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 Massage Channel (TMC), link, and shape points
    - GPS locations, speed limit, number of lanes, curvature, altitude, link length, shape distance

- **Real-time traffic data**
  - We have been downloading traffic data covering all roads in Michigan and Ohio, 24/7,
    - 463MB/day of real-time traffic flow and incident/day for states of Michigan and Ohio
  - weather condition data
  - V2V communication data

- **Large collections of real-world driving data**
  - Hundreds and thousands of trips from different drivers
  - 368 channels of signals
    - Vehicle, Environmental, Physiological, video
  - For a trip of less than one hour, the signal data size is about 400MB.
Prediction of Origin-Destination Traveling Speed Profile

A highly representative speed profile will be beneficial to:

- advanced vehicle energy management
  - It is important to have accurately predicted speed profile available at the beginning of the trip
- 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.

In ATIS:
- estimate accurate trip time
- provide useful information to driver for route selection

Fuel economy (and optimal energy management strategy) depends on driver demand, i.e. the drive cycle.
What is Trip Modeling?

Mathematically, trip modeling problem is to generate a sequence of predicted traveling speed from an Origin to a Destination (O_D)

\[ TM_{O_D} = \{ F(x_p | t_0, \Omega, \bar{v}_0) | p = 1, ..., N \} \]

- \( F \) is a system that predicts the traffic speed at location \( x_p \) at time \( t_{t[0,p]} \)
  - is the predicted time needed to travel from the origin \( O \) to \( x_p \),
- \( F \) makes the traffic speed prediction at time \( t_0 \), the trip starting time, based on the following information
  - \( \Omega \): a set of traveling points along the chosen route, which can be
    - a sequence of traffic sensor locations along the chosen route
    - A sequence of defined geographic locations along the chosen route
  - \( \bar{v}_0 \) is a sequence of traffic speed at time \( t_0 \) provided by the traffic sensors at location points in \( \Omega \).
Traffic Sensor and Route

- **TMC (Traffic Message Channel):** Real-time TMC traffic data in freeway.
- When the driver enters the origin and destination, a route is generated.
- 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.
- Example of distance distribution of Traffic sensors along a route in California Interstate I-405N.
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 the traffic conditions vary depending on time, day, locations, and 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
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, $t_0$.

- Speed Prediction Neural Network Systems (SPNNS)
- Dynamic Traversing Speed Profile (DTSP) algorithm.
Speed Prediction Neural Networks

- Speed Prediction Neural Network System (SPNNS) is designed to:
  - Predict spot traffic speed at every sensor location along a given route: $t_0 + j\Delta t$, $j=1, \ldots$
  - Classify traffic congestion into two levels, congested vs. non-congested,
  - Separate two different categories of traveling day.

**SPNNS**

- **Mon. ~Friday NNs**
  - $C_1_{-SPNN}$
  - **Congested NNs:** $C_1_{-SPNN_{1j}}$, where $j=1, \ldots 6$
  - **Non-Congested NNs:** $C_1_{-SPNN_{2j}}$, where $j=1, \ldots 6$

- **Sat. ~Sunday NNs**
  - $C_2_{-SPNN}$
  - **Congested NNs:** $C_2_{-SPNN_{1j}}$, where $j=1, \ldots 6$
  - **Non-Congested NNs:** $C_2_{-SPNN_{2j}}$, where $j=1, \ldots 6$
**Architecture of SPNNS**

- **SPNNS is a collection of speed prediction NN**, $C_{k-SPNN_{ij}}$, $i=1,...$

Traffic data acquired by the traffic sensor at location $p$ and the time, $t_0$

Speed Prediction Neural network, $C_{k-SPNN_{ij}}$

Predicted Speed for $j\Delta t$ later, $\Delta t=5$min, $j=1,...$. 

**NN inputs:**
- Traffic information speed $\nu$ and flow $q$ at time $t_0$

**Speed Prediction NN**

$C_{k-SPNN_{ij}}$

Predicted spot speed at the sensor location $p$: $\nu(p, t_0+j\Delta t)$
Based on the output from the selected neural networks, a time-space traffic prediction map is generated.
Time-Space Speed Prediction Map and Dynamic Traversing Speed Profile

\[\dot{v}(1, t_0 + \Delta t) \quad \dot{v}(2, t_0 + \Delta t) \quad \ldots \quad \dot{v}(p, t_0 + \Delta t) \quad \ldots \quad \dot{v}(N, t_0 + \Delta t)\]

\[\dot{v}(1, t_0 + 2\Delta t) \quad \dot{v}(2, t_0 + 2\Delta t) \quad \ldots \quad \dot{v}(p, t_0 + 2\Delta t) \quad \ldots \quad \dot{v}(N, t_0 + 2\Delta t)\]

\[\dot{v}(1, t_0 + j\Delta t) \quad \dot{v}(2, t_0 + j\Delta t) \quad \ldots \quad \dot{v}(p, t_0 + j\Delta t) \quad \ldots \quad \dot{v}(N, t_0 + j\Delta t)\]

\[\dot{v}(1, t_0 + M\Delta t) \quad \dot{v}(2, t_0 + M\Delta t) \quad \ldots \quad \dot{v}(p, t_0 + M\Delta t) \quad \ldots \quad \dot{v}(N, t_0 + M\Delta t)\]
System Evaluation with Real Traffic Data

- **Evaluation with real traffic data collected by the California Freeway Performance Measurement System.**
  - Traffic data collected in real time at 5 minute intervals from traffic sensors.

- **The route used in our experiments: California Interstate I-405N**
  - The busiest and most congested freeway in the United States.
  - 26.462 miles long, contains 52 traffic sensors, and has the posted speed limit of 65 mph.

- 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.
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-I-405-N route.
- The training data set for the SPNNS contains 673,920 data recordings (D=45 days)

Evaluation of SPNNS

Test data:
- March 1, 2010 to December 30, 2010 except days used in training the SPNNS.
  - 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
  - $t_0 + 5$, $t_0 + 10$, $t_0 + 15$, $t_0 + 20$, $t_0 + 25$, and $t_0 + 30$ minutes.
Performance Analysis of Speed Prediction NNs

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

RMSE: Root Mean Square Error
Baseline: Traffic Sensor Data at the trip starting time
Performances comparison of Spot Speed Prediction NNs with published work

<table>
<thead>
<tr>
<th></th>
<th>Prediction</th>
<th>Prediction Interval</th>
<th>Testing Data</th>
<th>Performances: MARE(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al. [24]</td>
<td>spot speed</td>
<td>10 min</td>
<td>6 days (2 links traffic data)</td>
<td>-10 min: 9.08%</td>
</tr>
<tr>
<td>Hussein Dia</td>
<td>spot speed</td>
<td>5 min, 10 min</td>
<td>30% of 2days (3 sections over 5 hour period on 2 days -5000 data points)</td>
<td>Speed -5 min:10% -15 min:16%</td>
</tr>
<tr>
<td></td>
<td>travel time</td>
<td></td>
<td></td>
<td>oTravel Time - 7%</td>
</tr>
<tr>
<td>Zheng et al.</td>
<td>spot flow</td>
<td>15 min</td>
<td>4 days (one sensor location)</td>
<td>-15 min :6.10%</td>
</tr>
<tr>
<td>Vanajakshi et al.[30]</td>
<td>spot speed</td>
<td>2 min 4 min … 60 min</td>
<td>1 day data (one sensor location - 720 data points)</td>
<td>-4 min: 4.5% -30 min: 9.8%</td>
</tr>
<tr>
<td>UMD SPNNS</td>
<td>spot speed</td>
<td>5 min 10 min … 30 min</td>
<td>260 days(52 sensor data – 3.89 million data points)</td>
<td>- 5 min: 2.83% - 10 min: 4.28% - 15 min: 5.49% - 20 min: 6.79% - 25 min: 7.66% - 30 min: 8.74%</td>
</tr>
</tbody>
</table>
Speed Profile Prediction by ITMS

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

Traveling distance (mile)

Velocity (mile per hour)

Driver at $t_0=8:05$ am

5 min Prediction

10 min Prediction

15 min Prediction

20 min Prediction

25 min Prediction

30 min Prediction
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 (288*15 days).

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

- UMD ITMS predict the traffic speed profile for a given freeway route based on sensor data in Freeway only

- UMD ITMS performance is 5.89% better than the performance of the baseline
Trip Starting time:
Friday, 10/22/2010 7:55 pm

Speed Profile generated at 10/22/2010 7:55pm

RMSE

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPNNS_2A: (all weekday NNs)</td>
<td>6.74 mph</td>
</tr>
<tr>
<td>SPNNS_2W: (Mon-Friday SPNNS)</td>
<td>3.92 mph</td>
</tr>
</tbody>
</table>

Note: Benchmark speed profile is generated with real measured sensor data.
Trip Starting time:
Friday, 10/22/2010 7:55 pm

Note: Baseline is generated with TMC data at the trip starting time.
Data driven personalized driving systems

- **Door-to-Door personal driving speed prediction**
  - Traffic sensor acquired at time t0
  - Predicted traffic flow along the traveling route at the reaching time, \( t_i \),
  - Historical driver’s driving speed profile
  - Roadway type and road geometry

- **Driver personal route prediction**
  - Predicting traveling route at the beginning of trip and
  - Incorporating contextual knowledge
    - Traffic congestion level
    - Incidents
    - Weather data
    - Day/time of travel
  - Adaptive learning
    - Driver departs from predicted route:
      - Learning from rejection
    - Driver changes routes due to seasonal changes
      - Detect obsolete knowledge and incrementally learn new knowledge
• Making personalized driving decisions in a connected dynamic traffic environment
  • Route generation and dynamic adaptation
  • Making personalized smart driving operations in a network of connected vehicles
    – Lane changes, over-pass, acceleration, deceleration, distance keeping,
  • Making a personalize operation while keeping equilibrium in a network of vehicles,
• Thank you