Predicting COVID-19 pandemic by spatio-temporal graph neural networks: A New Zealand's study

Bach Nguyen[1], Truong Son Hy[2], Long Tran-Thanh[3], and Nhung Nghiem[4]

[1] FPT Software AI Center, [2] Indiana State University, [3] University of Warwick, [4] University of Otago

Main contributions

- ATMGNN allows modeling of relationships between time series in evolving epidemic cases map
- Comprehensive analysis of forecasting models for epidemic prediction with detailed case study
- Complete pipeline from data gathering in New Zealand to preliminaries and prediction models
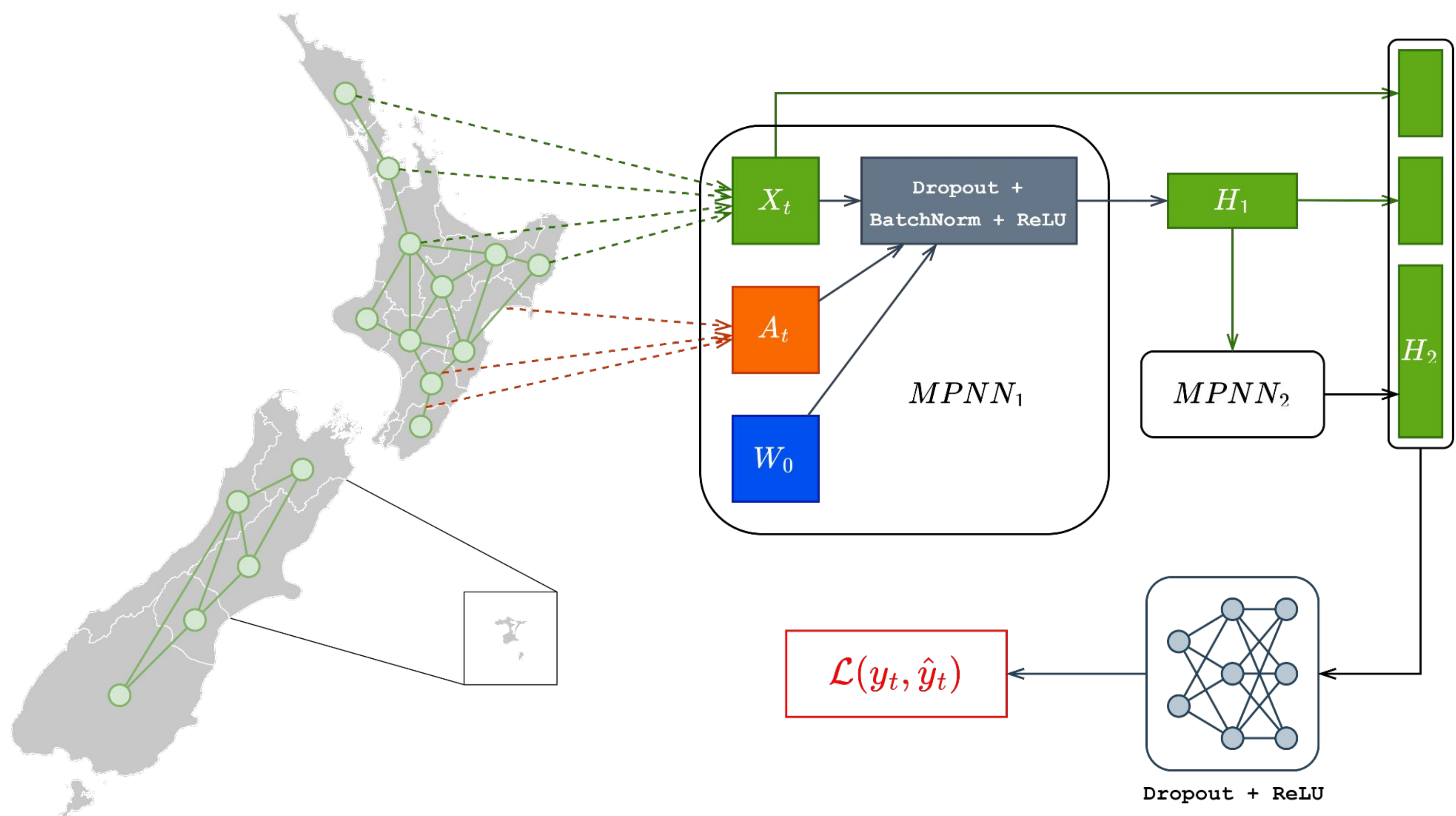
ATMGNN

Figure 1: **Attention Temporal Multiresolution Graph Neural Network** architecture on the graph representation of New Zealand. Note that dotted green arrows represent the extraction of historical case counts as node features, and dotted orange arrows represent the geospatial location between two regions extracted as edge features.

➜ **Neighborhood aggregation scheme in ATMGNN**

$H^{(k)} = (h_1^{(k)}, h_2^{(k)}, ..., h_n^{(k)})^\top$ denotes the matrix arrangement of all nodes in the disease graph, $\tilde{A}$ denotes normalized graph Laplacian.

$$H^{(k)} = \sigma(\tilde{A}H^{(k-1)}W^{(k)})$$

➜ **Time-series processing** Two solutions:
- LSTM-based series processing for low data count. Optimal for case with scarce disease data.
- Self-attention-based series processing for generics.

➜ **Multiresolution hierarchy learning on map**

Multiresolution Graph Neural Networks as graph variational autoencoder, implement learnable clustering mechanisms to construct a hierarchy of coarsening graphs. The <u>coarsened weighted adjacency</u> matrix $\tilde{A} \in \mathbb{N}^{k \times k}$ is defined as:

$$\tilde{A}_{ij} = \begin{cases} \frac{1}{2}\sum_{u,v \in V_i} A_{uv}, & \text{if } i = j, \\ \sum_{u \in V_i, v \in V_j} A_{uv}, & \text{if } i \neq j, \end{cases}$$

Experiments

- Equal/better performance than other spatio-temporal models
- Lower performance decay over long time windows
- ATMGNN has better out-of-distribution prediction performance than other highlighted models

Model	Next 3 Days			Next 7 Days			Next 14 Days			Next 21 Days		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
AVG	247.20	325.92	-3.22	258.95	340.95	-3.58	277.16	362.98	-4.22	292.23	379.98	-4.87
AVG_WINDOW	80.88	111.15	0.76	104.09	142.37	0.55	144.88	196.63	-0.02	176.82	238.39	-0.79
LAST_DAY	118.81	158.56	0.64	73.65	102.09	0.84	120.99	164.78	0.47	156.17	211.44	-0.08
LIN_REG	182.46	284.61	0.31	213.53	336.56	-0.01	272.77	440.60	-0.77	335.95	551.23	-1.81
GP_REG	331.43	471.17	-0.89	332.08	472.45	-0.98	325.55	464.20	-0.97	322.08	460.23	-0.96
RAND_FOREST	98.97	152.96	0.80	72.85	111.69	0.89	112.02	168.81	0.74	140.77	210.73	0.59
XGBOOST	109.68	165.36	0.77	**68.51**	**105.91**	**0.90**	108.45	165.17	0.75	137.45	208.13	0.60
PROPHET	119.32	642.78	-0.24	148.58	770.55	-1.50	222.01	526.58	-0.59	292.54	407.66	-0.17
ARIMA	132.49	534.26	0.14	155.44	523.57	-0.15	204.51	472.95	-0.28	239.06	423.17	-0.26
LSTM	186.86	242.62	-0.97	168.43	222.65	-0.39	140.69	192.39	0.38	128.04	182.35	0.59
MPNN	80.33	110.75	0.84	87.45	121.23	0.79	121.41	168.34	0.53	153.62	210.69	0.15
MGNN	80.87	111.67	0.83	89.77	124.56	0.74	125.30	172.46	0.46	156.25	213.55	0.06
MPNN+LSTM	**75.25**	**104.64**	**0.86**	85.14	117.92	0.84	**88.28**	**121.71**	**0.85**	99.85	137.74	**0.83**
ATMGNN	77.49	106.96	0.86	86.85	119.68	0.84	90.43	124.89	0.84	101.87	140.33	0.82

GitHub: https://github.com/HySonLab/pandemic_tgnn   TGL @ NeurIPS2023   TruongSon.Hy@indstate.edu