

시-공간 그래프 모델을 이용한 자전거 대여 예측

박장우*

국립순천대학교 인공지능공학부 교수

Prediction for Bicycle Demand using Spatial-Temporal Graph Models

Jangwoo Park*

Professor, Dept. of Artificial Intelligent Eng., Suncheon National University

요약 시간-공간적 의존성을 모두 고려하는 방법으로 그래프 신경망과 순환 신경망을 함께 사용하는 연구가 많이 진행되고 있다. 특히 그래프 신경망은 새롭게 활발히 연구되고 있는 분야이다. 서울시 자전거 대여 서비스(일명 따릉이)는 서울시 곳곳에 대여소를 갖추고 있으며 각 대여소에서 대여 정보가 충실하게 기록되어 있는 시계열 자료이다. 각 대여소의 대여 정보는 시간에 따른 주기성을 보이는 시간적인 특성을 갖추고 있으며, 지역적인 특성도 대여 현황에 큰 영향을 미치리라고 생각된다. 지역적 상관관계는 그래프 신경망을 이용하여 잘 이해할 수 있다. 이 연구에서는 서울시 자전거 대여 서비스의 시계열 데이터를 그래프로 재구성하고 그래프 신경망과 순차 신경망을 결합한 대여 예측 모델을 개발하였다. 시간에 따른 주기성과 같은 시간 특성과 지역적인 특성 및 각 대여소의 중요도 정도를 고려하였다. 대여소의 중요도 정도는 대여량 예측에 중요한 인자로 사용됨을 확인하였다.

주제어 : 그래프 신경망, 순환 신경망, 시-공간 모델

Abstract There is a lot of research on using a combination of graph neural networks and recurrent neural networks as a way to account for both temporal and spatial dependencies. In particular, graph neural networks are an emerging area of research. Seoul's bicycle rental service (aka Daereungi) has rental stations all over the city of Seoul, and the rental information at each station is a time series that is faithfully recorded. The rental information of each rental station has temporal characteristics that show periodicity over time, and regional characteristics are also thought to have important effects on the rental status. Regional correlations can be well understood using graph neural networks. In this study, we reconstructed the time series data of Seoul's bicycle rental service into a graph and developed a rental prediction model that combines a graph neural network and a recurrent neural network. We considered temporal characteristics such as periodicity over time, regional characteristics, and the degree importance of each rental station.

Key Words : Graph Neural Network; Recurrent Neural Network, Spatial-Temporal Model

1. 서론

서울시의 자전거 대여 서비스(일명 따릉이)[1]는 삼천

개 가까운 대여소를 서울 전역에 갖추고 있다. 각 대여소에서 자전거의 대여 수는 많은 차이가 있고, 대여 수량을 정확히 예측하는 것은 원활한 대여 서비스를 위해서 중

*교신저자 : 박장우(jwpark@scnu.ac.kr)

접수일 2023년 10월 23일 수정일 2023년 12월 19일 심사완료일 2023년 12월 21일

요하다. 서울시 자전거 대여 서비스와 관련된 데이터는 충분히 잘 수집되어 제공되고 있으므로 여러 다양한 연구에서 활용되고 있다. 특히 시계열 데이터를 정리하고 추세를 확인하기 위한 연구, 자전거 대여 수 예측을 위한 기계학습 모델 개발, 딥러닝 학습 모델 개발 등에 활발하게 사용되고 있다[2-4]. 그러나 자전거 대여소의 위치 정보를 이용한 그래프 신경 망을 이용하는 연구는 많지 않은 실정이다. 자전거 대여 데이터 세트는 위치 정보와 시간 정보를 결합한 그래프 데이터로 표현하고 분석할 수 있다.

시간-공간 상의 그래프로 표현 가능한 데이터 중에서 가장 활발한 연구가 진행되고 있는 것은 도심의 자동차 통행 속도와 통행량 예측이 이다[5-8]. 도심의 통행량은 시간에 따라서 변화하며 주위의 연결된 도로의 통행량에도 많은 영향을 받는다. 따라서, 공간상의 상관관계는 자동차 통행량 및 통행 속도 예측에 큰 영향을 미친다. 여러 관측소에서 측정된 시계열 데이터는 지역 사이의 상관관계를 고려한 그래프로 표현할 수 있다.

서울시의 자전거 대여 서비스의 데이터도 자동차 통행을 예측하는 것과 유사한 구조를 갖추고 있다. 서울시 여러 곳의 대여소에서 대여 정보가 기록되고 있다. 자전거 대여량은 시간에 따른 고유한 특성을 보이며, 대여소들 사이의 상관관계도 각 대여소에서 대여량 예측에 영향을 미칠 것이다. 따라서 서울시 자전거 대여량을 예측하는 것은 그래프 데이터 세트를 분석하는 중요한 자료가 되며, 그래프 신경망 모델을 연구하는 데 유용할 것이다.

본 연구에서는 서울시 자전거 대여 서비스의 시계열 데이터를 그래프로 구성하고, 그래프 신경망과 순환 신경망을 이용하여 시간-공간 의존성을 해석하는 방법에 대하여 설명한다. 시간-공간 그래프 모델을 이용하여 여러 대여소에 자전거 대여량을 예측하는 모델을 제시하고, 예측 모델에서 여러 대여소 사이의 지역적인 의존성 시간적인 의존성과 함께 고려한다. 논문에서 자전거 대여량 데이터로부터 그래프를 구성하고, 시간-공간 의존성을 모델링하기 위한 그래프 신경망과 순환 신경망을 결합하여 대여 예측 모델을 만들었다.

2. 관련 연구

2.1 GCN(Graph Convolutional Network)

그래프는 에지와 노드로 이루어져 있고, 서로 밀접하

게 연결된 관계망을 묘사하기 위하여 사용된다. 사회 관계망, 3차원 이미지, 분자 구조 등을 표현하는 데 유용하다. $G=(V,E)$ 를 그래프라고 하자. $V=\{v_1, v_2, \dots, v_n\}$ 이고 n 은 노드의 수이다. $E=\{e_{ij} \mid v_i, v_j \in V\}$ 는 에지 집합이다. 노드 v_i 의 특징 벡터는 $x_i \in \mathbb{R}^F$ 이고, $e_{ji} \in \mathbb{R}^D$ 는 노드 j 에서 i 로 연결된 에지 특징이다. GNN은 노드와 그래프의 표현을 학습하는 것이다. GNN은 노드의 가장 가까운 이웃 혹은 이웃의 이웃들의 노드 표현을 통합하여 노드의 표현을 업데이트한다. l -번째 레이어에서 노드 v_i 의 표현을 $h_i^{(l)}$ 로 표현하자. 가장 첫 번째 레이어의 경우 $h_i^{(0)} = x_i$. 일반적인 GNN에서 반복되는 과정을 AGGREGATE-COMBINE으로 표현한다[9, 10].

$$\begin{aligned} a_i^{(l)} &= \text{AGGREGATE}^{(l)}(\{h_j^{(l-1)} : j \in N(v_i)\}), \\ h_i^{(l)} &= \text{COMBINE}^{(l)}(h_i^{(l-1)}, a_i^{(l)}) \end{aligned}$$

GCN[11]은 딥러닝의 핵심적인 방법인 컨볼루션 연산과 가중치 공유를 그래프에 적용한 것이다. (l)-번째 레이어의 i -노드의 특징은 i -노드의 이웃인 노드 ($j \in N(i)$)에 의하여 계산된다. 여기에서 $N(i)$ 는 노드 i 의 이웃 노드들이다.

$$\begin{aligned} a_i^{(l)} &= \text{AGGREGATE}^{(l)}(\{h_j^{(l-1)} : j \in N(v_i)\}) \\ &= \sum_{j \in N(i)} \frac{1}{\sqrt{\deg(i)} \sqrt{\deg(j)}} W^{(l-1)} h_j^{(l-1)} + b^{(l-1)} \\ h_i^{(l)} &= \text{COMBINE}^{(l)}(h_i^{(l-1)}, a_i^{(l)}) \\ &= \sigma(W^{(l-1)} h_i^{(l-1)} + a_i^{(l)}) \end{aligned}$$

식에서 σ 는 비선형 함수이다. $\deg(i)$ 는 노드 i 의 node degree이고, 각 에지 $(j, i) \in E$ 에 대하여 정규화 계수는 $1/\sqrt{\deg(i)} \sqrt{\deg(j)}$ 이다. 식에서 보듯이 새로운 값으로 갱신되는 과정에서 컨볼루션 연산과 같이 지역적인 정보를 이용하여 가중치를 공유하게 된다. 각 노드의 특징을 계산하기 위해서는 각 노드에 연결된 이웃 노드의 특징값이 필요하게 된다. 이러한 연결성의 정보는 이웃 행렬(Adjacency Matrix)을 이용한다.

2.2 GAT(Graph Attention Network)

GAT[12, 13]는 그래프 신경망에 트랜스포머에서 이용하는 어텐션 개념을 도입한 것이다. GAT는 노드와 이웃 쌍을 동시에 계산할 수 있어서 효율적이다. GAT는 GCN과는 달리 정규화 계수를 어텐션 메카니즘을 이용하여 계산한다.

$$h_i^{(l)} = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} W^{(l)} h_j^{(l-1)} \right)$$

여기에서 α_{ij} 는 어텐션 스코어이고 노드 i 의 중요성을 결정한다.

$$\alpha_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in N_i} \exp(e_{ik})}$$

e_{ij} 는 어텐션 계수이고 각 노드 쌍에 대하여 다음과 같이 계산한다. 어텐션 계수는 노드 j 가 노드 i 의 이웃인 경우에만 계산된다[12].

$$e_{ij} = \text{LeakyReLU}(a^T [H_i W \parallel H_j W])$$

여기서 a 는 가중 벡터이고 LeakyReLU는 비선형 활성화 함수이다. \parallel 는 concatenation 기호이다.

자기-어텐션의 안정성을 위하여 멀티-헤드 어텐션을 적용하였다. 즉, K 개의 헤드에 대하여 위 계산을 독립적으로 수행하고 이를 합친다.

$$H_i^{(l+1)} = \parallel_{k=1}^K \sigma \left(\sum_{j \in N(i)} \alpha_{ij}^k W^k H_j^l \right)$$

2.3 순환신경망(Recurrent Neural Network)

시계열 데이터와 같은 순차적인 데이터를 처리하기 위하여 RNN(Recurrent Neural Network)이 사용된다. RNN은 시간 스텝 t 의 상태(h_t)는 이전 시간 스텝의 상태(h_{t-1})과 현재 값(x_t)를 입력으로 자신의 상태를 업데이트한다. RNN의 순환적인 구조 때문에 출력은 이전 모든 시간 스텝의 입력에 대한 함수가 된다. 시간 스텝이 길어지게 되면 RNN 셀을 거치면서 정보 일부를 잃게 되는 장기 의존성 문제가 발생한다. RNN의 역전파 방법인 BPTT(Back Propagation Through Time)는 모든 시간 스텝마다 처음부터 끝까지 역전파를 수행한다. 따라서 타임 스텝이 큰 경우 RNN은 매우 깊은 네트워크가 되며, 그레디언트 소실(vanishing gradient) 혹은 그레디언트 폭주(exploding gradient)현상을 경험하기 쉽다. 이러한 RNN의 어려움을 극복한 대표적인 구조로 LSTM(Long Short Term Memory)[14]과 GRU(Gated Recurrent Unit)[15]이 있다.

LSTM에서는 상태를 두 개의 벡터(h_t, c_t)로 나눈다. h_t 는 단기 상태, c_t 는 장기 상태라고 할 수 있다. LSTM의 핵심은 장기 상태에서 기억할 부분, 삭제할 부분, 입력에서 받아들일 부분을 게이트를 통하여 학습한다는 것이다. 이에 해당하는 장기 상태에서 기억과 망각을 망각

게이트(forget gate)를 이용하고, 새로운 입력을 추가할 부분은 입력 게이트(input gate)를 통하여 학습한다. 이렇게 학습된 장기 상태는 tanh 함수를 통과하고 출력 게이트(output gate)에서 단기 상태와 결합하여 출력을 만든다.

GRU는 LSTM의 간소화된 버전이라고 할 수 있다. LSTM의 두 상태(h_t, c_t)를 하나의 벡터로 합쳐서 표현하였다. 하나의 게이트 제어기를 이용하여 망각 게이트와 입력 게이트를 제어한다. 또한 GRU는 출력 게이트가 없고 전체 상태 벡터가 각 시간 스텝마다 출력되고 이전 상태의 출력을 제어하기 위하여 새로운 게이트 제어기가 있다.

3. 문제 분석 및 예측 모델

3.1 문제 설명 및 그래프

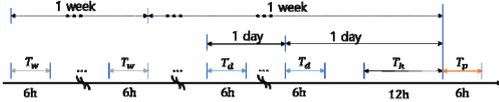
사용한 자전거 데이터는 2022년 3월부터 2022년 11월까지의 10개월의 데이터이다[21]. 자전거 대여 정보 시계열 데이터는 그래프 $G = (V, E, A)$ 로 표현할 수 있다. 여기에서 V 는 각 대여소를 표시하는 n 개의 노드의 집합이고, E 는 노드 사이의 연결 관계를 표시하는 에지, 그리고 A 은 대여소들 사이의 이웃 행렬이다. v_i 와 v_j 가 V 의 원소이고 $e_{ij} \in E$ 이면, A_{ij} 는 1이고 그렇지 않으면 0이다. 각 시간 스텝에서 그래프는 동적 특성 행렬 $X \in \mathbb{R}^{N \times F}$ 을 갖는다. 예측 모델 f 는 그래프와 T_h 스텝의 과거 데이터를 이용하여 향후 T_p 스텝의 값을 예측한다. 자전거 대여량을 예측하는 문제는 도로망에서 자동차의 교통량을 예측하는 것처럼 공간-시간 예측 문제라고 할 수 있다.

$$[X^{(t-T_h):t}, G] \xrightarrow{f} Y^{(t+1):(t+T_p)}$$

여기에서 $X^{(t-T_h):t} \in \mathbb{R}^{N \times F \times T_h}$, $X^{(t-T_h):t} \in \mathbb{R}^{N \times F \times T_h}$. 자전거 대여량 데이터 세트에서 차원 F 는 현재의 대여량, 24시간 전의 대여량, 1주일 전의 대여량 등 세 가지가 고려되었다. 그림 1에서 대여량을 이용하여 노드 특성 행렬을 구성하는 방법을 설명하였다.

자전거 대여 수는 시간적인 영향을 많이 받는다. 매일 매일의 주기성과 일주일의 주기성이 두드러진다. 따라서 하루 24시간과 일주일 7일을 임베딩하여 사용하였다. 각 대여소의 대여량을 시간 단위로 정리하면, 많은 곳은 400 이상의 대여가 있기도 하지만 대부분 10번 이하이다.

즉, 대여 횟수는 먹금수 분포 혹은 지수 분포를 갖는다. 따라서 대여량을 먹금수 변환[17, 18]을 하여 정규 분포에 가깝게 변환하였다.



[Fig. 1] Node features of graphs

노드들 사이의 연결 즉 에지는 대여소 사이의 거리와 가우시안 커널을 이용하는 방법이 있다. 이 방법은 자동차 통행량 예측을 위하여 많이 사용되고 있다. 또한 같은 경로를 이용한 빈도수를 이용하여 에지를 만들 수도 있다. 여기에서는 두 가지 방법을 모두 사용하였다. 두 방법 사이에는 결과 미치는 영향이 거의 없었고 거리를 이용하는 방법이 약간 좋은 결과를 얻을 수 있었다. 두 노드 사이의 에지 특징값은 다음과 같이 계산한다[4, 5].

$$w_{ij} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) & d_{ij} > k \\ 0 & \text{otherwise} \end{cases}$$

여기에서 d_{ij} 는 노드 v_i 와 v_j 사이의 거리, σ 는 거리의 표준 편차, k 는 미리 설정된 임계값이다.

그래프 해석에서 각 노드의 중요성은 노드마다 다르다. 예를 들어 사회 관계망에서 셀럽들은 그렇지 않은 사람들보다 관계망 해석에 큰 영향을 미칠 것이다. 따라서 노드의 중요성을 인코딩하는 것은 필요하다[9]. 노드에 연결된 이웃 노드들이 많다면 중요한 노드라고 할 수 있다. 자전거 대여 서비스에서도 대여소가 이웃한 대여소가 많다면 대여 및 반납에 영향을 미칠 것이다. 연구에서는 노드 중요성을 고려하기 위하여 node degree 즉, 노드에 직접 연결된 노드의 수를 인코딩하여 고려하였다.

3.2 그래프 모델 만들기

매시간 하나의 그래프가 만들어지고 각 그래프의 노드 v_i 는 특징 $x_i \in R^{T_h \times F}$ 을 갖는다. 과거 대여량 데이터와 함께 시간적인 특성을 고려하기 위하여 24개의 시각(하루)과 7개의 요일(일주일)을 임베딩하여 활용하였다. 노드의 중요성을 고려하기 위하여 중심성 인코딩을 추가하였다. 중심성 인코딩은 각 노드의 in-degree와 out-degree를 임베딩하여 사용하였다. 즉, GNN 또는 RNN의 입력에 단순히 노드 특징값을 입력한 것이 아니라 노드 특징을 선형 변환하여 d -차원으로 확장하고, 각

24개의 시간과 7개의 요일은 임베딩을 통하여 d -차원의 벡터로 임베딩하였다. 각 노드의 in-degree와 out-degree 역시 임베딩을 통하여 d -차원의 실숫값으로 변환하였다. 이렇게 변환된 값들은 모두 더해져 모델의 입력으로 제공된다.

$$h_i^{(0)} = f(x_i) + z_{hour} + z_{weekday} + z_{deg^-(v_i)} + z_{deg^+(v_i)}$$

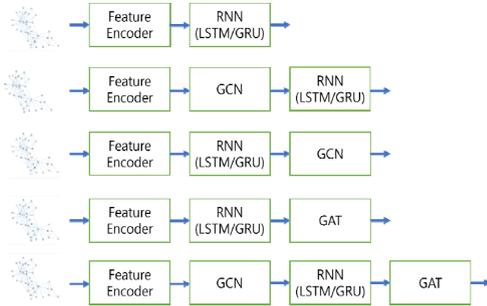
여기에서 $f(x_i)$ 는 노드 특징의 선형 변환, z_{hour} 와 $z_{weekday}$ 는 시간과 요일의 임베딩 벡터, $z_{deg^-(v_i)}$ 와 $z_{deg^+(v_i)}$ 는 in-degree와 out-degree의 임베딩 벡터[9]이다. 이렇게 입력 신호를 변환하는 것을 특징 인코더(feature encoder)라고 하였다.

시-공간의 상관관계를 모델링하는 많은 방법이 제시되고 있다. 그래프 구조를 이용하여 공간적인 의존성을 설명하고 시계열의 순차적인 성질은 RNN(Recurrent Neural Network)를 이용한다. RNN으로 LSTM 또는 GRU 등이 많이 사용된다. 공간상의 상관관계를 모델링하는 경우 GCN, GAT 등을 사용한다. 시공간 그래프 모델링은 다양한 분야에서 이용되고 있으며 여러 가지 변형된 구조들이 제안됐다.

본 연구에서는 GCN(혹은 GAT)와 LSTM(혹은 GRU)와 결합하여 시-공간 그래프 모델링을 구현하였다. 일반적으로 먼저 그래프를 GCN을 통과시킨 후 얻어진 결과를 LSTM에 입력하는 순서를 따른다. 즉, 그래프의 공간적인 의존성을 GCN을 이용하여 모델링한 후 시간적인 의존성을 파악하는 순서이다 교통 체증과 같은 데이터의 경우 공간상의 이웃 의존성이 매우 크기 때문에 이러한 모델이 우수한 결과를 얻고 있다고 생각한다. 특히 [19, 20] 등에서 GRU 혹은 LSTM 상태와 입력을 위하여 각 시간 샘플 그래프에 GCN을 우선 적용하는 형태를 사용하였다. 따라서 모델링에 사용된 과거 시간 샘플 수와 같은 수의 GCN이 사용된다.

그러나 자전거 대여의 경우 이웃의 대여량이 해당 대여소의 대여량에 미치는 영향이 매우 크지 않은 듯하다. 그러나 대여소가 여러 이웃과 가깝게 연결되어 이동할 수 있는 이웃이 많은 경우 이것을 고려하는 것은 중요하다. 공간상의 의존성과 시간적인 관계의 중요성을 알아 보기 위하여 GCN과 LSTM의 조합은 그림 2처럼 구성하였다. 첫 번째 모델은 특징 인코더를 통하여 인코딩한 후에 LSTM에 직접 입력하는 경우이다. 나머지 모델은 GNN과 RNN을 조합하였다. 시간 의존성을 먼저 모델링(LSTM/GCN)하고 다음에 공간 모델링(GCN, GAT)을

하는 모델을 구성하였고, 다음으로 공간 모델링을 먼저 한 후 시간 모델링을 하는 방법, 시간 모델링 전과 후에 공간 모델링 하는 방법을 비교하였다.



[Fig. 2] Experimental models

4. 결과 검토

자전거 대여 데이터를 그래프로 구성하고 훈련/검증/테스트 데이터로 나누었다(각각 70%, 20%, 10%). 각 노드의 특징(과거 대여량)과 시간 정보, 각 노드의 degree를 특징 인코더를 이용하여 변환하였다. 각 노드 degree가 예측 결과에 미치는 영향을 확인하기 위하여 특징 인코더에서 degree 임베딩을 포함한 경우와 그렇지 않은 경우로 나누어 살펴보았다. 즉, 각 노드에 연결된 이웃한 노드의 수는 그 노드의 중요성을 설명하는 중요한 지표이다. 표 1은 node degree를 고려한 경우와 그렇지 않은 경우의 결과이다. 표에서 사용한 모델은 간단한 LSTM이다. 성능지표로는 MAE(Mean Absolute Error)와 MSLE(Mean Square Log Error)를 사용하였다. 표에서 알 수 있듯이 node degree 임베딩을 고려한 경우가 우수하다. 즉, 각 노드에 연결된 이웃 노드의 수는 모델의 정확성에 중요한 역할을 한다.

<Table 1> The effect of node degree on the prediction accuracy

	With node degree	Without node degree
MAE	2.149	2.918
MSLE	0.052	0.103

Node degree의 중요성은 모든 모델에서 거의 같게 나타나고 있어서, 앞으로 설명하는 결과는 node degree를 포함한 것이다. 표 2는 각 모델의 계산된 MAE이다. 과거 12시간의 데이터를 이용하여 향후 6시간까지의 대

여량을 시간 별로 예측하였다. 표시한 오차는 대여수이다. 사용된 모델은 표에 표시되어 있다. 여기에서 RW는 현재의 값으로 미래의 6시간의 대여량을 예측하는 간단한 모델이고 HA는 모델의 훈련에 사용한 12시간 데이터의 평균을 예측값으로 사용한 것이다. GCN을 LSTM 앞에 적용한 모델에 비하여 LSTM 결과 후에 적용한 모델이 다소 좋은 결과를 보였다. 한편 GCN 보다 GAT를 적용한 것이 더 작은 MAE를 나타내었다. Node-degree 임베딩 후에 단지 LSTM을 적용한 간단한 모델이 가장 좋은 결과를 보인다. 즉, 자전거 데이터 세트의 경우 특정 노드에서 자전거 대여량의 예측은 시간 의존성을 고려하는 것이 특히 중요하다는 점을 설명한다. 또한 이웃의 자전거 대여량이 특정 대여소의 자전거 대여량에 미치는 영향은 크지 않다고 생각된다. 특히, 자전거 대여소와 연결된 이웃의 수(노드의 중요성)는 모델의 정확성에 큰 영향을 미치고 있다.

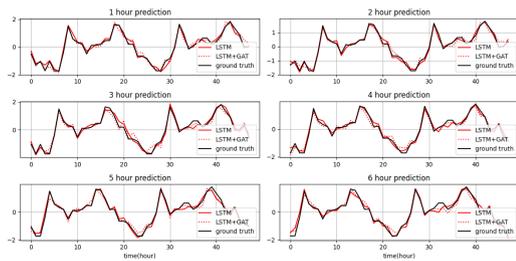
표 2에서 확인할 수 있듯이 가까운 시각인 한 시간 후보다 두 시간 혹은 세 시간의 후의 예측 오차가 더 적게 나타났다. 여섯 시간 후의 예측은 다른 시간에 비하여 다소 좋지 못하였다. 장기 예측을 위한 정교한 모델의 구축이 필요하다고 생각된다. 장기 특성을 반영하기 위하여 일간 주기성, 주간 주기성 등을 좀 더 길게 고려하는 것도 하나의 방법이라고 생각한다. GRU 모델을 사용한 경우의 결과도 같이 나타내었다. GRU보다 LSTM을 사용하면 우수한 MAE를 얻었다. GRU를 사용하였을 때는 GCN을 사용한 경우가 가장 좋은 결과를 얻을 수 있었다.

그림 3은 특정 대여소에서 테스트 데이터 세트의 예측 결과를 나타내었다. 48시간 동안 예측한 결과를 6시간 동안 각 시간에 대하여 표시하였다. 결과는 LSTM만을 사용한 경우와 LSTM과 GAT를 사용한 결과를 같이 비교하였다. 결과 그림으로부터 둘 사이의 차이는 별로 크지 않음을 알 수 있다. 또한 앞의 표 2에서도 알 수 있듯이 6시간 동안의 각 시각에서 예측 오차가 크지 않다.

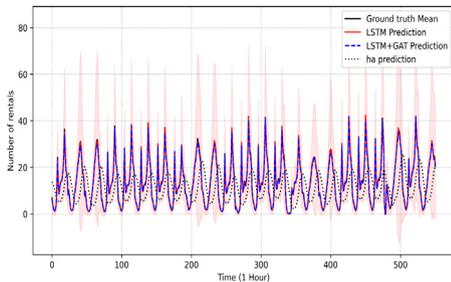
그림 4는 각 대여소의 평균 대여량을 시간에 대하여 표시하였다. 계산 결과는 훈련에 사용한 훈련 데이터와 검증 데이터가 아닌 테스트 데이터 셋에 대하여 구하였다. 그래프에는 실제 측정값, LSTM모델로 예측한 결과, LSTM과 GAT를 이용한 모델의 결과를 표시하였다. 모델들은 대여소의 평균값을 잘 설명하고 있으며, 대여량의 시간적인 주기성 등을 잘 추적하고 있음을 확인할 수 있다. 간단한 LSTM 모델이 그래프 신경망을 도입한 모델에 비하여 다소 우수한 특성을 얻을 수 있었다.

<Table 2> MAEs for 6 hour predictions

	1h	2h	3h	4h	5h	6h
LSTM	2.15	1.99	1.97	2.01	2.16	2.63
LST+GCN	2.69	2.41	2.37	2.46	2.72	3.19
LSTM+GAT	2.54	2.26	2.20	2.26	2.45	2.93
GCN+LSTM	3.14	2.98	2.96	3.03	3.22	3.54
GCN+LSTM+GCN	3.43	3.21	3.16	3.24	3.40	3.69
GCN+LSTM+GAT	3.50	3.33	3.29	3.35	3.53	3.79
GRU	3.34	3.21	3.20	3.25	3.33	3.52
GRU+GCN	3.17	3.01	2.98	3.04	3.17	3.56
GRU+GAT	3.65	3.50	3.47	3.46	3.51	3.69
GCN+GRU	3.92	3.89	3.88	3.88	3.92	4.01
GCN+GRU+GCN	3.93	3.88	3.82	3.82	3.89	4.03
GCN+GRU+GAT	3.94	3.86	3.82	3.81	3.86	4.01
RW	6.74	8.84	10.5	11.9	13.2	14.3
HA	10.9	11.8	12.5	12.9	13.3	13.4



[Fig. 3] Predictions for 2 days at given stations



[Fig. 4] Mean demand prediction for test dataset

5. 결론

자전거 데이터는 지역 사이와 의존성을 보이고 있으며, 시간적인 주기성 등 시간적인 변화 특성이 뚜렷하다. 본 연구 서울시 자전거 대여 시계열 자료를 그래프 형식으로 변환하여 시간-공간상의 상관관계를 모델링하였다. 시간-공간 그래프 모델을 사용하여 각 대여소의 대여량 예측하였다. 자전거 대여소의 공간적인 특성인 각 대여소와 연결된 이웃 대여소의 수가 대여 예측에 큰 영향을 미쳤다. 이웃 대여소의 특징값들이 정확한 예측에 미치는 영향을 확인하기 위하여 그래프 신경망인 GCN과

GAT를 도입하였다. 시간 의존성을 고려하기 위한 LSTM과 GRU를 이용하였다. LSTM 혹은 GRU 앞에서 공간적 성질을 고려하는 것보다 뒤에서 고려하는 것이 조금 더 긍정적인 결과를 보였다.

서울시 자전거 대여 서비스는 매우 잘 측정된 시계열 데이터이고 여러 연구에서 활발히 사용되고 있다. 자전거 대여 데이터를 활용한 데이터 분석은 여러 방법으로 제시되고 있으며, 대여량의 예측 모델 연구도 활발하다. 또한, 자전거 대여 데이터 세트는 지역-시간 상관관계를 고려하는 그래프 데이터와 그래프 신경망을 연구하는 중요한 자료이다. 자전거 대여소와 자전거 도로 등에 속도 통행량을 측정할 수 있는 IoT 장비를 활용한다면 보다 정확한 통행량 예측과 대여량 예측이 가능할 것이고 자전거가 더욱더 스마트한 교통수단의 일환이 될 수 있으리라고 기대한다.

REFERENCES

- [1] Seoul Bike, <https://www.bikeseoul.com/>
- [2] V.E.Sathishkumar and Y.Cho. "A rule-based model for Seoul Bike sharing demand prediction using weather data." *European Journal of Remote Sensing* 53, vol.sup1., pp.166-183, 2020.
- [3] V.E.Sathishkumar, J.Park, and Y.Cho. "Seoul bike trip duration prediction using data mining techniques." *IET Intelligent Transport Systems*, Vol.14, No.11, pp.1465-1474, 2020.
- [4] V.E.Sathishkumar, J.Park, and Y.Cho. "Using data mining techniques for bike sharing demand prediction in metropolitan city." *Computer Communications* 153, pp.353-366, 2020.
- [5] C.Zhang, J.J.Q.Yu, and Y.Liu. "Spatial-temporal graph attention networks: A deep learning approach for traffic forecasting." *IEEE Access* 7, 166246-166256, 2019.
- [6] X.Kong., W.Xin, X.Wei, P.Bao, J.Zhang and W.Lu. "STGAT: Spatial-temporal graph attention networks for traffic flow forecasting." *IEEE Access* 8, 134363-134372., 2020.
- [7] B.Yu, H.Yin, and Z.Zhu. "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting." *arXiv preprint arXiv:1709.04875*, 2017.
- [8] J.Zhu, Y.Song, L.Zhao and H.Li "A3t-gcn: Attention temporal graph convolutional network for traffic forecasting." *ISPRS International Journal of Geo-Information*, 10, Vol.7, p.485, 2021.
- [9] C.Ying, T.Cai, S.Luo, S.Zheng, G.Ke, D.He, T.Shen

and T.Y.Liu, "Do transformers really perform badly for graph representation?" *Advances in Neural Information Processing Systems* 34, 28877-28888., 2021.

- [10] Pytorch-Geometric, <https://pytorch-geometric.readthedocs.io/en/latest/index.html>
- [11] T.N.Kipf, and M.Welling. "Semi-supervised classification with graph convolutional networks." *arXiv preprint arXiv:1609.02907*, 2016.
- [12] P.Veličković, G.Cucurull, A.Casanova, A.Romero, P.Lio, and Y.Bengio "Graph attention networks." *arXiv preprint arXiv:1710.10903*, 2017.
- [13] S. rody, U.Alon, and E.Yahav. "How attentive are graph attention networks?" *arXiv preprint arXiv:2105.14491*, 2021.
- [14] R.C.Staudemeyer and E.R.Morris. "Understanding LSTM-a tutorial into long short-term memory recurrent neural networks." *arXiv preprint arXiv:1909.09586*, 2019.
- [15] K.Cho, B.Merrienboer, C.Caglar, D.Bahdanau, F.Bougares, H.Schwenk and Y.Bengio "Learning phrase representations using RNN encoder-decoder for statistical machine translation." *arXiv preprint arXiv:1406.1078*, 2014.
- [16] S.Guo, Y.Lin, N.Feng, C.Song, and H.Wan "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting." *Proceedings of the AAAI conference on artificial intelligence.*, Vol.33., No.1., pp.922-929, 2019.
- [17] Power transform, Wikipedia, https://en.wikipedia.org/wiki/Power_transform
- [18] sklearn, PowerTransformer, <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.PowerTransformer.html>
- [19] Y.Seo, M.Defferrard, P.Vandergheynst and X. resson "Structured sequence modeling with graph convolutional recurrent networks." *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December, 13-16*, pp.362-373, 2018.
- [20] J. CHen, X.Wang, and X.Xu. "GC-LSTM: Graph convolution embedded LSTM for dynamic network link prediction." *Applied Intelligence*, pp.1-16., 2022.
- [21] Seoul Open Data Square, <https://data.seoul.go.kr/>

박 장 우(Jangwoo Park)

[정회원]



- 1987년 2월 : 한양대학교 공과대학 전자공학과(공학사)
- 1989년 2월 : 한양대학교 대학원 전자공학과(공학석사)
- 1993년 8월 : 한양대학교 대학원 전자공학과(공학박사)
- 1995년 3월 ~ 현재 : 순천대학교 공과대학 인공지능 공학부 교수

<관심분야>

사물인터넷, 그래프 신경망, 딥러닝