e represents the dimension of each entity vector, and each row in the matrix represents the word vector of an entity; and the relationship matrix structure is $m \times d_r$, where m represents the number of relationships and d_r represents the dimension of each relation vector. For the knowledge graph trained by TransE, if two vectors are extracted from the entity matrix and the relationship matrix, respectively, and L_2 operation is performed on these two vectors, the obtained result will be similar to the vector of another entity in the entity matrix. Finally, the existing triples (h, r, t) in the

knowledge graph are represented in the form of word vectors. Figure 4 shows the basic principle of the TransE model.

4.2. Traffic Knowledge Graph Pretraining. The traffic knowledge graph contains a large amount of traffic semantic information related to the operation, scheduling, and management of the metro network, such as traffic events that affect subway traffic flow, bad weather, and information about the metro system itself (metro network structure and POI). This information can be collectively referred to as external factors affecting subway traffic flow. In order to utilize these external factors in the traffic flow prediction model, the knowledge graph representation learning method is adopted to pretrain the traffic knowledge graph, so as to obtain the traffic semantic representation expressing the influence of external factors. In the real world, metro traffic flow will be disturbed by different external factors at different times, and the dynamic characteristics of external factors need to be fully considered in pretraining.

As shown in Figure 5, the TransE model is used to pretrain the subgraphs at different times. According to half an hour as a time period, the subknowledge graph under each time period is extracted from the metro knowledge graph. Table 3 shows the main entities and relationships in the pretrained knowledge graph, which represent the external factors that we incorporate in implementing traffic predictions. Each subknowledge graph consists of two parts; one part is the static road network information and POI. The other part is dynamic traffic events and weather information. Through the representation learning of each subknowledge graph, the semantic representations (KG_t) of 325 metro station entities in Beijing under different time periods are obtained, and the dimension of each semantic representation is 30 dimensions. When using knowledge representation technology to obtain the representation of station entities, the external factors affecting metro operation will also be mapped to the low-dimensional vector space at the same time, and the relationship between station entities and external factor entities also follows the semantic rules of the TransE model. Therefore, the semantic information of external factors is also included in the representation results. The formula for knowledge graph pretraining is shown below:

$$KG_t = \text{TransE}(h_t, r_t, t_t), \quad (2)$$

where, KG_t represents the semantic representation of 325 metro stations at time t , h_t are the station entities, r_t and t_t represent relationships and tail entities related to station entities at time t .

4.3. Spatial-Temporal Modeling of Traffic Flow. Modeling the spatiotemporal characteristics of metro traffic flow is a key task to achieving subway traffic flow prediction. In order to effectively learn the spatiotemporal characteristics of metro traffic flow, this paper uses graph convolutional neural networks and temporal convolutional neural networks to model spatial and temporal features, respectively. The specific

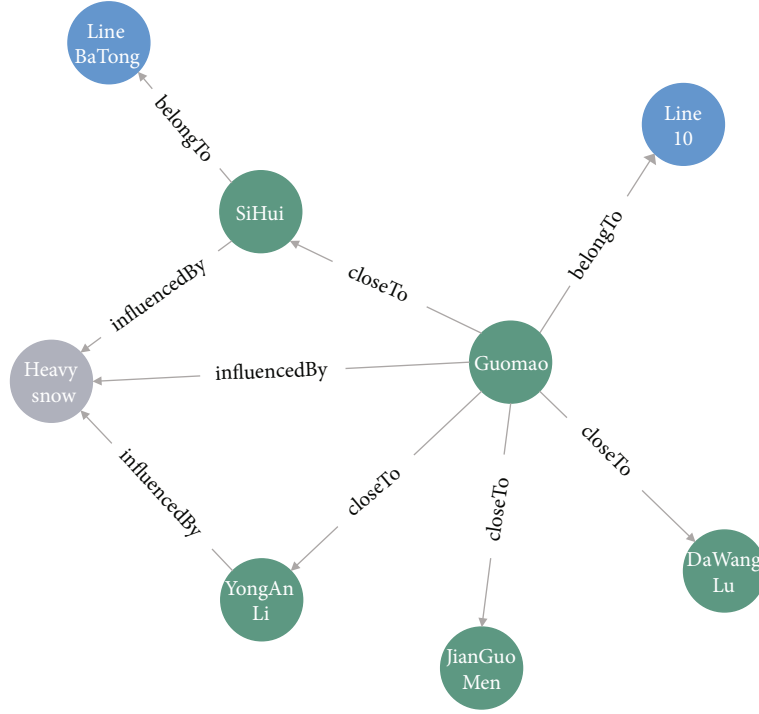


FIGURE 3: Metro traffic knowledge graph stored in graph database.

usage of graph convolutional neural network is described in detail below.

Metro networks generally have a graph structure. Convolutional neural networks used in the image field are generally used to extract local spatial features of data in Euclidean space. Graph Neural Network (GCN) is more suitable for data in nonEuclidean space, and GCN can capture more realistic spatial features of metro traffic flow. The frequency domain-based graph convolution operation requires the use of the Laplacian matrix of the graph:

$$L = I_N - D^{-1/2} A D^{-1/2}, \quad (3)$$

where, I_N represents the identity matrix of order N , A is the adjacency matrix of the orbital network graph, and D is the degree matrix of the orbital graph calculated from the adjacency matrix. Therefore, the definition of the graph convolution operation on the graph signal of the orbital network using the convolution kernel is:

$$\Theta *_g X = \Theta(L)X = \Theta(U\Lambda U^T)X = U\Theta(\Lambda)U^T X, \quad (4)$$

where, $*_g$ represents the operator of the graph convolution operation, L is the Laplacian matrix of the orbital graph, U means the matrix composed of the eigenvectors of the matrix L , and Λ represents the diagonal matrix composed of the eigenvalues of L . Due to a large number of metro stations, the decomposition of the matrix L in the above formula requires a large amount of calculation. Therefore, this paper uses the Chebyshev polynomial approximation graph

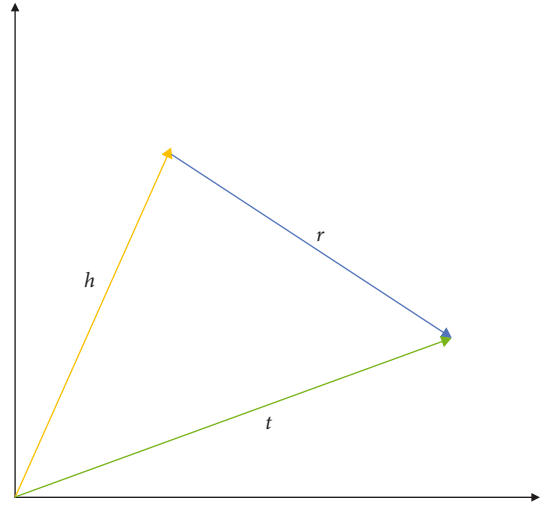


FIGURE 4: Knowledge graph representation learning based on TransE model.

convolution kernel to solve the graph convolution operation:

$$\Theta *_g X = \Theta(L)X \approx \sum_{k=0}^{K-1} \theta_k T_k(\tilde{L})X. \quad (5)$$

θ_k represents the k -dimensional coefficient vector of the Chebyshev polynomial. The recursive formula of this polynomial is: $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$ and $T_0(x) = 1$, $T_1(x) = x$. In the above formula, $\tilde{L} = (2/\lambda_{\max})L - I_N$, and λ_{\max} is the maximum value of the eigenvalues of the matrix L . The Chebyshev polynomial is used to complete the

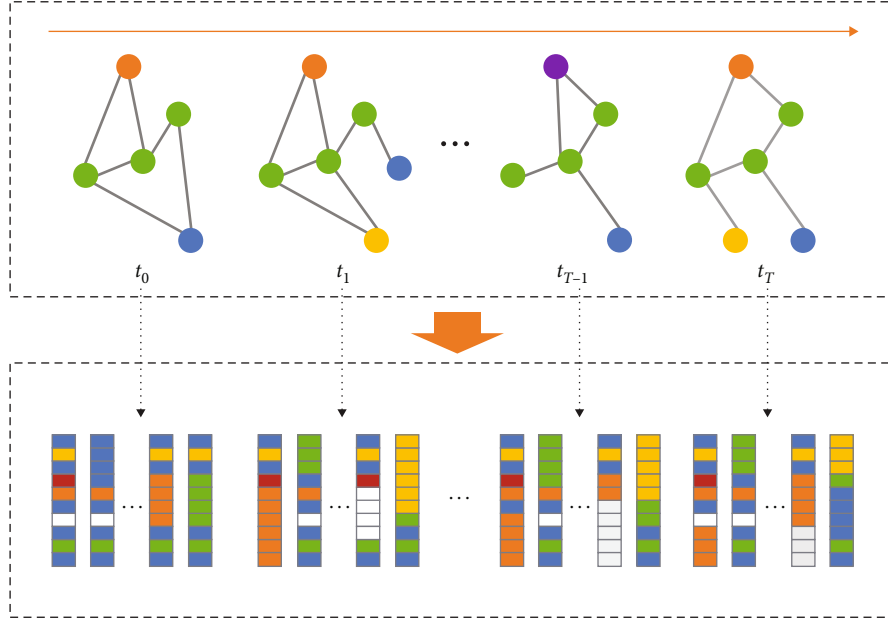


FIGURE 5: Traffic knowledge graph pretraining.

TABLE 3: Main entities and relation in the pretrained traffic knowledge graph.

Core entity	Relation	Related entity	Type
Station	closeTo	Station	Static entity
Station	belongTo	Line	Static entity
Station	belongTo	Area	Static entity
Station	closeTo	POI	Static entity
Station	influencedBy	Weather	Dynamic entity
Station	influencedBy	Event	Dynamic entity

approximation of the convolution operation on the graph, which reduces the computational complexity of the graph convolution operation. The spatial modeling model used in this chapter is based on ChebNet [36] simplification to obtain first-order Chebyshev graph convolution, and its formula is:

$$\dot{Z}_t = \delta \left(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} Z_t W \right) \mathbf{0}. \quad (6)$$

In the formula, Z_t is the input feature of graph convolution, the adjacency matrix $\tilde{A} = A + I_N$ and the degree matrix $\tilde{D} = \sum_j \tilde{A}_{ij}$ are added to the node's own feature.

The recurrent neural network is generally used to model the traffic flow time series in the traffic flow prediction problem. However, the prediction of each state in the recurrent neural network is closely dependent on the previous state, and the computational complexity is high. In many studies, time-series data is regarded as one-dimensional structured data, and it is more computationally efficient to use convolutional neural networks to extract time-series features on time-series data. This section uses Gated Temporal Convolutional Network (Gated TCN) to extract and model tempo-

ral features in metro traffic flow. A Gated TCN consists of two layers: a 1D convolutional, gated linear unit GLU that learns temporal properties. Its calculation formula is as follows:

$$\Gamma_{*t} X = P \odot \sigma(Q) \in R^{(M-K_t+1) \times C_o}. \quad (7)$$

The width of the convolution kernel of the one-dimensional convolution is K_t , and the temporal convolution network performs convolution operations on the K_t neighbors of the input traffic flow, and each operation shortens the traffic flow sequence by $K_t - 1$. The input of each track site is represented as a sequence of length M , the number of channels is C_i , and the input is denoted as $X \in R^{M \times C_i}$. The convolution kernel is $\Gamma \in R^{K_t \times C_i \times 2C_o}$, the input is convolved to obtain the output $[PQ] \in R^{(M-K_t+1) \times (2C_o)}$, P and Q are the input of the gated linear unit (GLU), \odot represents the Hadamard product operation, and $\sigma(Q)$ is the activation function.

4.4. Traffic Flow Prediction with Knowledge Graph. After pretraining the traffic knowledge graph to obtain the semantic representation of the metro station, the semantic vector reflecting the influence of external factors needs to be utilized in the prediction model. As shown in Figure 6, the traffic flow sequence data input passes through the spatiotemporal feature modeling module in the upper part, which captures the temporal dynamics and spatial correlation of metro traffic flow. After the knowledge graph representation learning pretraining, the dynamic semantic representation reflecting the influence of external factors is obtained within a time period of half an hour, and the gated fusion mechanism is used to fuse the output of the spatiotemporal feature modeling module with the dynamic semantic representation. Finally, the prediction results of metro

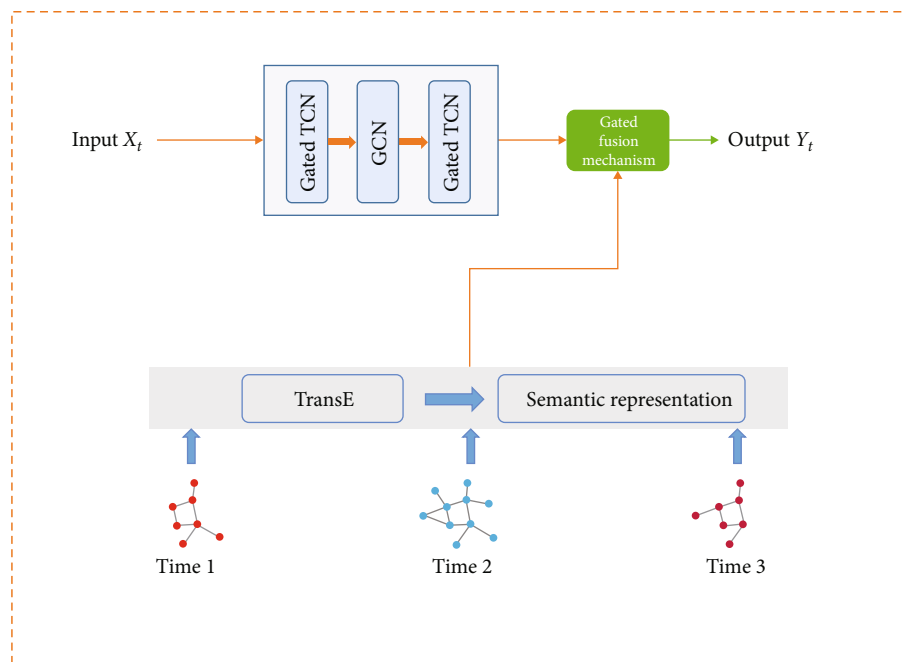


FIGURE 6: Metro traffic flow prediction model based on knowledge graph representation learning.

TABLE 4: Traffic flow prediction results on the BJMF15 dataset.

Baselines		SVR	LSTM	GCN	TGCN	STGCN	KGR-STGNN
Prediction steps	Evaluation indicators						
12	MAE	22.280	21.883	21.747	11.827	8.365	8.211
	RMSE	48.789	42.368	45.178	20.928	17.326	16.296
15	MAE	25.306	25.753	23.211	12.023	9.632	8.996
	RMSE	55.425	47.978	47.284	21.171	20.318	17.568
18	MAE	27.677	27.949	25.110	13.168	11.135	9.624
	RMSE	60.380	52.644	51.440	23.107	22.413	18.998
21	MAE	29.832	29.381	26.309	13.113	13.569	11.403
	RMSE	64.463	55.423	53.655	23.427	26.988	22.989

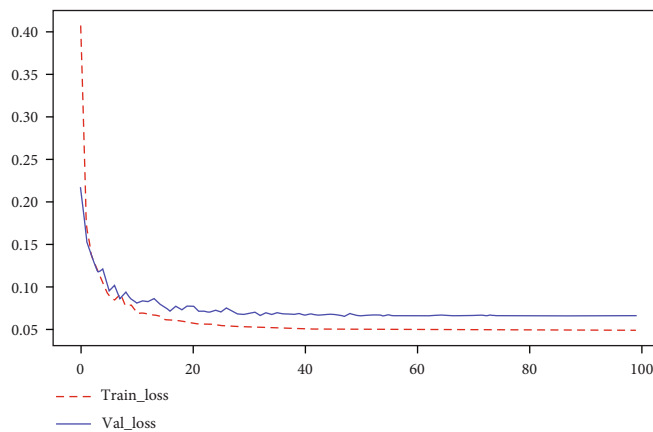


FIGURE 7: The change of loss function of metro traffic flow prediction model based on knowledge graph representation learning.

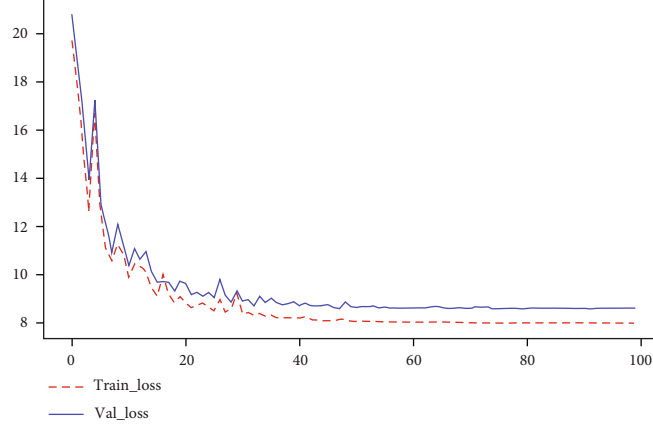


FIGURE 8: The change of MAE of metro traffic flow prediction model based on knowledge graph representation learning.

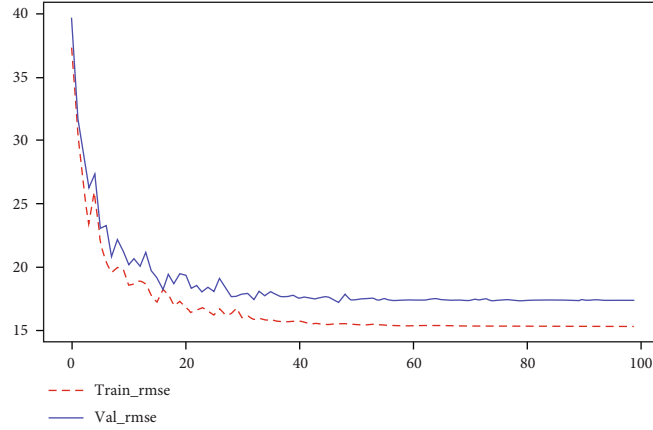


FIGURE 9: The change of RMSE of metro traffic flow prediction model based on knowledge graph representation learning.

traffic flow under the influence of external factors are obtained.

The spatiotemporal feature modeling part consists of two Gated TCN layers and one spatial GCN layer, and the spatial GCN layer in the middle undertakes two Gated TCN layers. The calculation formula of this part is as follows:

$$X'_t = \Gamma_{1*_{\tau}} \text{ReLU} \left(\Theta_{*g} \left(\Gamma_{0*_{\tau}} X_t \right) \right). \quad (8)$$

$\Gamma_{0*_{\tau}}$ and $\Gamma_{1*_{\tau}}$ are the temporal convolution kernels of the two Gated TCN layers, Θ_{*g} represents the spatial graph convolution at the middle layer Convolution kernel, and $\text{ReLU}(\cdot)$ is the activation function.

The semantic representation vector obtained every half an hour is defined as KG_t , The gated fusion mechanism is used in the prediction model to fuse spatiotemporal features and semantic features. During fusion, the time of the semantic representation vector is guaranteed to be consistent with the time of the spatiotemporal feature modeling output, and

its calculation formula is as follows:

$$Y_t = X'_t \otimes \tanh(KG_t). \quad (9)$$

5. Experiment

5.1. Experiment Settings. The experiments in this chapter need to use the constructed metro knowledge map, so this chapter conducts experiments on the Beijing metro passenger flow dataset (BJMF15), and the Beijing metro knowledge graph is also used in the experiment. The experiments evaluate the prediction performance of the knowledge graph representation learning-based traffic flow prediction model (KGR-STGNN) on the dataset BJMF15.

BJMF15: Passenger flow data collected by Beijing Metro System in 2015. The unprocessed raw data includes daily passenger travel record data at 325 stations on 13 lines of the Beijing Metro in August. The data consists of passenger travel records, including entry routes, entry stations, entry times, exit routes, exit stations, and exit times. Taking every five minutes as a time step, the metro passenger flow data is obtained by processing the original passenger swiping card data.

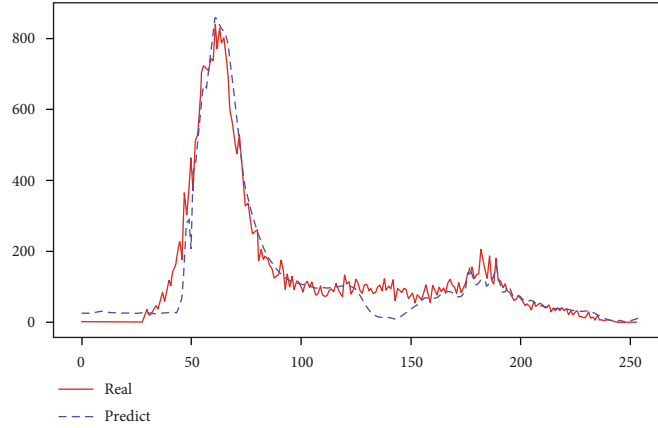


FIGURE 10: Prediction results of Pingguoyuan Station of Metro Line 1 on August 29, 2015.

In practical experiments, the complete BJMF15 dataset is divided into training, validation, and test sets. The split ratio of the dataset is 80%, 10%, and 10%, respectively. Different from the prediction of short-term traffic flow in Chapter 2, the impact of external factors on traffic flow needs to be better reflected in a longer time. The experiments in this chapter use historical traffic data of 12 consecutive time steps (1 hour) to predict 12-time steps (60 minutes), 15-time steps (75 minutes), 18-time steps (90 minutes), and 21-time steps (105 minutes) data.

In the KGR-STGNN model, the learning rate is set to 0.001, the size of each batch of training data, that is, the batch size is set to 32, and the number of training epochs is 100. Experiments on graph convolutional layers set a dropout of $p=0.3$. All experiments are performed on an Nvidia RTX 2080ti server with 11GB of video memory, and the server operating system is Linux.

5.2. Evaluating Indicator. In order to compare the prediction accuracy of the proposed model and the basic model, accurate numerical indicators are needed as the evaluation basis. This chapter evaluates the forecasting accuracy of all methods using two commonly used traffic forecasting metrics, including:

(1) Mean Absolute Error (MAE), which is a basic measure that reflects the actual situation of prediction accuracy

$$MAE(x, \hat{x}) = \frac{1}{N} \sum_{i \in N} |x_i - \hat{x}_i|. \quad (10)$$

(2) Root Mean Square Error (RMSE), which is used to measure the deviation of the predicted value from the true value

$$RMSE(x, \hat{x}) = \sqrt{\frac{1}{N} \sum_{i \in N} (x_i - \hat{x}_i)^2}, \quad (11)$$

where, N represents the number of stations in rail transit, x represents the actual traffic flow, and \hat{x} represents the predicted traffic flow.

5.3. Baselines. In order to verify the effectiveness of the Metro Traffic Flow Prediction Model (KGR-STGNN) based on knowledge graph representation learning, this paper uses the following models as benchmarks:

- (1) *SVR (Support Vector Regression Machine)*. A method for predicting future traffic flow using a support vector machine
- (2) *LSTM (Long Short-Term Memory Network)*. A time-series prediction model that models the long-term temporal characteristics of traffic-flow sequences
- (3) *GCN (Graph Neural Network)*. A prediction method to mine the spatial correlation of traffic flow
- (4) *GRU (Gated Recurrent Unit)*. Similar to the LSTM structure, part of the structure is optimized, and the amount of model parameters is reduced
- (5) *TGCN*. The graph convolutional neural network is used to learn the spatial characteristics of the node traffic flow. The gated recurrent unit is used to model the temporal characteristics of the traffic flow sequence, and then the two are combined to obtain a spatiotemporal traffic flow prediction model
- (6) *STGCN*. A traffic flow prediction method that uses convolutional neural networks to learn spatiotemporal characteristics. In terms of spatial characteristics, a graph convolutional neural network is used to obtain spatial dependencies between nodes, and one-dimensional convolution is used to model traffic flow sequences in terms of temporal characteristics

5.4. Prediction Performance. The traffic flow prediction model based on knowledge graph representation learning proposed in this paper is tested on the BJMF15 dataset and compared with five benchmark methods. In the experiments where the prediction step size is set to 12 steps, 15 steps, 18 steps, and 21 steps, the experimental results of each prediction method are shown in Table 4. It can be seen from the results that the traffic flow prediction model (KGR-STGNN)

based on knowledge graph representation learning proposed in this paper has better prediction performance. Among them, the prediction results of TGCN, STGCN, and KGR-STGNN are quite different from those of SVR, LSTM, and GCN. The first three models all model the temporal and spatial characteristics of metro traffic flow, proving the spatial and temporal modeling of traffic flow data effectiveness. Compared with STGCN, the traffic flow prediction model (KGR-STGNN) based on knowledge graph representation learning proposed in this paper has similar spatiotemporal characteristic learning methods. Compared with STGCN, KGR-STGNN uses the representation of the influence of external factors obtained through knowledge graph representation learning. With the increase in the prediction step size, the MAE evaluation indicators of the KGR-STGNN model are 8.211, 8.996, 9.624, and 11.403, respectively, and the prediction performance is improved by 1.84%, 6.60%, 13.57%, and 15.96%, respectively, compared with the results of the STGCN model. The experimental results show that adding external factor representation can improve the prediction model more and more with the increase in prediction time. As the prediction time becomes longer, the influence of external factors and traffic knowledge on the traffic flow becomes more obvious, proving the KGR-STGNN model's effectiveness in predicting medium and long-term metro traffic flow.

Figures 7–9 are the changes of the loss function, MAE, and RMSE of the metro traffic flow prediction model based on knowledge graph representation learning during the training and validation of the model. The red curve in the figure represents the change curve in the training case, and the blue curve represents the change curve in the validation case. It can be seen from the above three figures that the loss function and prediction result of the subway traffic flow prediction model based on knowledge graph representation learning can decrease rapidly and stably and finally tend to be stable and reach the model convergence state. This shows that the design of the prediction model in this paper is reasonable, which can effectively reduce the prediction error and achieve good prediction results.

In order to visually show the prediction effect of the KGR-STGNN model, this paper selects the traffic flow prediction results of Pingguoyuan Station of Metro Line 1 on August 29, 2015, for visualization to compare the gap between the actual value and the predicted value. As shown in Figure 10, the predicted results are very close to the real values, especially the curves in the peak passenger flow period have basically overlapped. The prediction results also prove that the model proposed in this paper can predict the traffic flow.

6. Conclusions

There are a large number of external factors in traffic data that affect traffic flow. Using the information contained in external factors in traffic flow forecasting can effectively improve the forecasting effect. To this end, this paper proposes a metro traffic flow prediction model based on knowledge graph representation learning to build the model. In

order to model the influence of these external factors, the model learns the representation of the metro knowledge graph and obtains a semantic representation that contains information about the external factors that affect the metro operation. In addition, the model adopts time-gated convolutional network and graph convolutional network to model traffic flow's spatiotemporal features. A gated fusion mechanism is used to fuse the semantic representation of the subway network with the output of the spatiotemporal feature learning module. Finally, the spatiotemporal feature representation of subway traffic flow data considering the influence of external factors is obtained. Experiments on Beijing metro traffic flow data show that the prediction model proposed in this chapter outperforms other benchmark models in metro traffic flow prediction, especially long-term traffic flow prediction. The experimental results verify that the knowledge representation obtained through representation learning can effectively improve the performance of metro traffic flow prediction.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 62072015, U19B2039, U1811463, 61632006, 61876012, and 61902053; the National Science Foundation of Beijing under Grant 4172003; and the China Scholarship Council under Grant 201806540008. The authors recognise those individuals who provided help during the research and preparation of the manuscript.

References

- [1] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to WordNet: an on-line lexical database," *International Journal of Lexicography*, vol. 3, no. 4, pp. 235–244, 1990.
- [2] S. Amit, *Introducing the knowledge graph*, America: Official Blog of Google, 2012.
- [3] B. Abu-Salih, "Domain-specific knowledge graphs: a survey," *Journal of Network and Computer Applications*, vol. 185, no. 5, article 103076, 2021.
- [4] F. Bond and A. Bond, "GeoNames Wordnet (GeoWN): extracting wordnets from GeoNames," in *Proceedings of the 10th Global Wordnet Conference*, pp. 387–393, Poland, 2019.
- [5] M. Rotmensch, Y. Halpern, A. Tlimat, S. Horng, and D. Sontag, "Learning a health knowledge graph from electronic medical records," *Scientific Reports*, vol. 7, no. 1, pp. 1–11, 2017.
- [6] T. Grainger, K. AlJadda, and M. Korayem, "The semantic knowledge graph: a compact, auto-generated model for real-

- time traversal and ranking of any relationship within a domain,” in *Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics (dsaa)*, pp. 420–429, Montreal, QC, Canada, 2016.
- [7] J. Z. Pan, G. Vetere, and J. M. Gomez-Perez, *Exploiting Linked Data and Knowledge Graphs in Large Organisations*, Springer, Heidelberg, 2017.
- [8] F. L. Li, H. Chen, and G. Xu, “AliMeKG: domain knowledge graph construction and application in e-commerce,” in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pp. 2581–2588, 2020.
- [9] R. T. Muppalla, S. Lalithsena, and T. Banerjee, “A knowledge graph framework for detecting traffic events using stationary cameras,” in *Proceedings of the 2017 ACM on Web Science Conference*, pp. 431–436, New York, USA, 2017.
- [10] N. Zhang, S. Deng, H. Chen et al., “Structured knowledge base as prior knowledge to improve urban data analysis,” *ISPRS International Journal of Geo-Information*, vol. 7, no. 7, p. 264, 2018.
- [11] R. Keller, “Building a knowledge graph for the air traffic management community,” in *Proceedings of the 2019 World Wide Web Conference*, pp. 700–704, Raleigh, North Carolina, USA, 2019.
- [12] G. Zhou and F. Chen, “Urban congestion areas prediction by combining knowledge graph and deep spatio-temporal convolutional neural network,” in *Proceedings of the 4th International Conference on Electromechanical Control Technology and Transportation*, pp. 105–108, Guilin, China, 2019.
- [13] J. Zhu, X. Han, H. Deng et al., “KST-GCN: a knowledge-driven spatial-temporal graph convolutional network for traffic forecasting,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 1–11, 2022.
- [14] X. R. Sun and Y. Meng, “Identifying traffic events from Weibo with knowledge graph and target detection,” *Data Analysis and Knowledge Discovery*, vol. 4, no. 12, pp. 136–147, 2020.
- [15] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, “A survey on knowledge graphs: representation, acquisition, and applications,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 2, pp. 494–514, 2022.
- [16] T. Mikolov, I. Sutskever, and K. Chen, “Distributed representations of words and phrases and their compositionality,” in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2, pp. 3111–3119, Montreal, Canada, 2013.
- [17] A. Bordes, N. Usunier, and A. Garcia-Duran, “Translating embeddings for modeling multi-relational data,” in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2, pp. 2787–2795, Montreal, Canada, 2013.
- [18] Z. Wang, J. Zhang, J. Feng, and Z. Chen, “Knowledge graph embedding by translating on hyperplanes,” *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, vol. 28, no. 1, pp. 1–8, 2014.
- [19] Y. Lin, Z. Liu, and M. Sun, “Learning entity and relation embeddings for knowledge graph completion,” *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, vol. 29, no. 1, pp. 2181–2187, 2015.
- [20] J. Y. Campbell and S. B. Thompson, “Predicting excess stock returns out of sample: can anything beat the historical average?,” *The Review of Financial Studies*, vol. 21, no. 4, pp. 1509–1531, 2008.
- [21] M. Mohammad, R. Hashem, and M. Zahi, “Short-term prediction of traffic volume in urban arterials,” *Journal of Transportation Engineering*, vol. 121, no. 3, pp. 249–254, 1995.
- [22] E. Zivot and J. Wang, *Vector autoregressive models for multivariate time series*, Modeling Financial Time Series with S-Plus®, 2006.
- [23] C. H. Wu, C. C. Wei, and D. C. Su, “Travel-time prediction with support vector regression,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 276–281, 2004.
- [24] S. Sun, C. Zhang, and G. Yu, “A bayesian network approach to traffic flow forecasting,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 124–132, 2006.
- [25] X. L. Zhang, H. E. Guo-Guang, and L. U. Hua-Pu, “Short-term traffic flow forecasting based on k-nearest neighbors non-parametric regression,” *Journal of Systems Engineering*, vol. 24, no. 2, 2009.
- [26] Z. Zhene, P. Hao, and L. Lin, “Deep convolutional mesh rnn for urban traffic passenger flows prediction,” in *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/ SCALCOM/UIC/ATC/CBDCOM/ IOP/SCI)*, pp. 1305–1310, Guangzhou, China, 2018.
- [27] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [28] F. Rui, Z. Zuo, and L. Li, “Using lstm and gru neural network methods for traffic flow prediction,” in *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, Wuhan, Hubei, China, 2016.
- [29] J. Gehring, M. Auli, and D. Grangier, “Convolutional sequence to sequence learning,” in *International Conference on Machine Learning*, pp. 1243–1252, Sydney, Australia, 2017.
- [30] B. Yu, H. Yin, and Z. Zhu, “Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting,” in *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, Stockholm, Sweden, 2018.
- [31] S. Fang, Q. Zhang, and G. Meng, “Gstnet: global spatial-temporal network for traffic flow prediction,” in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, pp. 2286–2293, Macao, China, 2019.
- [32] Z. Wu, S. Pan, and G. Long, “Graph wavenet for deep spatial-temporal graph modeling,” in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, Macao, China, 2019.
- [33] F. Yu and V. Koltun, *Multi-scale context aggregation by dilated convolutions*, in ICLR, 2016.
- [34] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *5th International Conference on Learning Representations*, Singapore, 2017.
- [35] L. Zhao, Y. Song, and C. Zhang, “T-gcn: a temporal graph convolutional network for traffic prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2019.
- [36] M. Defferrard, X. Bresson, and P. Vandergheynst, “Convolutional neural networks on graphs with fast localized spectral filtering,” in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, pp. 3844–3852, Barcelona, Spain, 2016.