Examining Travelers Who Pay to Drive Slower in the Katy Managed Lanes

Presented by: Farinoush Sharifi
Negative Travel Time Saving in the Managed Lanes

11% of paid trips
Uneconomical Managed Lane Trips

What are the commonalities among these trips?
What causes these trips?
Can we predict these trips?
What are the characteristics of these trips?

Cumulative Percentage of Trips

Time Saved (minutes)
Research Purpose

The main goal of this study is to look into U-ML trips, search for commonalities among these trips, establish some insight into this travel choice, and find the most relevant factors.

To maintain this goal, this research will:

• explore ML trips, and specifically U-ML trips, and their characteristics
• identify the most important variables affecting U-ML trips
• investigate into the way these variables impact U-ML trips
• estimate a model to predict U-ML trips.
Overview

- Background
- Dataset
- Methodology
- U-ML Trip Identification
- Sampling & Resampling
- Data Analysis
- Results
- Conclusion
Managed Lanes

- Restricting vehicle eligibility
- Limiting facility access
- Collecting variably priced tolls

- Congestion reduction
- Higher speed
- Saving travel time
Katy Managed Lanes

- 12-mile section of I-10
- City of Katy to Downtown Houston
- At least 4 GPLs (free lanes)
- 2 MLs (HOT lanes)
- 3 toll plazas

* Source: HCTRA Website
Toll Pricing

- Time of day
  1. HOVs during times other than Monday to Friday, 5 am to 11 am and 2 pm to 8 pm
  2. SOVs

<table>
<thead>
<tr>
<th>Dates</th>
<th>Direction</th>
<th>Time of Day</th>
<th>Toll at Eldridge</th>
<th>Toll at both Wilcrest and Wirt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept 7, 2013, to Today</td>
<td>Westbound</td>
<td>Peak: 4-6 pm weekdays</td>
<td>$3.20</td>
<td>$1.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shoulder: 3-4 &amp; 6-7 pm weekdays</td>
<td>$2.10</td>
<td>$1.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Off-peak: all other times</td>
<td>$0.40</td>
<td>$0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High Peak: 7-8 am weekdays</td>
<td>$3.20</td>
<td>$1.90</td>
</tr>
<tr>
<td></td>
<td>Eastbound</td>
<td>Low Peak: 8-9 am weekdays</td>
<td>$2.60</td>
<td>$1.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High Shoulder: 6-7 am weekdays</td>
<td>$2.10</td>
<td>$1.20</td>
</tr>
</tbody>
</table>
Data Sources

- **TxDOT AVI Sensors**
  - 38 sensors
  - Transponder IDs
  - Detection Time

- **HCTRA AVI Sensors**
  - 12 sensors
  - Transponder IDs
  - Detection Time
  - Toll Paid

- **TxDOT Lane Closure Data**
  - Lane Closure
  - Incident
  - Toll Paid

- **NCDC by NOAA**
  - Precipitation
Final Dataset

• Almost 3 years of data
• Covering 7,013,587 paid ML trips
• Includes the alternate GPL trips

Variables of Study

✓ Day of the trip
✓ Time of the trip
✓ Route of the trip
✓ Length of the trip
✓ Safety
✓ Toll
✓ Traffic flow
✓ Travel behavior
✓ Trip frequency
Methodology

- U-ML Trip Identification
- Sampling & Resampling
- Primary Data Analysis
- Final Data Analysis
U-ML Trip Identification

- **Binary:** faster / slower
- **Multiclass:** faster / slower / too close to zero difference ( |RTTD|= 0.05)

\[ TTD = (TT_{ML} - TT_{GPL}) \]

\[ RTTD = \frac{(TT_{ML} - TT_{GPL})}{TT_{ML}} \]

\[ uneco_{binary} = \begin{cases} 
1 \text{ if } (TTD \geq 0) & \text{(Uneconomical ML trip)} \\
0 \text{ if } (TTD < 0) & \text{(Economical ML trip)}
\end{cases} \]

\[ uneco_{multiclass} = \begin{cases} 
1 \text{ if } (RTTD > 0.05) & \text{(Uneconomical ML trip)} \\
0.5 \text{ if } (-0.05 \leq RTTD \leq 0.05) & \text{(Middle ML trip)} \\
0 \text{ if } (RTTD < -0.05) & \text{(Economical ML trip)}
\end{cases} \]
Sampling & Resampling

- Sample set: Almost 1 million ML trips
- Training set (80%) and Test set (20%)
- Dataset is **imbalanced**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Type</th>
<th>U-ML</th>
<th>Middle ML</th>
<th>E-ML</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binary</strong></td>
<td>Imbalanced</td>
<td>11.20%</td>
<td>--</td>
<td>88.80%</td>
</tr>
<tr>
<td></td>
<td>Under-sampled</td>
<td>50.05%</td>
<td>--</td>
<td>49.95%</td>
</tr>
<tr>
<td><strong>Multiclass</strong></td>
<td>Imbalanced</td>
<td>5.10%</td>
<td>16.47%</td>
<td>78.43%</td>
</tr>
<tr>
<td></td>
<td>Under-sampled</td>
<td>33.33%</td>
<td>33.33%</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

Balancing training set; Under-Sampling
Modeling

Random Forest
- Accurate model
- Variables of importance
- Large datasets

Logistic Regression
- Simple interpretation of variables
- Traditional trip choice model
- Variable’s impact
Results

• Evaluation of Models: *Area Under ROC Curve (AUC)*
  
• Binary Random Forest (BRF) Model: **AUC = 0.7543**
• Multiclass Random Forest (MRF) Model: **AUC = 0.7520**
• Binary Logistic Regression (BLR) Model: **AUC = 0.6360**

*Random Forest* is more *accurate* than Logistic Regression.
*Binary classification* or *multiclassification* are *not significantly different*.
*BRF model* is more *efficient* than MRF model.
*BLR model* is still *useful*...
Results

- High ML Traffic Flow
- Low GPL Traffic Flow
- Small Toll Rate
- Large Travel Time Variability

Higher Likelihood of Negative Travel Time Saving or Uneconomical ML Trip
Results

- Start Location
- End Location

Sensor pairs with lower number of ML trips has a higher probability of these uneconomical trips.

Most and Least Likely Routes for U-ML Trips
(Red arrows show the most likely routes, and green arrows show the least likely routes)
Conclusion

- Provides a better ML trip prediction in the future
- Various potential factors; rain and blockage
- Probe data: a sample of trips
- No demographic data: wealth and income
- Only focusing on Katy Managed Lanes
- Initiative study on unexpected lane choice behavior
Thank You!

Contact:

Farinoush Sharifi
Texas A&M Transportation Institute

f-sharifi@tti.tamu.edu
(979) 845-7871