

Graph Neural Network for Time-series Forecasting

이종우
FRE Lab

'24.04.26

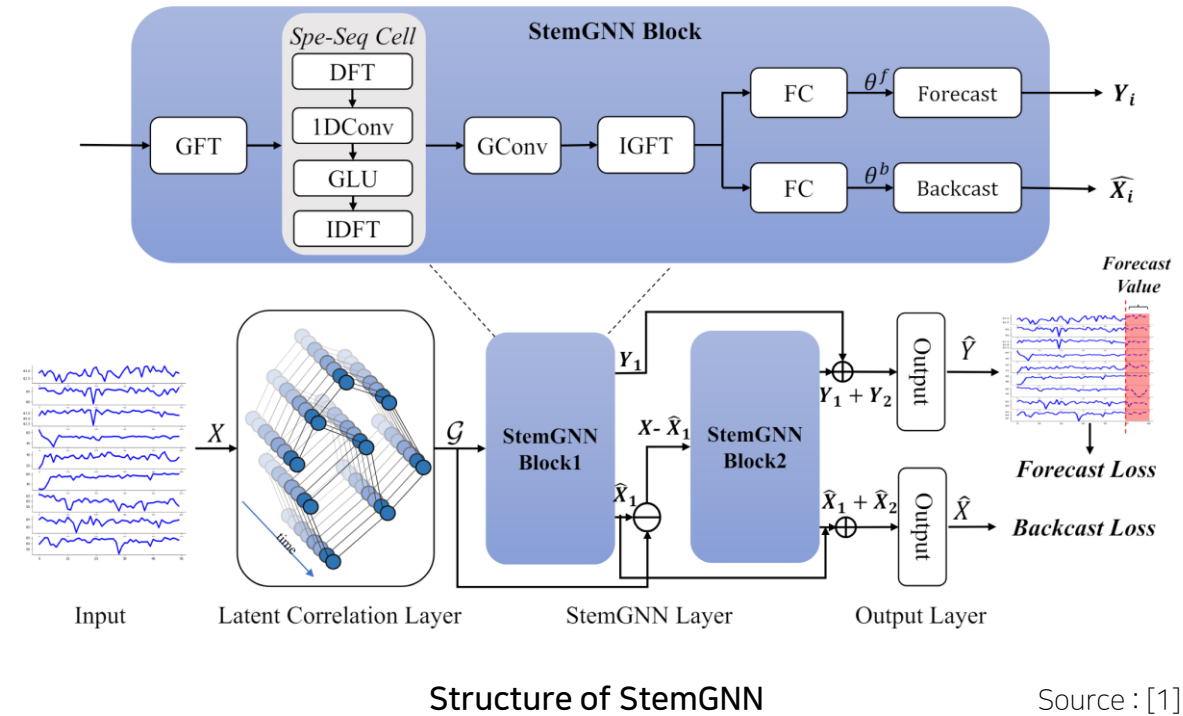
1. Spectral temporal graph neural network for multivariate time-series forecasting [1]
2. Multivariate time series forecasting with graph neural networks[3]
3. Conclusion

1-1. Spectral temporal graph neural network for multivariate time-series forecasting : Abstract

- Spectral Temporal Graph Neural Network(StemGNN)을 제시
- Spectral domain에서 inter-series correlation과 temporal dependency를 같이 파악

- 1) Inter-series correlation by Graph Fourier Transform
- 2) Temporal dependency by Discrete Fourier Transform

- Convolution과 sequential 모듈을 통해 spectral representation에서 pattern 파악 및 forecasting을 수행
- Inter-series correlation을 스스로 학습, pre-defined information이 필수적이지 않음



1-2. Process of StemGNN

- Latent correlation layer를 통해 Data로부터 graph structure와 weight matrix를 생성
 - 다수의 시계열 데이터로부터 latent correlation 학습을 위해 self-attention 방식 적용
- Graph Fourier Transform(GFT)를 통해 graph를 Spectral domain의 matrix 상태로 변환
- Discrete Fourier Transform(DFT)를 통해 univariate 시계열을 frequency domain으로 변환
- 1D convolution과 GLU layer를 통해 frequency domain에서 feature pattern 파악
- InverseDFT를 통해 frequency domain을 spectral domain으로 변환, graph convolution(GConv)를 수행
- InverseGFT를 통해 graph를 spectral domain에서 spatial, temporal domain으로 변환
- GLU와 FC sub layer로 구성된 output layer를 통해 예측값 반환
- Forecasting output Y_i 는 future value estimation을 위해서, backcasting output \hat{X}_i 는 representation power를 향상시키기 위해 auto-encoding과 궤를 같이하는 방식의 학습을 위해 사용됨

1-3. Problem Definition & Result

- Problem definition

- Multivariate temporal graph $\mathcal{G} = (X, W)$ 는 $X \in \mathbb{R}^{N \times T}$ (N은 시계열의 수(=node 수), T는 timestamp)와 adjacency matrix $W \in \mathbb{R}^{N \times N}$ 으로 구성. 앞의 K timestamp 데이터를 통해 t시점 이후의 H time stamp를 예측하는 Forecasting model F 를 구하고자 함 $\hat{X}_t, \hat{X}_{t+1}, \dots, \hat{X}_{t+H-1} = F(X_{t-K}, \dots, X_{t-1}; \mathcal{G}; \Phi)$

- Result

- METR-LA, PEMS-BAY, PEMS03·04·07·08(Traffic), Solar, Electricity(Energy), ECG5000, Covid-19(Health) 데이터 셋에서 기존의 FC-LSTM, SFM, N-BEATS etc. 보다 좋은 성능을 보임

- Ablation Study

- Temporal dependency를 학습하는 Spe-Seq Cell을 제외한 경우 가장 큰 성능 차이를 보임

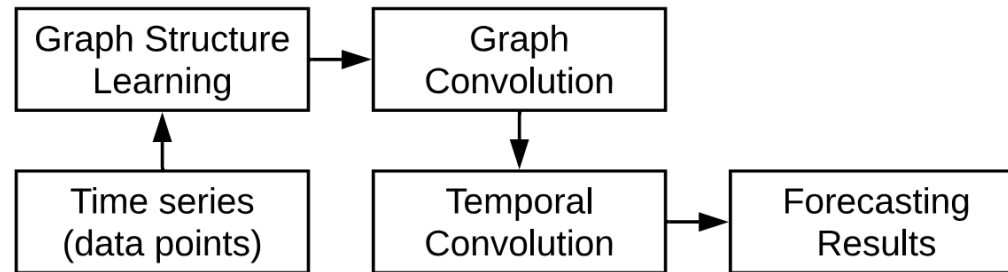
	StemGNN	w/o LC	w/o Spe-Seq Cell	w/o DFT	w/o GFT	w/o Residual	w/o Backcasting
MAE	2.144	2.158	2.612	2.299	2.237	2.256	2.203
RMSE	4.010	4.017	4.692	4.170	4.068	4.155	4.077
MAPE(%)	5.010	5.113	6.180	5.336	5.222	5.230	5.130

Results for ablation study of the PEMS07 dataset

Source : [1]

2-1. Multivariate time series forecasting with graph neural networks : Abstract

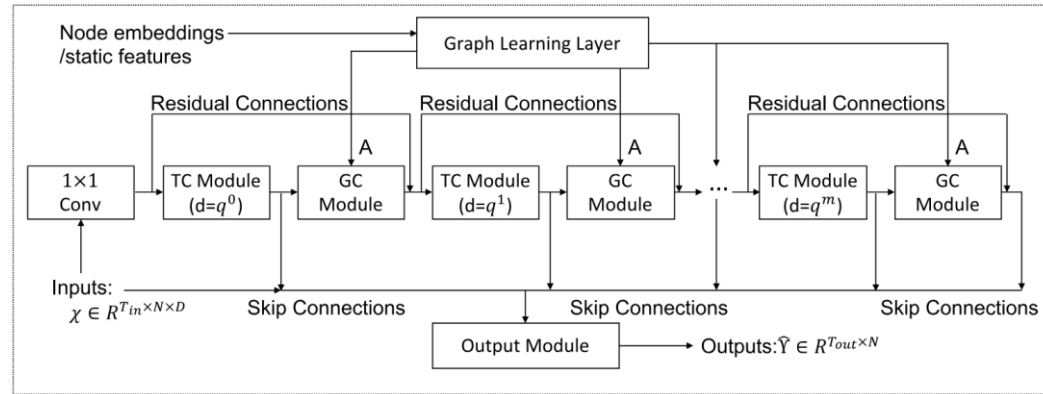
- Graph structure learning layer를 통해 시계열 데이터로부터 Sparse graph adjacency 행렬 추출
- Graph convolution module에서 변수간 Spatial dependency를 Directed graph 형태로 파악
- Temporal convolution module을 통해 temporal pattern 파악
- MTGNN 모델은 multivariate 시계열 데이터 모델링과 내부 그래프 구조 학습을 동시에 진행 가능
 - 모델 학습의 최적화를 위해 curriculum learning strategy와 시계열 데이터를 subgroup으로 나누어 학습하는 방법을 적용



Proposed Framework

Source : [3]

2-1. Process of MTGNN



Framework of MTGNN

Source : [3]

- Graph learning layer를 통해 시계열 데이터 간 hidden 관계를 파악하여 adjacency matrix를 학습하며 uni-directional relationship을 추출하도록 디자인됨
- Graph convolution module에서는 노드들과 그 이웃들의 정보를 통해 spatial dependency를 추출하며 2개의 mix-hop propagation layer로 구성되어 노드의 inflow & outflow 정보를 각각 학습 후 더함
- Temporal convolution module은 1D convolution filter를 통해 sequential pattern을 파악하는 dilated inception layer 2개로 구성되며 한 layer는 tanh가 filter로써, 다른 layer는 sigmoid가 다음 모듈로 pass되는 정보의 양을 control하는 gate로써의 역할을 수행하며 이 둘을 곱함

2-3. Problem definition & Result

- Problem definition
 - $z_t \in \mathbb{R}^N$ 은 time step t 때 차원이 N 인 multivariate variable이며 historical P time step $X = \{z_{t_1}, z_{t_2}, \dots, z_{t_p}\}$ 가 관측되었을 때 Q step 이후의 값이나 sequence $Y = \{z_{t_{t+Q}}\}$ or $\{z_{t_{P+1}}, z_{t_{P+2}}, \dots, z_{t_{P+Q}}\}$ 를 예측
 - X 로부터 Y 를 l_2 regularization이 적용된 absolute loss를 최소화하며 mapping하는 함수 f 모델링이 목표
- Result
 - Single-step(Traffic, Solar-Energy, Electricity, Exchange-Rate)에서는 대부분의 데이터에서 기존 방식들보다 우수한 성능을 보임. Exchange-Rate의 경우 TPA-LSTM에 소폭 밀리는 성능을 보임
 - Multi-step(METR-LA, PEMS-BAY; Traffic)에서는 METR-LA데이터에서는 기존 방식들에 비해 우수한 성능을 보이나 PEMS-BAY에서는 그렇지 않음
- Ablation study
 - Graph convolution module을 linear layer로 대체한 경우 가장 성능 하락폭이 큰 결과를 보임

3. Conclusion

- 두 논문 모두 그래프를 형성함에 있어 사전정보가 필수적이지 않으며 다변량 시계열 데이터를 그래프 신경망을 통해 모델링 및 예측한다는 공통점이 있으나 학습 방법에 있어 상이한 모습을 보임
- StemGNN은 Graph Convolutional Network_[2]을 적용하여 Spatial & temporal domain을 Fourier Transform을 통해 Spectral domain으로 변환하여 학습
- Latent correlation layer를 통해 학습된 Graph(Correlation matrix)는 학습 초기에 한번만 학습됨
- MTGNN은 Spatial & temporal domain에서 그대로 학습을 진행하며 각 노드관 관계는 Undirected가 아닌 Directed graph의 형태로 학습됨
- Graph learning layer를 통해 학습된 Adjacency matrix는 학습이 진행됨에 따라 update됨
- 두 논문의 인용 수에는 차이가 있으나 적절한 학습 방법을 선택하여 개선 및 발전시키는 것이 옳다고 사료됨

- 1) Cao, D., Wang, Y., Duan, J., Zhang, C., Zhu, X., Huang, C., ... & Zhang, Q. (2020). Spectral temporal graph neural network for multivariate time-series forecasting. *Advances in neural information processing systems*, 33, 17766-17778.
- 2) Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
- 3) Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020, August). Connecting the dots: Multivariate time series forecasting with graph neural networks. *In Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 753-763).