北京航空航天大学计算机学院
School of Computer Science and Engineering, Beihang University

多变量时序预测的公平性

领读人：寄家豪

北京航空航天大学计算机学院
先进计算机应用技术教育部工程研究中心

arxiv论文（2023.01）

Learning Informative Representation for Fairness-aware Multivariate Time-series Forecasting: A Group-based Perspective

Hui He, Qi Zhang, Shoujin Wang, Kun Yi, Zhendong Niu, and Longbing Cao, *Senior Member, IEEE*

Hui He received the M.E. degree from University of Shanghai for Science and Technology, Shanghai, China in 2020. She is currently pursuing the Ph.D. degree at Institute of Engineering Medicine, Beijing Institute of Technology, Beijing, China. Her current research interests focus on multivariate time-series analysis and knowledge services.

机器学习中的公平性问题

公平性定义

- **Fairness** in machine learning refers to the various attempts at correcting algorithmic bias in automated decision processes based on machine learning models.
- Decisions made by computers after a machine-learning process may be considered ***unfair*** if they were based on variables considered **sensitive**.
- Examples of these kinds of variable include gender, ethnicity, sexual orientation, disability and more.

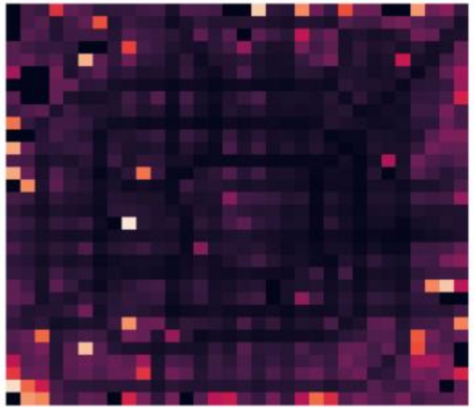
一个模型对于不同的样本的*拟合效果的一致*程度。如果一个模型在某些样本上效果很好，另一些样本上效果不好，那么这个模型的公平性就比较差。

公平性案例

人脸识别

信用贷款

推荐系统



智能交通

一个模型对于不同的样本的*拟合效果的一致*程度。如果一个模型在某些样本上效果很好，另一些样本上效果不好，那么这个模型的公平性就比较差。

公平性趋势

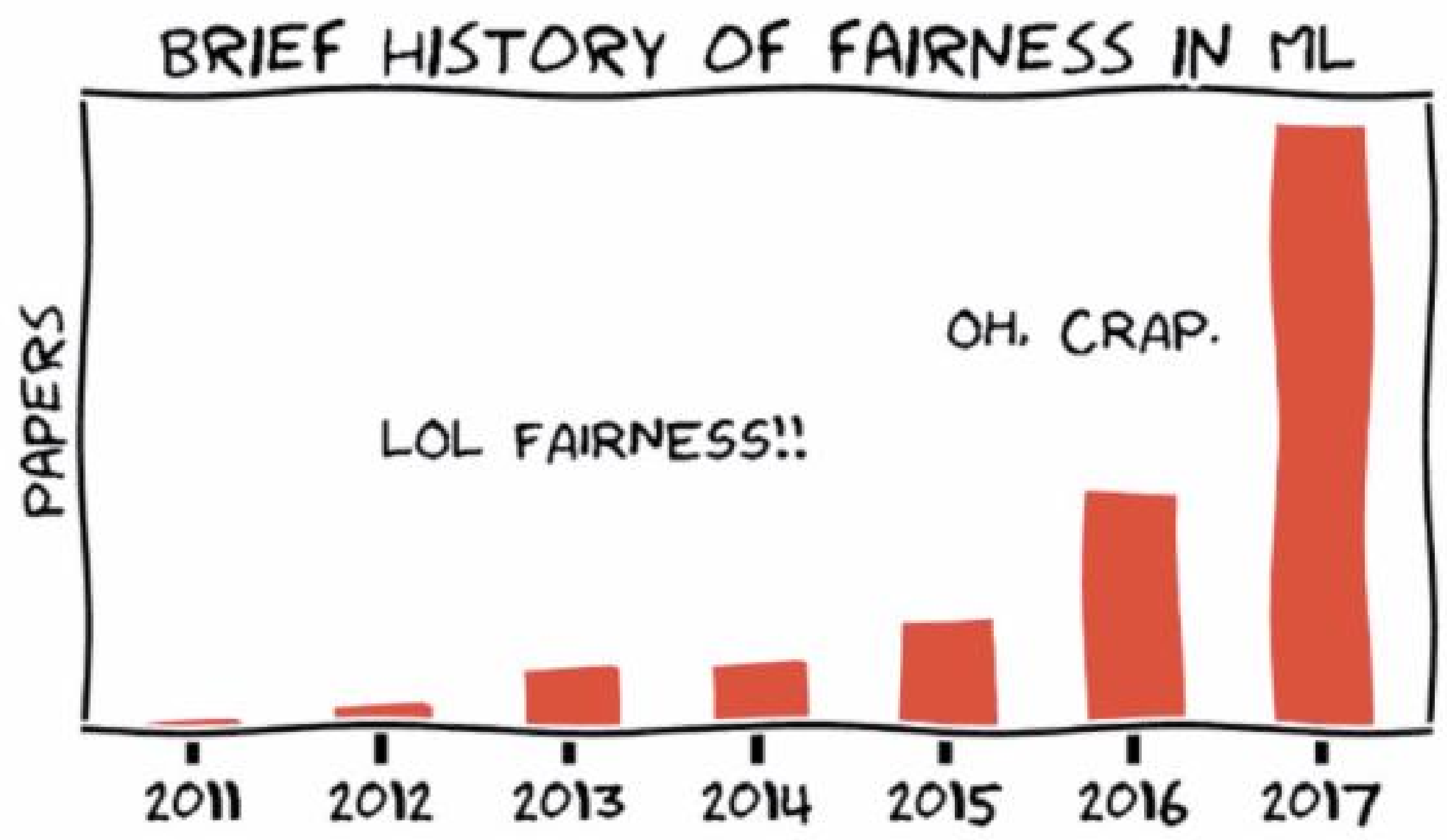


Fig1. The number of publications on fairness from 2011 to 2017

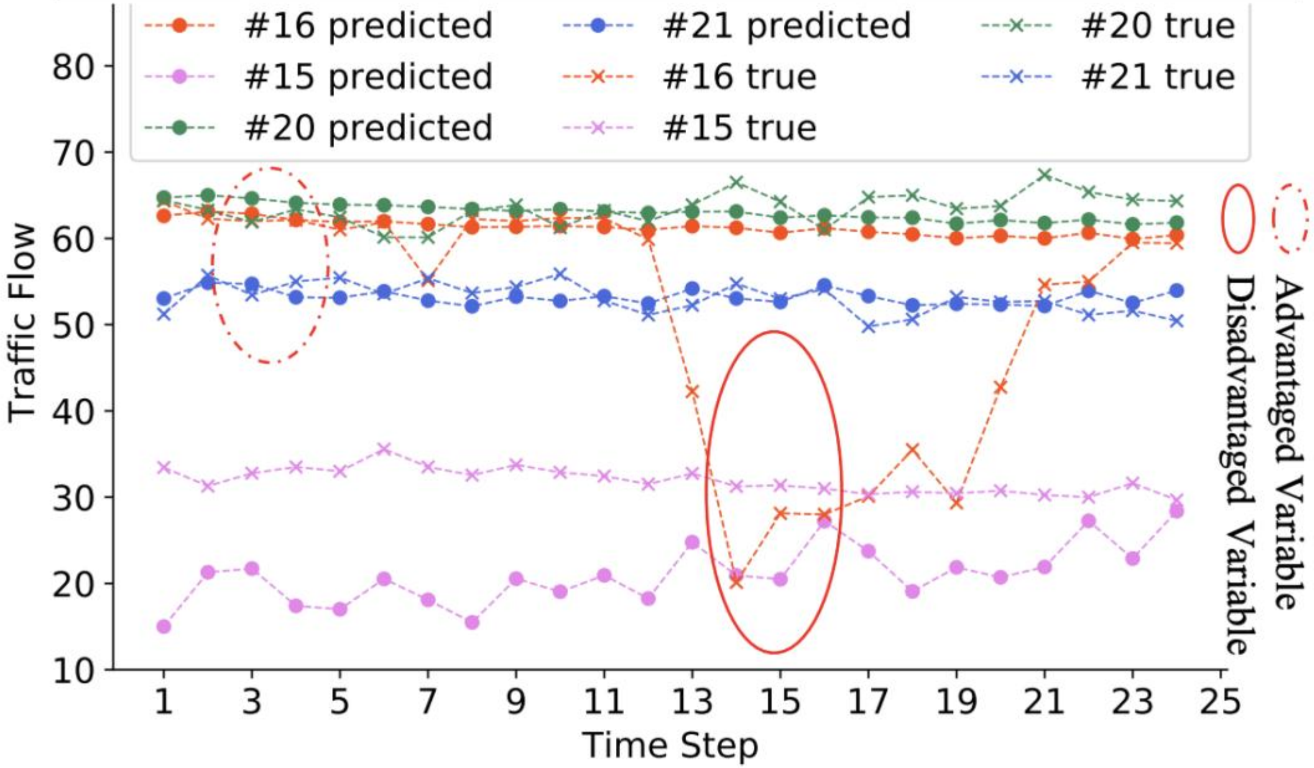
(Narayanan, 2018)

MTS预测中的公平性问题

MTS预测的公平性

- 多源时序预测的公平性
  - 模型是否对各个**变量**的预测效果都比较好?

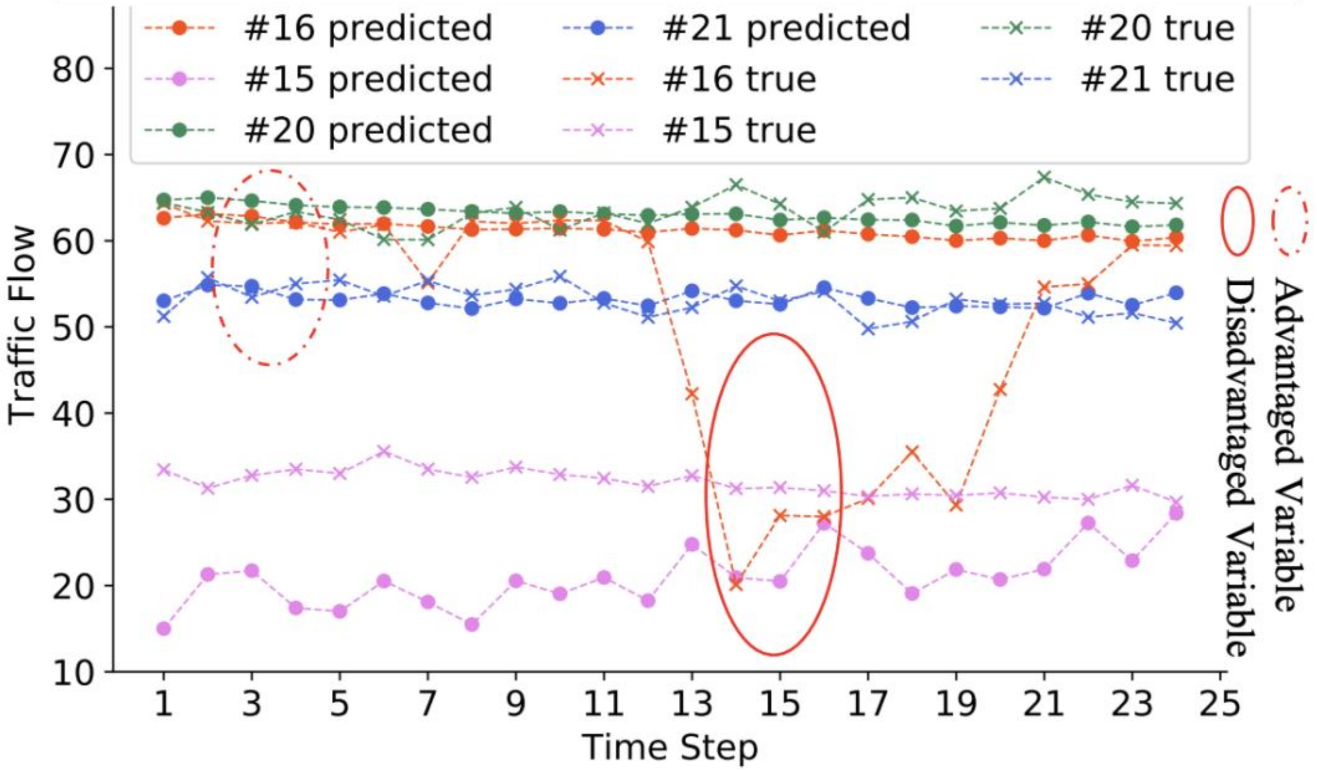
Methods	StemGNN	Informer	TS2Vec	MTGNN	AGCRN
VAR	97.0574	46.9595	119.3862	144.0949	122.7769



MTS预测的公平性

- 多源时序预测的公平性
  - 模型是否对各个**变量**的预测效果都比较好?
- 原因分析
  - **不同样本的特点不同，模型**在训练过程中可能**被某些样本的特点主导**

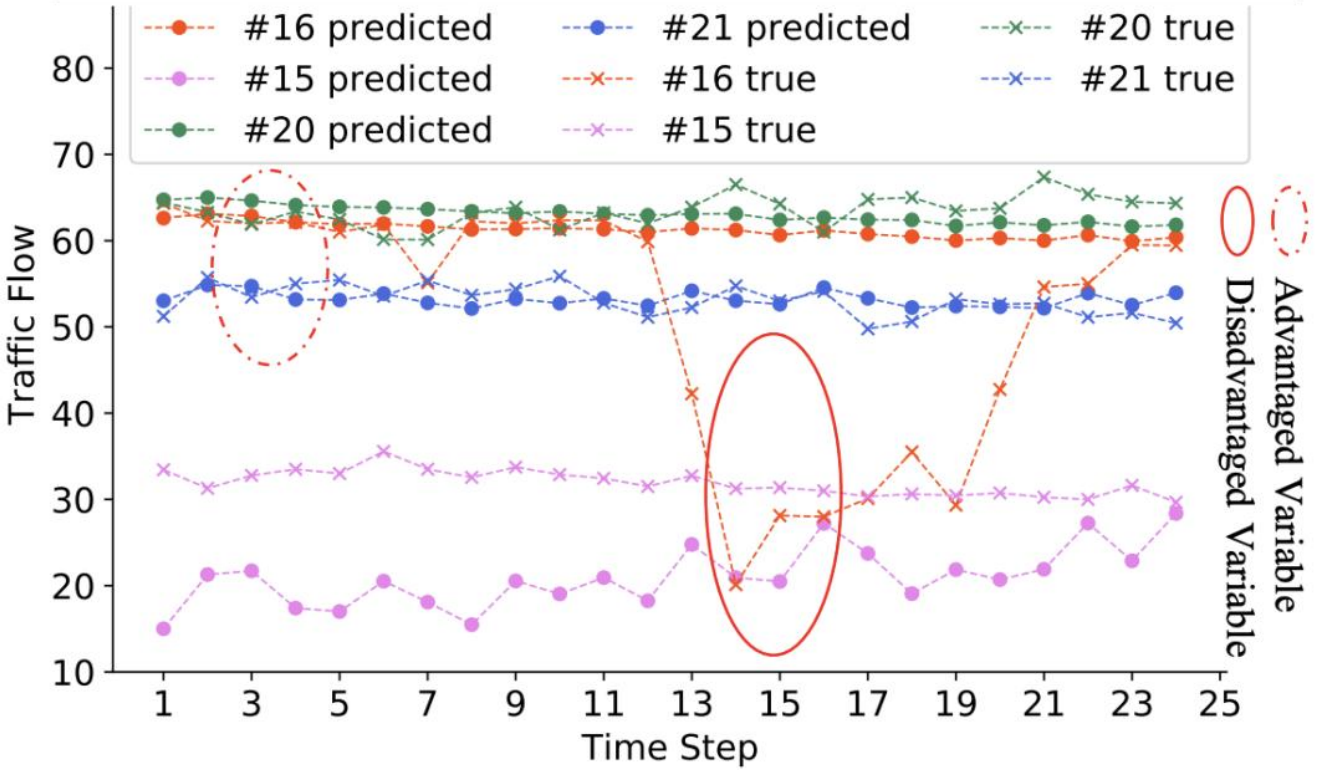
Methods	StemGNN	Informer	TS2Vec	MTGNN	AGCRN
VAR	97.0574	46.9595	119.3862	144.0949	122.7769



MTS预测的公平性

- 多源时序预测的公平性
  - 模型是否对各个**变量**的预测效果都比较好?
- 原因分析
  - **不同样本的特点不同，模型**在训练过程中可能**被某些样本的特点主导**

Methods	StemGNN	Informer	TS2Vec	MTGNN	AGCRN
VAR	97.0574	46.9595	119.3862	144.0949	122.7769



1、不同的变量，其序列**pattern**可能存在很大的差异

2、大部分序列都是平稳的，主导了模型的训练过程

解决方案

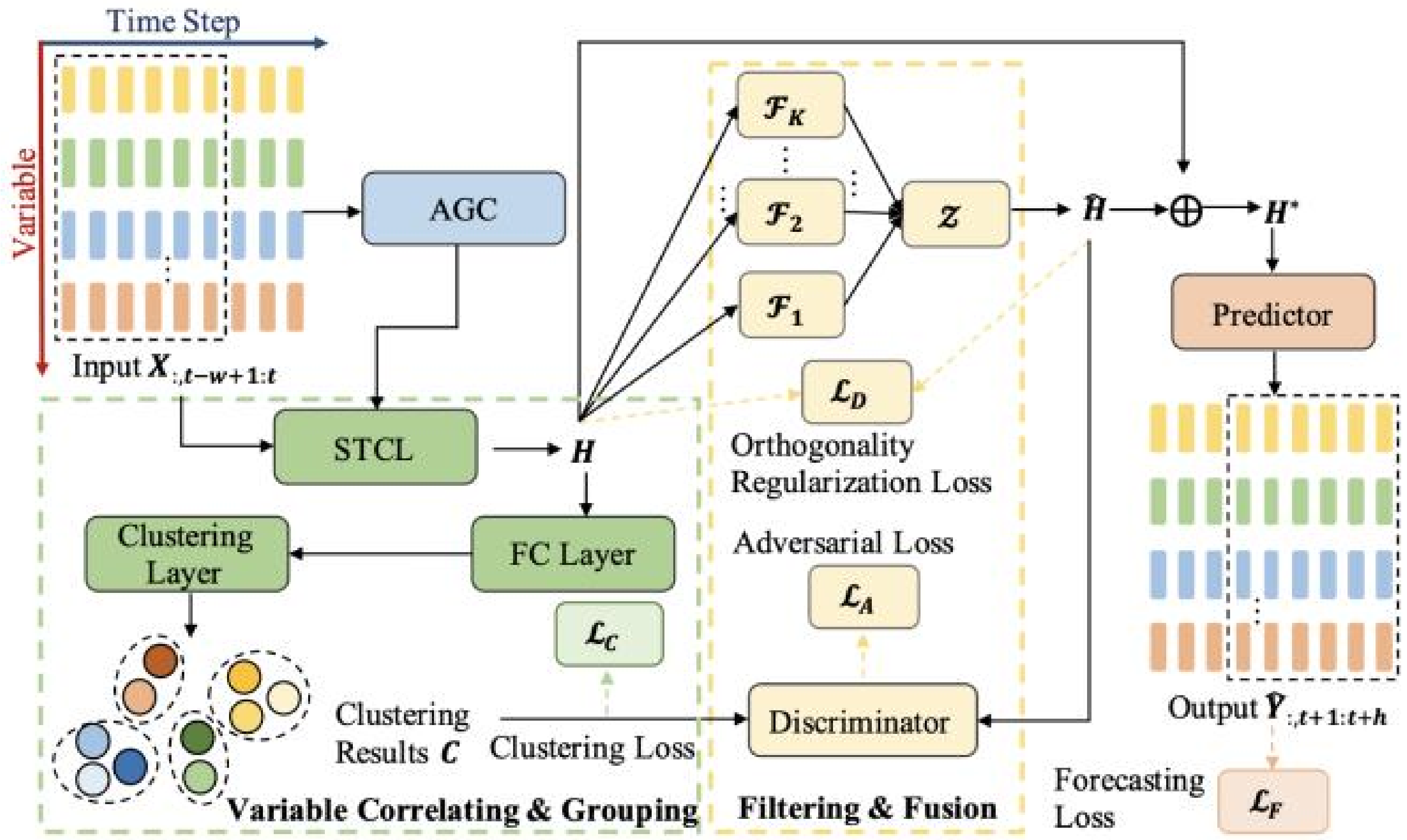
解决思路

- 原因分析
  - **不同样本的特点不同，模型**在训练过程中可能**被某些样本的特点主导**
- 解决思路
  - 将序列间的**共性与特性**分解开来，独立建模

Under the sliding forecasting setting with a fixed window size of $w \in \mathbb{N}^+$ and a sliding step of 1, we have the input, i.e., the observed values of all the N variables in w successive steps till the t^{th} time step, $\boldsymbol{X}_{:,t-w+1:t} = \{\boldsymbol{X}_{:,t-w+1}, \boldsymbol{X}_{:,t-w+2}, ..., \boldsymbol{X}_{:,t}\} \in \mathbb{R}^{w \times N}$. We articulate the research problem of MTS forecasting on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{M})$ to emphasize the spatio-temporal correlations simultaneously. The set of nodes $\mathcal{V}$ denotes input series $\boldsymbol{X}_{:,t-w+1:t}$, where $|\mathcal{V}| = N$ and each series/variable $\boldsymbol{X}_{i,:}$ corresponds to a node. $\mathcal{E}$ represents the set of edges, and $\boldsymbol{M} \in \mathbb{R}^{N \times N}$ is defined as the adjacent matrix. In addition, we evaluate the forecasting unfairness with the variance in forecasting errors of different variables, where the larger the variance is, the more unfair the forecasts are. Accordingly, the target is to accurately forecast based on the graph $\mathcal{G}$ the future sequence of $h \in \mathbb{N}^+$ steps $\boldsymbol{Y}_{:,t+1:t+h} = \{\boldsymbol{Y}_{:,t+1}, \boldsymbol{Y}_{:,t+2}, ..., \boldsymbol{Y}_{:,t+h}\}$ successive to the t^{th} time step through one forward procedure and guarantee a small forecasting error variance simultaneously.

模型框架

- 利用聚类的方法将多变量序列分组，得到不同组的表示
- 使用分解学习和对抗学习，从原始表示中剥离掉各个组特有的信息，得到公共的信息
- 通过前两步实现**公共信息**和**特有信息**的剥离，再基于这两个部分信息进行最终预测

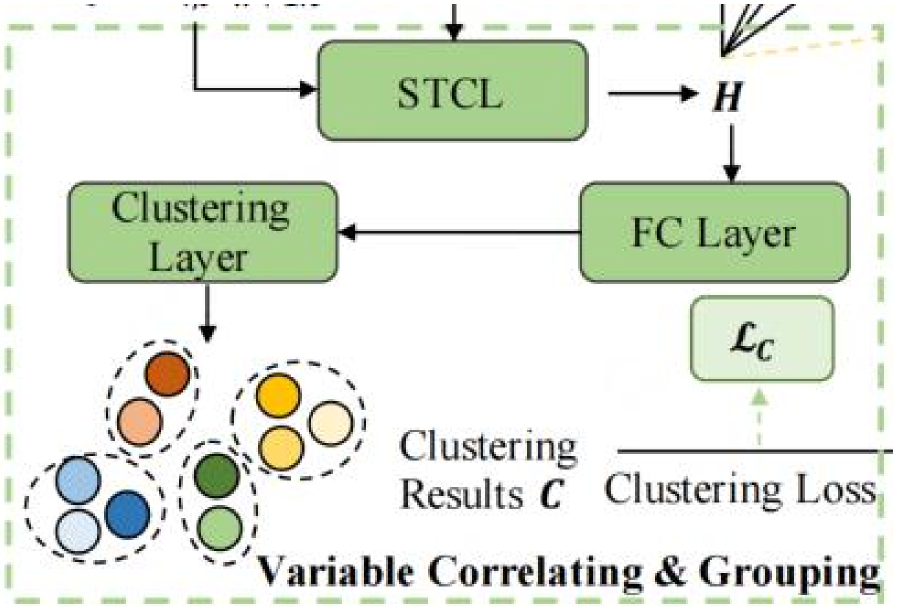


模型细节1：序列聚类

- K-Means聚类的优化目标为：
  - 最小化样本到样本中心的距离和
  - 其可以通过Spectral relaxation转化为以下损失函数[1]

$$\mathcal{L}_C = Tr(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}) - Tr(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{F}) \quad s.t. \boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} = \boldsymbol{I}$$

  - 其中，F是H的前k个奇异向量
- 前提：H是固定的
- EM迭代优化：
  - 固定H，SVD求F
  - 固定F，训练几轮H



[1] H. Zha et al. Spectral Relaxation for K-means Clustering

模型细节2：分解学习与对抗学习

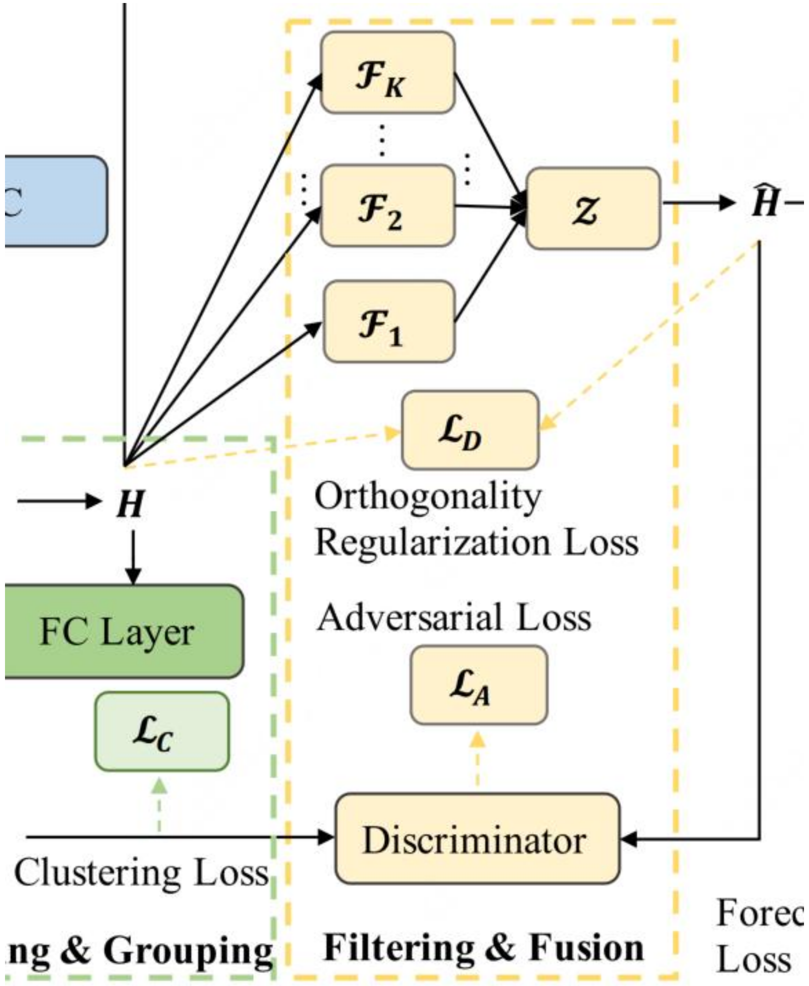
- 分解学习：过滤掉group相关的信息

$$\widehat{\boldsymbol{H}} = \mathcal{Z}(\mathcal{F}_1(\boldsymbol{H}), \mathcal{F}_2(\boldsymbol{H}), ...\mathcal{F}_K(\boldsymbol{H}))$$

  - F_k 是第k个group对应的过滤器（FC实现）
  - Z是合并函数，如取平均

- 对抗学习： $\mathcal{L}_A = \frac{1}{N}\sum_{i=0}^{N-1}\|\widehat{\boldsymbol{H}}_i - \boldsymbol{C}_i\|_2^2$ F、Z函数

- 正则项：推远group相关和group无关表征的距 $\mathcal{L}_D = \frac{1}{N}\sum_{i=0}^{N-1}\left|\frac{\widehat{\boldsymbol{H}}_i \cdot \boldsymbol{H}_i}{\|\widehat{\boldsymbol{H}}_i\| \cdot \|\boldsymbol{H}_i\|}\right|$



模型细节3：序列表示学习

- 自适应图构造**AGC**：

$$\widetilde{\boldsymbol{M}} = \delta(ReLU(\boldsymbol{E_M} \cdot \boldsymbol{E_M^{\mathrm{T}}}))$$

- 时空关系学习**STCL**：
  - 在**GRU**基础上，每个单元的计算引入了**GCN**模块
  - 类似于**DCRNN**（**Li et al,**

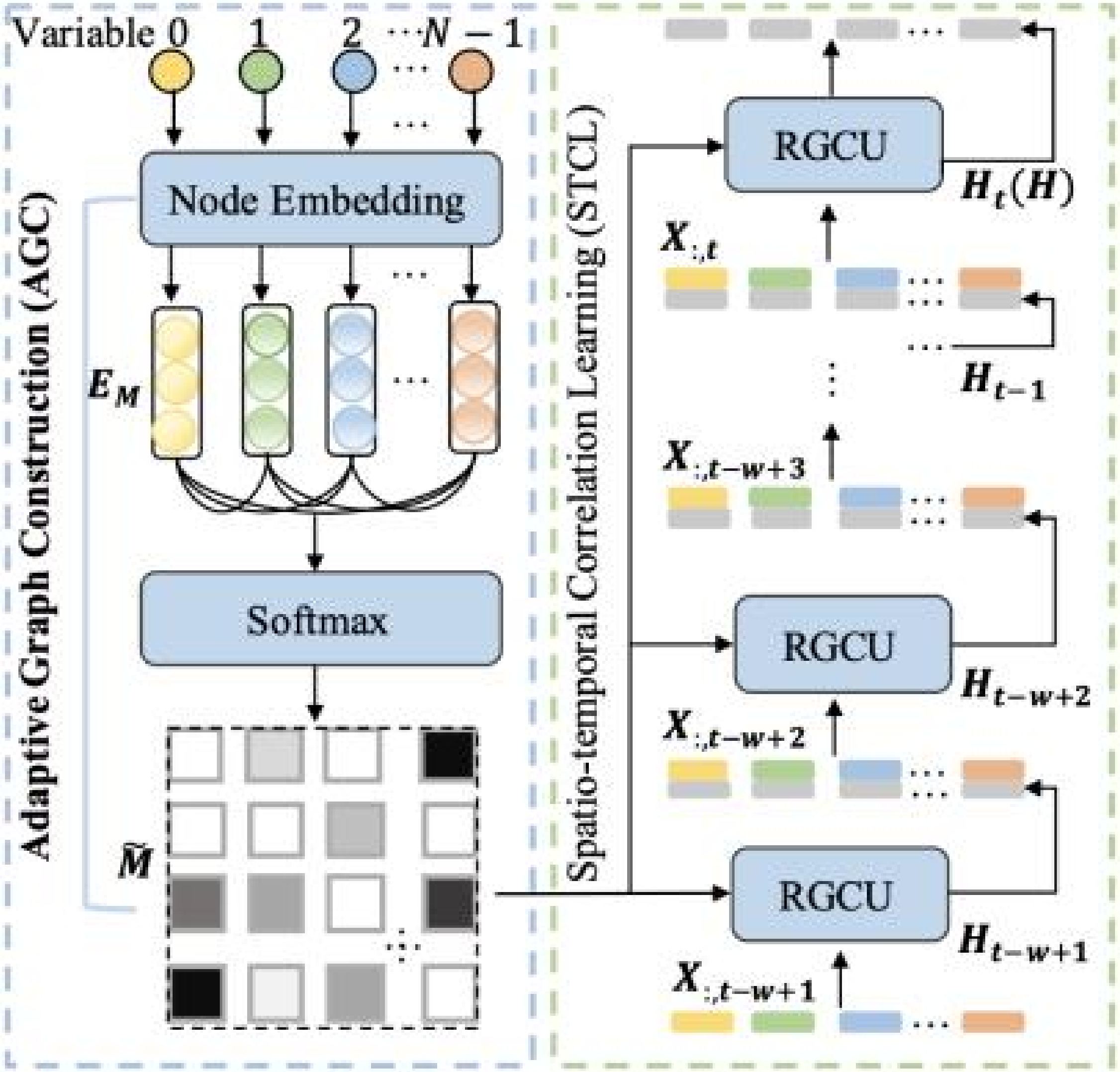
$$\widetilde{\boldsymbol{M}} = \delta(ReLU(\boldsymbol{E_M} \cdot \boldsymbol{E_M^{\mathrm{T}}}))$$

$$\boldsymbol{r}_t = \sigma(\widetilde{\boldsymbol{M}}[\boldsymbol{X}_{:,t}||\boldsymbol{H}_{t-1}]\boldsymbol{W_r} + \boldsymbol{b_r})$$

$$\boldsymbol{u}_t = \sigma(\widetilde{\boldsymbol{M}}[\boldsymbol{X}_{:,t}||\boldsymbol{H}_{t-1}]\boldsymbol{W_u} + \boldsymbol{b_u})$$

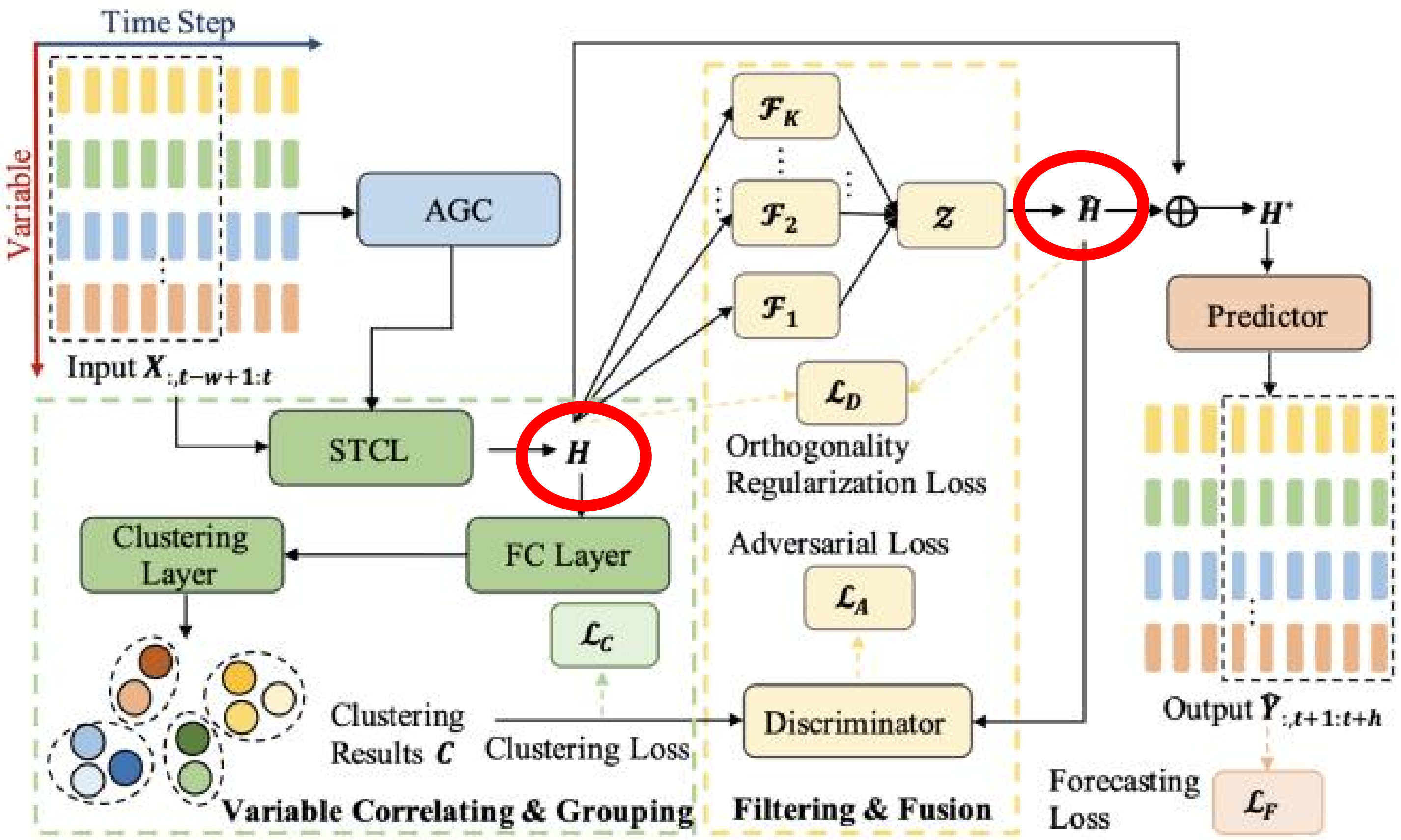
$$\boldsymbol{c}_t = \theta(\widetilde{\boldsymbol{M}}[\boldsymbol{X}_{:,t}||\boldsymbol{r} \odot \boldsymbol{H}_{t-1}]\boldsymbol{W_c} + \boldsymbol{b_c})$$

$$\boldsymbol{H}_t = \boldsymbol{u} \odot \boldsymbol{H}_{t-1} + (1 - \boldsymbol{u}) \odot \boldsymbol{c}_t$$



模型细节4：最终预测

- 联合group相关和group无关的表征做预测



模型训练

Algorithm 1 Adversarial Training for *FairFor*

Require: Hidden state $\boldsymbol{H}$, group results $\boldsymbol{C}$, network $\mathcal{R}$, filter and fusion functions $\mathcal{F}, \mathcal{Z} \subseteq \mathcal{R}$, discriminator $\mathcal{D}$ and cluster number K

1: Orthogonal initialize cluster indicator matrix $\boldsymbol{F}$
2: **for** each training iteration i **do**
3: $\widehat{\boldsymbol{H}} \leftarrow \mathcal{Z}(\{\mathcal{F}_k\}_{k\in[K]}(\boldsymbol{H}))$ by Eq. (10)
4: $\boldsymbol{H}^\star \leftarrow \boldsymbol{H} + \widehat{\boldsymbol{H}}$
5: Optimize $\mathcal{L}$ w.r.t $\boldsymbol{H}, \widehat{\boldsymbol{H}}, \boldsymbol{H}^\star, \boldsymbol{C}, \mathcal{R}$ with $\mathcal{D}$ being fixed
6: **if** $i\%3 == 0$ **then**
7: Update $\boldsymbol{F}$ by computing K-truncated SVD of $\boldsymbol{H}$
8: **end if**
9: $\mathcal{L}_A \leftarrow \mathbb{E}[||\widehat{\boldsymbol{H}} - \boldsymbol{C}||_2^2]$ by Eq. (11)
10: Optimize $\mathcal{L}_A$ w.r.t $\widehat{\boldsymbol{H}}, \boldsymbol{C}, \mathcal{D}$ with $\mathcal{R}, \boldsymbol{H}, \boldsymbol{H}^\star$ fixed
11: **end for**

实验结果

数据集与Baseline

Tasks	#Time Step	#Variable	Interval	Start Time
Traffic	10,392	963	1hour	1/1/2008
PeMSD7(M)	11,232	228	5min	7/1/2016
Solar-Energy	52,560	137	10min	1/1/2016
ECG5000	5,000	140	–	–

PeMSD7(M) [1] records the traffic flow data of the detectors in California. It includes 228 variables and 11,232 time steps at a 5-minute interval;

Solar-Energy [2] is collected from National Renewable Energy Laboratory (NREL) and records the solar power production in 2006. It includes 137 variables and 52,560 time steps at a 10-minute interval;

Traffic [3] is originally collected from the California Department of Transportation and describes the occupancy rate of different lanes in San Francisco highway. It contains 963 variables and 10,560 time steps with a 10-minute interval, where each observation is between 0 and 1;

ECG5000 [4] records 5,000 heartbeats randomly selected from a 20-hour long ECG downloaded from Physionet. It contains 140 variables and 5,000 time steps.

LSTNet [7]: integrates RNNs and CNNs to capture the short- and long-term temporal patterns respectively;

TPA-LSTM] [8]: embeds temporal pattern attention into RNNs to discover both relevant time series and time steps;

TS2Vec [54]: is a dilated CNN based universal framework to capture multi-scale contextual information in MTSs;

Informer [9]: is a Transformer-based model with *ProbSparse* self-attention mechanism and generative style decoder to predict long time series at one forward step;

Pyraformer [10]: is a Transformer-based model to simultaneously extract multiple ranges of temporal dependencies via the compact multi-resolution operation;

MTGNN [46]: combines graph learning, graph convolution and temporal convolution together to learn the spatial-temporal correlations without pre-defined graph structure;

StemGNN [12]: uses a GCN-based spectral network that can capture inter-series dependencies and temporal correlations jointly in the spectral domain;

AGCRN [4]: employs GCNs embedded with an adaptive node-specific pattern learning module to capture fine-grained inter-series relationships and uses RNNs to capture temporal patterns.

公平性结果

Dataset	Metric	LSTNet	TPA-LSTM	TS2Vec	Informer	Pyraformer	MTGNN	StemGNN	AGCRN	***FairFor***
PeMSD7(M)	VAR	23.6003	42.2336	28.4647	28.1218	48.4494	42.7716	26.5534	22.7220	**20.3152**
Solar-Energy	VAR	5.6880	10.1448	5.2330	5.2096	10.6149	9.7419	5.7216	5.6543	**5.1188**
Traffic	VAR	9.51e-4	8.07e-4	5.30e-4	4.27e-4	7.63e-4	6.43e-4	6.71e-4	5.59e-4	**4.17e-4**
ECG5000	VAR	0.5668	0.2577	0.2360	0.2257	0.2618	0.2255	0.1774	0.2467	**0.1732**

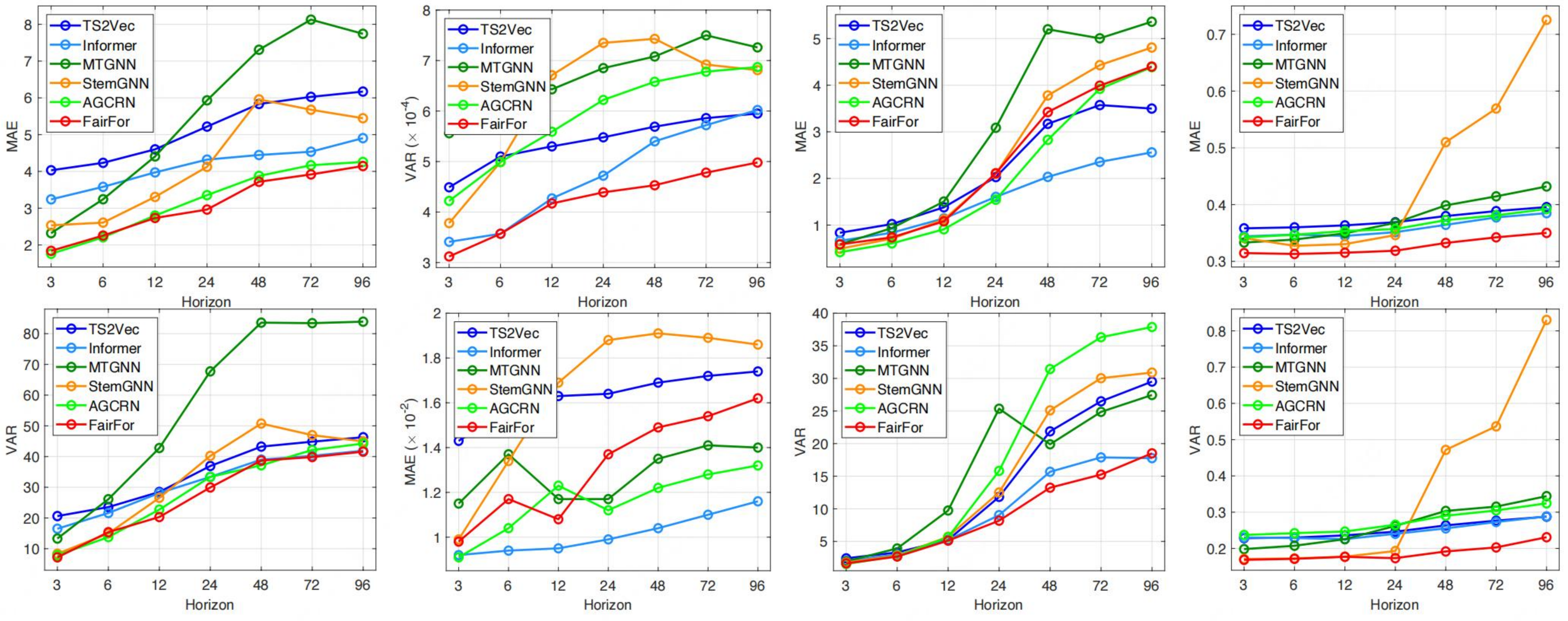
所有数据集预测公平性均有提升：**Improvement: 10.59%/1.74%/2.34%/2.37%**

预测精度结果

- 公平性可以提升准确率（一定程度上）

Dataset	Metric	LSTNet	TPA-LSTM	TS2Vec	Informer	Pyraformer	MTGNN	StemGNN	AGCRN	*FairFor*
PeMSD7(M)	MAE	3.2004	4.7573	4.6001	3.9728	4.2925	4.4072	3.3064	2.7983	**2.7336**
	RMSE	5.6566	7.8292	7.0445	6.6654	8.1646	7.5602	5.7927	5.4950	**5.2798**
	MAPE	0.0870	0.1932	0.1117	0.0957	0.1067	0.1127	0.0817	0.0694	**0.0678**
Solar-Energy	MAE	1.3030	2.0399	1.3810	1.1426	2.2316	1.5043	1.1205	**0.9090**	1.0807
	RMSE	2.7057	3.1682	2.6721	2.5702	3.9558	2.7029	2.6720	**2.5457**	2.5608
	MAPE	3.4220	3.4175	3.3694	3.3826	3.4080	3.3727	3.4096	**3.2905**	3.3584
Traffic	MAE	0.0220	0.0158	0.0163	**0.0095**	0.0445	0.0117	0.0169	0.0123	0.0108
	RMSE	0.0374	0.0317	0.0282	**0.0215**	0.0523	0.0279	0.0310	0.0260	0.0240
	MAPE	0.7736	0.4924	0.5630	**0.2482**	2.2674	0.3429	0.4768	0.2676	0.2669
ECG5000	MAE	0.4978	0.4268	0.3634	0.3448	0.3287	0.3493	0.3303	0.3534	**0.3152**
	RMSE	0.9022	0.6874	0.6067	0.5921	0.6110	0.5769	0.5352	0.6096	**0.5207**
	MAPE	0.9334	0.9561	0.9664	0.9481	1.0132	0.951	0.9176	0.9854	**0.8931**

预测精度结果



(a) PeMSD7(M) (b) Solar-Energy (c) Traffic (d) ECG5000

消融实验

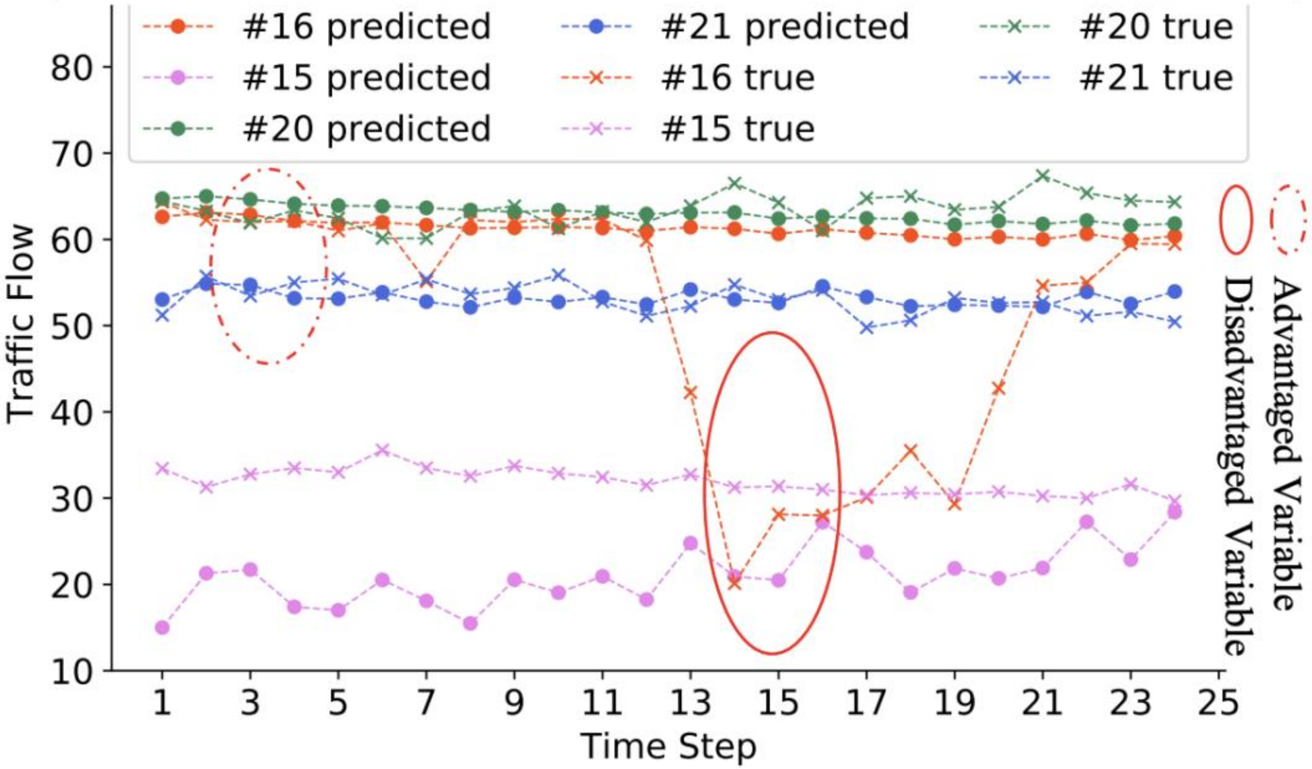
- 准确率是讨论公平性的基本，没有准确率讨论公平性的意义不大！

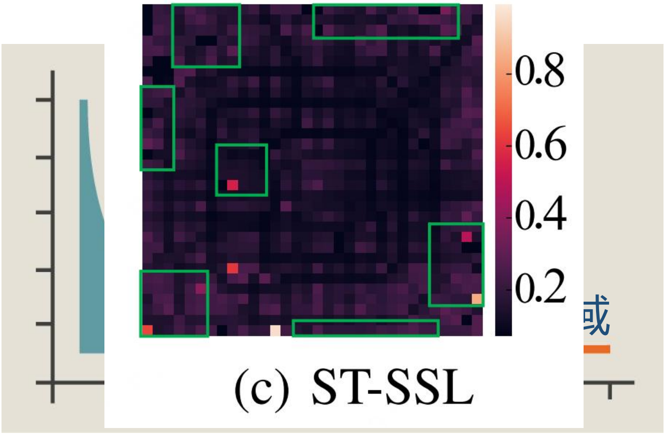
Models	PeMSD7(M)			Solar-Energy			Traffic			ECG5000		
	MAE	MAPE	VAR	MAE	MAPE	VAR	MAE	MAPE	VAR	MAE	MAPE	VAR
w/o STCL+AGC	4.2412	0.1018	30.2543	1.5797	3.4438	6.1499	0.0164	0.4758	5.93e-4	0.4500	0.9024	1.0522
w/o AL$\mathcal{D}$	2.8889	0.0689	20.0300	1.3456	3.4157	5.3100	0.0142	0.3541	4.38e-4	0.3159	0.8982	0.1744
w/o CL	2.8522	0.0689	20.3052	1.2181	3.3994	5.2194	0.0145	0.3471	4.47e-4	0.3161	0.8976	0.1749
w/o ORL	2.8497	0.0690	20.3152	1.1736	3.3865	5.2112	0.0138	0.3784	4.56e-4	0.3159	0.8960	0.1747
FairFor	**2.7983**	**0.0678**	**19.9953**	**1.0807**	**3.3826**	**5.1188**	**0.0123**	**0.3429**	**4.27e-4**	**0.3152**	**0.8931**	**0.1732**

一点思考

- 公平性(Fairness)与鲁棒性(Robustness)的区别与联系

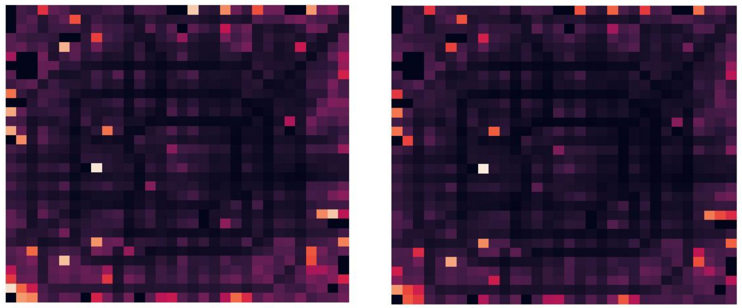
Methods	StemGNN	Informer	TS2Vec	MTGNN	AGCRN
VAR	97.0574	46.9595	119.3862	144.0949	122.7769





(c) ST-SSL

长尾分布导致模型偏向热门区域


(a) ST-ResNet (b) AGCRN

[1] J. Ji et al. Spatio-Temporal Self-Supervised Learning for Traffic Forecasting

Take-home

- 时空+公平性研究是一片蓝海，大家快来一起探索!
- 公平性可以一定程度上提升准确率
- 准确率是讨论公平性的基本
- 通过拆分不变和变化的表征可以提升公平性

感谢各位！