Graph Neural Controlled Differential Equations for Traffic Forecasting

Jeongwhan Choi, Hwangyong Choi, Jeehyun Hwang, Noseong Park

Department of Artificial Intelligence, Yonsei University

{jeongwhan.choi, hwangyong753, hwanggh96, noseong}@yonsei.ac.kr

Overview

- We present the method of spatio-temporal graph neural controlled differential equation (**STG-NCDE**).
- We design two NCDEs for learning the temporal and spatial dependencies of traffic conditions and combine them into a single framework.
- **STG-NCDE** is robust to the irregularity of the temporal sequence.
- Our large-scale experiments with 6 datasets and 20 baselines clearly show the efficacy of the proposed method.

Spatio-temporal Data

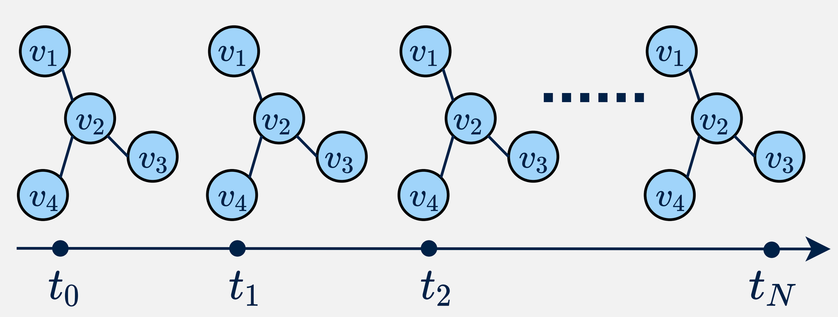

Figure 1: Example of spatio-temporal graph data

- The spatio-temporal graph data frequently happens in real-world applications.
- For instance, the traffic forecasting task is one of the most popular problems in the area of spatio-temporal processing.

Neural Controlled Differential Equatnios (NCDEs)

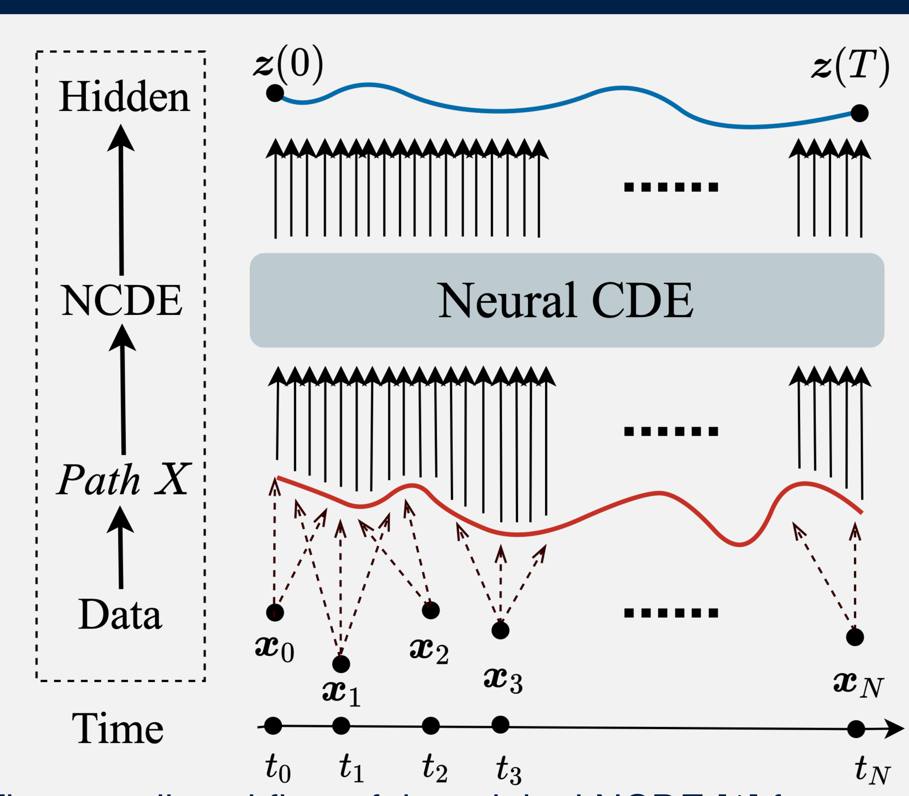

Figure 2: The overall workflow of the original NCDE [1] for processing time-series.

- NCDEs generalize RNNs in a continuous manner and can be written as follows:

$$z(T) = z(0) + \int_0^T f(z(t); \theta_f) dX(t) \tag{1}$$

$$= z(0) + \int_0^T f(z(t); \theta_f) \frac{dX(t)}{dt} dt, \tag{2}$$

- It has not been studied yet how to combine the NCDE technology and the graph convolutional processing technology.
- We integrate them into a single framework to solve the spatio-temporal forecasting.

Overall Design

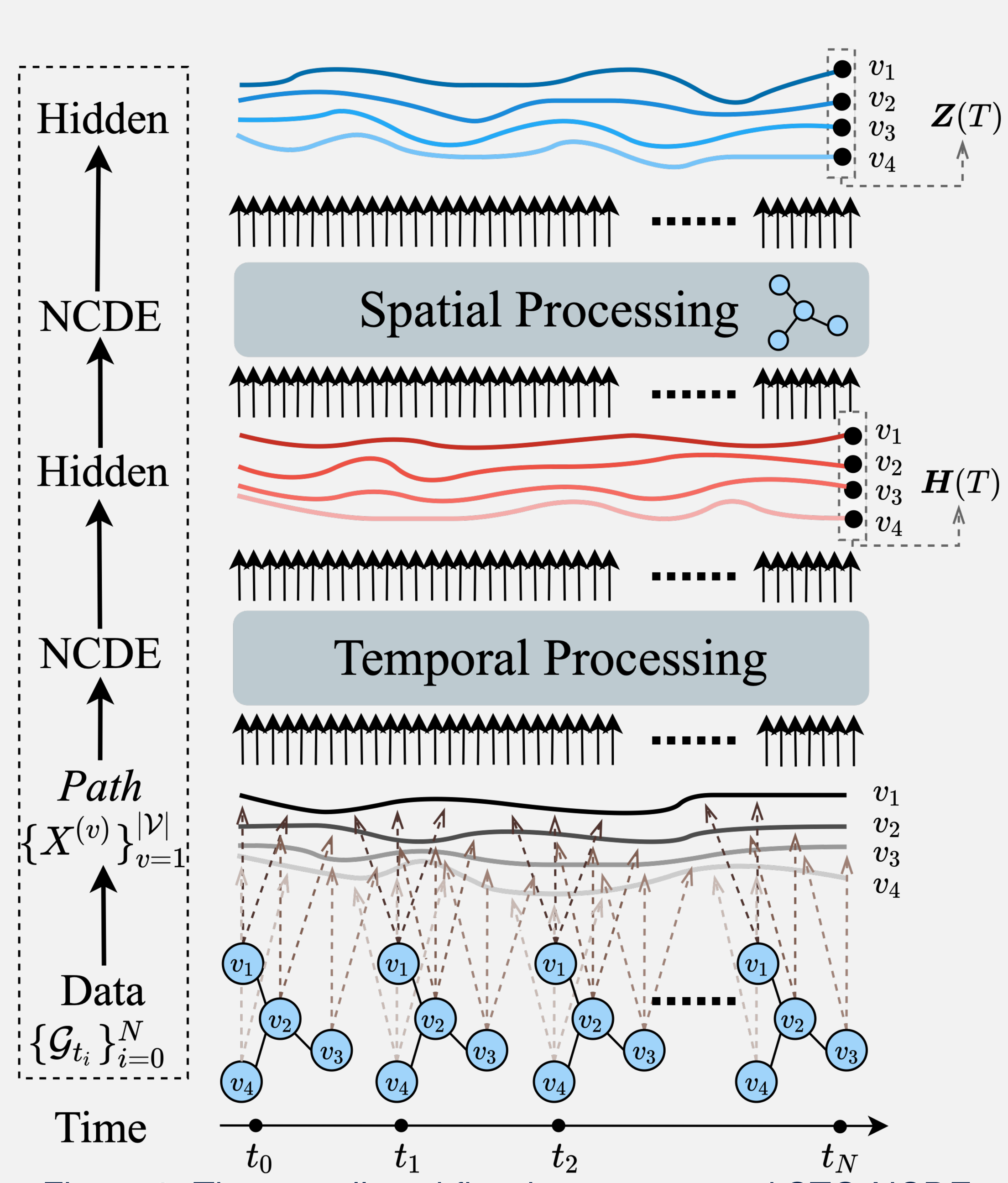

Figure 3: The overall workflow in our proposed STG-NCDE.

- Our proposed **STG-NCDE** consists of two NCDEs.

Graph Neural Controlled Differential Equations

Temporal processing.

- The first NCDE for the temporal processing can be written as follows:

$$\boldsymbol{h}^{(v)}(T) = \boldsymbol{h}^{(v)}(0) + \int_0^T f(\boldsymbol{h}^{(v)}(t); \boldsymbol{\theta}_f) \frac{dX^{(v)}(t)}{dt} dt, \tag{3}$$

  where $\boldsymbol{h}^{(v)}(t)$ is a hidden trajectory (over time $t \in [0, T]$) of the temporal information of node v.
- Eq. (3) can be rewritten as follows using the matrix notation:

$$\boldsymbol{H}(T) = \boldsymbol{H}(0) + \int_0^T f(\boldsymbol{H}(t); \boldsymbol{\theta}_f) \frac{d\boldsymbol{X}(t)}{dt} dt, \tag{4}$$

  where $\boldsymbol{X}(t)$ is a matrix whose v-th row is $X^{(v)}$.

Spatial processing.

- The second NCDE starts for its spatial processing as follows:

$$\boldsymbol{Z}(T) = \boldsymbol{Z}(0) + \int_0^T g(\boldsymbol{Z}(t); \boldsymbol{\theta}_g) \frac{d\boldsymbol{H}(t)}{dt} dt, \tag{5}$$

  where the hidden trajectory $\boldsymbol{Z}(t)$ is controlled.
- After combining Eqs. (4) and (5), we have the following single equation which incorporates both the temporal and the spatial processing:

$$\boldsymbol{Z}(T) = \boldsymbol{Z}(0) + \int_0^T g(\boldsymbol{Z}(t); \boldsymbol{\theta}_g) f(\boldsymbol{H}(t); \boldsymbol{\theta}_f) \frac{d\boldsymbol{X}(t)}{dt} dt. \tag{6}$$

Experiments

Traffic forecasting on real-world datasets.

Table 1: Forecasting error on PeMSD3, PeMSD4, PeMSD7 and PeMSD8

Model	PeMSD3			PeMSD4			PeMSD7			PeMSD8		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
STGCN	17.55	30.42	17.34%	21.16	34.89	13.83%	25.33	39.34	11.21%	17.50	27.09	11.29%
DCRNN	17.99	30.31	18.34%	21.22	33.44	14.17%	25.22	38.61	11.82%	16.82	26.36	10.92%
GraphWaveNet	19.12	32.77	18.89%	24.89	39.66	17.29%	26.39	41.50	11.97%	18.28	30.05	12.15%
ASTGCN(r)	17.34	29.56	17.21%	22.93	35.22	16.56%	24.01	37.87	10.73%	18.25	28.06	11.64%
MSTGCN	19.54	31.93	23.86%	23.96	37.21	14.33%	29.00	43.73	14.30%	19.00	29.15	12.38%
STG2Seq	19.03	29.83	21.55%	25.20	38.48	18.77%	32.77	47.16	20.16%	20.17	30.71	17.32%
LSGCN	17.94	29.85	16.98%	21.53	33.86	13.18%	27.31	41.46	11.98%	17.73	26.76	11.20%
STSGCN	17.48	29.21	16.78%	21.19	33.65	13.90%	24.26	39.03	10.21%	17.13	26.80	10.96%
AGCRN	15.98	28.25	15.23%	19.83	32.26	12.97%	22.37	36.55	9.12%	15.95	25.22	10.09%
STFGNN	16.77	28.34	16.30%	20.48	32.51	16.77%	23.46	36.60	9.21%	16.94	26.25	10.60%
STGODE	16.50	27.84	16.69%	20.84	32.82	13.77%	22.59	37.54	10.14%	16.81	25.97	10.62%
Z-GCNETs	16.64	28.15	16.39%	19.50	31.61	12.78%	21.77	35.17	9.25%	15.76	25.11	10.01%
STG-NCDE	**15.57**	**27.09**	**15.06%**	**19.21**	**31.09**	**12.76%**	**20.53**	**33.84**	**8.80%**	**15.45**	**24.81**	**9.92%**

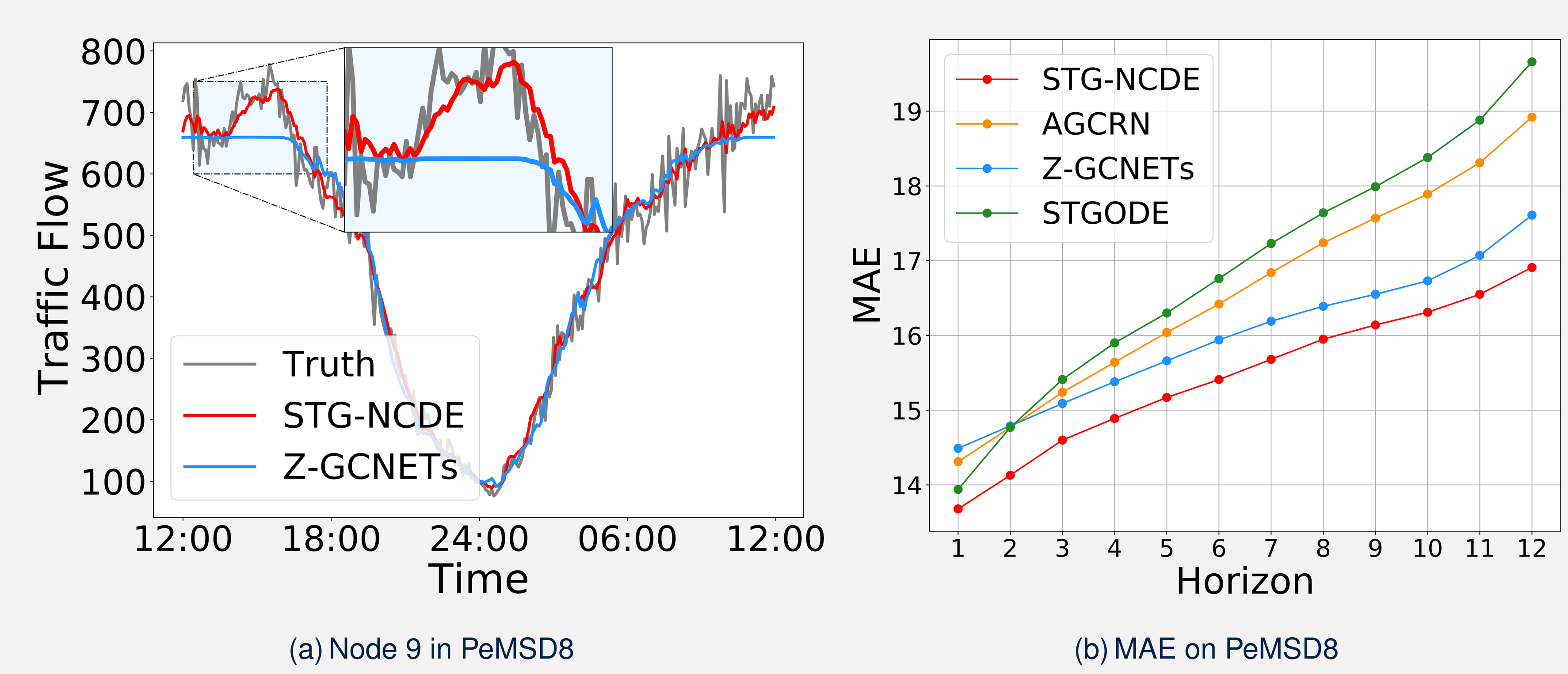

Figure 4: (a)Traffic forecasting visualization. (b)Prediction error at each horizon.

Irregular traffic forecasting.

Table 2: Forecasting error on irregular PeMSD4

Model	Missing rate	MAE	RMSE	MAPE
	10%	19.36	31.28	12.79%
STG-NCDE	30%	19.40	31.30	13.04%
	50%	19.98	32.09	13.48%

Conclusion

- We presented a **STG-NCDE** model to perform traffic forecasting.
- **STG-NCDE** shows the best accuracy in benchmark datasets and robustness in irregular settings.
- We believe that the combination of NCDEs and GCNs is a promising research direction for spatio-temporal processing.

References

[1] Patrick Kidger, James Morrill, James Foster, and Terry Lyons. Neural Controlled Differential Equations for Irregular Time Series. *NeurIPS*, 2020.