# When Will You Arrive?
# Estimating Travel Time Based on Deep Neural Networks

Dong Wang
Duke University

# Travel Time Estimation (TTE)

**TTE is a long-standing and critically important topic in the area of Intelligent Transportation Systems**

- Dispatch taxis to passengers in shortest time
- Better planning the routes, avoiding congested roads
- Help to alleviate urban traffic congestion

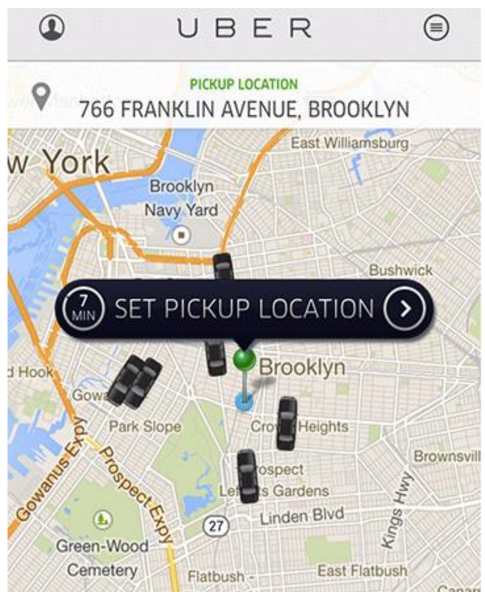
Bing maps

高德地图 amap.com

滴滴

Google Maps

UBER






Achieved 3/1926 in DataCastle Competition 2017; AAAI18

# Inputs

- Objective

  **Given**: 1. path (sequence of locations)

  2. start time    3. driver(optional)

  **Estimate**: the travel time

- Data

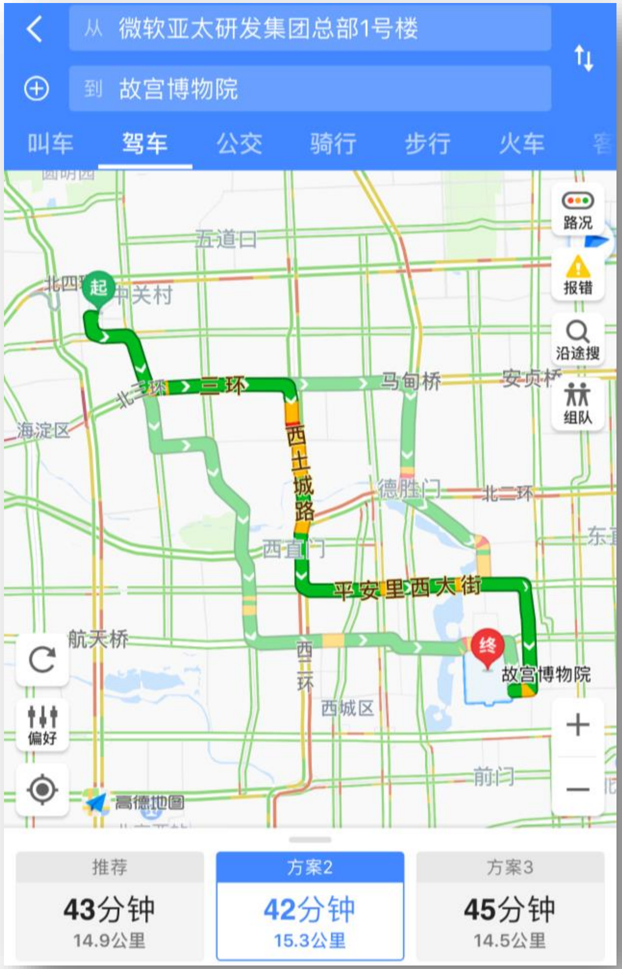
  GPS trajectory:

  – a sequence of GPS points

  – GPS point: latitude, longitude, timestamp, driver ID(optional).

  – Sample a GPS point each 200m~400m

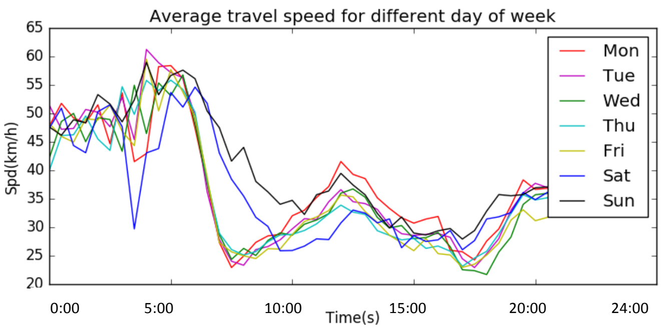
  – We use the timestamp as the ground truth

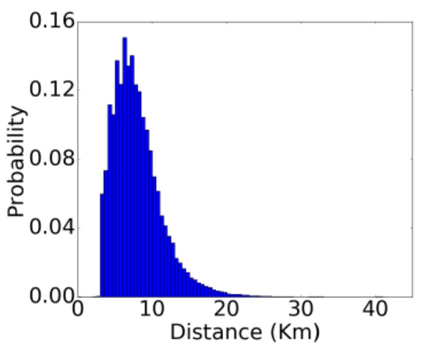


# Challenges

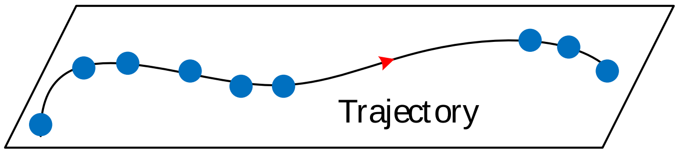
- Balance between collective vs individual
  - Estimate individual roads or the entire path
- Data imbalance problem
  - Over 80% GPS records fall on <20% roads
- Diverse influences
  - Complex spatial and temporal correlation
  - External factors

    day of the week, starting time, driver, distance
- Variable length
  - Trajectory lengths vary from 2km to 41km.

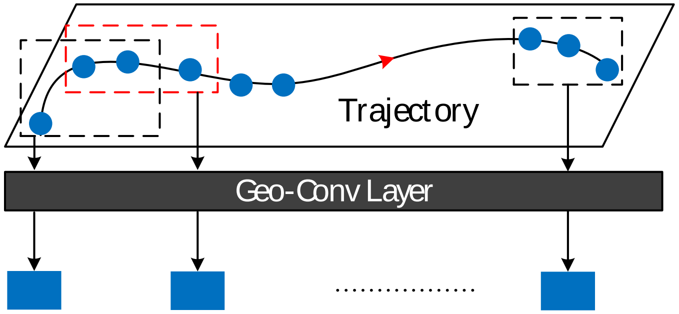



# DeepTTE Architecture

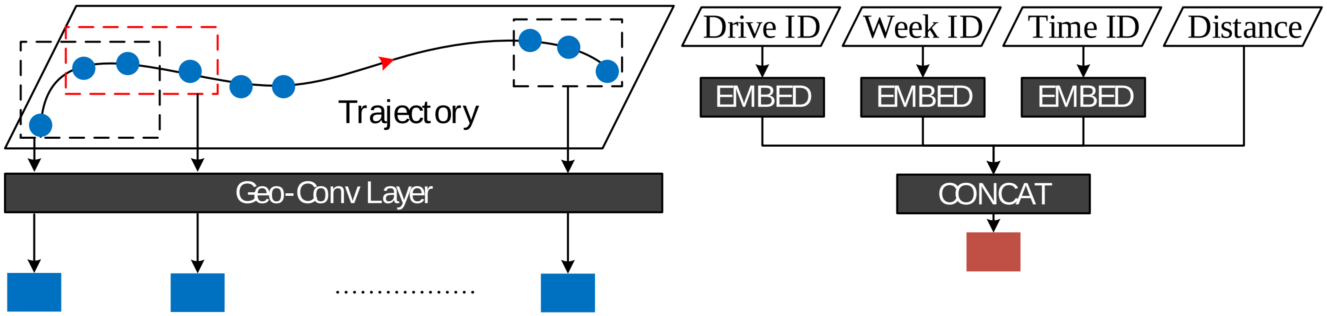


# DeepTTE Architecture



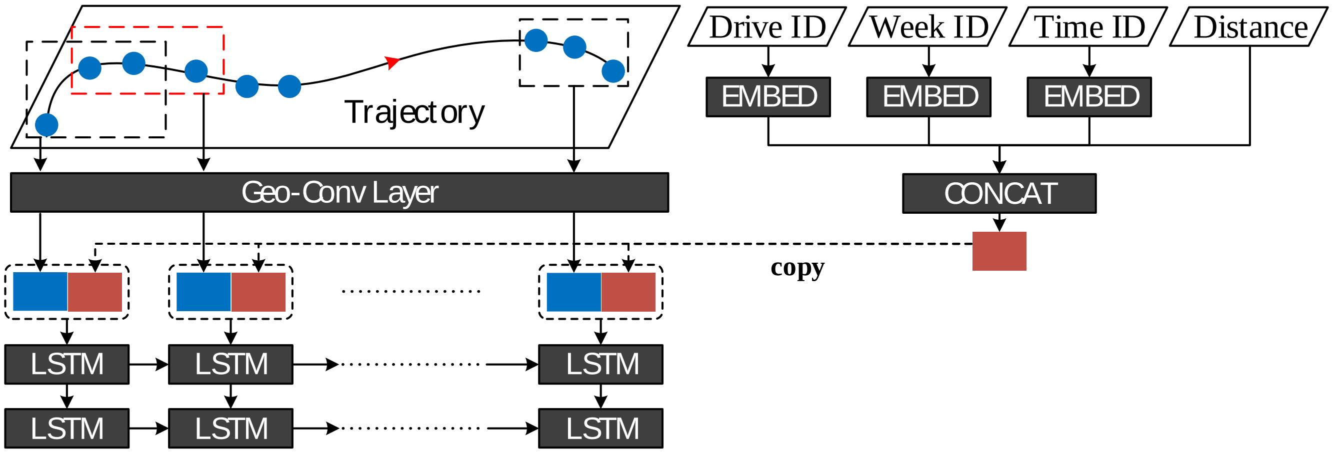

- Capture spatial dependencies

# DeepTTE Architecture



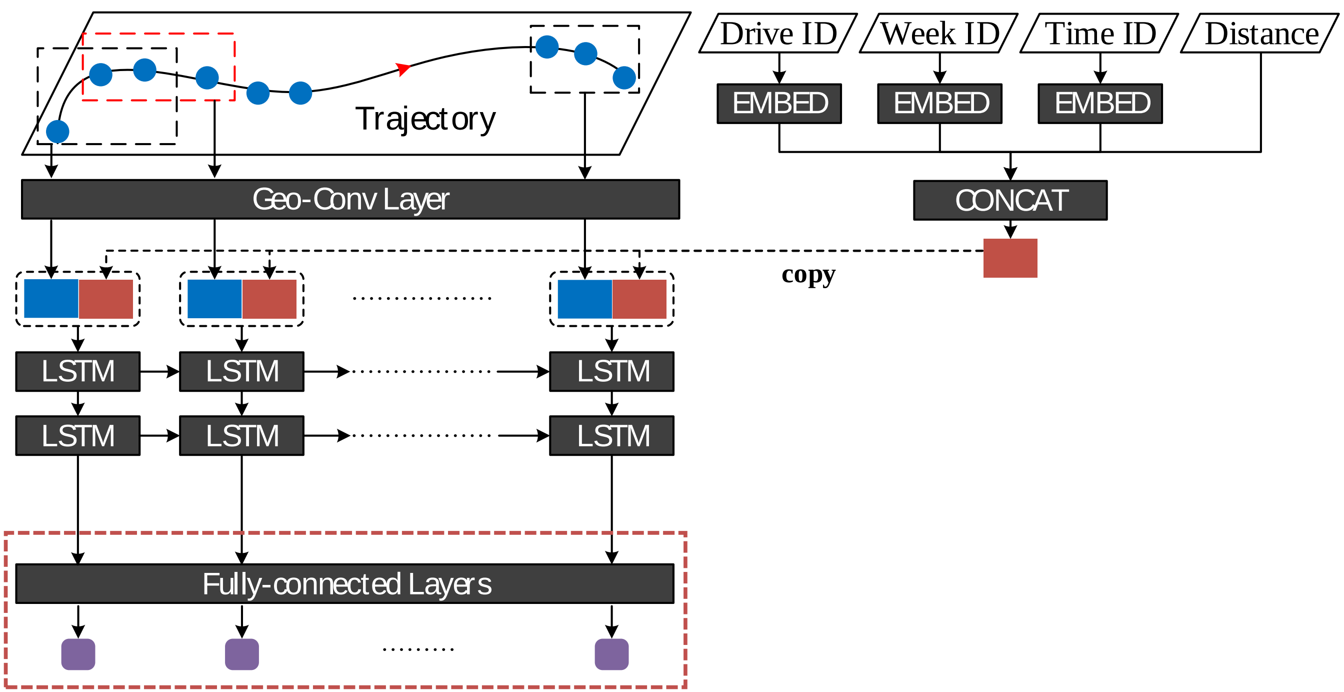

- Capture spatial dependencies
- Handle external factors & share similar pattern

# DeepTTE Architecture




- Capture spatial dependencies
- Handle external factors & share similar pattern
- Learn temporal dependencies

# DeepTTE Architecture



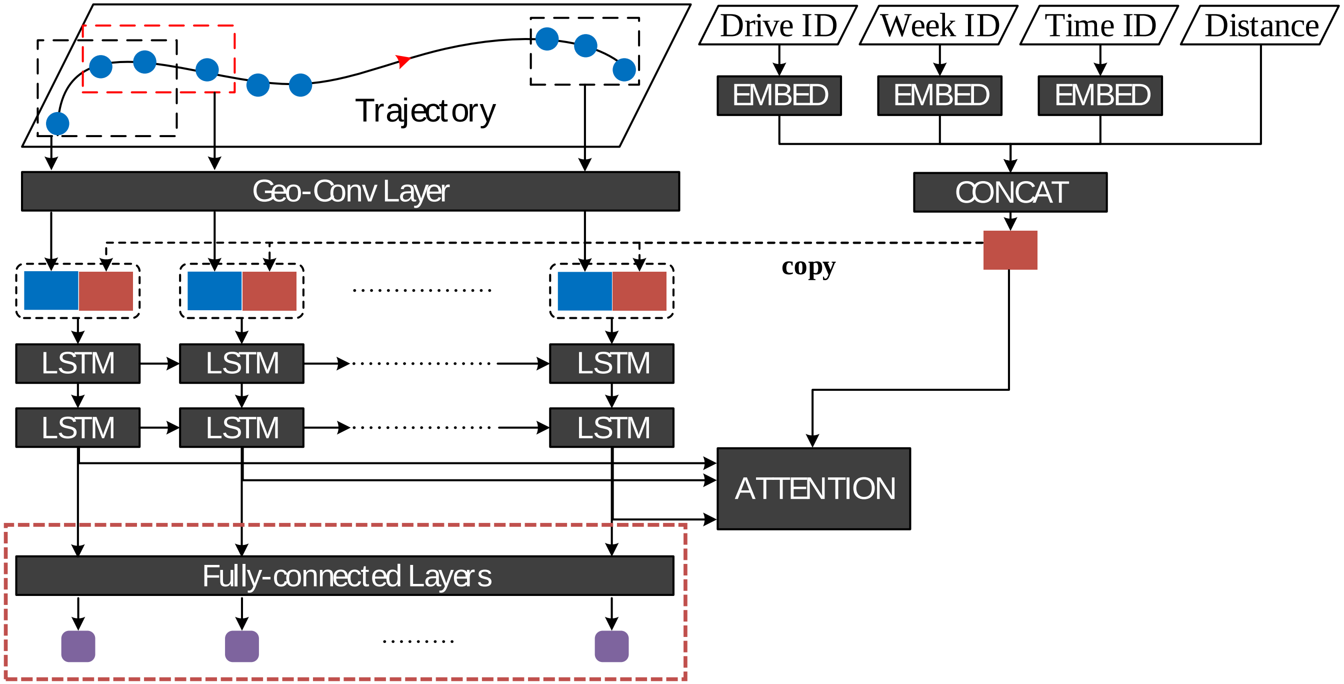

TTE for each local path

- Capture spatial dependencies
- Handle external factors & share similar pattern
- Learn temporal dependencies
- Address imbalance data problem

Individual:
Local error may accumulate

Collective:
Data sparse

# DeepTTE Architecture



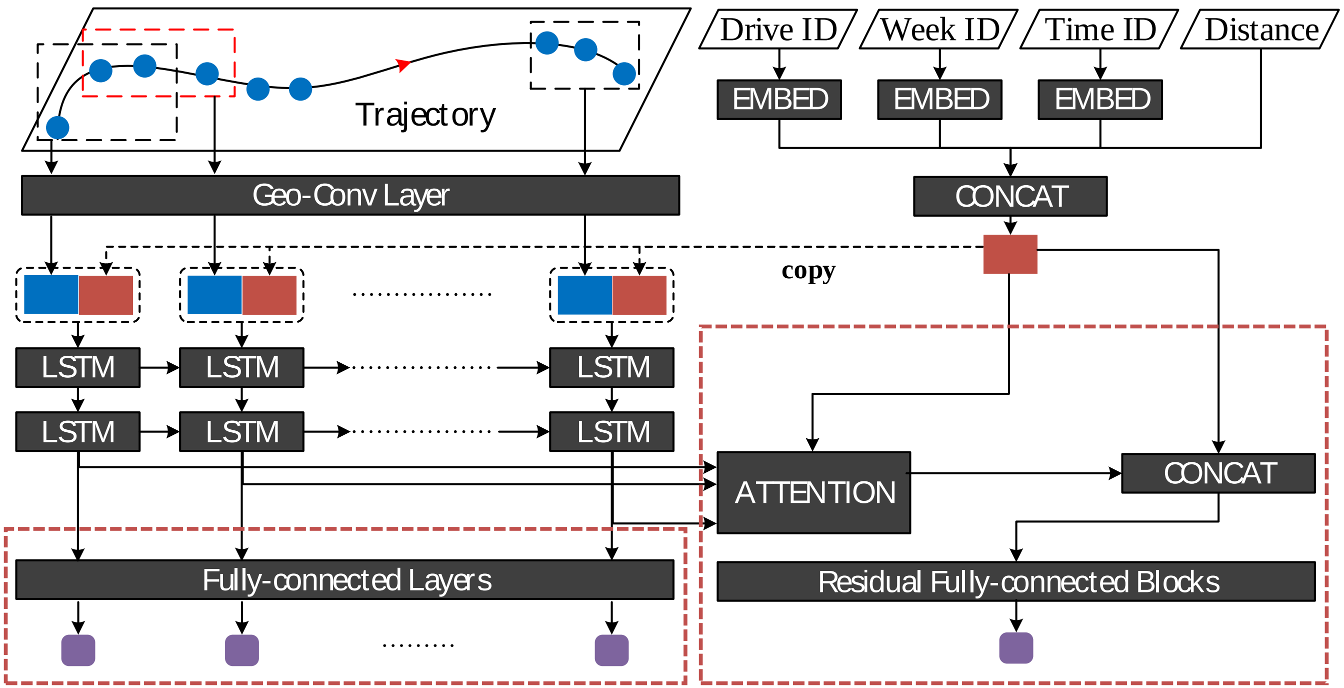

TTE for each local path

- Capture spatial dependencies
- Handle external factors & share similar pattern
- Learn temporal dependencies
- Address imbalance data problem
- Learn weights for different local path

Individual:
Local error may accumulate

Collective:
Data sparse

# DeepTTE Architecture




TTE for each local path

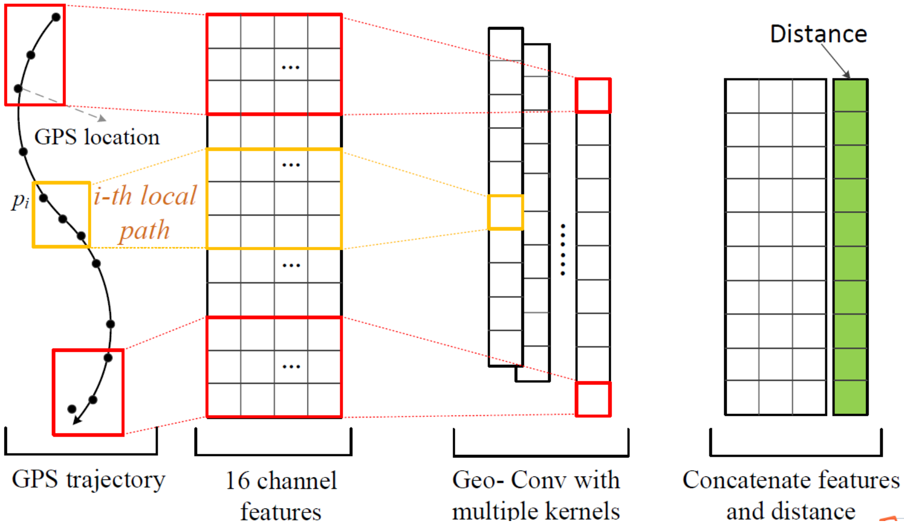
TTE for the entire path

- Capture spatial dependencies
- Handle external factors & share similar pattern
- Learn temporal dependencies
- Address imbalance data problem
- Learn weights for different local path
- Get a better estimation result

Individual:
Local error may accumulate

Collective:
Data sparse

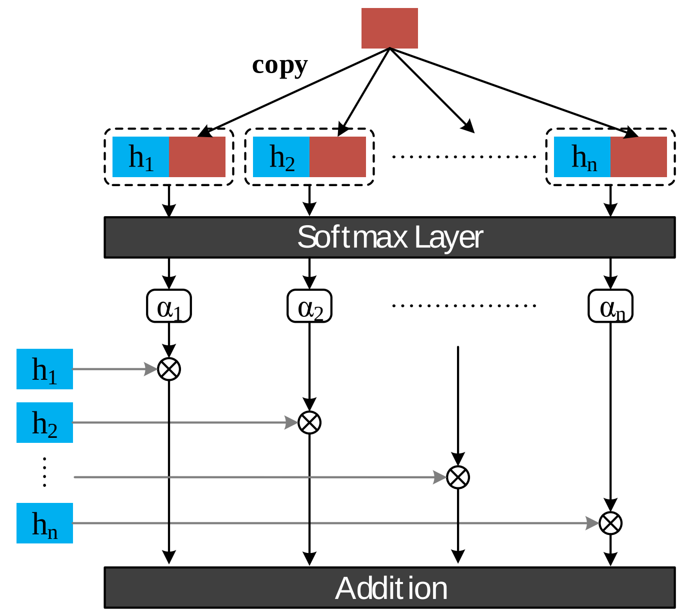
# Geo-Conv Layer




- Transforms the raw GPS sequence to a series of feature maps.
- Capture spatial correlations of each local path

# Attention Mechanism


Local path with more intersections or in extremely congested need more attention.



# EXPERIMENT

Data Description

- 1.4 billion GPS records of 14,864 taxis in Oct. 2014 in Chengdu.
- Total number of trajectories: 9,653,822. (60GB)
- Use the last 7 days (from 24th to 30th) as the test set
- 5-folds cross validation on the rest data
- Pytorch 2.0 + GeForce 1080 GPU

# Baseline methods

- **AVG**:

  total distance divided by the average speed for related weekID and timeID) **28.18%**

- **TEMP[1]**:

  collective estimation base on neighbor trajectories **22.82%**

- **Gradient Boosting:** **20.32%**
- **AttrTTE**:

  attribute component with one layer MLP **17.63%**

- **DeepTTE** **11.92%**

[1] A Simple Baseline for Travel Time Estimation using. Large-Scale Trip Data. Hongjian Wang, Zhenhui Li, Yu-Hsuan Kuo, Dan Kifer. SIGSpatial 2016.

# Effects of Geo-conv layer and LSTM

- Without Geo-conv layer: 12.74 %
- Without LSTM layers: 15.06%
- With both Geo-conv and LSTM: **11.92%**