Intelligent Trip Modeling for Prediction of Origin-Destination Traveling Speed Profile

-- an application of data Science in transportation systems

Yi L Murphey, Professor
Department of Electrical and Computer Engineering
Associate Dean for Graduate Education and Research
College of Engineering and Computer Science
University of Michigan-Dearborn
Dearborn, MI 48128
yilu@umich.edu
Big Data Challenges in Intelligent Transportation Systems (ITS)

- Data Complexity in ITS
  - Volume, Variety, Velocity, Veracity
  - Road Geometry data
    - Traffic Message Channel (TMC), link, and shape points
    - GPS localizations, speed limit, number of lanes, curvature, altitude, link length, shape points, etc.
  - Weather condition data
  - Communication data
  - Traffic flow and incident data for states of Michigan and Ohio 24/7, 463MB/day
  - Real-time traffic data
  - Large collections of real-world driving data
    - V2V communication data
    - Weather condition data
    - Communication data
  - Real-time traffic data
  - Large collection of real-world driving data

For a trip of less than one hour, the signal data size is about 400MB.

- Vehicle, Environmental, Physiological, video
  - 368 channels of signals
  - Hundreds and thousands of trips from different drivers
  - Road Geometry data
  - Traffic Message Channel (TMC), link, and shape points
  - Weather condition data
  - Communication data
  - Traffic flow and incident data for states of Michigan and Ohio 24/7, 463MB/day
  - Real-time traffic data
  - Large collection of real-world driving data
  - V2V communication data
  - Weather condition data
  - Communication data
  - Real-time traffic data
  - Large collection of real-world driving data
A highly representative speed profile will be beneficial to advanced vehicle energy management. It is important to have an accurately predicted speed profile available at the beginning of the trip for many applications in Advanced Traveler Information Systems (ATIS).

- Provides the driver with a realistic estimation of traveling time.
- Allows the driver to make more informed decisions.
- Provides the driver with a realistic estimation of traveling time at the beginning of the trip.

Fuel economy (and optimal energy management) depends on driver demand, i.e., the drive cycle.

For route selection:
- Provide useful information to driver
- Estimate accurate trip time

I n ATIS

Driver demand, i.e., the drive cycle

<table>
<thead>
<tr>
<th>Speed(mph)</th>
<th>Distance(mile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>60</td>
<td>80</td>
</tr>
</tbody>
</table>


A highly representative speed profile will be beneficial to advanced vehicle energy management.
What is Trip Modeling?

Mathematically, trip modeling problem is to generate a sequence of travel time at time $t_0$ provided by the traffic sensors at location points in $\mathbb{R}$. A sequence of defined geographic locations along the chosen route can be:

- A sequence of traffic sensor locations along the chosen route
- A sequence of traffic sensor locations along the chosen route
- A set of traveling points along the chosen route, which can be

Following information:

- $F$ is a system that predicts the traffic speed at location $x$ at time $t_0$, the trip starting time, based on the

$$\{N(t) = d\left| \left( \begin{array}{c} 0 \rightarrow \Omega \rightarrow 0 \rightarrow d \end{array} \right) \right| dt \} = \sigma_0 (O \rightarrow D)$$

$N(t)$ is the predicted traveling speed from an origin to a destination $O \rightarrow D$.
Intelligent Systems Lab

Traffic Message Channel (TMC): Real-Time Traffic Data in Freeway

Section

TMC (Traffic Message Channel) : Real Time Traffic Data in Freeway

Example of distance distribution of Traffic sensors along a route in California

A route is represented by the sequence of sections where traffic sensors are located along the route.

In general, traffic sensors are spaced unevenly, traffic data are sampled at different frequencies.

When the driver enter the origin and destination, a route is generated.

Traffic Sensor and Route
Current methods for generating a speed profile for a given route are mostly based on posted speed limits. Dynamic prediction of the speed profile for a selected route is essential because traffic conditions vary depending on time, day, locations, etc.

Limitation of Current Methods

- Accidental events, etc.
Objective of Research

- Develop accurate and computationally effective trip modeling algorithms
  - Accurately predicting driver-specific vehicle speed in real-time along a user-selected route

Intelligent Systems Lab

ADASRP Geo Info

Nokia Traffic Info

Bluetooth

UMD

Selected route

Traffic Knowledge Base

Traffic Statistics along route

Selected route

Static info

Dynamic info

Coordinates

UMD Trip Modeling System

TDI Speed profile at 8:00 in 11/7/2012
UMD developed the ITMS (Intelligent Trip Modeling System) to predict traffic speed on the freeway based on traffic sensor data available at the trip starting time, $t_0$. The approach for speed profile prediction includes:

- **Dynamic Traversing Speed Profile (DTSP)** algorithm.
- **Speed Prediction Neural Network System (SPNNS)**

The system takes into account:

- Road geometric data
- Historical traffic data
Speed Prediction Neural Networks

- Predict spot traffic speed at every sensor location along a given route at \( t_0 + j\Delta t \)
- Separate two different categories of traveling day
- Classify traffic congestion into two levels: congested vs. non-congested
- Designated to SPNNS (Speed Prediction Neural Network System)
SPNNS is a collection of speed prediction NN, $C_k^{SPNN}$, $i,j \in 1, \ldots$. SPNNS is a collection of speed prediction NN, $C_k^{SPNN}$, $i,j \in 1, \ldots$.
Based on the output from the selected neural networks, a time-space traffic prediction map is generated.
Time-Space Prediction Map and Dynamic Traversing Speed Profile
System Evaluation with Real Traffic Data

Intelligent Systems Lab

- Traffic data collected in real time at 5 minute intervals from traffic sensors.
- Traffic data collected by the California Freeway Performance Measurement System.
- The route used in our experiments: California Interstate I-405N
- 26.462 miles long, contains 52 traffic sensors, and has the posted speed limit of 65 mph.
- The busiest and most congested freeway in the United States.
- This route was chosen because of its interesting traffic dynamics such as recurrent congestion during morning peak hours, free flow and ramp metering are all present in this route.

Intelligent Systems Lab
Traffic Data used in Evaluation of SPNNS

Training:
SPNNS were trained using 1.5 months of traffic data – September 1, 2010 to October 15, 2010 from the California-I405-N route. The training data set for the SPNNS contains 673,920 data recordings (D=45 days).

Test data:
- Testing data set contains 3,893,760 data recordings (D=260 days).
- Speed prediction capability of SPNNS was evaluated at the following future time instances:
  - \( t_0 + 5 \)
  - \( t_0 + 10 \)
  - \( t_0 + 15 \)
  - \( t_0 + 20 \)
  - \( t_0 + 25 \)
  - \( t_0 + 30 \) minutes.

Evaluation of SPNNS

The test data set for the SPNNS contains 673,920 data recordings (D=45 days). The training data set used is the data from September 1, 2010 to October 15, 2010 from the California-I405-N route.

Traffic Data used in Evaluation of SPNNS

Intelligent Systems Lab

College of Engineering and Computer Science

University of Michigan-Dearborn
Performance Analysis of Speed Prediction NNs

Baseline: Traffic Sensor Data at the trip starting time
RMSE: Root Mean Square Error

Tested with 260 days of real traffic data: 3.89 million data points
Route: California I405 N (52 sensor data)

Speed Prediction NNs

Improvement over Baseline

Baseline: Traffic Sensor Data at the trip starting time
RMSE: Root Mean Square Error

Tested with 260 days of real traffic data: 3.89 million data points

Performance of Speed Prediction Neural Networks
### Performances comparison of Spot Speed Prediction NNs with published work

<table>
<thead>
<tr>
<th>Method</th>
<th>Predicted Intervals</th>
<th>Testing Data</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al. [24]</td>
<td>10 min</td>
<td>30 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>Hussein Dia [2]</td>
<td>5 min</td>
<td>10 min, 15 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>[30]</td>
<td></td>
<td></td>
<td>Spot Flow</td>
</tr>
<tr>
<td>Vanajakshi et al.</td>
<td>15 min</td>
<td>15 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>Zheng et al. [29]</td>
<td>30 min</td>
<td>60 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>UMD SPNNS</td>
<td>5 min, 10 min, 15 min</td>
<td>5 min, 10 min, 15 min</td>
<td>Spot Speed</td>
</tr>
</tbody>
</table>

#### Lee et al. [24]
- **Interval:** 10 min
- **Data points:** 3.89 million
- **Sensor data:** 260 days

#### Hussein Dia [2]
- **Interval:** 5 min, 10 min, 15 min
- **Data points:** 30% of 2 days
- **Data:** 5,000 observations
- **Sections:** 5 hours over 2 days

#### Vanajakshi et al.
- **Interval:** 15 min
- **Data points:** 3 sections over 5 hours

#### Zheng et al. [29]
- **Interval:** 30 min
- **Data points:** 60 min
- **Sensor data:** 4 days

#### UMD SPNNS
- **Interval:** 5 min, 10 min, 15 min, 20 min, 25 min, 30 min
- **Data points:** 3.89 million
- **Sensor data:** 260 days

#### Performances: MARE (%)

<table>
<thead>
<tr>
<th>Method</th>
<th>Interval</th>
<th>Testing Data</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al. [24]</td>
<td>10 min</td>
<td>30 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>Hussein Dia [2]</td>
<td>5 min</td>
<td>10 min, 15 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>[30]</td>
<td></td>
<td></td>
<td>Spot Flow</td>
</tr>
<tr>
<td>Vanajakshi et al.</td>
<td>15 min</td>
<td>15 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>Zheng et al. [29]</td>
<td>30 min</td>
<td>60 min</td>
<td>Spot Speed</td>
</tr>
<tr>
<td>UMD SPNNS</td>
<td>5 min, 10 min, 15 min</td>
<td>5 min, 10 min, 15 min</td>
<td>Spot Speed</td>
</tr>
</tbody>
</table>
Speed Profile Prediction by ITMS

Intelligent Transportation Systems Lab

Traveling distance (mile)

Distance (mile)

Traveling time (min)

Velocity (mile per hour)

Driver at t=8:05 am

Speed Profile generated at Tuesday 10/19/2010 08:05 am

17
Evaluation of ITMS Speed profile generation

Route #1: California I 405 N

Total 2,880 trips are evaluated: California PEMS data contains 288 trips per day.

UMD ITMS performance is 5.89% better than the baseline.

UMD ITMS predicts the traffic speed profile for a given freeway route based on sensor data.

UMD ITMS performance is 3.45% better than the baseline.

Error Comparison

<table>
<thead>
<tr>
<th>DTD Trip Modeling</th>
<th>DTD Base Line</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE (mph)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>3.4</td>
<td>0.00</td>
</tr>
<tr>
<td>3.45</td>
<td>3.45</td>
<td>0.00</td>
</tr>
<tr>
<td>3.5</td>
<td>3.5</td>
<td>0.00</td>
</tr>
<tr>
<td>3.55</td>
<td>3.55</td>
<td>0.00</td>
</tr>
<tr>
<td>3.6</td>
<td>3.6</td>
<td>0.00</td>
</tr>
<tr>
<td>3.65</td>
<td>3.65</td>
<td>0.00</td>
</tr>
<tr>
<td>3.7</td>
<td>3.7</td>
<td>0.00</td>
</tr>
<tr>
<td>3.75</td>
<td>3.75</td>
<td>0.00</td>
</tr>
<tr>
<td>3.8</td>
<td>3.8</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Intelligent Systems Lab

Speed Profile Generation by ITMS: example

Trip Starting time: Friday, 10/22/2010 7:55 pm

Benchmark: Speed profile is generated with real measured sensor data.

SPNNS_2A (all weekday NNs)
SPNNS_2W (Mon-Friday SPNNS)
SPNNS_2A (No weekday categorization)

SPNNS_2W (Mon-Friday SPNNS)

Trip Modeling by NN prediction (All weekday NNs)
Trip Modeling by NN prediction (Friday NNs)
Benchmark (Sensor data)

<table>
<thead>
<tr>
<th>RMSE</th>
<th>3.92 mph</th>
<th>6.74 mph</th>
</tr>
</thead>
</table>

Note: Benchmark speed profile is generated with real measured sensor data.
Note: Baseline is generated with TMC data at the trip starting time.

Baseline: Trip starting time TMC information

Benchmark (Sensor data)

Trip modeling by NN prediction (Friday NNS)

Baseline: 7.79 mph

SPNNS_2W (Mon-Friday SPNNS): 3.92 mph

Benchmark

Baseline:

Trip starting time TMC information

Baseline

Trip starting time: 10/22/2010 7:55 pm

SPNNS_2W: 7.79 mph

Baseline: TMC at Trip starting time

RMSE

Intelligent Systems Lab

College of Engineering and Computer Science

UNIVERSITY OF MICHIGAN-DEARBORN
Other data science research in intelligent transportation systems

Intelligent Transportation Systems Lab

UMD Traffic Data Collection

ADASRP Geo Info

Backup Server

Cloud Server

HERE Traffic Info

Driving speed prediction
Traffic congestion prediction
Traffic routing
Mobility pattern
Driving speed prediction
Traffic routing
Traffic congestion prediction
Mobility pattern
Start time
Route coordinates
Location A to any Location B
Driver enter the route from Any Location

Vehicle-to-Vehicle (V2V)

Weather data (any historical)
Incident data in MI & OH (09/01/2011 to current)
Individual Driving Data & matching geo data
Incident data in MI & OH (09/01/2011 to current)
Traffic Sensor data in MI & OH (09/01/2011 to current)
UMD mobility database

Traffic pattern and weather data (any historical)

Other data science research in intelligent transportation systems

College of Engineering and Computer Science
University of Michigan-Dearborn
Data driven personalized driving systems

Intelligent Systems Lab

Door-to-Door personalized driving prediction

- Traffic congestion level
- Incorportating contextual knowledge
- Predicting traffic along the traveling route at the reaching time, \( t_i \)
- Predicted traffic flow along the traveling route at the reaching time, \( t_i \)
- Traffic sensor acquired at time \( t_0 \)
- Historical driver's driving speed profile
- Roadway type and road geometry
- Day/time of travel
- Weather data
- Incidents
- Traffic congestion level
- Driver departs from predicted route:
- Learning from rejection
- Adaptive learning
- Driver changes routes due to seasonal changes
- Detect obsolete knowledge and incrementally learn new knowledge
- Driver changes routes due to seasonal changes
- Learning from rejection

Driver personal route prediction

- Door-to-Door personalized driving prediction

College of Engineering and Computer Science
University of Michigan-DeARBORN
Building personalized driving assistant systems

Intelligent Systems Lab

Intelligent Systems Lab

Making personalized driving assistant systems through machine learning

Making personalized driving decisions in a connected dynamic traffic environment

- Lane changes, over-pass, acceleration, deceleration,

- Making personalized smart driving operations in a network of connected vehicles

- Making personalized operation while keeping distance keeping

Making personalized driving operations in a network of connected vehicles

Making personalized driving operations in a network of connected vehicles

Route generation and dynamic adaptation

Making personalized driving decisions in a connected dynamic traffic environment
Thank you