Introduction to Traffic Modeling & Prediction

Suining He
Department of Computer Science and Engineering
University of Connecticut
suining.he@uconn.edu

Overview

- Smart transportation
- Traffic Prediction Modeling
- Case Studies: Ride Sharing

Smart Transportation

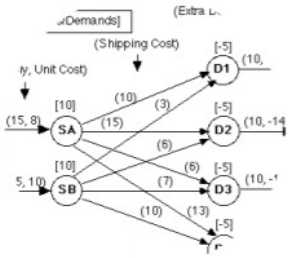
Engineering

Policy Making

Planning

Transportation

Design



Science

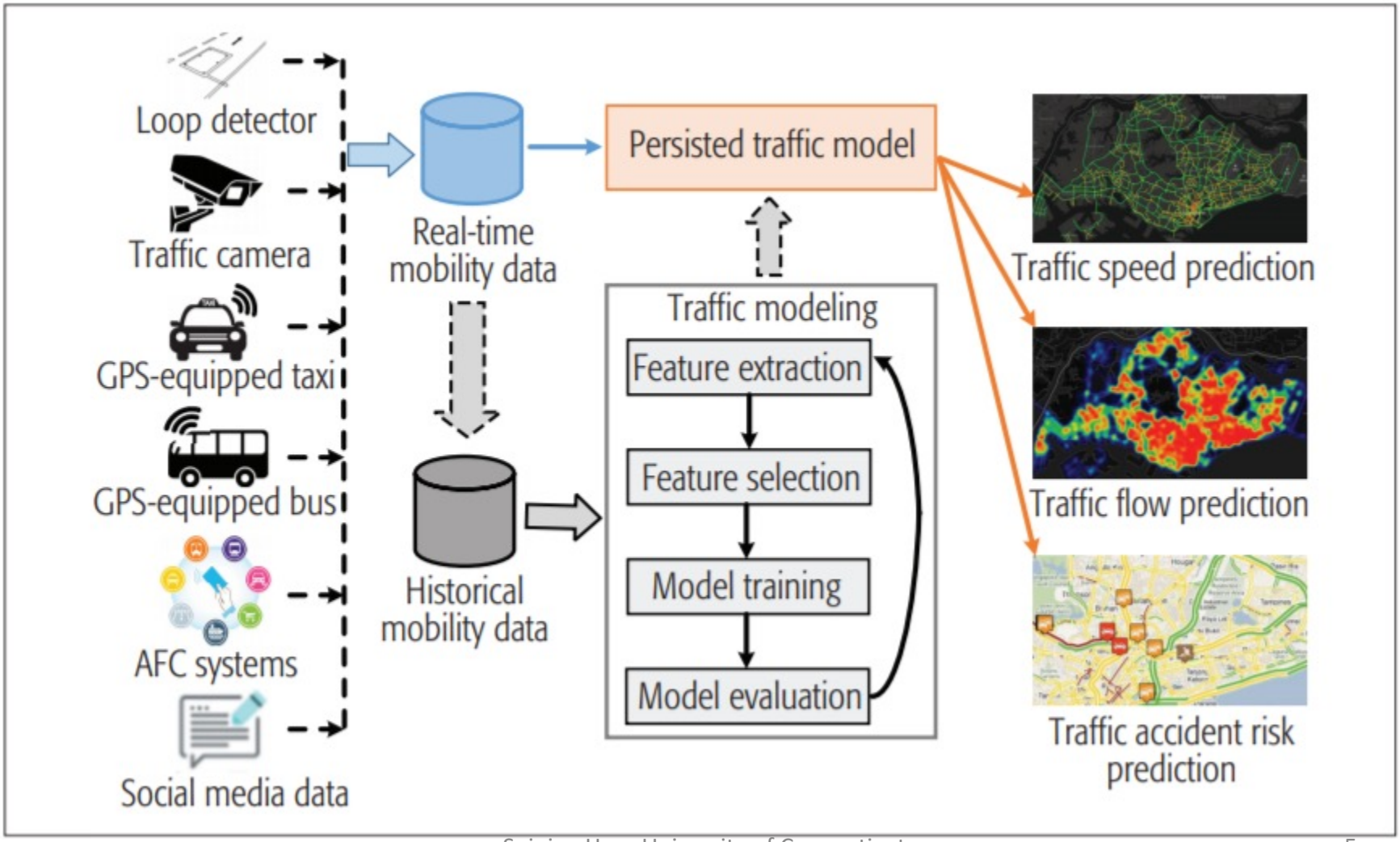
Management

.....

Cutting Edge Issues & Players

- Issues:
  - Cooperative vehicle-highway systems
  - Ride sharing
  - Multi-modal transportation
- Smart transportation system
  - Smart vehicles
  - Smart infrastructures
  - Smart travelers

Traffic Prediction Modeling

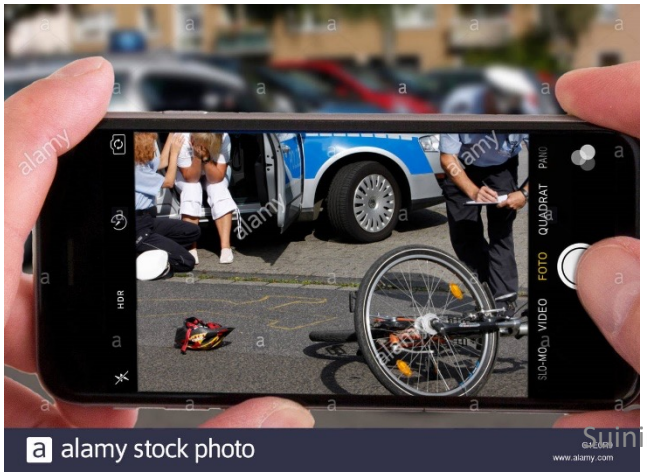


Mobility Data Collection

- Road infrastructures
  - Loop detectors, traffic cameras, radars
  - Electronic Toll Collection (ETC)
  - Sparse coverage
- Ubiquitous sensing and intelligent transportation systems
  - Mobile devices with GPS enabled
  - Automatic fare collection (subways/buses/taxis)

Other Data Sources

- Accident reports
- Social networking
- Cellphone data
- Crowdsourcing system



Problems in Traffic Prediction

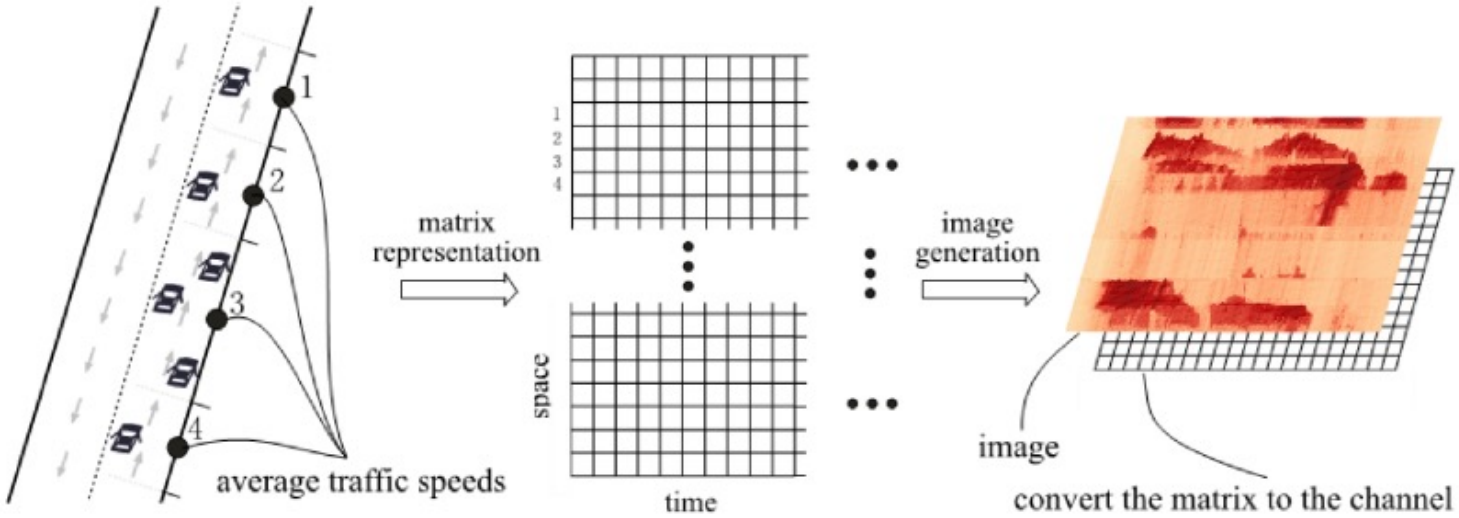
Category	Involved data sources	Desired output
Traffic speed prediction	Infrastructures, GPS-equipped vehicles	Average traffic speed (or congestion level)
Traffic flow prediction	Infrastructures, AFC systems	Total number of objects passing through a road/region
Traffic accident risk prediction	Infrastructures, AFC systems, social media data, historical accident reports	Accident risk probability for each road/region

Liu, Zhidan, et al. "Urban traffic prediction from mobility data using deep learning." *IEEE Network* 32.4 (2018): 40-46.

Traffic Speed Prediction

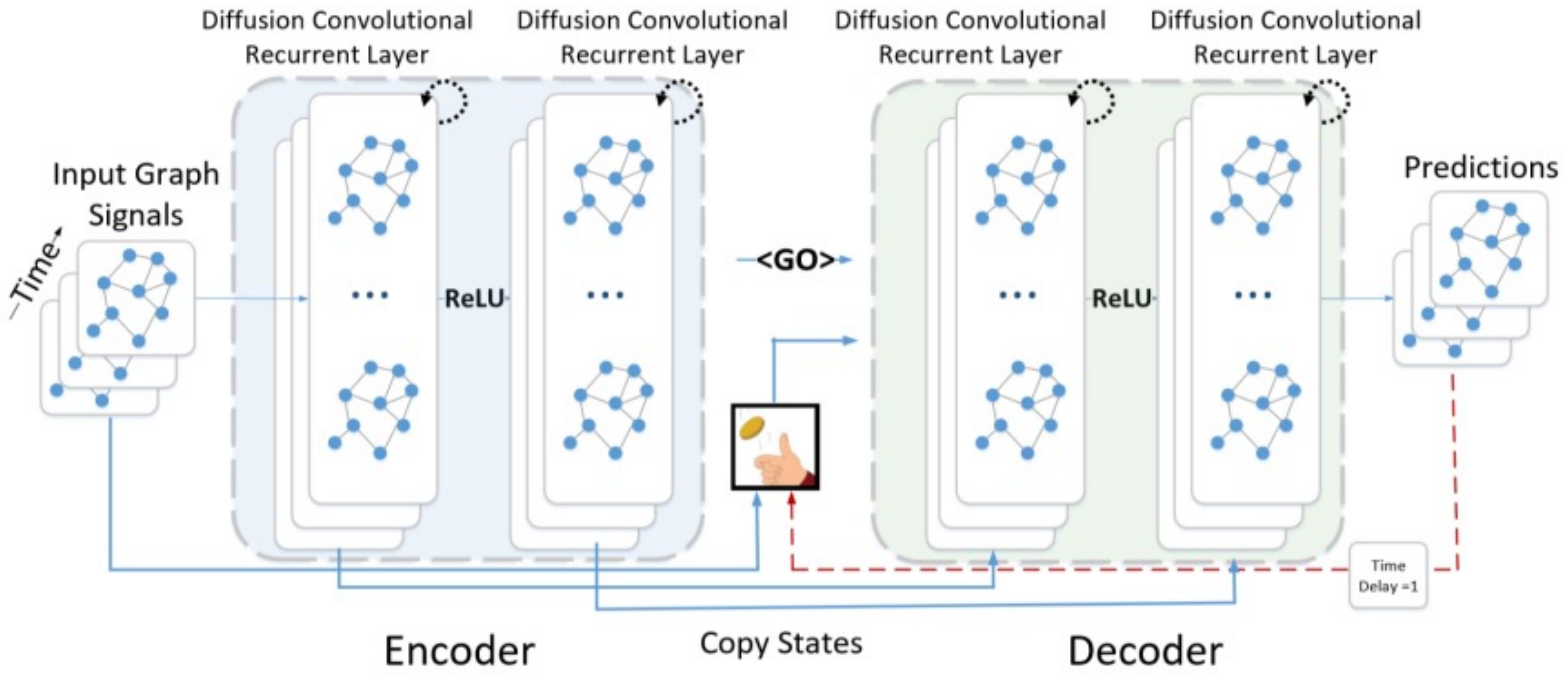
- General definition:
  - Average travelling speed of all sampling vehicles on a given road segment
  - Fine-grained
- Input:
  - Loop detectors and cameras
  - GPS-equipped vehicles
- Output:
  - Future traffic speeds
  - Congestion levels (slow, normal and fast)

Image-based Traffic Forecasting



- Model traffic speeds in different locations as a matrix
- Apply convolution to model the spatio-temporal dependency

Graph-based Traffic Prediction



- Spatial dependency: diffusion convolution on graph
- Temporal dependency: augmented recurrent neural network

Traffic Flow Prediction

- General definition:
  - Total number of target objects (vehicles or humans) that pass through an area during a period
  - Coarse-grained (in/out-flows)
- Input:
  - Flows through an area (a road segment or a region)
  - Mobility data from infrastructures and AFC systems
- Output:
  - Movements of crowds
  - Traffic distributions

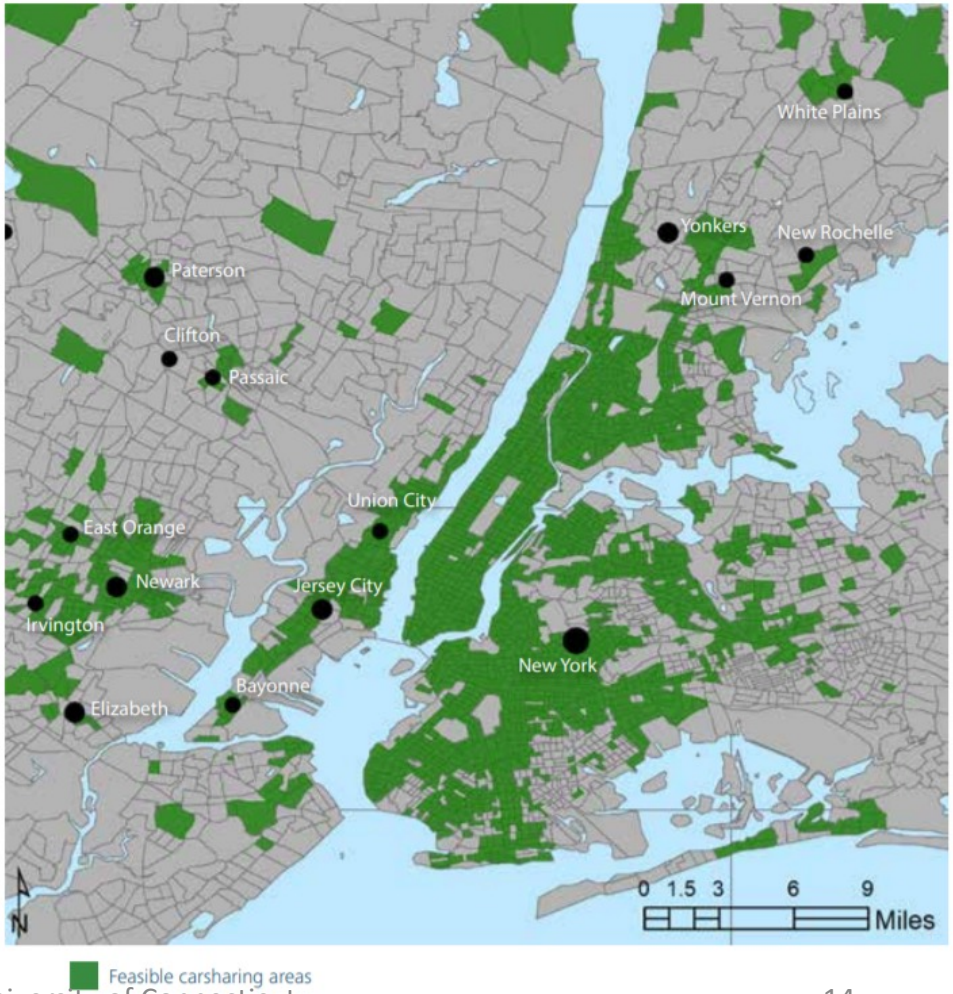
Traffic Accident Risk Prediction

- General definition:
  - How likely traffic accidents might occur on a road/region
  - Fine/coarse-grained
- Input:
  - Mobility data (current traffic conditions and human mobility)
  - Historical accident reports
- Output:
  - Likelihood of traffic accidents

Case Studies: Ride Sharing

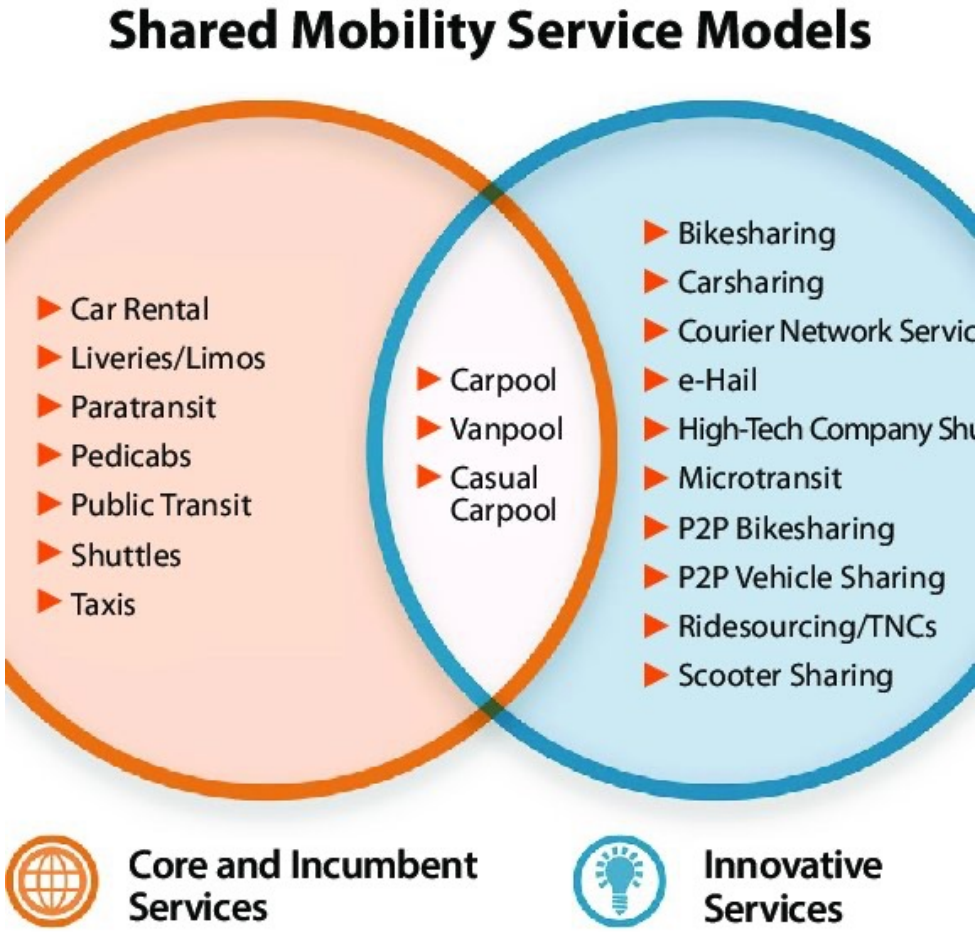
- Using personal automobiles wisely
  - Environment, gasoline, traffic
  - Congestions: $78 billion waste in 2007 in US
- Low private car occupancy rates

Figure 7. Carsharing feasibility, New York metropolitan area



Ride Sharing Business Models

- 1600s: Taxi industry
- 1891: Invention of taximeter
- 1970s: Carpooling
- 2000s: Carsharing
- 2010s: Peer-to-peer ride sharing



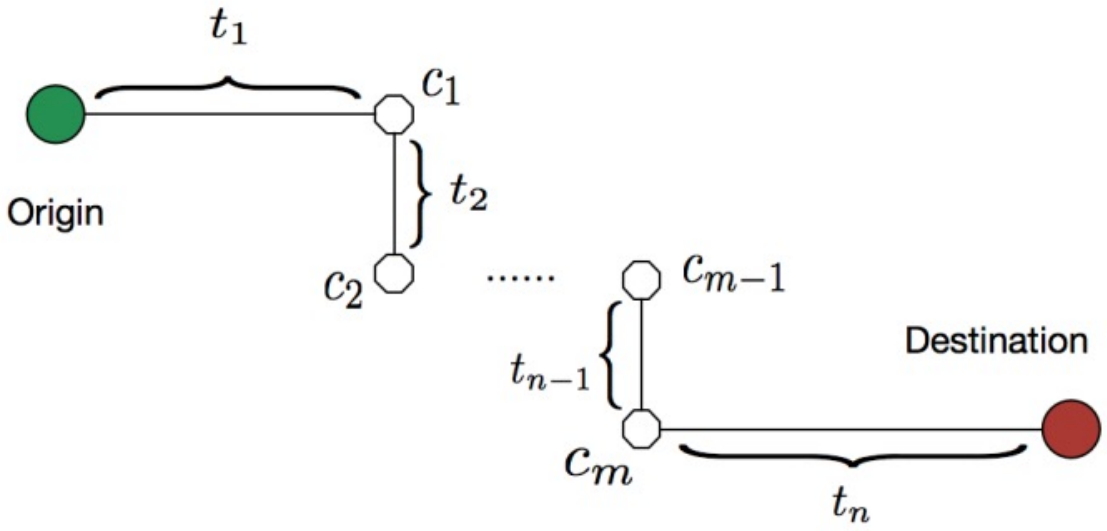
Estimated Time of Arrival (ETA)

- Essential for ride sharing
  - Route planning, order dispatching, pricing
- General approach:
  - Based on the speed by which it has covered the distance traveled so far
- Designs
  - How to measure spee
  - Given input locations
  - Unexpected events
  - Dynamic

ETA Models

- Additive models
  - Rule-based additive models: explicitly modeling the segments in a path
  - Aggregating the time of sub-paths
  - ML models for the sub-path problem
- Global models:
  - Formulating ETA as a regression problem: Learning to estimate the travel times
  - Simple regression model and deep learning
- Path-free models (if path is not available)

Simple Additive Model



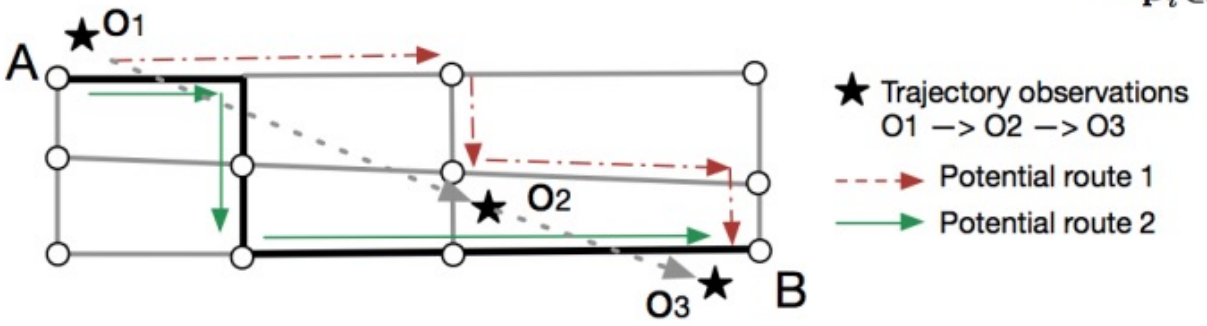
$$ETA = \sum_{i=1}^{n} t_i + \sum_{j=1}^{m} c_j$$

road segments crossings

- Simple rules based on physical structure of road network
- Popular solution in digital map industry
- Challenges: precise speed estimation and error accumulation

Aggregating Sub-paths

$$\mathcal{N}(\mathbf{q}) = \{\mathbf{p}_i \in \mathcal{D} | dist(o_i, o_q) \leq \tau \text{ and } dist(d_i, d_q) \leq \tau\}, \qquad \widehat{t_q} = \frac{1}{|\mathcal{N}(\mathbf{q})|} \sum_{\mathbf{p}_i \in \mathcal{N}(\mathbf{q})} w_i t_i.$$



- The path can be decomposed into sub-paths
- Time of each sub-path is inferred from the history
- ETA: summation

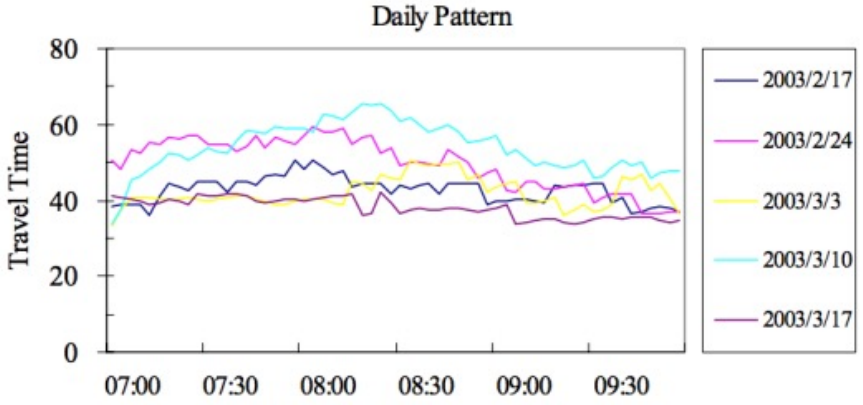
Time Series Prediction

- Regression (like support vector regression)
- Deep learning: recurrent neural network

selected routes

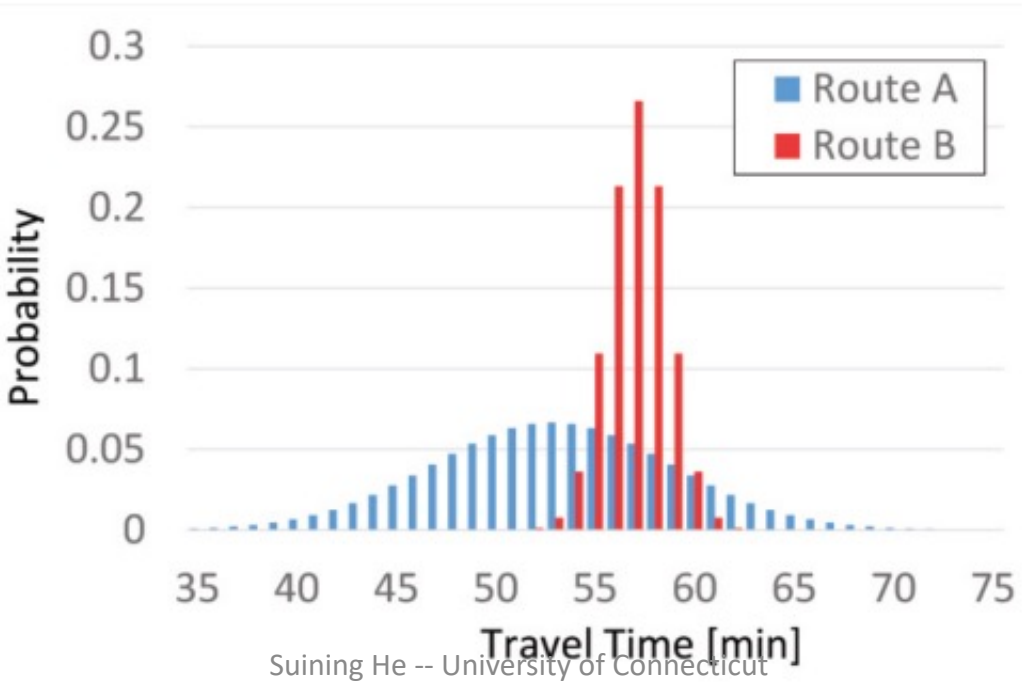
f(t-n),..,f(t-1) → f(t)

f(t-n),..,f(t-m) → f(t)



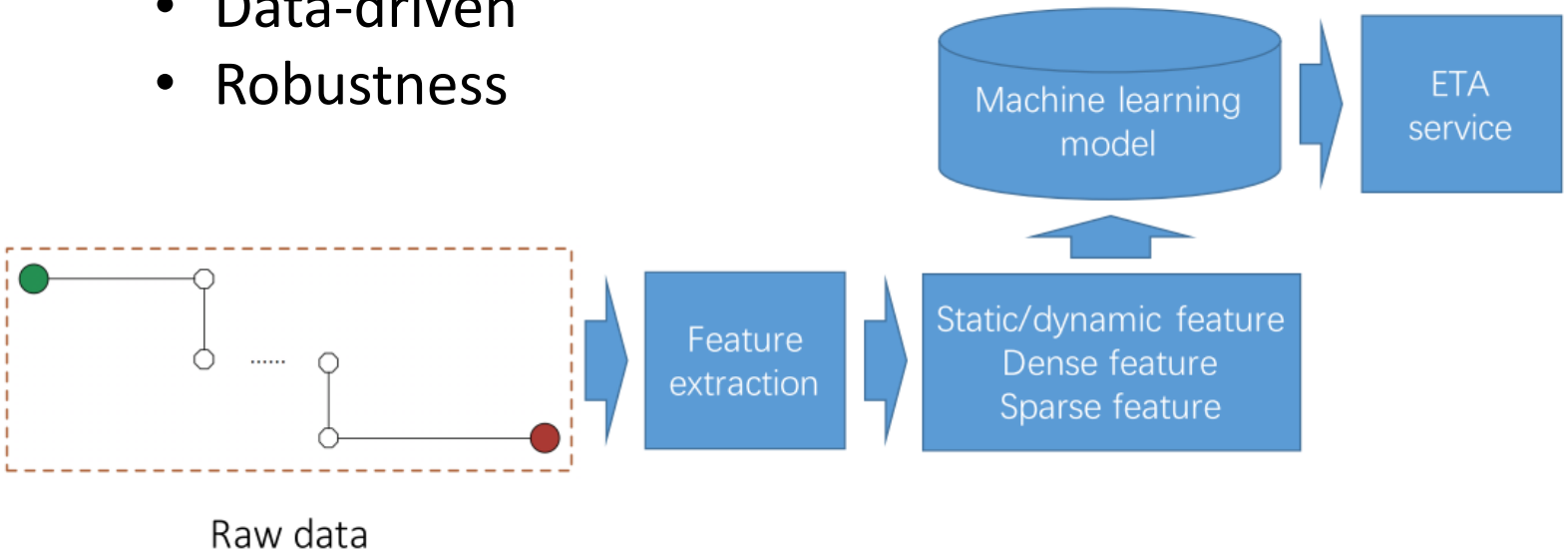
Probabilistic ETA

- Output a distribution of ETA
- Variance of travel time
- Gaussian distribution for link travel time

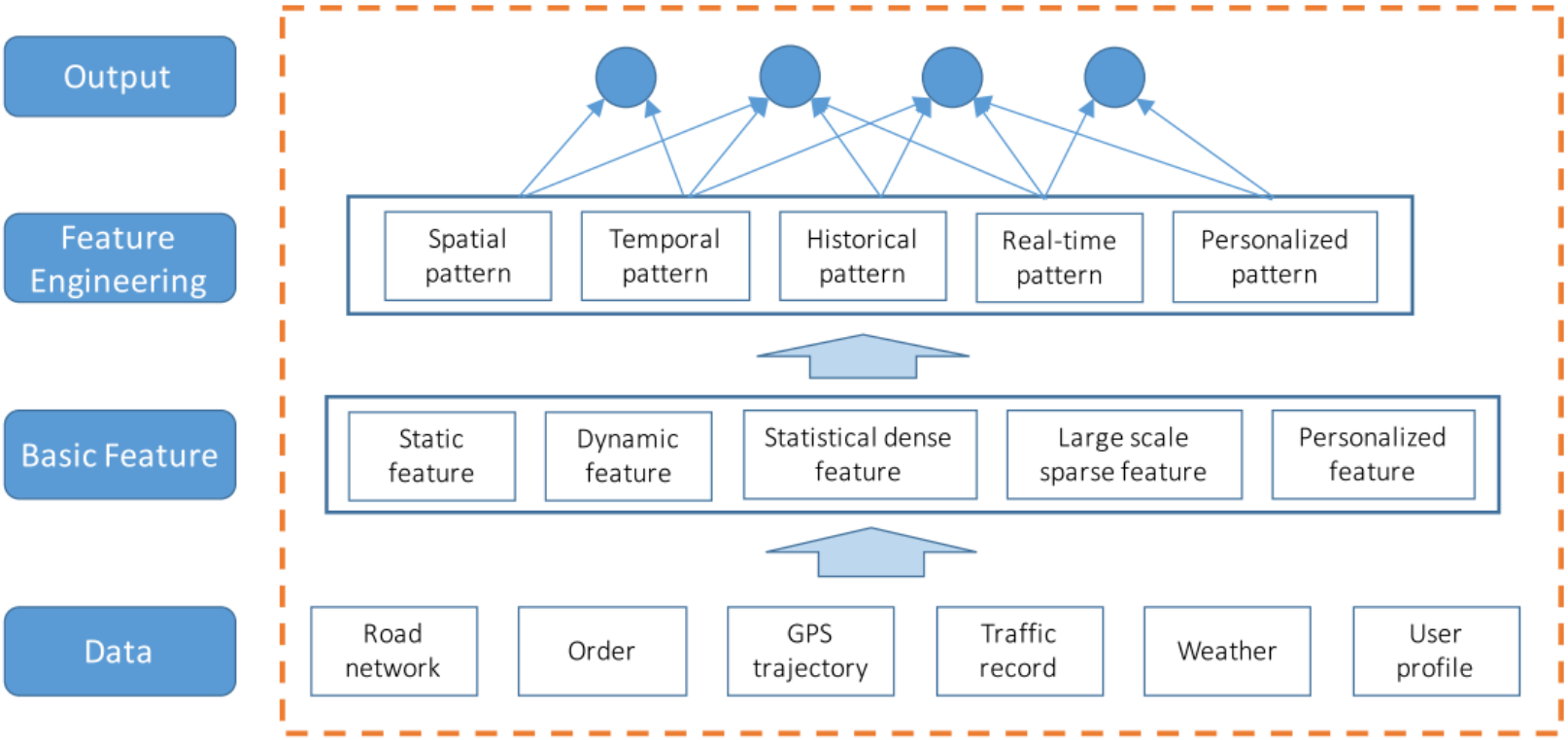


Learning to Estimate Travel Time

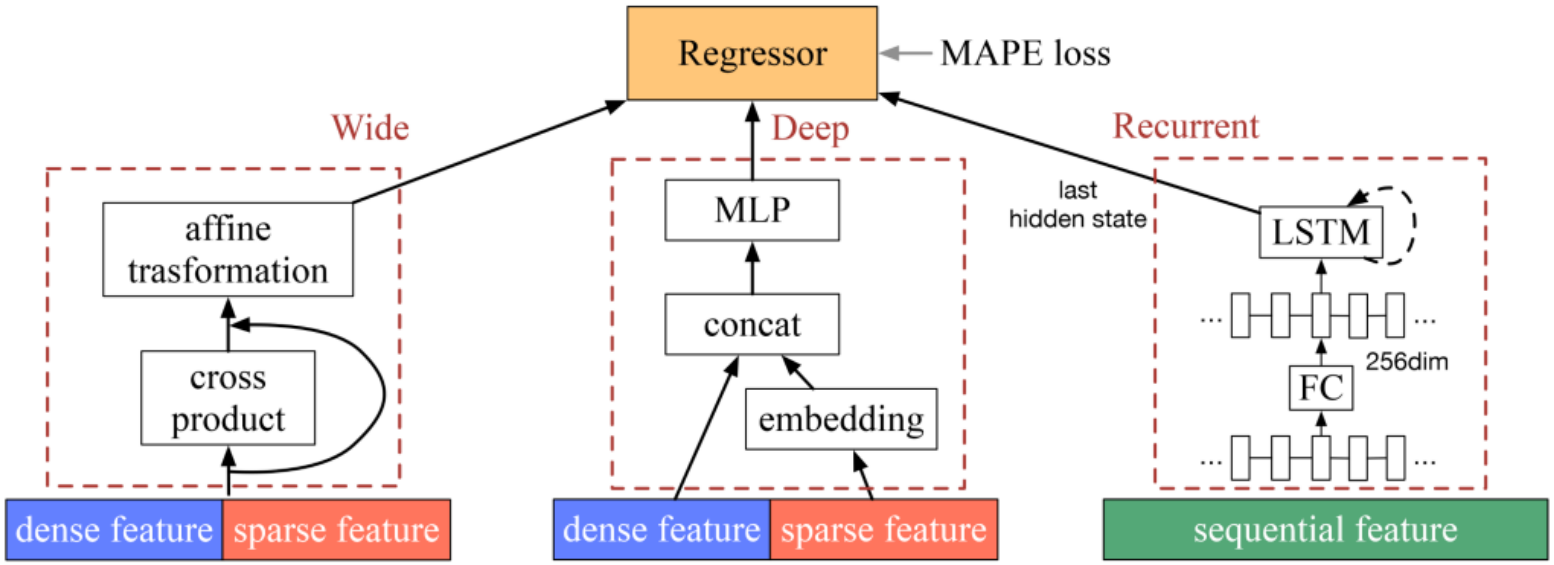
- Big data + machine learning
  - High accuracy
  - Data-driven
  - Robustness



Comprehensive Modeling



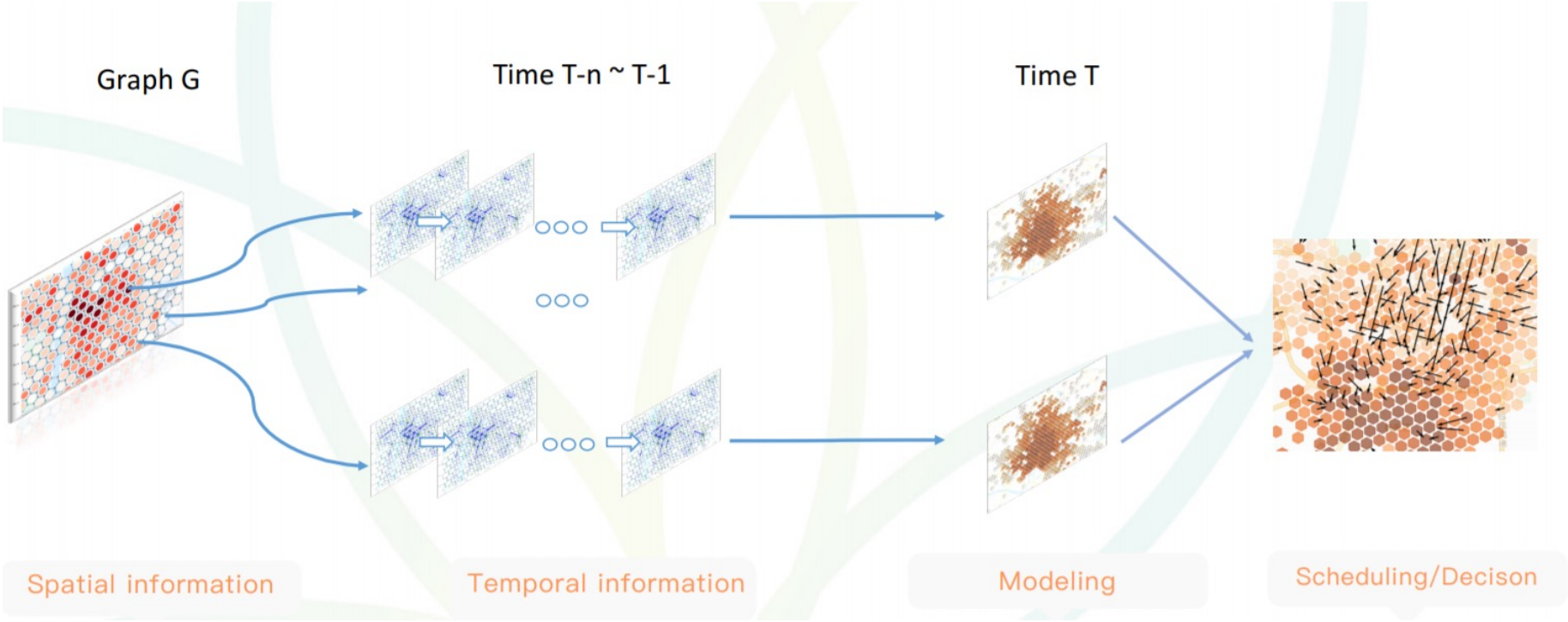
Wide-Deep-Recurrent Network



- Wide: dense features (road network/traffic)
- Deep: sparse features (personalized/orders)
- Recurrent: sequential features (time-related)

Learning to Estimate the Travel Time, KDD 2018

Spatio-Temporal Data Mining



Summary

- Smart transportation
- Traffic Prediction Modeling
- Case Studies: Ride sharing