

基于交通流 ST_SSL 的改进工作

摘要

在本次项目当中，我所复现的工作是基于交通流 ST_SSL，其提出了一个 ST_SSL 的模型框架，综合时间与空间完成了建模，并结合图神经网络实现了鲁棒交通流当中的交通流出量与流入量的任务。具体的，它根据对比学习与自监督分类任务分别对时间与空间进行建模，综合图神经网络编码的 loss 值定义了一个多重目标的损失函数；并且在训练过程基于图数据的拓扑结构以及属性进行了数据增强，丰富了自监督学习的数据。最后在我的复现下，所有数据集的实验基本都能够达到与原文略低的效果。并且，受到近期在图神经网络当中兴起的预训练架构思维的影响，我在后续工作中为 ST_SSL 增加了预训练部分，并且其架构的训练下，得到了大多数指标比原文复现更好的效果，特别的在 BJTaxi 以及 NYCTaxi 数据集当中取得了全面提升的结果。

关键词：ST_SSL；图神经网络；预训练架构；数据增强

1 引言

在本章当中，根据我自己的选题以及复现任务，我将对选题背景、选题依据以及选题意义分别进行详细的描述。

1.1 选题背景

随着全球城市化的加速和人口的增长，城市面临着日益复杂的交通管理挑战。城市中的交通流量不断增加，人们对快速、便捷的交通方式的需求导致了更多车辆上路，进一步增加了交通流量的复杂性，其中不仅仅是包括私人轿车，无论是出租车还是当前新兴的共享行业如共享单车与共享电动车都将成为交通流量的组成部分。无论是为了更好的维持交通的正常运行还是维持交通的可持续发展，都需要有效的方法来预测和管理这些流量。

1.2 选题依据

当前图神经网络的应用也逐渐流入交通流的预测任务当中，我的研究方向也是正好是有关图神经网络方向，最近对图神经网络的前沿技术有着较多的了解；并且在鲁棒性交通流的预测任务的调研工作当中，发现其虽有很多研究者，但基于图神经网络所解决问题的研究还没有完全与图神经网络的前沿研究同步，所以希望通过将近期所了解的关于图神经网络的前沿技术迁移入交通领域当中，希望寻求一个更好的解决方案以及预测效果，在检验自身的学习的同时，丰富自己在科研上的探索经验。

1.3 选题意义

通过对交通流信息的预测任务，可以更好地解决当今社会上的交通问题，比如（1）通过准确预测交通流量，城市可以更好地规划和优化交通系统。这包括合理安排道路和交叉口的设计，制定最佳交通信号控制策略，以及更有效地管理交通拥堵；（2）通过提前预测交通拥堵区域和时间，城市管理者可以采取的措施来减少拥堵，例如提供替代交通路线、调整公共交通服务和实施交通管制措施；（3）预测交通流量有助于更好地利用交通资源，包括公共交通工具、停车场和道路网络。

2 相关工作

在本章节当中，我将对当前图神经网络以及图神经网络在交通流建模上的应用分别进行详细的概括与描述，对其发展历程进行相应的分析。

2.1 图神经网络

图神经网络在近期已经得到了很多的研究关注，而关于图神经网络的研究也得到了飞速的发展，从 kipf 等人 [4] 提出的图卷积网络，到加入了注意力机制的 [7] 图注意力网络。后续由于图数据的稀缺以及标注难度大，研究者又进而提出了结合了生成模型的 Gpt-gnn [1] 网络。除此以外，随着图神经网络的飞速发展，越来越多的研究者希望借用这个利用非同类似图像以及文本信息数据的图数据神经网络来解决更多领域内的问题，比如有关社交网络的分析、推荐系统、生物信息学还有本次复现任务所研究的关于时间图的交通领域。

2.2 鲁棒交通流预测

跨不同空间区域、不同时间段的鲁棒交通流预测对于推进智能交通系统 [3] 至关重要，其中交通预测旨在从过去的交通观测值中预测交通量（例如，给定时刻每个区域的流入和流出）。尽管当前图神经网络通过对 GCN 以及 GAT 等网络在区域交通流的预测上已经有了较好的效果，但当前工作仍存在有两个问题。第一个问题是缺乏对不同区域的交通分布倾斜所表现出的空间异质性的建模。以图 1 (a) 为例，A 和 B 是北京的两个现实世界区域，具有不同的城市功能，即居住区和交通枢纽。从图 1 (b) 中我们可以观察到它们截然不同的交通流分布。然而，大多数现有模型忽略了这种空间异质性，并且很容易偏向交通量较高的热门区域，这使得它们不足以学习高质量的全市交通模式表示。虽然一些研究试图捕获不同区域上具有多个参数集的异质流量分布 [6]，但所涉及的大参数可能会导致倾斜分布流量的次优问题数据。

此外，当前的流量预测方法使用所有时间段的共享参数空间对时间动态进行建模，这很难精确地保留潜在嵌入空间中的时间异质性。在现实生活场景中，不同区域的交通模式随着时间的推移而变化，例如从早上到晚上，这导致了如图 1 (c) 所示的时间异质性。

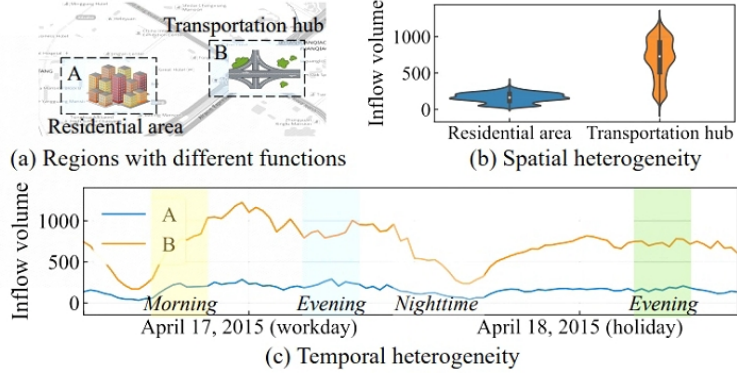


图 1. 时空信息

3 本文方法

在本章当中，我将对复现论文模型实现细节进行详细的介绍与分析，针对 ST_SSL 模型的不同阶段任务以及损失函数进行详细的介绍。

3.1 本文方法概述

首先，本文主要的框架可见图 2，其中 (a) 是 ST_SSL 的整体算法流程；(b) 是基于空间维度的建模；(c) 是基于时间维度的建模。具体的，我们将在后续章节对 ST_SSL 的具体流程进行详细介绍。

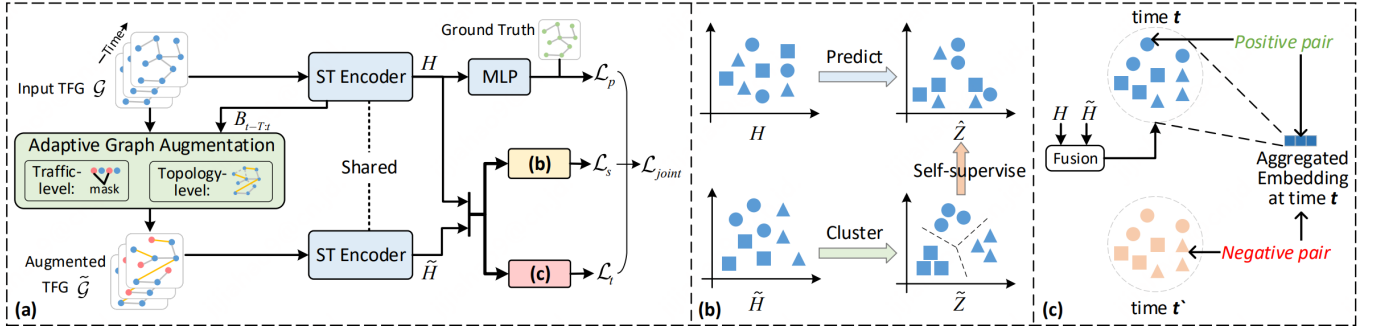


图 2. ST_SSL 整体流程

为了后续更好的进行文章方法的讲解，在这里先进行相关定义的介绍：

交通流图 (TFG, Traffic Flow Graph): 交通流图被定义为: $G = (V, E, A, X_{t-T:t})$ ，其中 V 代表区域空间集合（也可以看作是节点集），其长度为 $|V| = N$ ； E 是连接两个节点的边集。而 $A \in \mathbb{R}^{N \times N}$ 为 TFG 的邻接矩阵；交通张量 $X_{t-T:t} : t \in \mathbb{R}^{T \times N \times 2} = (X_{t-T}, \dots, X_t)$ 代表过去 T 个时间步长的城市交通流出与流入数据。

任务定义: 给定截至当前时间步的历史交通流量图 G ，我们旨在学习一个预测函数来准确估计未来 $t+1$ 时间步长所有区域的交通量，即 $X_{t+1} \in \mathbb{R}^{N \times 2}$ 。

3.2 Spatio-Temporal Encoder

在 ST_SSL 任务当中, 文章提出了一种新的基于时空的图神经网络编码模型 ST Encoder, 其对于时序流量模式的编码中, 采用了具有门控机制 [8] 的沿时间维度的一维因果卷积。具体地, 我们的时间卷积 (TC) 将交通流张量作为输入, 并为每个区域输出一个时间感知的嵌入:

$$(B_{t-T_{out}}, \dots, B_t) = TC(X_{t-T}, \dots, X_t) \quad (1)$$

其中 $B_t \in \mathbb{R}^{N \times D}$ 表示在时间步长上的区域嵌入矩阵 t 。第 n 行 $B_{t,n} \in \mathbb{R}^D$ 为区域 r_n 的 embedding 值, D 代表嵌入的维度, T_{out} 为 TC 编码器在卷积操作后输出的嵌入序列的长度。

除此以外, 还加入了一个基于图消息传递机制的空间卷积 (SC) 编码器:

$$E_t = SC(B_t, A) \quad (2)$$

在这之中 A 是图 \mathbb{G} 的区域矩阵, 在 SC 编码器的作用后, 可以获得加入区域信息的 Embedding $(E_{t-T_{out}}, \dots, E_t)$ 。

ST 编码器采用“三明治”块的结构, 其中 $TC \rightarrow SC \rightarrow TC$ 为每个单独的块。通过叠加多个块, 多次卷积后可以得到时间维度为 T 的嵌入矩阵 (H_{t-T}, \dots, H_t) 序列。基于 ST 编码器的嵌入传播和聚合后, 时间维度 T 降为零, 我们为 ST 编码器生成最终的嵌入矩阵 $H \in \mathbb{R}^{N \times D}$, 其中每一行 $h_n \in \mathbb{R}^D$ 表示区域 r_n 的最终嵌入。

3.3 在 TFG 中的数据增强方案

这里介绍对 TFG 数据增强的方案, 分为两个阶段, 一个是基于流量的数据增强, 一个是基于拓扑结构的数据增强。

3.3.1 基于流量的数据增强

文章受到 [5] 的数据增强策略的启发, 在构建的交通张量 $X_{t-T:t}$ 设置了个增强算子, 以此适应每个区域学习的时间感知交通模式的依赖性。具体的, 文章通过伯努利分布获取到掩码概率 $\rho_{\tau,n} \text{Bern}(1 - p_{\tau,n})$, 其中 $\rho_{\tau,n}$ 值越高代表其区域 r_n 被掩码的概率越大。基于流量的增强数据被表示为 $\tilde{X}_{t-T:t}$ 。

3.3.2 基于拓扑结构的数据增强

拓扑结构的数据增强其实也就是通过增加节点之间连接的边的方式增加数据的数量, 而在交通流中就是增强各个节点之间的联系与依赖关系。这里提到一个新的概念异质度程度 $q_{m,n}$, 其反应的是两个节点之间的差异性, 具体计算方法可见下:

$$q_{m,n} = \frac{u_m^T}{\|u_m\| \|u_n\|} \quad (3)$$

对于两个不相邻区域 r_m, r_n , 其掩码概率 $\rho_{m,n} \text{Bern}(1 - q_{m,n})$ 同样通过伯努利分布获取, 其中低异质度程度 $q_{m,n}$ 的两个区域将会对其进行增加边的操作。

3.4 基于 SSL 的空间维度建模

在基于空间维度的建模任务当中，首先文章通过分析区域之间的差距，将其划分为如郊区、购物中心、交通枢纽等区域类型，并通过如此定义进行聚类处理，进而获取到相应的区域类别编码 $\{c_1, \dots, c_k\}$ 。其中聚类过程可以定义为 $\tilde{z}_{n,k} = c_k^T \tilde{h}_n$ ，其中 $\tilde{h}_n \in \mathbb{R}^D$ 是 TFG 的区域 r_n 的编码， $\tilde{z}_{n,k}$ 表示的是区域 r_n 在第 k 类的相关性得分。为了提出一个可学习的自监督任务，文章将空间建模中增加了其对区域类别的预测任务，其中预测任务的 loss 定义为：

$$l(h_n, \tilde{z}_n) = \sum_k \tilde{z}_{n,k} \log \frac{\exp(\hat{z}_{n,k}/\gamma)}{\sum_j \exp(\hat{z}_{n,j}/\gamma)} \quad (4)$$

其中 γ 是控制 softmax 层输出平滑的 temperature 参数，整体所有区域自监督的目标定义为：

$$L_s = \sum_{n=1}^N l(h_n, \tilde{z}_n) \quad (5)$$

3.5 基于 SSL 的时间维度建模

进一步，文章为时间维度同样设置了一个基于自监督学习 (SSL) 的任务，将时间异质性注入到时间感知区域 embedding 当中。其先通过融合增强的与原始的基于时间的区域 embeddings：

$$v_{t,n} = w_1 \odot h_{t,n} + w_2 \odot \tilde{h}_{t,n} \quad (6)$$

其中 \odot 代表哈达玛积， w_1 与 w_2 为可学习参数，基于此可以得到在 t 时段的城市级别的表达 s_t (其中 ρ 是 sigmoid 函数)：

$$s_t = \rho\left(\frac{1}{N} \sum_{n=1}^N v_{t,n}\right) \quad (7)$$

在不同的时间维度上，文章将同一个时间段的交通图编码作为自监督任务当中的正样本对，对不同时间段的交通图编码作为自监督任务当中的负样本对。于是文章提出了基于对比学习的方法来帮助模型对时间维度的信息进行学习，其损失定义如下：

$$L_t = -\left(\sum_{n=1}^N \log g(v_{t,n}, s_t) + \sum_{n=1}^N \log(1 - g(v_{t',n}, s_t))\right) \quad (8)$$

3.6 整体建模与损失函数定义

除了上述基于空间与时间的任务建模外，ST_SSL 还有一个用于转换预测结果的 MLP 层，其预测未来时间交通流的方式如下：

$$\hat{x}_{t+1,n} = MLP(h_n) \quad (9)$$

其中 $\hat{x}_{t+1,n}$ 是预测结果，其优化方式是通过最小化下面的损失函数：

$$L_p = \sum_{n=1}^N \lambda \left| x_{t+1,n}^{(0)} - \hat{x}_{t+1,n}^{(0)} \right| + (1 - \lambda) \left| x_{t+1,n}^{(1)} - \hat{x}_{t+1,n}^{(1)} \right| \quad (10)$$

其中 $x_{t+1,n}^{(0)}$ $x_{t+1,n}^{(1)}$ 为交通流入量和流出量的真实标签， λ 是平衡流入与流出量的一个参数。

模型整体的损失函数可以由上述所有任务损失函数联合定义：

$$L_{joint} = L_p + L_s + L_t \quad (11)$$

4 复现细节

在本章节中，我将对我在进行 ST_SSL 模型进行复现过程中的遇到的问题以及解决方法，还有并在复现过后仍对代码进行了进一步创新，结合预训练框架尝试对模型效果增加进一步的提升。

4.1 与已有开源代码对比

本次复现参考了论文 [2] 所提供的源代码<https://github.com/Echo-Ji/ST-SSL>，当前代码对比于开源代码有了进一步的改进，增添了更多的数据集处理，并根据创新点增添了新的代码量。具体的，我在基于开源代码中增加了预训练阶段，借此希望通过先不学习时空相关信息，而只是将区域数据集作为图数据进行训练，于是我采用同当前图神经网络当中的预训练架构对 ST_SSL 模型进行了修改，希望借此以获取更好的实验结果。其中用于评估的两个指标 MAE,MAPE 分别为平均绝对误差与平均百分比误差，其均为反应预测误差的指标，且都是越小代表模型效果预测能力越强，其具体计算如下：

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (12)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \right) \quad (13)$$

在本次复现工作中，我完成了 4 个数据集的复现工作，其具体的复现效果可见表1。

表 1. 实验结果

数据集	实验设置	MAE(in)	MAPE(in)	MAE(out)	MAPE(out)
NYCBike1	原文	4.94	5.26	23.69	24.60
	复现	5.01	5.36	24.13	24.18
NYCBike2	原文	5.04	4.71	22.54	21.17
	复现	5.05	4.73	22.33	21.07
BJTaxi	原文	11.31	11.40	15.03	15.19
	复现	11.46	11.54	15.19	15.35
NYCTaxi	原文	11.99	9.78	16.38	16.86
	复现	12.76	10.24	17.87	17.33

4.2 实验环境搭建

在实验中，我采用的核心代码库以及其版本可见表2。

具体的，我所采用的数据集为原文所采用的四个有关交通流的数据集，其分别为 NY-CBike1、NYCBike2、BJTaxi、NYCTaxi，其分别是关于在地区当中自行车以及出租车的流动

表 2. 版本信息

函数库	版本
<i>numpy</i>	1.21.2
<i>pandas</i>	1.3.5
<i>pyyaml</i>	6.0
<i>torch</i>	1.10.1

关系。其数据集的统计方法是通过将一个地区划分为 $m \times n$ 个网格区域，随后通过划分时间区间以统计各个网格区域当中的自行车或出租车的数量，借此构成了有关时间序列的交通流数据集，具体信息可见表3。

表 3. 数据集信息介绍

数据类型	共享单车		出租车 GPS	
数据集	NYCBIKE1	NYCBIKE2	NYCTAXI1	BJTAXI
时间间隔	1 HOUR	30 MIN	30 MIN	30 MIN
节点	16×8	10×20	10×20	32×32
自行车/出租车	6.8k+	2.6m+	22m+	34k+

4.3 创新点

随着图神经网络近期的快速发展，更多有关其他领域中的优秀技术也被不断将概念迁移至图神经网络的研究当中，就比如基于预训练架构的图神经网络学习，当前正是图神经网络所研究的热门方向。但在交通流的研究当中，当前并没有什么研究将预训练架构进行结合。所以，我希望通过结合预训练的架构尝试增强对鲁棒交通流的预测的效果。新的模型架构图可见图3。

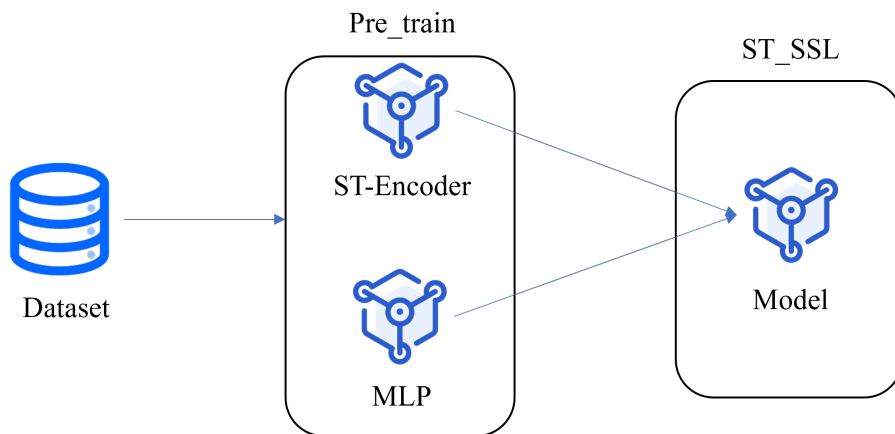


图 3. 改进后模型框架图

在预训练阶段，我旨在增强图编码器对图的表征能力，故在预训练阶段，我并不考虑有关 ST_SSL 当中的有关时间和空间的建模，而重在完成对图数据编码的图编码器以及最终进行交通流预测的 MLP 层的模块进行初始化训练。在预训练阶段所训练得到的 MLP 参数与 ST-Encoder 的参数将全部继承给后续 ST_SSL 下游任务的训练。基于这样的方式，我们可以在 ST_SSL 初始阶段就可以获得一个表征能力较成功的 ST-Encoder，并且在 ST_SSL 阶段训练不将 ST-Encoder 部分参数进行冻结，而是按照与原文相同流程进行进一步结合时空建模的训练任务一同训练，进而希望借此获取更好的模型交通流预测效果。

5 实验结果分析

在对比自己的创新点完成后与原文复现结果可以看出，在部分数据集当中，加入了预训练阶段的实验结果会比复现效果有所提升，达成了实验的目的。具体结果可见表格4。

表 4. 实验结果

数据集	实验设置	$MAE(in)$	$MAPE(in)$	$MAE(out)$	$MAPE(out)$
<i>NYCBike1</i>	原文	4.94	5.26	23.69	24.60
	复现	5.01	5.36	24.13	24.18
	加入预训练	4.97	5.24	24.47	24.25
	提升	↑	↑	↓	↓
<i>NYCBike2</i>	原文	5.04	4.71	22.54	21.17
	复现	5.05	4.73	22.33	21.07
	加入预训练	5.09	4.68	22.44	20.99
	提升	↓	↑	↓	↑
<i>BJTaxi</i>	原文	11.31	11.40	15.03	15.19
	复现	11.46	11.54	15.19	15.35
	加入预训练	11.21	11.29	15.16	15.27
	提升	↑	↑	↑	↑
<i>NYCTaxi</i>	原文	11.99	9.78	16.38	16.86
	复现	12.76	10.24	17.87	17.33
	加入预训练	12.06	9.73	16.28	16.78
	提升	↑	↑	↑	↑

在表格中可以看出，加入了预训练后在 NYCBike1、BJTaxi 以及 NYCTaxi 数据集中，模型效果表现更好，甚至有些超过了原文所提的效果，事实证明加入预训练阶段的实验对于交通流的预测是有实际效果的，可以有更多的研究价值。

6 总结与展望

在此次复现以及改进过程中，我完成了对 ST_SSL 原文实验结果的复现，其复现结果大致能达到文章所提及的结果。除此以外，我还对实验进行了相应的改进，基于图神经网络的新兴的预训练架构为其增加了相应的预训练阶段实验，最终改进后的结果相较复现结果在大部分指标中有了相应的提升，特别的在 BJTaxi 和 NYCTaxi 上有着全面提升的效果。

除此以外我发现在当前预训练阶段的任务设置过于简便，没有过多的对于交通流数据的相关考量，虽然有些数据集结果有所提升，但收益甚微，所以我认为在未来还能对预训练阶段进行进一步的考量，看能不能有更多的预训练任务选择，进而达到更好的模型效果。

参考文献

- [1] Ziniu Hu, Yuxiao Dong, Kuansan Wang, Kai-Wei Chang, and Yizhou Sun. Gpt-gnn: Generative pre-training of graph neural networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1857–1867, 2020.
- [2] Jiahao Ji, Jingyuan Wang, Chao Huang, Junjie Wu, Boren Xu, Zhenhe Wu, Junbo Zhang, and Yu Zheng. Spatio-temporal self-supervised learning for traffic flow prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 4356–4364, 2023.
- [3] Jiahao Ji, Jingyuan Wang, Zhe Jiang, Jingtian Ma, and Hu Zhang. Interpretable spatiotemporal deep learning model for traffic flow prediction based on potential energy fields. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 1076–1081, 2020.
- [4] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [5] Mengzhang Li and Zhanxing Zhu. Spatial-temporal fusion graph neural networks for traffic flow forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 4189–4196, 2021.
- [6] Zheyi Pan, Zhaoyuan Wang, Weifeng Wang, Yong Yu, Junbo Zhang, and Yu Zheng. Matrix factorization for spatio-temporal neural networks with applications to urban flow prediction. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM '19*, page 2683–2691, New York, NY, USA, 2019. Association for Computing Machinery.
- [7] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [8] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*, 2017.