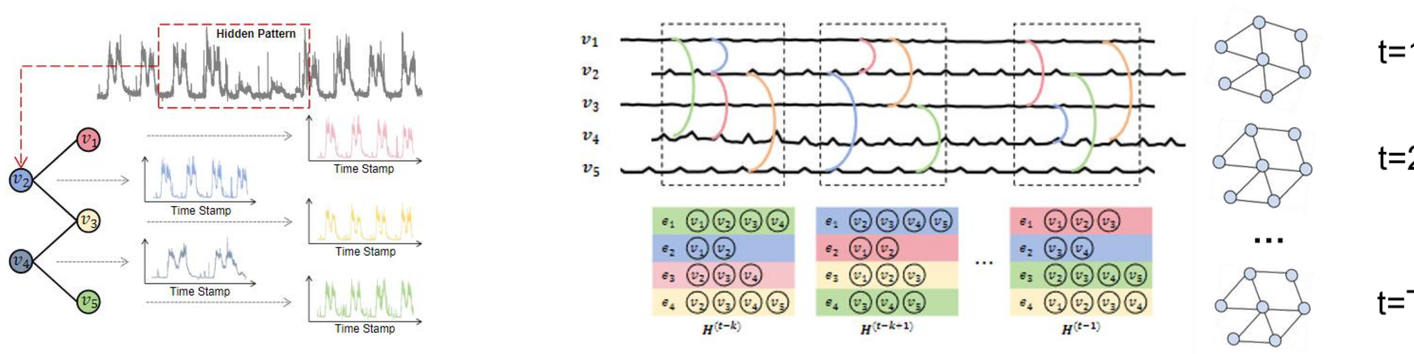
Decoupling Spatio-Temporal Prediction: When Lightweight Large Models Meet Adaptive Hypergraphs

Jiawen Chen, Qi Shao, Duxin Chen, Wenwu Yu

Background and Chanlleges



Classic LLMs fail to learn temporal and spatial feature information simultaneously

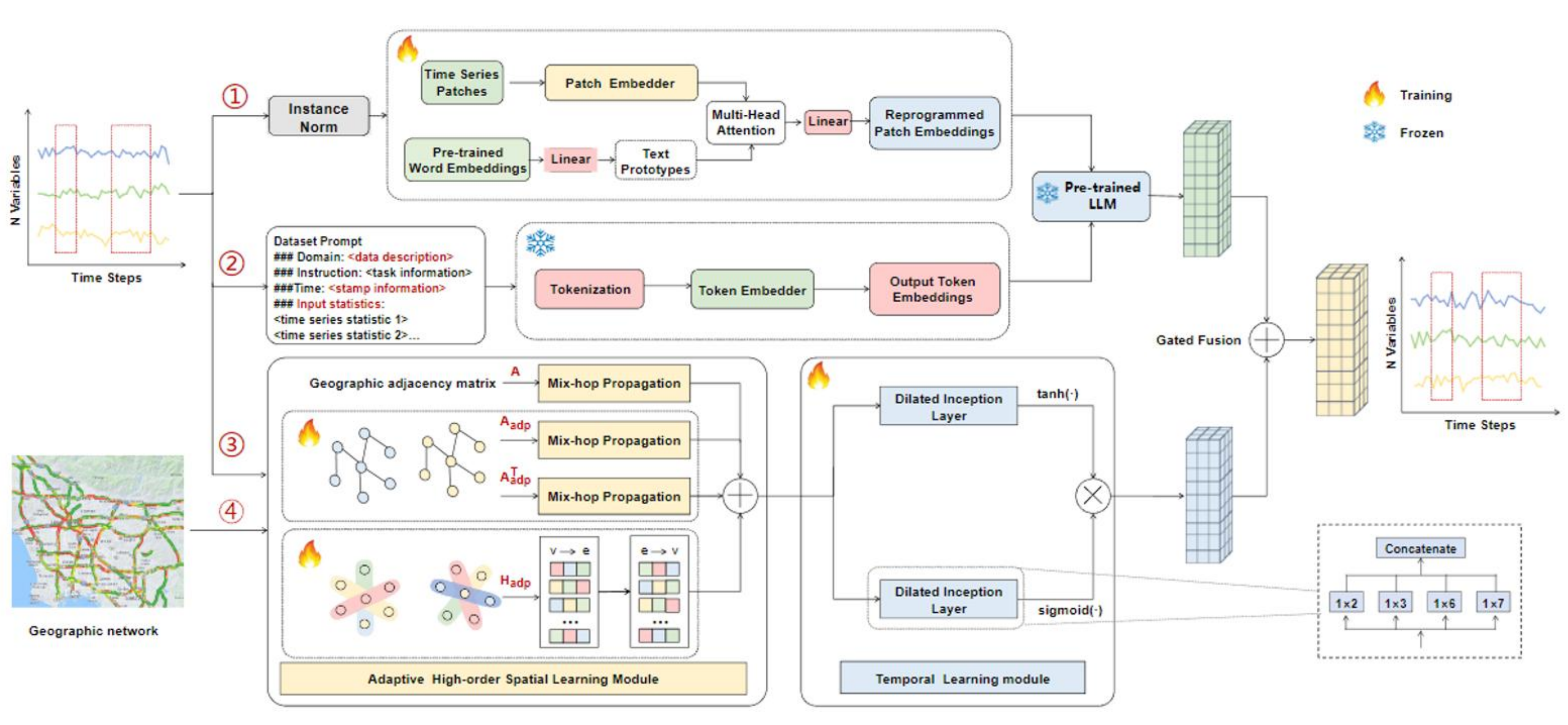
- Focus on temporal patterns, neglecting spatial dependencies.
- High computational cost for spatio structures modeling.

Classical spatio-temporal models fail to learning latent or high-order spatial coupling interactions

- Static graphs hard to model dynamic spatial changes.
- Only pairwise spatial relations are typically considered.

Method

Spatio-Temporal Modeling Prediction with LLMs



Framework of SHT-SepNet

Contributions

- Lightweight spatio-temporal separation framework (STH-SepNet) fuses textual cues with high-order spatial dependencies .
- Adaptive hypergraph structure for spatial modeling that dynamically captures intricate dependencies via effective order modeling.
- State-of-the-art performance and single-GPU efficiency by experiments, confirming the framework's readiness for real-world deployment.

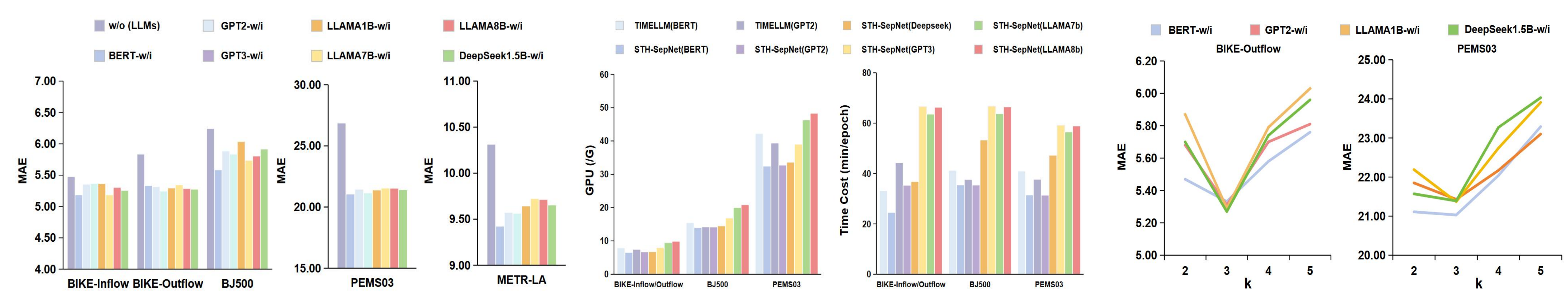
Results and Conclusion

Model	BIKE-Inflow		BIKE-Outflow		PEMS03		BJ500		METR-LA	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Autoformer (NIPS, 2021)[30]	7.01	17.52	7.19	17.75	44.87	70.84	10.79	16.06	12.47	20.04
Informer (AAAI, 2021)[34]	8.25	20.37	9.21	21.50	33.72	52.15	7.58	11.96	14.50	20.35
FEDformer (PMLR, 2022)[29]	6.28	16.30	6.56	16.67	35.00	50.84	10.77	15.99	12.35	18.79
DLinear (AAAI, 2023) [23]	5.71	15.49	5.82	15.36	45.30	66.81	8.55	13.49	10.90	17.31
TimesNet (ICLR, 2023)[35]	5.54	15.41	5.56	15.18	37.54	62.99	8.67	13.96	10.22	18.29
PatchTST (ICLR, 2023)[36]	5.53	15.39	5.63	15.23	48.42	78.24	8.79	14.28	10.13	18.27
iTransformer (ICLR, 2024)[26]	6.05	16.39	6.15	16.69	43.63	70.61	9.01	14.32	10.15	18.36
TIMELLM (ICLR, 2024)[37]	6.81	16.72	6.93	16.30	32.62	49.77	7.25	11.58	12.36	18.53
AdaMSHyper (NIPS, 2024)[38]	6.72	16.91	7.04	17.14	33.49	50.37	7.41	11.60	12.51	18.60
AGCRN (NIPS, 2020)[39]	6.64	16.14	6.77	16.36	33.14	54.88	6.32	12.81	11.39	23.15
ASTGCN (AAAI, 2019)[40]	6.66	15.87	6.26	14.48	30.65	53.96	6.34	11.34	10.54	22.76
MSTGCN (TNSRE, 2021)[41]	5.91	14.11	6.04	14.24	29.57	47.97	5.62	11.15	10.17	20.24
MTGNN (KDD, 2020)[1]	6.16	14.80	5.93	13.93	29.04	50.32	5.86	10.91	9.98	21.23
STGODE (KDD, 2021)[42]	6.77	15.93	6.82	15.50	33.39	54.16	6.44	12.14	11.48	22.85
STSGCN (AAAI, 2020)[14]	6.73	15.89	6.58	15.36	34.23	58.07	6.40	12.03	11.07	22.79
STGCN (IJCAI, 2018)[43]	7.08	15.72	7.36	16.11	36.02	53.44	6.73	12.62	12.38	22.55
GMAN (AAAI, 2020)[44]	6.73	15.60	6.94	15.84	33.96	53.02	6.41	12.40	11.69	22.37
STAEformer (CIKM, 2024)[27]	5.97	14.57	6.17	14.70	29.62	48.03	5.79	10.42	9.91	21.17
STD-MAE (IJCAI, 2024) [45]	6.13	14.87	6.21	14.37	30.40	48.38	5.92	11.49	10.52	23.11
STH-SepNet (Ours)	**5.18**	**14.40**	**5.33**	**14.23**	**21.03**	**34.17**	**5.58**	**9.77**	**9.42**	**16.41**

Model		BIKE-Outflow		PEMS03	
		MAE	RMSE	MAE	RMSE
BERT	TIMELLM	6.74	16.13	32.68	50.39
	STH-SepNet-Static	6.34	16.41	29.53	48.94
	STH-SepNet-GNN	5.47	14.36	21.11	34.52
	STH-SepNet	**5.33**	**14.23**	**21.03**	**34.17**
GPT2	TIMELLM	6.93	16.30	32.63	49.83
	STH-SepNet-Static	6.53	16.61	30.03	49.07
	STH-SepNet-GNN	5.68	14.48	21.85	35.78
	STH-SepNet	**5.31**	**14.24**	**21.43**	**35.01**
GPT3	TIMELLM	7.01	16.63	34.85	51.89
	STH-SepNet-Static	6.79	16.71	30.64	49.16
	STH-SepNet-GNN	5.66	13.90	21.27	34.78
	STH-SepNet	**5.24**	**14.16**	**21.13**	**34.69**
LLAMA1B	TIMELLM	7.10	16.74	34.06	51.59
	STH-SepNet-Static	6.27	16.40	29.76	48.63
	STH-SepNet-GNN	5.87	14.80	22.19	35.87
	STH-SepNet	**5.29**	**14.20**	**21.37**	**34.92**
LLAMA7B	TIMELLM	6.95	16.41	34.17	52.47
	STH-SepNet-Static	6.73	16.87	30.24	49.92
	STH-SepNet-GNN	5.91	14.94	21.64	35.24
	STH-SepNet	**5.34**	**14.31**	**21.52**	**35.17**
LLAMA8B	TIMELLM	7.02	16.55	35.72	51.97
	STH-SepNet-Static	6.85	16.64	30.47	49.98
	STH-SepNet-GNN	5.70	14.27	21.57	35.23
	STH-SepNet	**5.28**	**14.20**	**21.51**	**35.19**
DeepSeek1.5B	TIMELLM	6.94	16.25	33.19	50.28
	STH-SepNet-Static	6.79	16.37	30.26	48.81
	STH-SepNet-GNN	5.74	14.55	21.73	35.47
	STH-SepNet	**5.27**	**14.19**	**21.39**	**34.96**

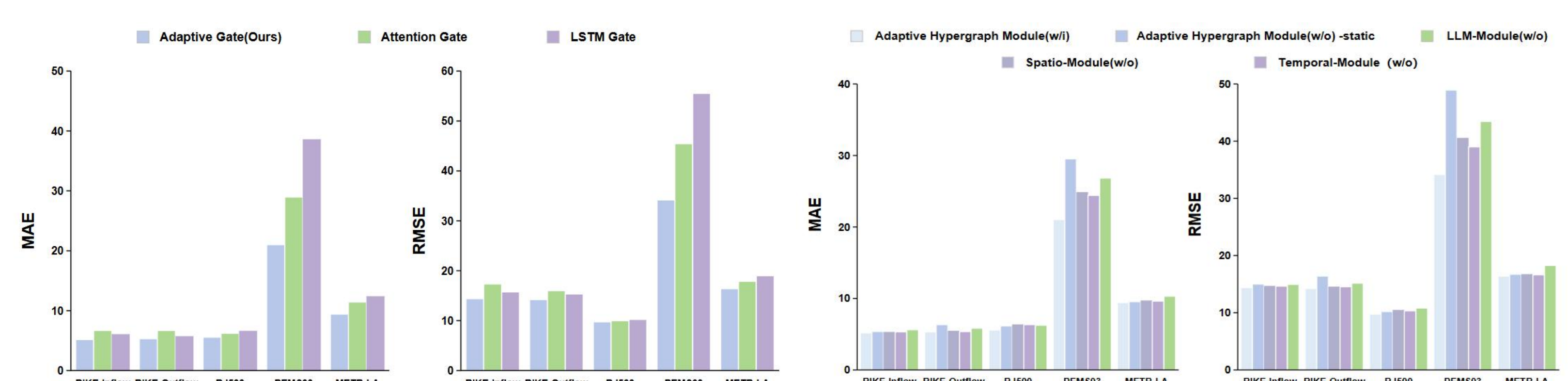
✓ STH-SepNet decreases errors by 28.8 % on drifting data and running 23.8 % better than LLM baselines

✓ Adaptive hypergraphs dynamically reweight to capture sudden spatial drift, outperform massive models with a lightweight backbone, and consistently excel across architectures via higher-order interactions.



✓ Lightweight LLMs (BERT,GPT2) achieve competitive performance while maintain stability and efficiency

✓ STH-SepNet (BERT) outperforms TIMELLM and larger STH-SepNet in GPU usage, training speed



✓ Adaptive gate fusion achieves superior MAE/RMSE

✓ Adaptive hypergraph module shows obvious improvements

Dataset		2	3	4	5
BIKE-Outflow	BERT	**13.36**	14.23	14.89	15.37
	GPT2	14.48	**14.24**	15.29	15.58
	LLAMA1B	14.80	**14.20**	15.54	16.19
	DeepSeek1.5B	14.55	**14.19**	15.46	16.05
PEMS03	BERT	34.52	**34.17**	35.81	37.84
	GPT2	35.78	**35.01**	36.22	37.74
	LLAMA1B	35.87	**34.92**	37.16	39.07
	DeepSeek1.5B	35.47	**34.96**	38.03	39.27

✓ Excessive parameter scaling is not essential for effective temporal modeling

✓ Effective order (k=3), STH-SepNet model achieves smallest MSE