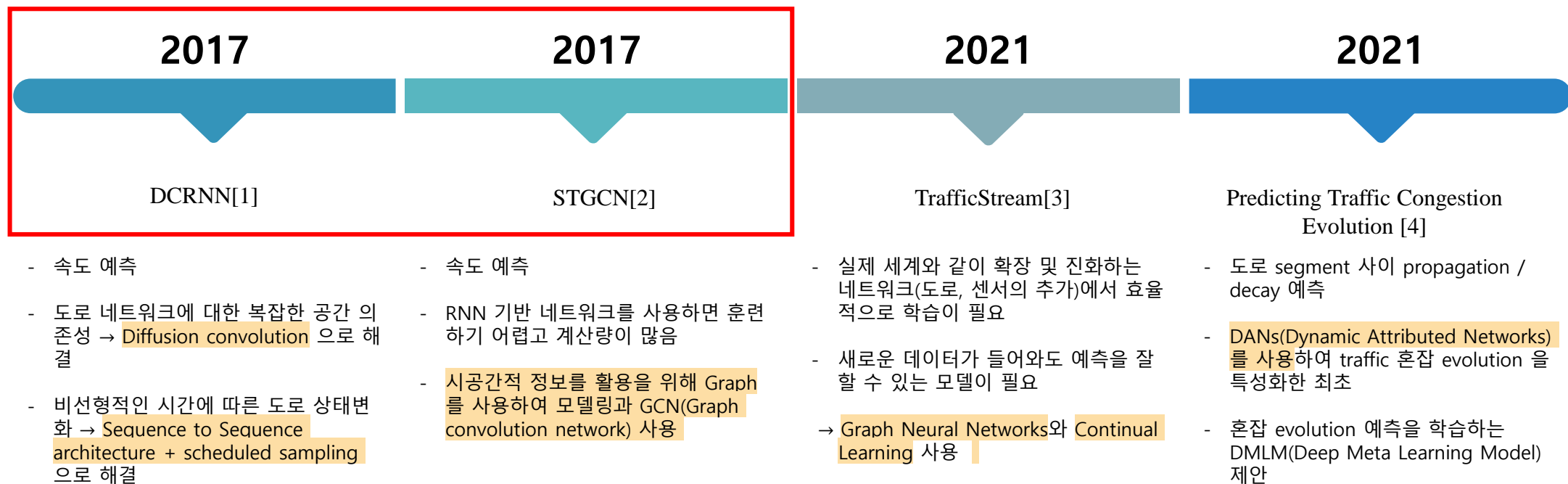


ETRI – 용역 과제

Traffic Congestion Propagation Prediction

2022.09.19
데이터 인텔리전스 연구실
김정선, 한승훈, 최민규

- **Timeline**
- **Dataset**
- **T-map Dataset**
- **Others**
- **ISSUE**



[1] Li, Yaguang, et al. "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting." arXiv preprint arXiv:1707.01926 (2017).

[2] Yu, Bing, Haoteng Yin, and Zhanxing Zhu. "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting." arXiv preprint arXiv:1709.04875 (2017).

[3] Chen, Xu, Junshan Wang, and Kunqing Xie. "TrafficStream: A Streaming Traffic Flow Forecasting Framework Based on Graph Neural Networks and Continual Learning." arXiv preprint arXiv:2106.06273 (2021).

[4] Sun, Yidan, et al. "Predicting Traffic Congestion Evolution: A Deep Meta Learning Approach." IJCAI. 2021.

- Others

- DCRNN, STGCN 에 초점을 두어 코드 실행

	DCRNN	STGCN	MW-TGC[5]
METR-LA	✓	✓	
PEMS-BAY	✓		
PeMSD7(M)		✓	
PeMSD7(L)		✓	
Seoul_speed		✓	✓

- **METR-LA**

- 2012/03/01 ~ 2012/06/30 4개월 데이터, 207개 센서, Los Angeles
- 70% training, 20% test, 10% validation
- #Nodes 207 #edges 1,722 #Features 2

- **PEMS-BAY**

- 2017/01/01 ~ 2017/05/31 6개월 데이터, 325개 센서, California
- 70% training, 20% test, 10% validation
- #Nodes 325 #edges 2,694 #Features 2

- Seoul_speed

- 2018/04/01 ~ 2018/04/30 1개월 데이터, 서울
- Urban-core, Urban-mix으로 나눔
- Urban-core
 - #Nodes 304 #edges 696
 - 평균 속도가 27.63km/h, 표준 편차가 11.25km/h, 제한 속도가 60km/h
 - 동일한 제한 속도가 있는 도로 구간만
 - 서울 강남
- Urban-mix
 - #Nodes 1,007 #edges 2,295
 - 평균 속도는 30.69km/h, 표준 편차는 16.91km/h, 제한 속도 40km/h ~ 100km/h
 - Urban-core를 확장하여 나타냄. 제한 속도가 여러 개 있음.
 - 청담대교에서 동작대교까지 이어지는 도시의 동서단을 연결하는 고속도로, 도시 간선도로, 골목길, 교량

- Seoul_speed, STGCN 에 적용

- STGCN 은 .h5 필요하여 만들어 적용
- batch size , epochs 50 (default settings)

```
epoch 35 , train loss: 0.30953302172144137 , validation loss: 0.35728273389639553
epoch 36 , train loss: 0.3085315256052895 , validation loss: 0.35660389305411516
epoch 37 , train loss: 0.3081978736216385 , validation loss: 0.3577447897279179
epoch 38 , train loss: 0.30781919424423504 , validation loss: 0.3571863586043978
epoch 39 , train loss: 0.307518810124736 , validation loss: 0.35735274246939414
epoch 40 , train loss: 0.3071801368615085 , validation loss: 0.3569398713070196
epoch 41 , train loss: 0.3064530464143425 , validation loss: 0.35756768692623486
epoch 42 , train loss: 0.30623709183463527 , validation loss: 0.3571367266920063
MAE: 2.769977759410074 , MAPE: 0.11421313306142994 , RMSE: 4.00048234611768157275
```

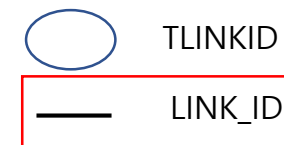
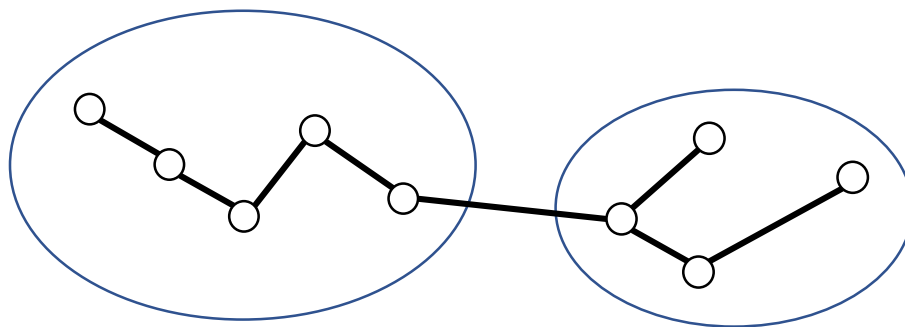
- Seoul_speed, DCRNN 에 적용

- DCRNN에 weight 부분, 마지막 평가하는 부분 추가하여야함

```
Epoch: 98 Train Loss: 0.4928283512592316 Valid Loss: 5.566777071802598 Test Loss: 5.743278108168546
Epoch: 99 Train Loss: 0.5054105397902037 Valid Loss: 5.550193680435468 Test Loss: 5.735393145631324
```

- T-map 지도 형상 Dataset

- .dbf 지리 사상의 속성 정보를 제공하는 dBASE 파일(table)



LINK_ID를 예측하는게 목표
LINK_ID 연결 관계 고민 중
TLINKID만 먼저 속도 예측하기 진행

IDXNAME	LINK_ID	TLINKIDP1	TLINKIDN1	ST_ND_ID	ED_ND_ID	LENGTH	ST_DIR	ED_DIR
55270000	1			8	10	52	283	103
55270000	10	10416802	10416801	34	54	98	148	328
55270000	100			223	233	54	32	212

LinkShape.dbf 일부분

- TLINKID 속도 예측하기 위해 사용한 Dataset

- tsdlinkturnmaster_20211221.csv (TSD Network (Topology))
 - tsdlinkid 연결 관계
- 2021.07.01 ~ 2021.09.30: ts_202107-09.zip (교통정보 이력 데이터)
 - 날짜, tsdlinkid, 속도

fromtsdlinkid	totsdlinkid
1000041	1000061
1000041	1000071
1000042	1000051
1000051	16005862
1000051	16007221
1000052	1000041
1000061	1000042

tsdlinkturnmaster_20211221.csv

```
202107012355,16005421,15
202107012355,16005422,22
202107012355,16005431,55
202107012355,16005441,37
202107012355,16005451,29
202107012355,16005461,36
202107012355,16007182,77
202107012355,16007191,55
202107012355,16007192,52
202107012355,16007202,30
```

ts_202107-09.zip

- 초기 결과

- tsdlinkid에 대한 속도 예측
- STGCN 이용 (default setting)
- 예제로 일부분 영역을 지정하여 진행 (node = 38, edge = 50)
- 영역에 나와있는 tsdlinkid의 연결관계를 tsdlinkturnmaster_20211221.csv에서 찾아서 진행



- 초기 결과 이유

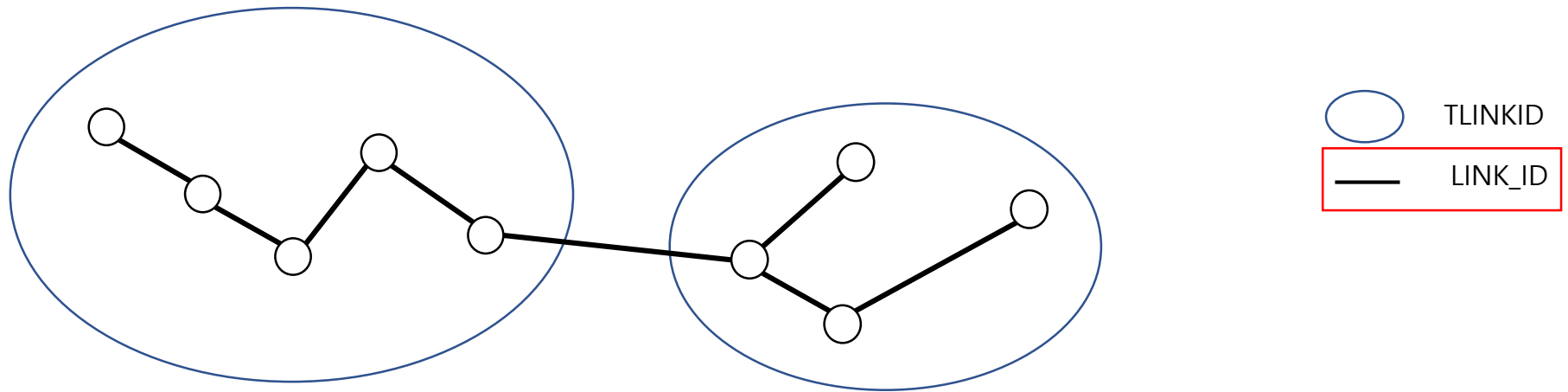
- 거리에 대한 정보가 없어, weight를 1로 적용
- Weight : $W_{ij} = \exp(-\frac{dist(v_i, v_j)^2}{\sigma^2})$, if $dist(v_i, v_j) \leq k$, otherwise 0

사용하는 방법으로
다시 세팅

```
epoch 44 , train loss: 0.5618260246746384 , validation loss: 0.9516385979559724
epoch 45 , train loss: 0.5603646284238789 , validation loss: 0.9521081979896296
epoch 46 , train loss: 0.5569836747349572 , validation loss: 0.956722533396605
epoch 47 , train loss: 0.5559610458207734 , validation loss: 0.9570730472209215
epoch 48 , train loss: 0.5547721346528235 , validation loss: 0.9534911910136763
epoch 49 , train loss: 0.5537293919150836 , validation loss: 0.9628197485532631
epoch 50 , train loss: 0.5528204001892079 , validation loss: 0.9659050840263317
test loss: 0.7880204428457519
MAE: 6.8352286614601745 , MAPE: 2253141.128272355 , RMSE: 9.773024610217222
(traffic) root@437c250b158e:/workspace/cmg/stgcn#
```

- T-map 지도 형상 Dataset

- .dbf 지리 사상의 속성 정보를 제공하는 dBASE 파일(table)



IDXNAME	LINK_ID	TLINKIDP1	TLINKIDN1	ST_ND_ID	ED_ND_ID	LENGTH	ST_DIR	ED_DIR
55270000	1			8	10	52	283	103
55270000	10	10416802	10416801	34	54	98	148	328
55270000	100			223	233	54	32	212

LinkShape.dbf 일부분

NODE_ID	X	Y
8	127.136484	36.1666666
34	127.206062	36.1702333
223	127.220566	36.2075456

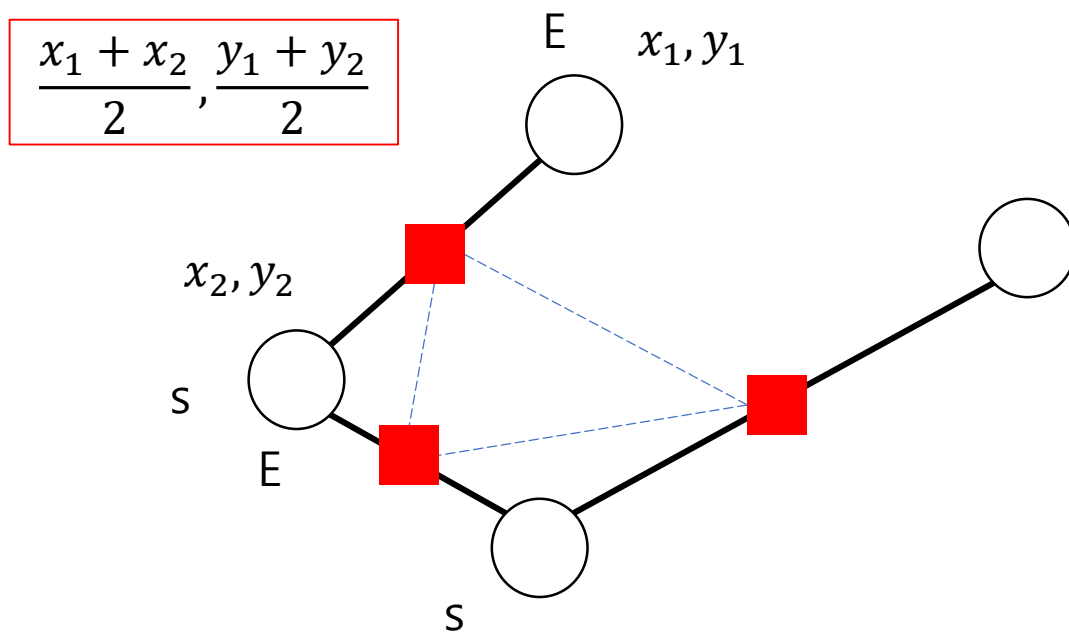
NodeShape.dbf 일부분

이용하여 weight 적용
& LINK_ID 에 대하여 예측

- 데이터의 거리 계산 과정

- weight를 줄 때, 거리 — —

— LINK_ID
— — 거리 계산



- re-setting 결과

- LINK_ID에 대한 속도 예측
- STGCN 이용 (default setting)
- 예제로 일부분 영역을 지정하여 진행 (node = 260, edge = 9407)

- 초기 결과 원인


- Linkshape.dbf 에서 TLINKIDP1, TLINKIDN1 둘 다 있을 경우 TLINKIDP1의 속도를 가지고 와서 적용 (TLINKIDP1, TLINKIDN1 에 따라 속도가 다르기 때문)
- weight에서 k setting

```
epoch 48 , train loss: 0.7295432684617432 , validation loss: 0.7437714364324862  
epoch 49 , train loss: 0.7289762179195731 , validation loss: 0.7400819376179072  
epoch 50 , train loss: 0.728909807373001 , validation loss: 0.7413112190488267  
test loss: 0.7622599793681701  
MAE: 6.428070479820044 , MAPE: 82564.79180091713 , RMSE: 9.446369732031599
```

re-setting STGCN 결과

- 데이터 경진대회

- Dataset : T-map mobility, kakao mobility



The 40th Anniversary International Conference of the Korean Society of Transportation (KST) & 10th International Public Transportation Forum (IPTF)
K-Transport, Present & Future
28th Wed - 30th Fri, September 2022
LG Sciencepark, Seoul, Korea

INVITATION CONFERENCE COMMITTEE RESEARCH ANTHOLOGY CONFERENCE VENUE PROGRAM **DATA COMPETITION**

DATA COMPETITION

KOREAN SOCIETY OF TRANSPORTATION DATA COMPETITION

- Hosted by KAKAO MOBILITY and TMAP MOBILITY (alphabetically ordered).
- Predict real-world transportation phenomenon.
- Deal with high-quality data offered from the top mobility companies in Korea.
- Apply your machine learning techniques and prove yourself.

- 데이터 경진대회

- Dataset : T-map mobility
- PROBLEM OVERVIEW : 과거 교통정보를 사용하여, 미래 시간의 1시간 교통정보(truetraveltime) 예측
- TARGET SITES
 - 고속도로구간 : 남양교 ~ 평택 마도교 ~ 한국 시흥 고속도로 (tsdlinkid = 2594592)
 - 도시도로구간 : 뚝섬로 뚝도 시장 앞 사거리영동대교 북단 사거리까지 (tsdlinkid = 1018052)
- DATASET DESCRIPTION

Field name	Description
periodtime	교통 정보 수집 시간, yyyymmddhhmm
tsdlinkid	Link index
length	Link 길이, meter
offerspeed	Provided(Generated) 속도, km/h
offertraveltime	Provided(Generated) 이동시간, 초
truespeed	속도, km/h
truetraveltime	이동시간, 초

Field name	Description
perioddate	Date, yyyymmdd
dayofweek	요일
holiday	(0 : non-holiday, 1 : weekend or/and holiday)
description	Holiday Type

- Linkshape.dbf 에서 LINK_ID 의 방향을 알 수 있는지

감사합니다.