北京航空航天大學
BEIHANG UNIVERSITY

BIGSCITY
Beihang Interest Group on SmartCity

Deep Transfer Learning for City-scale Cellular Traffic Generation through Urban Knowledge Graph

Shiyuan Zhang, Tong Li, Shuodi Hui, Guangyu Li, Yanping Liang, Li Yu, Depeng Jin, Yong Li

Background

Applications of 5G

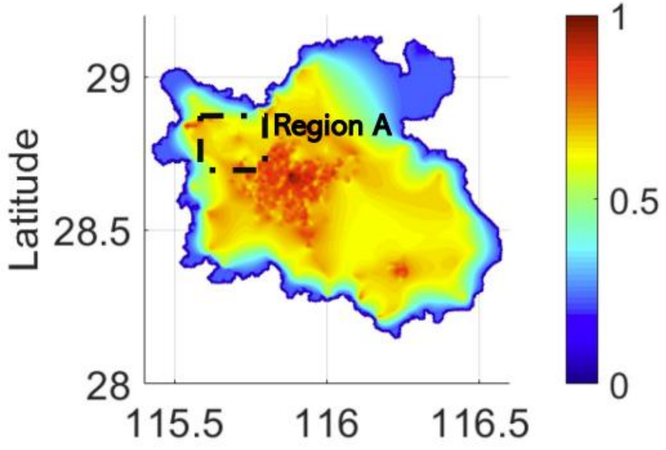
5G networks are recognized as providing solutions for all applications, including networked vehicles, the internet of things, augmented reality, virtual reality, super-high quality online videos, and many more customized services for subscribers.

The main problem is How to deploy 5G base stations

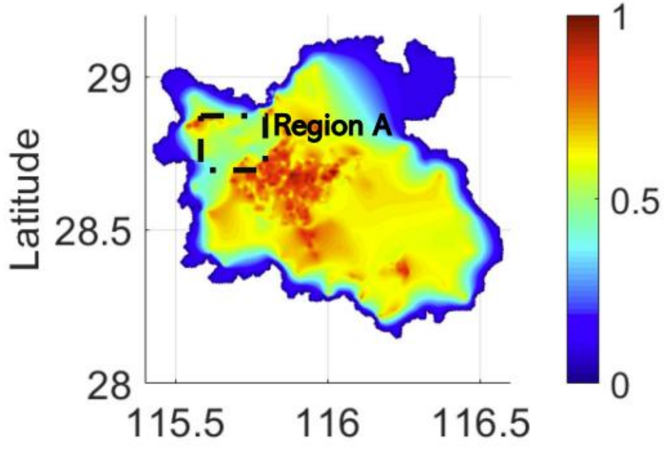
Background

The deployment of base stations traditionally relies on the experiences of experts. Communication engineers manually plan out the sites of base stations.

- Such a manual approach is limited by expensive labor costs and cannot find optimal solutions for large-scale areas, such as city-scale.
- Merely relying on human experiences can easily result in a high mismatch between human traffic demand and deployed base stations.



base station density



traffic volume

Missmatch!

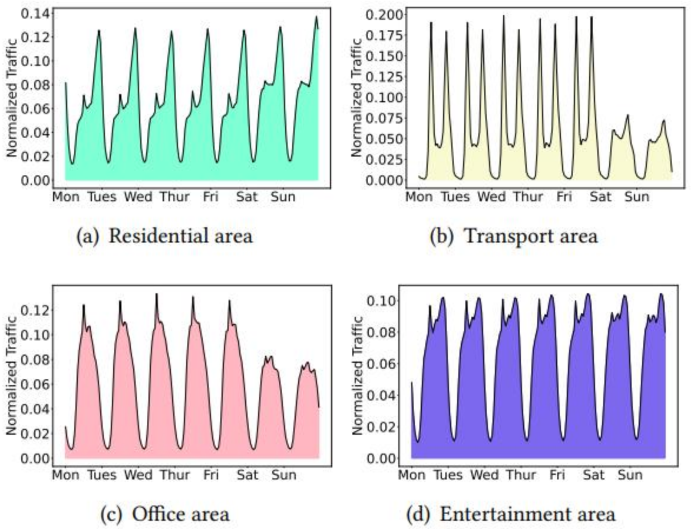
Background

Motivation

- Deploying base stations according to the estimated traffic load is a more practical approach.
- Generating or estimating cellular traffic load for newly deployed 5G base stations is challenging due to the lack of historical data.

Challenges

- How to build a bridge between the source city and the target city for cellular traffic generation.
- How to extract and represent relations between base stations.
- How to transfer temporal patterns of base stations' traffic.



Preliminaries

Urban Knowledge Graph

The urban knowledge graph takes urban content, such as base stations, point of interests (POIs), and regions, as entities where spatial and semantic dependencies are modeled as relations.

Probleam Defination

Given the historical traffic dataset of the source city and the urban knowledge graphs of the source city and the target city, our goal is to generate city scale traffic for the base stations located in the target city.

Input

$$G^{SRC} = (\varepsilon^{SRC}, R^{SRC}, F^{src}),$$

$$G^{TRG} = (\varepsilon^{TRG}, R^{TRG}, F^{TRG}),$$

$$X^{SRC} = \{X_i^{SRC}\}, i = 1, \ldots, N^{SRC}$$

ADAPTIVE

Output

$$X^{TRG} = \{X_i^{TRG}\}, i = 1, \ldots, N^{TRG}$$

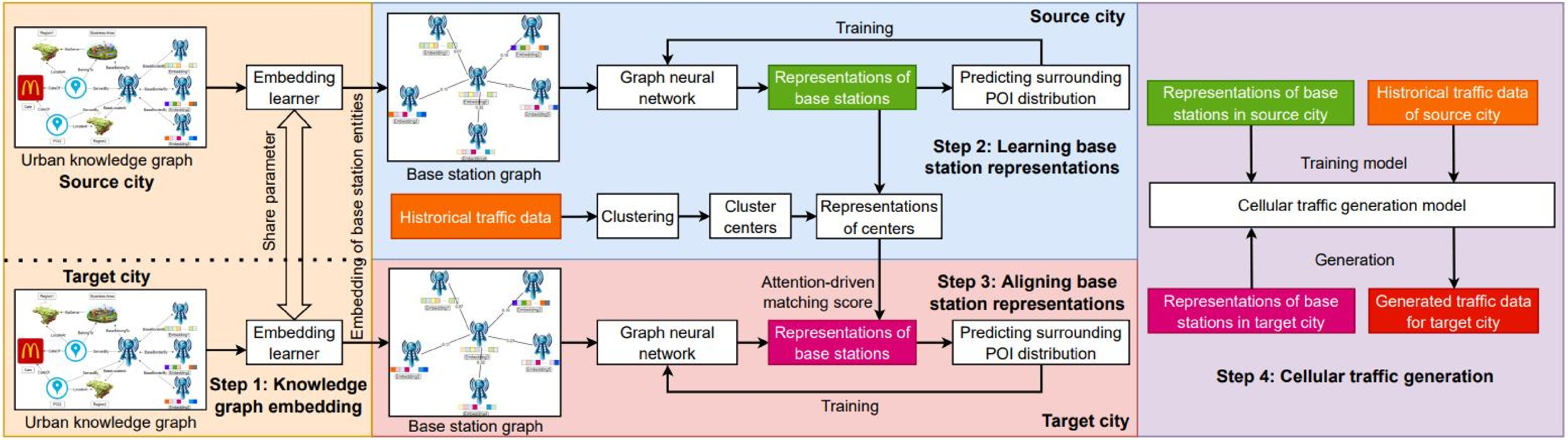
Framework Overview

(1) Knowledge graph embedding

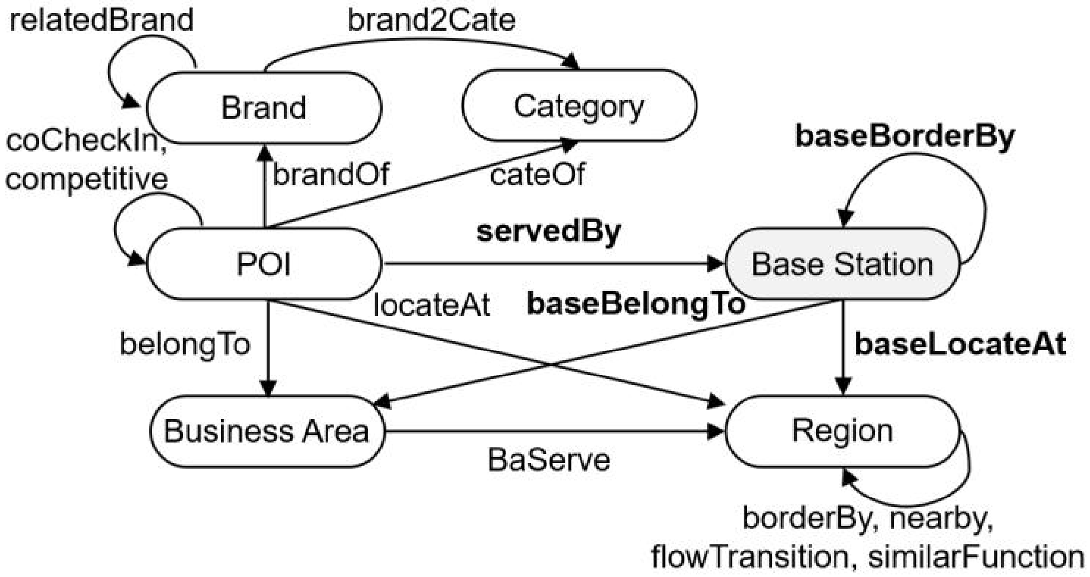
(2) Learning base station representations: GCN + L_{PD} Loss

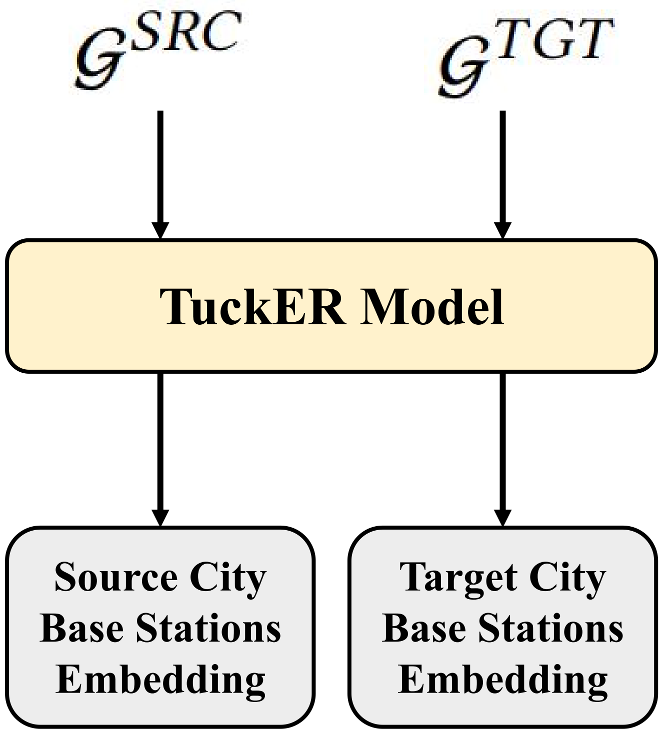
(3) Aligning base station representations: 2 types of Loss L_{PD} and $L_{Pattern}$

(4) Cellular traffic generation: a multi-generator structure



Knowledge Graph Embedding

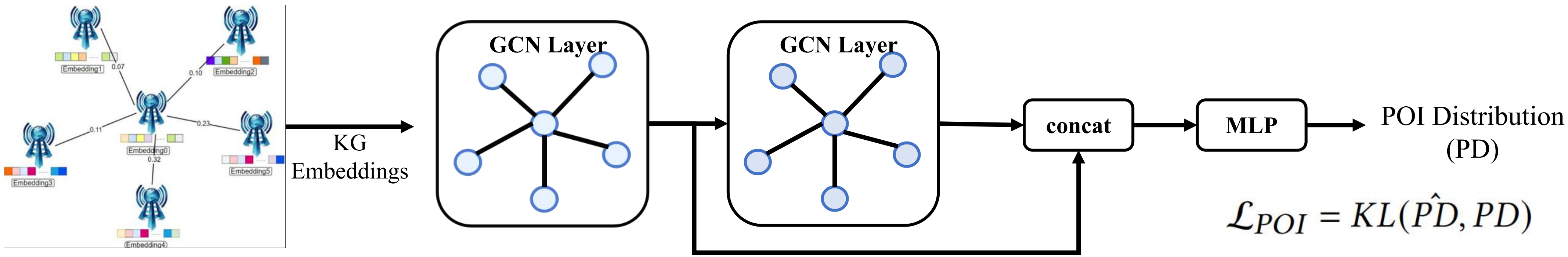




One base station is connected to other entities with four relations.

- One base station is located at a region.
- One base station belongs to a business area.
- A POI is served by a base station.
- A base station borders another base station.

Learning Base Station Representations



$$\mathcal{L}_{POI} = KL(\hat{PD}, PD)$$

- Top � closest base stations
- The edge weights

$$A_{BS}(i,j) = \frac{1/Dist(i,j)}{\max\limits_{i,j}\left(1/Dist(i,j)\right)}$$

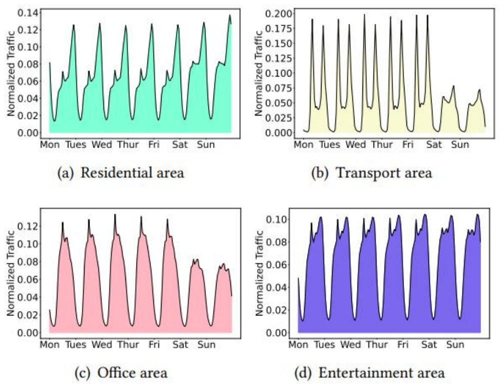
Two-layer GCN model

$$H_{BS}^{(l)} = \sigma\left(D_{BS}^{-\frac{1}{2}} A_{BS} D_{BS}^{-\frac{1}{2}} W^{(l)} H_{BS}^{(l-1)}\right)$$

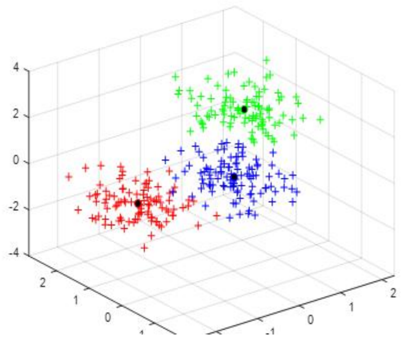
$$Z_{BS} = [H_{BS}^{(1)}, H_{BS}^{(2)}]$$

As for the source city, its base station represen tations are obtained by training GCN and MLP models with the POI similarity loss

Aligning Base Station Representations



Time Series of Source City



Source City Base Station Representation Cluster

$$\hat{z}_{BS,c} = \frac{1}{N_{BS,c}} \sum_{c_{BS_i}==c} z_{BS_i}^{SRC}$$

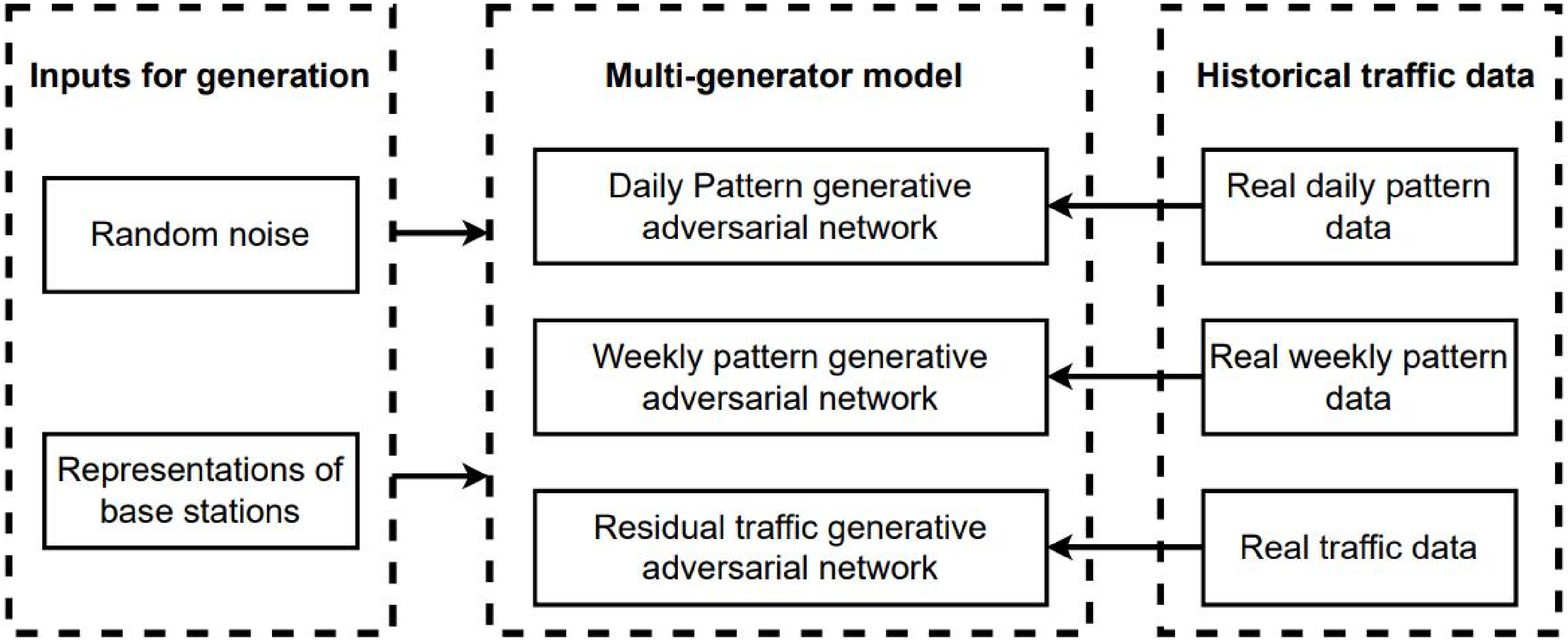
$$M(BS_i^{TGT}) = \frac{\exp(sim(z_{BS_i}^{TGT}, \hat{z}_{BS,k}))}{\sum_j^K \exp(sim(z_{BS_i}^{TGT}, \hat{z}_{BS,j}))}$$

$$k = \arg\max_j sim(z_{BS_i}^{TGT}, \hat{z}_{BS,j})$$

$$\mathcal{L}_{Pattern} = -\sum_i (1 - M(BS_i^{TGT})) \log(M(BS_i^{TGT}))$$

As for the target city, its base station representations are then obtained by training GCN and MLP models with both the POI similarity loss and the pattern matching loss

Cellular Traffic Generation



- We leverage a multi-generator structure to capture the daily pattern and weekly pattern of the traffic time series.
- We train the feature-enhanced generative adversarial network with Wasserstein loss based on the historical traffic data and representations of base stations in the source city.
- By feeding the noise and the target city's representations into the trained model, we can obtain the generated traffic data for the target city.

Experiments

Datasets

Beijing, Shanghai, Nanjing

Baselines

- TransGAN: A transformer-based GAN that combines a multi-scale discriminator to concurrently capture low-level textures and semantic contexts with a generator using transformer blocks that gradually enhance feature resolution.
- LSTM-based GAN: Two LSTMs are used as the generator and discriminator in constructing the GAN.
- TCN-based GAN: A GAN using temporal convolutional networks (TCNs) as the generator and discriminator for cellular traffic generation.

Experiments

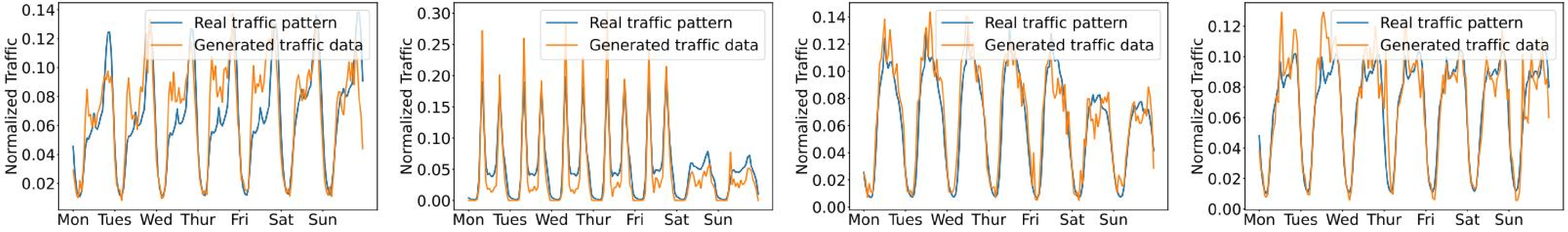
Overall Performance Evaluation (Shanghai -> Beijing, Nanjing)

Cities	Shanghai (Source) → Nanjing (Target)						Shanghai (Source) → Beijing (Target)					
Methods	Traffic Volume		First-order Difference		Daily Periodic Component		Traffic Volume		First-order Difference		Daily Periodic Component	
	JSD	Δ	JSD	Δ	RMSE	Δ	JSD	Δ	JSD	Δ	RMSE	Δ
Trans	0.5378	47.79%	0.1656	160.38%	0.0587	12.67%	0.5869	104.49%	0.1857	166.05%	0.1021	63.1%
Trans+ *PD*	0.5101	40.18%	0.1666	161.95%	0.0594	14.01%	0.5743	100.1%	0.1851	165.19%	0.1031	64.7%
Trans+ E_{BS}	0.5307	45.84%	0.1564	145.91%	0.0589	13.05%	0.5883	104.98%	0.1749	150.57%	0.1024	63.58%
Trans+ Z_{BS}	0.5140	41.25%	0.1540	142.14%	0.0566	8.64%	0.5759	100.66%	0.1575	125.64%	0.1023	63.42%
RNN	0.7294	100.44%	<u>0.0863</u>	35.69%	0.0567	8.83%	0.7103	147.49%	0.1846	164.47%	0.1054	68.37%
RNN+ *PD*	0.5914	62.52%	0.1328	108.81%	0.0638	22.46%	0.6613	130.42%	0.0944	42.41%	0.1079	72.36%
RNN+ E_{BS}	0.6226	71.09%	0.0931	46.38%	0.0523	0.38%	0.7026	144.81%	0.0766	9.74%	0.1010	61.34%
RNN+ Z_{BS}	0.5913	62.49%	0.1328	108.81%	0.0638	22.46%	0.6613	130.42%	0.0944	35.24%	0.1078	72.2%
TCN	0.7626	109.56%	0.1426	124.21%	0.1289	147.41%	0.5774	101.18%	0.1853	165.47%	0.1076	71.88%
TCN+ *PD*	0.5945	63.37%	0.1328	108.81%	0.1036	98.85%	0.4259	48.4%	0.0995	42.55%	0.1016	62.3%
TCN+ E_{BS}	0.7814	68.73%	0.1085	70.6%	0.0927	77.93%	0.8513	196.62%	0.0858	22.92%	0.0965	54.15%
TCN+ Z_{BS}	0.5674	55.92%	0.0963	51.42%	0.0847	62.57%	0.4133	44.01%	0.0844	20.92%	0.0841	34.35%
ADAPTIVE	0.7173	97.11%	0.1047	64.62%	<u>0.0516</u>	-0.96%	0.5703	98.71%	0.1454	108.31%	0.0890	42.17%
ADAPTIVE+ *PD*	0.5853	60.84%	0.0998	56.92%	0.0540	3.65%	0.5045	75.78%	**0.0679**	-2.72%	<u>0.0698</u>	11.5%
ADAPTIVE+ E_{BS}	<u>0.4985</u>	36.99%	0.0972	52.83%	**0.0470**	-9.79%	<u>0.3444</u>	20.0%	0.1782	155.3%	0.0712	13.74%
ADAPTIVE+ Z_{BS}	**0.3639**	0	**0.0636**	0	0.0521	0	**0.2870**	0	<u>0.0698</u>	0	**0.0626**	0

Experiments

Case Study

- ➢ The temporal patterns of the traffic time series generated by ADAPTIVE are consistent with real traffic.
- ➢ This verifies that the traffic temporal patterns are successfully transferred from the source city to the target city across different functional regions, demonstrating the effectiveness of the key designs of the knowledge graph module and attention-driven matching score.



(a) Residential area (b) Transport area (c) Office area (d) Entertainment area

Experiments

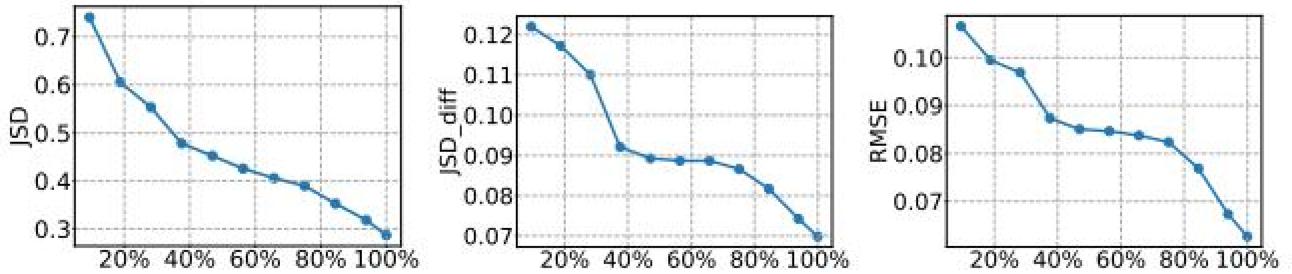
The traning data size

the model performance improves with the increase in the number of base stations in the training set. note: 40%, 80%.

Scale of urban knowledge graphs

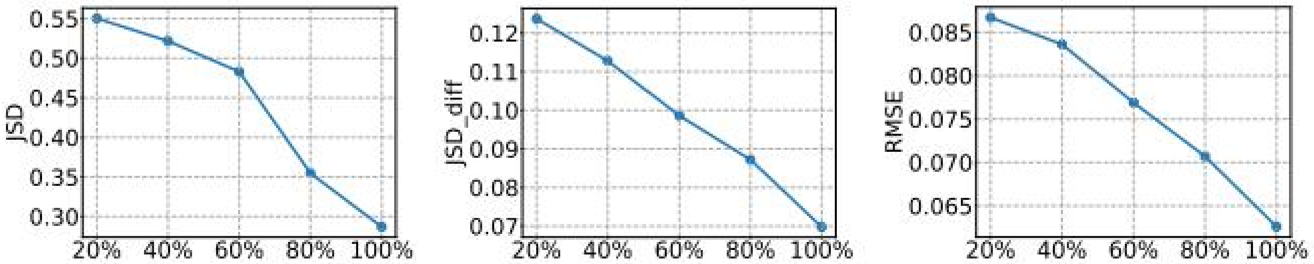
➢ changes with the scale of urban knowledge graphs by randomly removing the POI entities in the target city

➢ compared with the scale of historical data, environmental factors play a more critical role in the traffic generation task



(a) Traffic Volume (b) First-order Difference (c) Daily Periodic

Figure 7: Sensitivity to the training data size by changing the proportion of base stations in the training dataset.

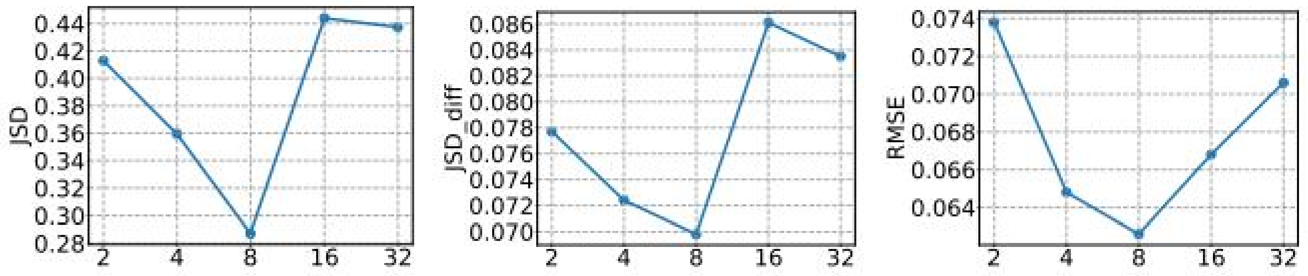


(a) Traffic Volume (b) First-order Difference (c) Daily Periodic

Figure 8: Sensitivity to the scale of knowledge graphs by changing the proportion of POIs in the knowledge graph.

Number of traffic pattern clusters

Too many clusters will make the difference between clustered patterns smaller and make it difficult for the model to learn the correct pattern.

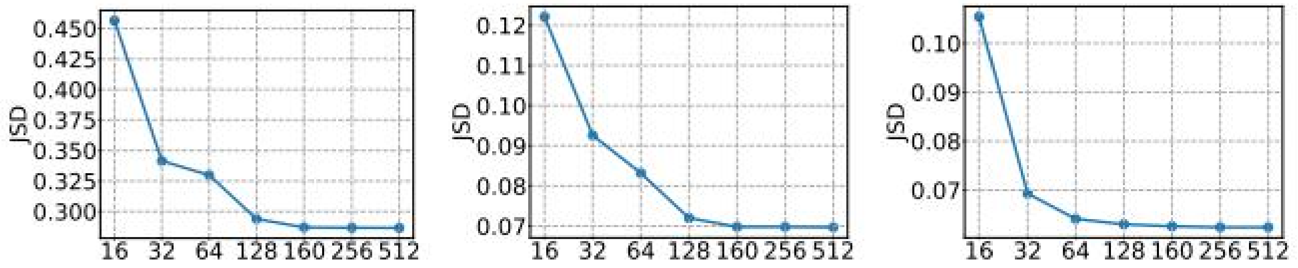


(a) Traffic Volume (b) First-order Difference (c) Daily Periodic

Figure 9: Sensitivity to the number of traffic pattern clusters.

Dimensions of base station representations

128-dimensional vectors can represent base stations in a city while containing information on the environmental factors of cities, spatial and environmental contextual relations between base stations, and traffic temporal patterns of base stations.



(a) Traffic Volume (b) First-order Difference (c) Daily Periodic

Figure 10: Sensitivity to the dimensions of base station representations.

Experiments

Cities	Beijing (Source) → Shanghai (Target)						Beijing (Source) → Nanjing (Target)					
Methods	Traffic Volume		First-order Difference		Daily Periodic Component		Traffic Volume		First-order Difference		Daily Periodic Component	
	JSD	Δ	JSD	Δ	RMSE	Δ	JSD	Δ	JSD	Δ	RMSE	Δ
Trans	0.4258	131.16%	0.1429	54.65%	0.0731	-0.14%	0.5310	58.89%	0.1656	161.67%	0.0819	94.54%
Trans+ *PD*	0.4040	119.33%	0.1409	52.49%	0.0740	1.09%	0.5100	52.6%	0.1666	161.83%	0.0594	41.09%
Trans+ E_{BS}	0.4303	133.6%	0.1277	38.2%	0.0730	-0.27%	0.5307	58.8%	0.1564	148.74%	0.0589	39.9%
Trans+ Z_{BS}	0.3979	116.02%	0.0988	6.93%	0.0734	0.27%	0.5134	53.62%	0.1540	154.1%	0.0574	36.34%
RNN	0.6584	257.44%	<u>0.0709</u>	-23.27%	0.0796	8.74%	0.7293	118.22%	0.0862	35.96%	0.0566	34.44%
RNN+ *PD*	0.5043	173.78%	0.1311	41.88%	0.0786	7.38%	0.6154	84.14%	0.1427	125.08%	0.0737	75.06%
RNN+ E_{BS}	0.5219	183.33%	0.1336	44.59%	0.0848	15.85%	0.7129	113.32%	0.0774	<u>22.08%</u>	0.0587	39.43%
RNN+ Z_{BS}	0.5042	173.72%	0.1311	41.88%	0.0786	7.38%	0.5914	76.96%	0.1322	108.52%	0.062	47.27%
TCN	0.5849	217.54%	0.1344	45.45%	0.0784	7.1%	0.8416	151.83%	0.1019	60.73%	0.0636	51.07%
TCN+ *PD*	0.4592	149.29%	**0.0669**	-27.6%	0.0784	7.1%	0.6013	79.92%	0.1097	73.03%	0.0636	51.07%
TCN+ E_{BS}	0.4252	130.84%	0.0979	5.95%	0.0780	6.56%	0.8242	146.62%	0.1443	127.6%	0.0537	27.55%
TCN+ Z_{BS}	0.3770	104.67%	0.0734	-20.56%	0.0784	7.1%	<u>0.4739</u>	41.8%	0.1023	61.36%	0.0636	51.07%
ADAPTIVE	0.3247	76.28%	0.1304	41.13%	<u>0.0641</u>	-12.43%	0.5561	66.4%	0.1710	169.72%	0.0517	22.8%
ADAPTIVE+ *PD*	0.6250	239.31%	0.1119	21.1%	**0.0457**	-37.57%	0.7109	112.72%	0.1444	127.76%	0.0363	-13.78%
ADAPTIVE+ E_{BS}	<u>0.2304</u>	25.08%	0.0729	-21.1%	0.0658	-10.11%	0.4951	48.14%	0.1311	106.78%	**0.0457**	8.55%
ADAPTIVE+Z_{BS}	**0.1842**	0	0.0924	0	0.0732	0	**0.3342**	0	**0.0634**	0	<u>0.0421</u>	0

Experiments

Cities	Nanjing (Source) → Beijing (Target)						Nanjing (Source) → Shanghai (Target)					
Methods	Traffic Volume		First-order Difference		Daily Periodic Component		Traffic Volume		First-order Difference		Daily Periodic Component	
	JSD	Δ	JSD	Δ	RMSE	Δ	JSD	Δ	JSD	Δ	RMSE	Δ
Trans	0.5869	185.46%	0.1857	125.64%	0.1021	89.78%	0.4173	165.46%	0.1326	79.67%	0.0713	-9.06%
Trans+ *PD*	0.6734	227.53%	0.1894	130.13%	0.1037	92.75%	0.3946	151.02%	0.1266	71.54%	0.0720	-8.16%
Trans+ E_{BS}	0.6814	231.42%	0.1803	119.08%	0.1012	88.1%	0.4319	174.75%	0.1022	38.48%	0.0712	-9.18%
Trans+ Z_{BS}	0.6627	222.32%	0.1653	100.85%	0.0987	83.46%	0.3879	146.76%	0.0924	25.2%	0.0696	-11.22%
RNN	0.7712	275.1%	0.1314	59.66%	0.0999	85.69%	0.5876	273.79%	0.0903	22.36%	0.0710	-9.44%
RNN+ *PD*	0.7016	241.25%	0.1049	27.46%	0.1088	102.23%	0.4418	181.04%	0.101	36.86%	0.0786	0.26%
RNN+ E_{BS}	0.7185	249.46%	0.0894	8.63%	0.1041	93.49%	0.4697	198.79%	0.1353	83.33%	0.0854	8.93%
RNN+ Z_{BS}	0.6016	192.61%	0.0944	14.7%	0.1078	100.37%	0.4416	180.92%	0.1011	36.99%	0.0717	-8.55%
TCN	0.6037	193.63%	0.1623	97.21%	0.1068	98.51%	0.5926	276.97%	0.1097	48.64%	0.0777	-0.89%
TCN+ *PD*	0.5012	143.77%	0.0867	5.35%	0.1076	100.0%	0.516	228.24%	0.0965	30.76%	0.0816	4.08%
TCN+ E_{BS}	0.8513	314.06%	0.0858	4.25%	0.0965	79.37%	0.5144	227.23%	0.1079	46.21%	0.0890	13.52%
TCN+ Z_{BS}	0.3855	87.5%	0.1218	48.0%	0.1076	100.0%	0.3682	134.22%	0.0876	18.7%	0.0748	-4.59%
ADAPTIVE	0.3406	65.66%	0.1635	98.66%	0.0711	32.16%	0.3430	118.19%	0.1406	90.51%	0.0633	-19.26%
ADAPTIVE+ *PD*	0.4311	109.68%	**0.0453**	-44.96%	0.0692	28.62%	0.5881	274.11%	0.0864	17.07%	**0.0454**	-42.09%
ADAPTIVE+ E_{BS}	0.3303	60.65%	0.0584	-29.04%	0.0711	32.16%	0.2218	41.09%	**0.0728**	-1.36%	0.0656	-16.33%
ADAPTIVE+ Z_{BS}	**0.2056**	0	0.0823	0	**0.0538**	0	**0.1572**	0	0.0738	0	0.0784	0

谢谢各位听众

敬请批评指正