Travel Time Estimation Based on Neural Network with Auxiliary Loss

Yunchong Gan, Haoyu Zhang, Mingjie Wang

GitHub Repo: https://github.com/Phimos/SIGSPATIAL-2021-GISCUP-2nd-Place-Solution

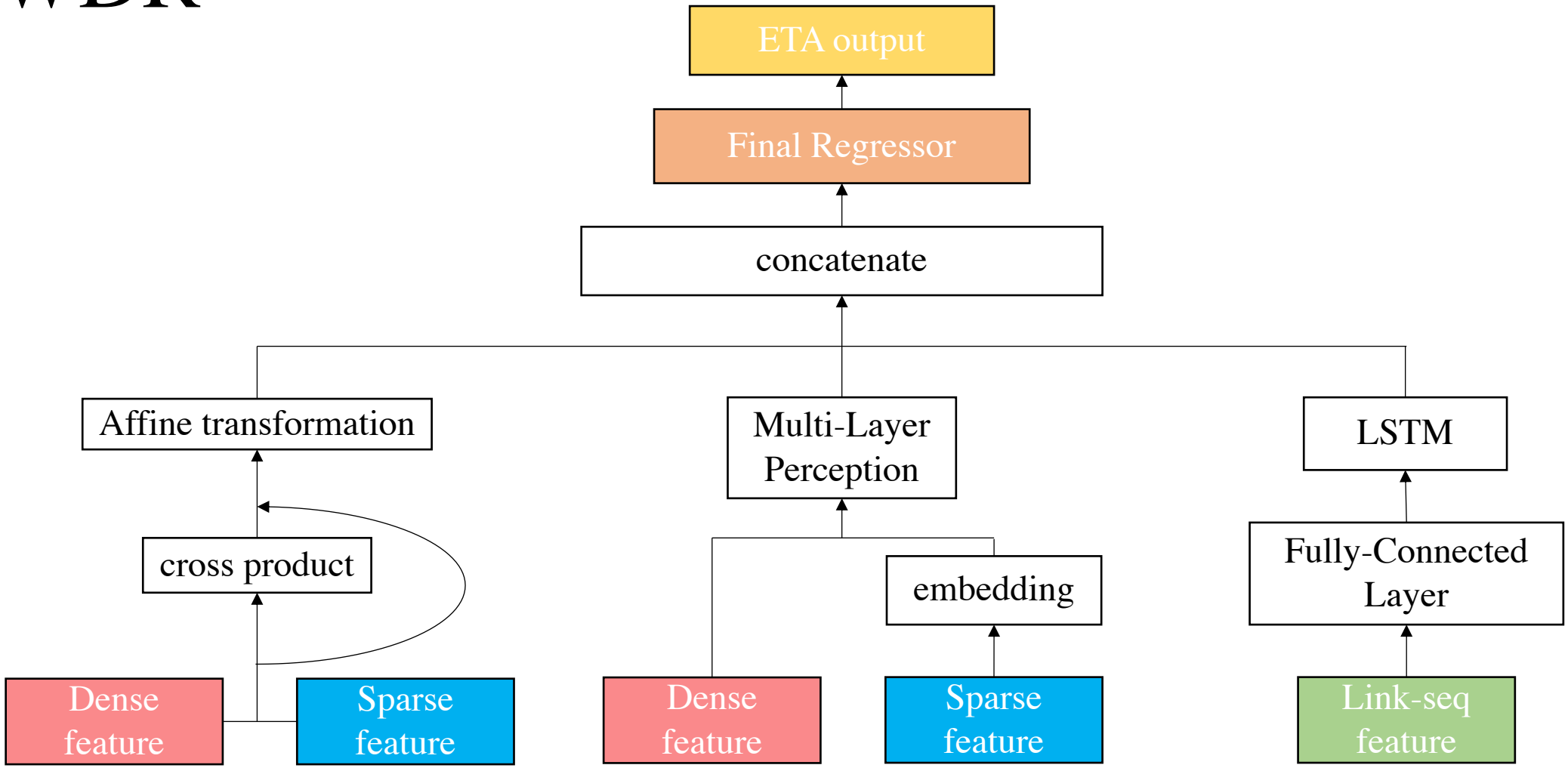
Contribution

- We propose a Wide-Deep-Double-Recurrent model with Auxiliary loss named WDDRA, which has excellent performances in the ETA problem.
- We show that auxiliary loss can reduce the influence of overfitting, which makes the training process more stable and generalization ability better.
- We show that K-fold bagging strategy can further improve the accuracy and robustness of our model in the ETA problem.

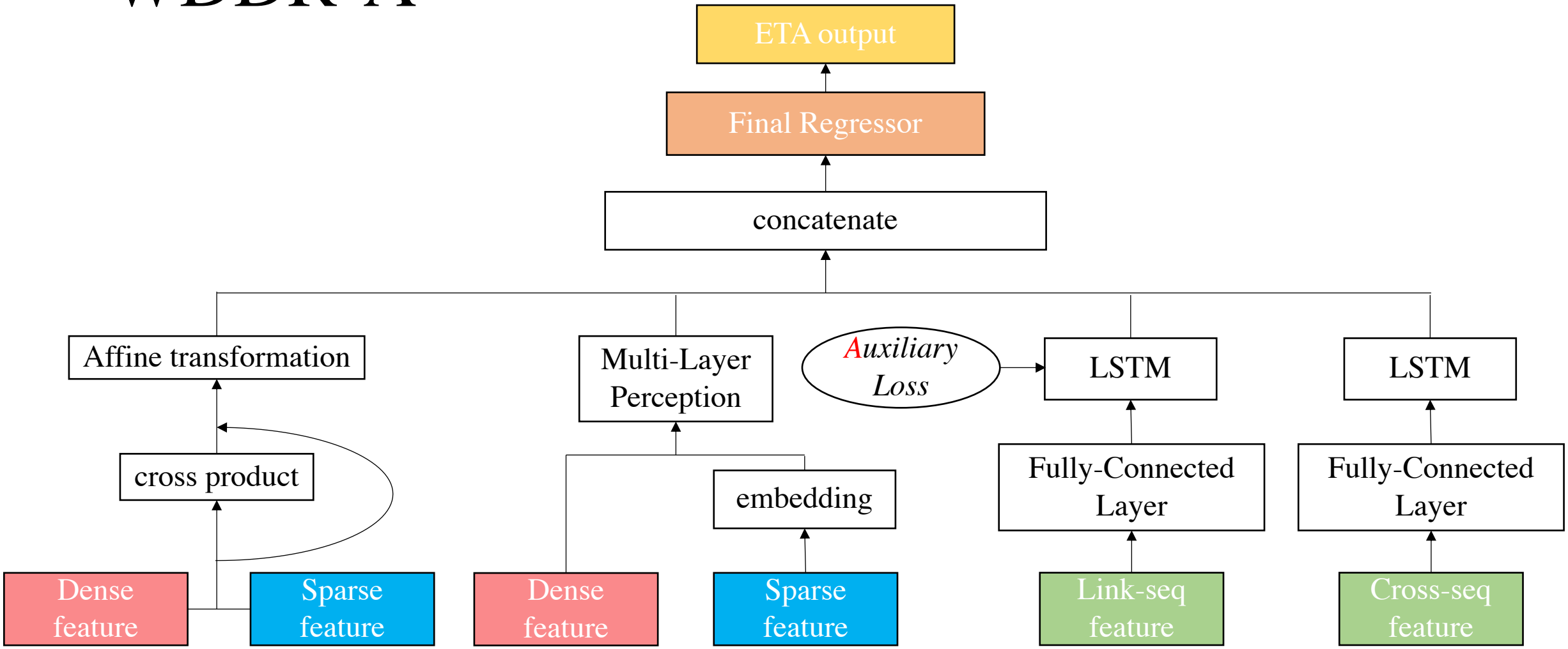
Problem setting

- Problem purpose: Accurately predict the trip duration time of each sample.
- Given features: The departure time, driver id, route links and average through time and traffic condition of each links.
- Dataset: The travel data of Shenzhen in china from August 1st to August 31st from DiDi platform.
- Metrics: MAPE $= \frac{1}{n}\sum_{i=1}^{n}\frac{|eta-ata|}{ata}$

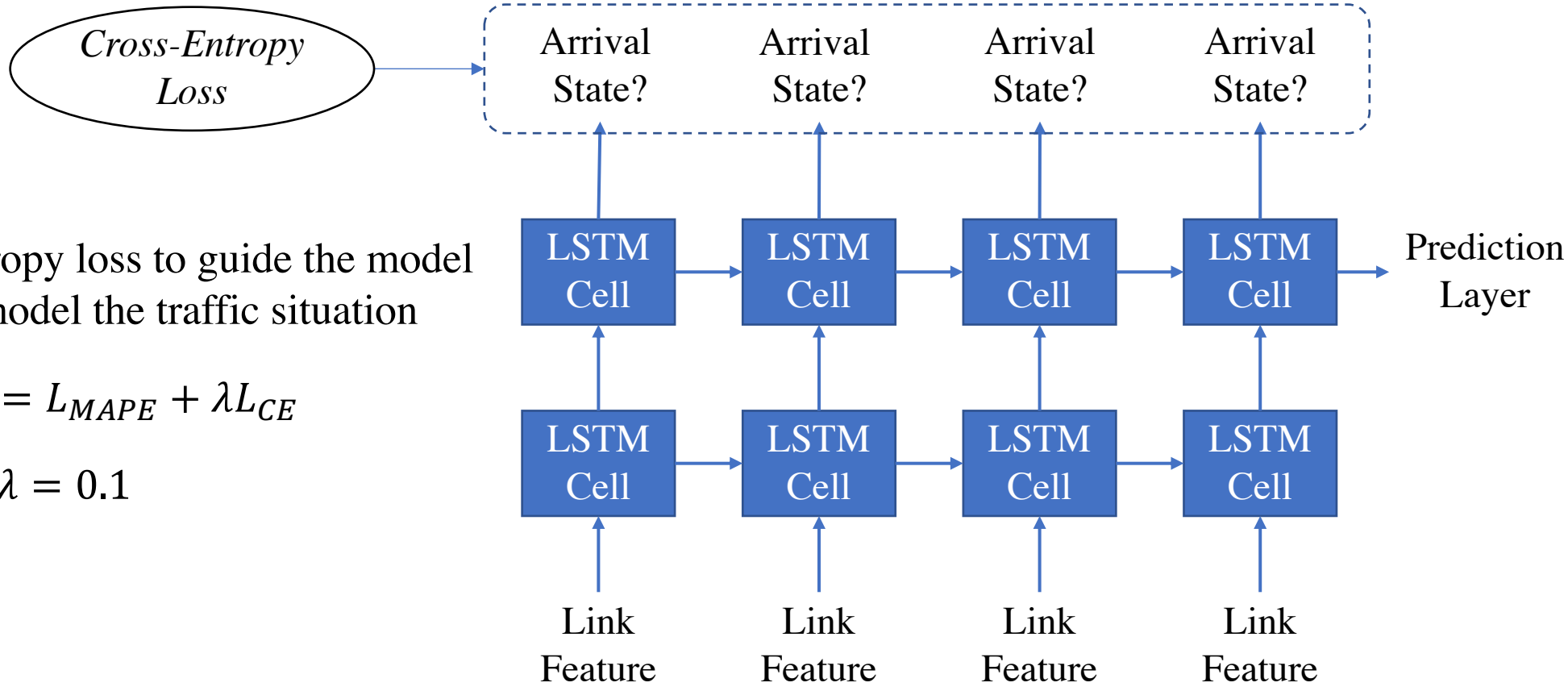
WDR


ETA output
Final Regressor
concatenate
Affine transformation
cross product
Dense feature
Sparse feature
Multi-Layer Perception
embedding
Dense feature
Sparse feature
LSTM
Fully-Connected Layer
Link-seq feature
Wide
Deep
Recurrent

WDDR-A


ETA output
Final Regressor
concatenate
Affine transformation
cross product
Dense feature
Sparse feature
Multi-Layer Perception
embedding
Dense feature
Sparse feature
Auxiliary Loss
LSTM
Fully-Connected Layer
Link-seq feature
LSTM
Fully-Connected Layer
Cross-seq feature
Wide
Deep
Recurrent-Link
Recurrent-Cross

Auxiliary loss



Use cross-entropy loss to guide the model to implicitly model the traffic situation

$$L = L_{MAPE} + \lambda L_{CE}$$

In our model, $\lambda = 0.1$

Cross Embedding

Link Feature

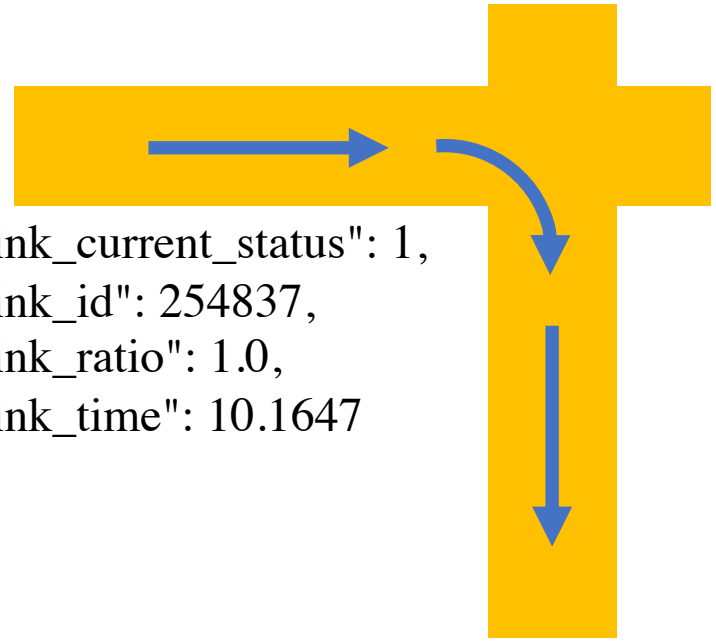
LinkID Embedding	status/ratio/time…

254837

LinkID Embedding	status/ratio/time…

448818

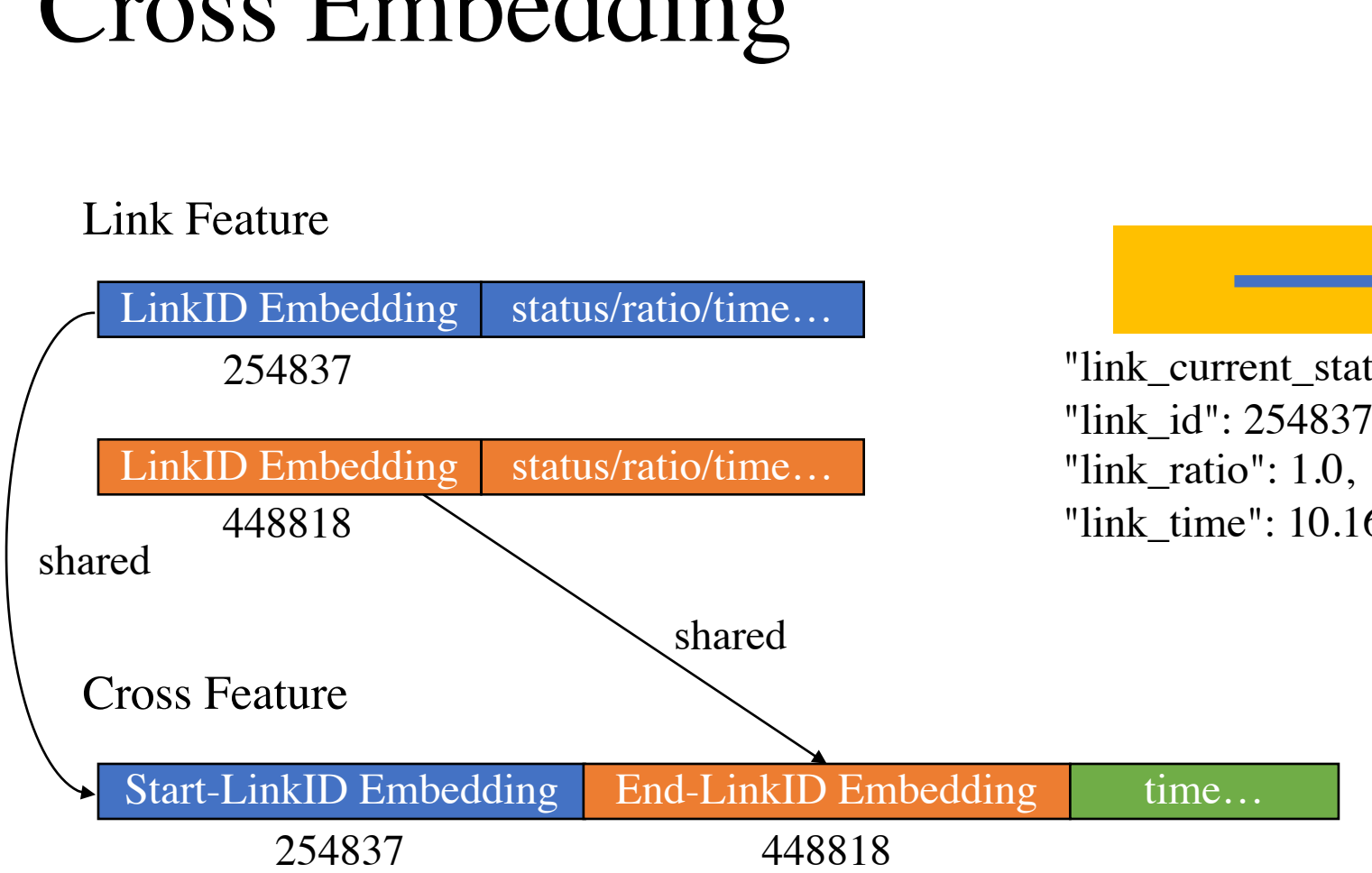
"cross_id": "254837_448818",
"cross_time": 35.0



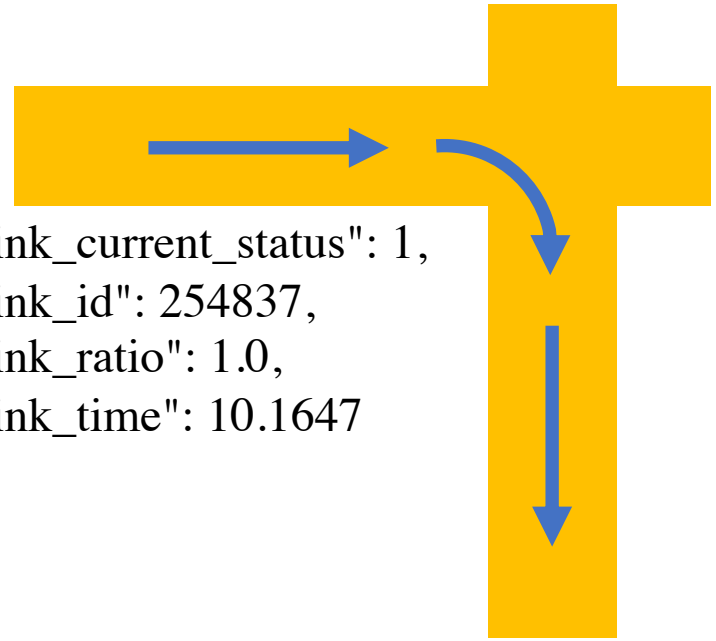
"link_current_status": 1,
"link_id": 254837,
"link_ratio": 1.0,
"link_time": 10.1647

"link_arrival_status": 1,
"link_current_status": 1,
"link_id": 448818,
"link_ratio": 1.0,
"link_time": 6.075

Cross Embedding



"cross_id": "254837_448818",
"cross_time": 35.0



"link_current_status": 1,
"link_id": 254837,
"link_ratio": 1.0,
"link_time": 10.1647

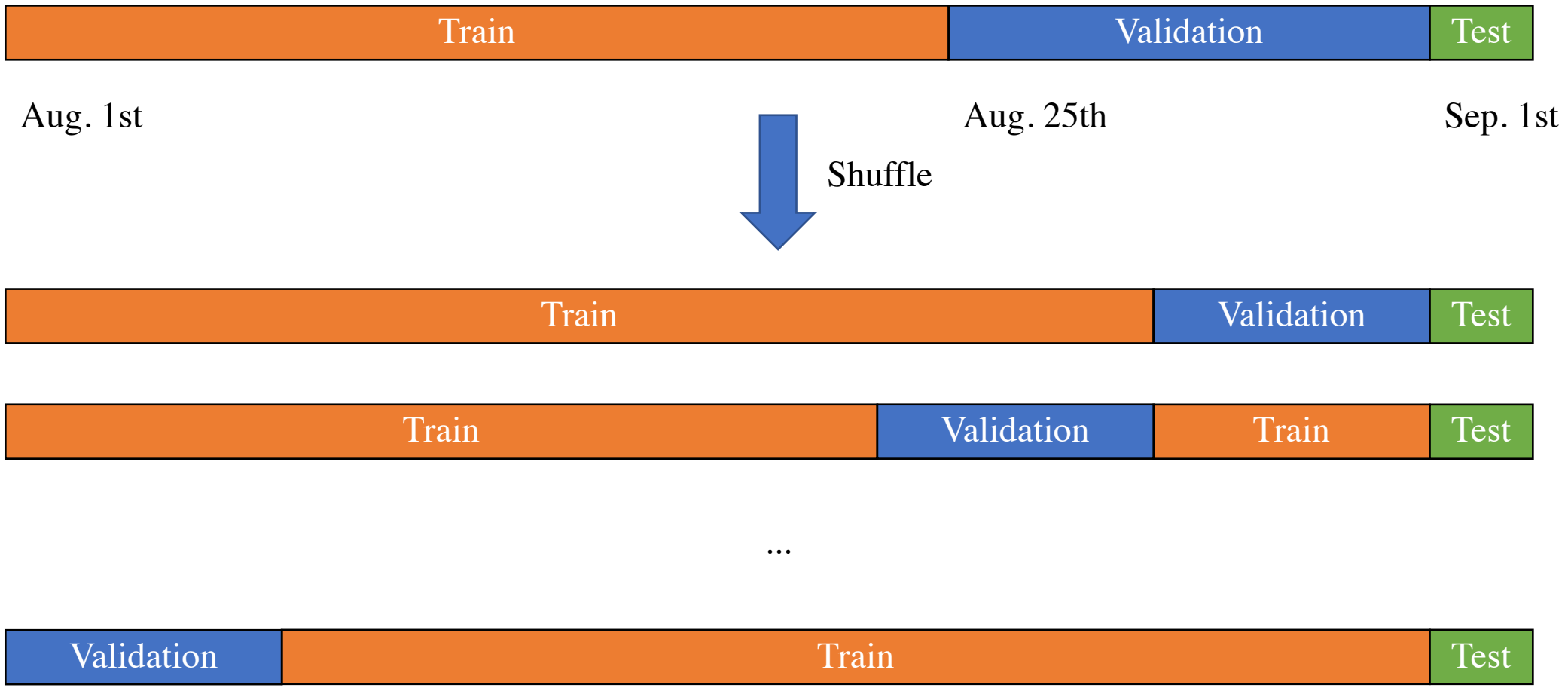
"link_arrival_status": 1,
"link_current_status": 1,
"link_id": 448818,
"link_ratio": 1.0,
"link_time": 6.075

Training

Hyper parameters

Batch size: 512
Optimizer: Adam
Learning rate: 1e-4
Epochs: 25
Time decay: 0.98

K-Fold Bagging



Numerical Result

Performance Comparison

Model	MAPE
WDR	12.103%
WDR-A	12.045%
WDDR	12.037%
WDDR-A	**12.000%**

The MAPE test on single-fold local CV

Improvement of K-fold Bagging

Bagging Strategy	MAPE
Single Fold	12.378%
5 Fold	12.161%
10 Fold	12.138%
Model Fusion (15 model)	**12.086%**

The MAPE test on test set

Note: A means auxiliary loss

Future Work

- Use graph neural network to model road network
- Use fine-grained weather information
- Add geographic information such as latitude and longitude, road length, etc.

Thank You