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

Presented by: Farinoush Sharifi

The Peak of Transportation
2018
Keystone Colorado
ite Joint Western & Texas District Meeting

Negative Travel Time Saving in the Managed Lanes

11% of paid trips

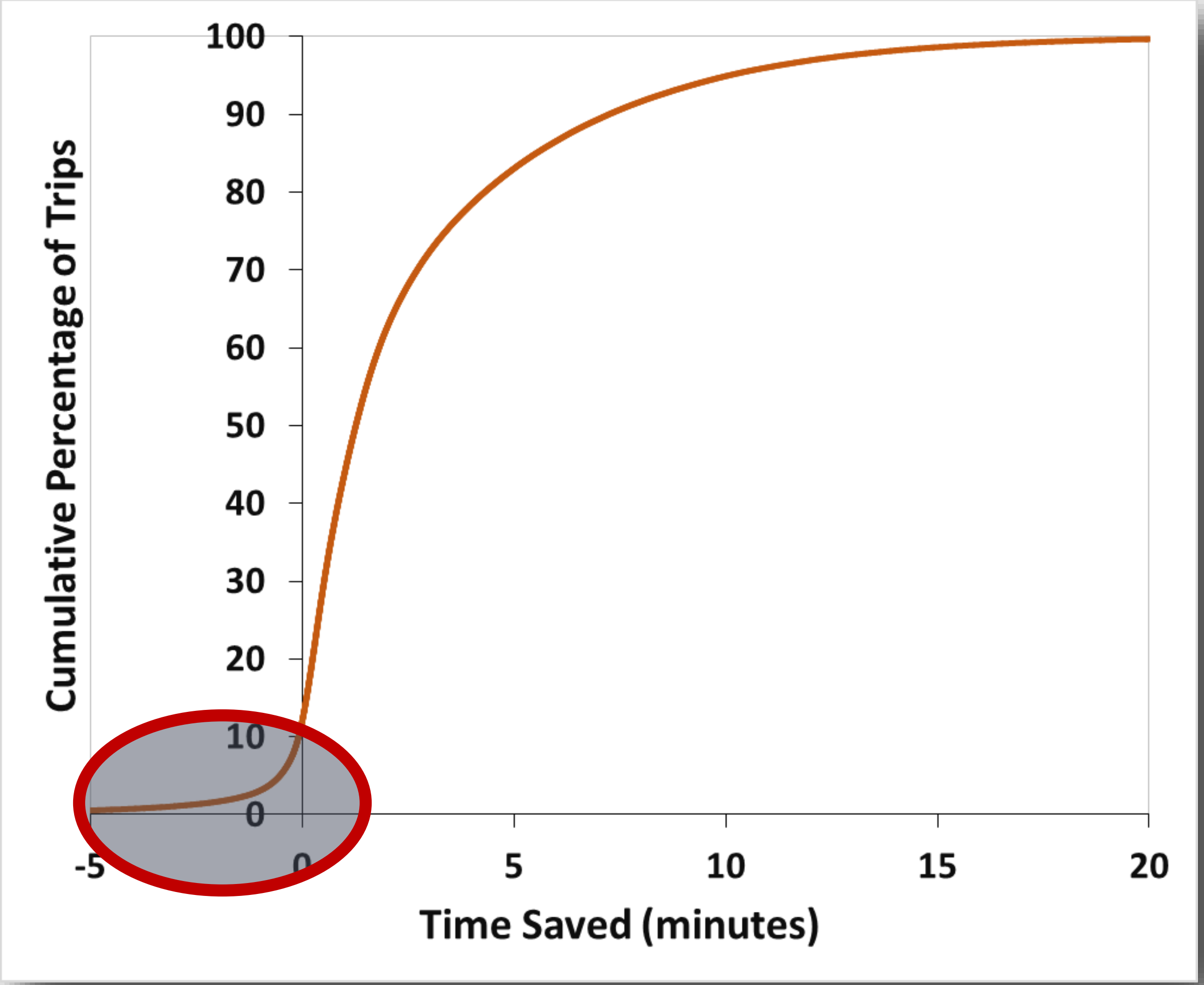
Uneconomical Managed Lane Trips

What are the commonalities among these trips?

What causes these trips?

Can we predict these trips?

What are the characteristics of these trips?



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

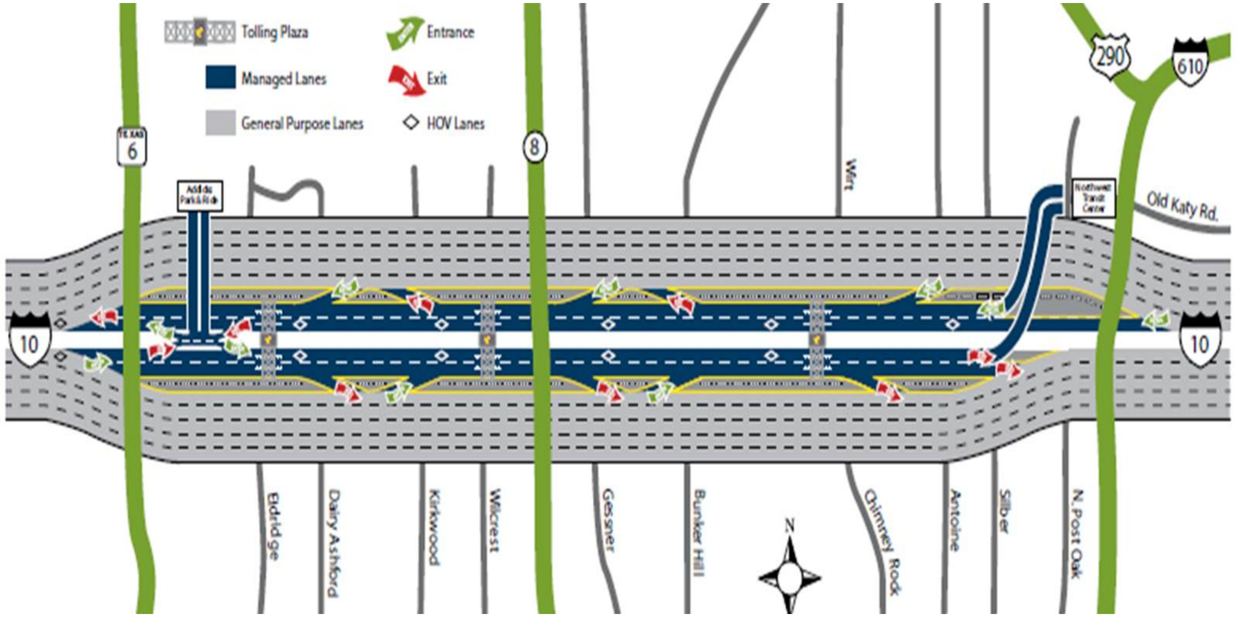
Managed Lanes

- 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]

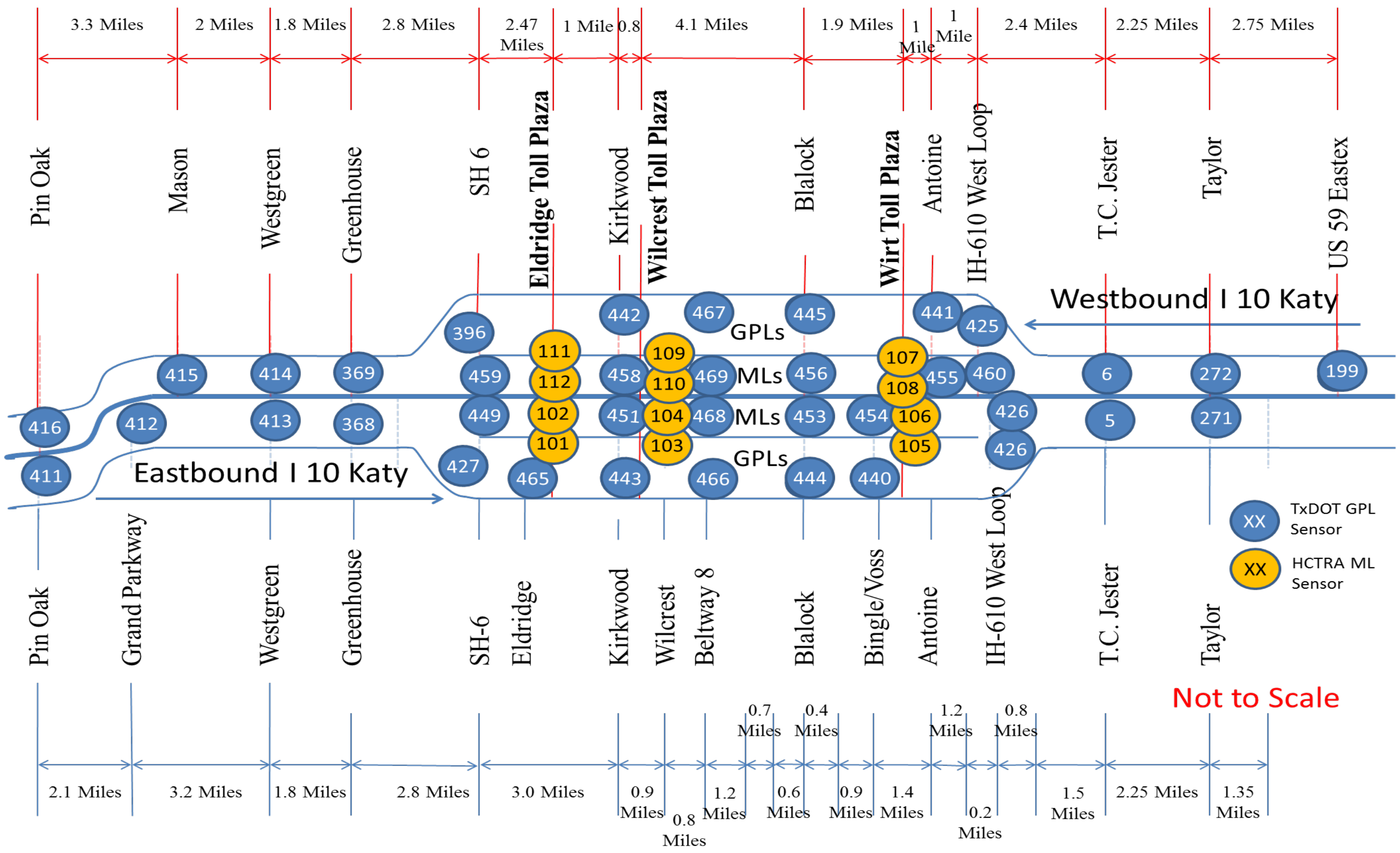
Toll Pricing

- Time of day
- Number of passengers

1. *HOVs during times other than Monday to Friday, 5 am to 11 am and 2 pm to 8 pm*
2. *SOVs*

Dates	Direction	Time of Day	Toll at Eldridge	Toll at both Wilcrest and Wirt
Sept 7, 2013, to Today	Westbound	Peak: 4-6 pm weekdays	$3.20	$1.90
		Shoulder: 3-4 & 6-7 pm weekdays	$2.10	$1.20
		Off-peak: all other times	$0.40	$0.30
	Eastbound	High Peak: 7-8 am weekdays	$3.20	$1.90
		Low Peak: 8-9 am weekdays	$2.60	$1.70
		High Shoulder: 6-7 am weekdays	$2.10	$1.20

Toll Collection



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
- **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)

$$TTD = (TT_{ML} - TT_{GPL}) \qquad RTTD = (TT_{ML} - TT_{GPL})/TT_{ML}$$

$$uneco_{binary} = \begin{cases} 1\ if\ (TTD \geq 0)\ (Uneconomical\ ML\ trip) \\ 0\ if\ (TTD < 0)\quad (Economical\ ML\ trip) \end{cases}$$

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

Sampling & Resampling

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

Balancing training set; Under-Sampling

Classification	Type	U-ML	Middle ML	E-ML
Binary	Imbalanced	***11.20%***	--	***88.80%***
	Under-sampled	50.05%	--	49.95%
Multiclass	Imbalanced	***5.10%***	***16.47%***	***78.43%***
	Under-sampled	33.33%	33.33%	33.33%

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)*

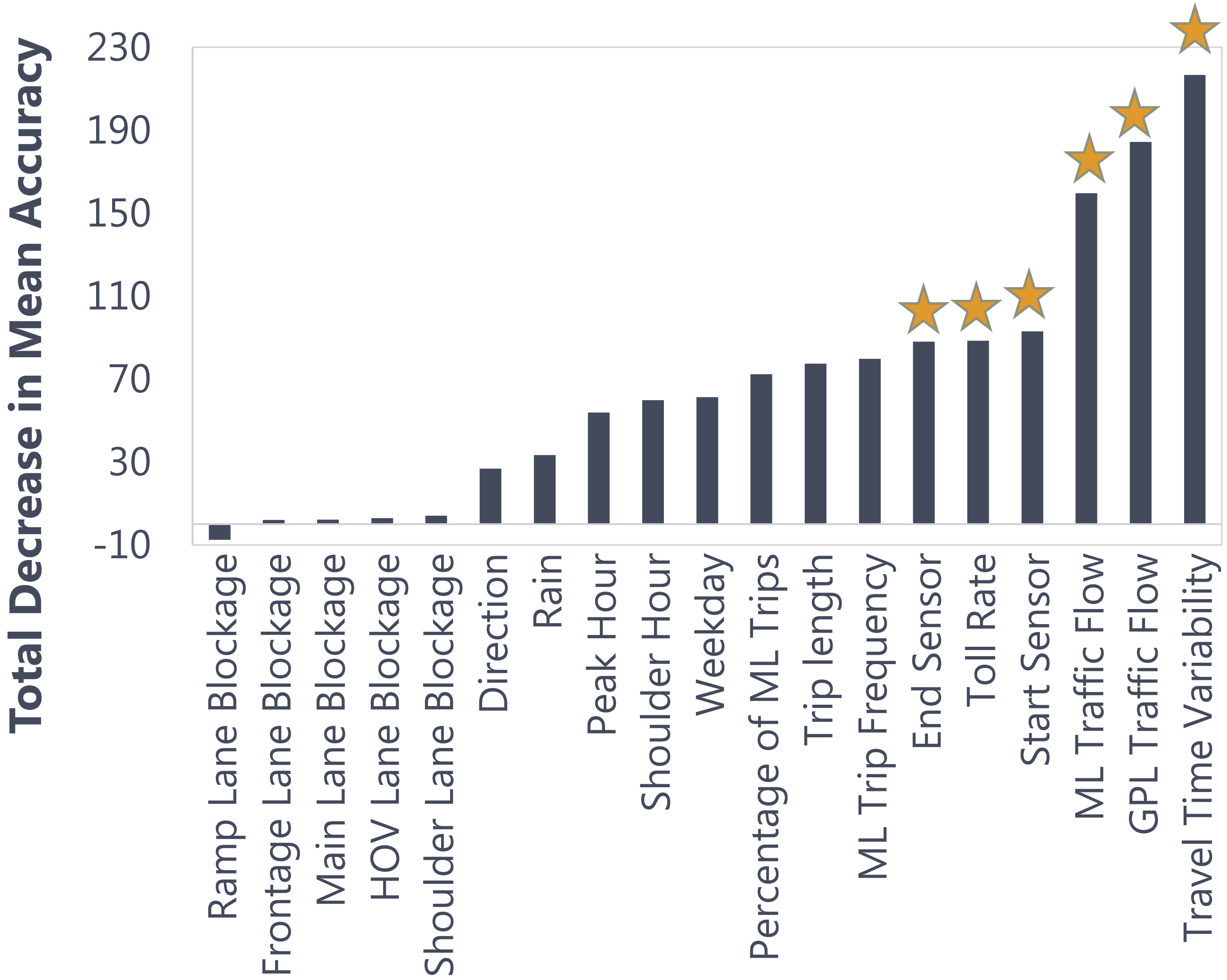
- Binary Random Forest (BRF) Model: **AUC = 0.7543**
- Multiclass Random Forest (MRF) Model: **AUC = 0.7520**
- Binary Logistic Regression (BLR) Model: **AUC = 0.6360**

*__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

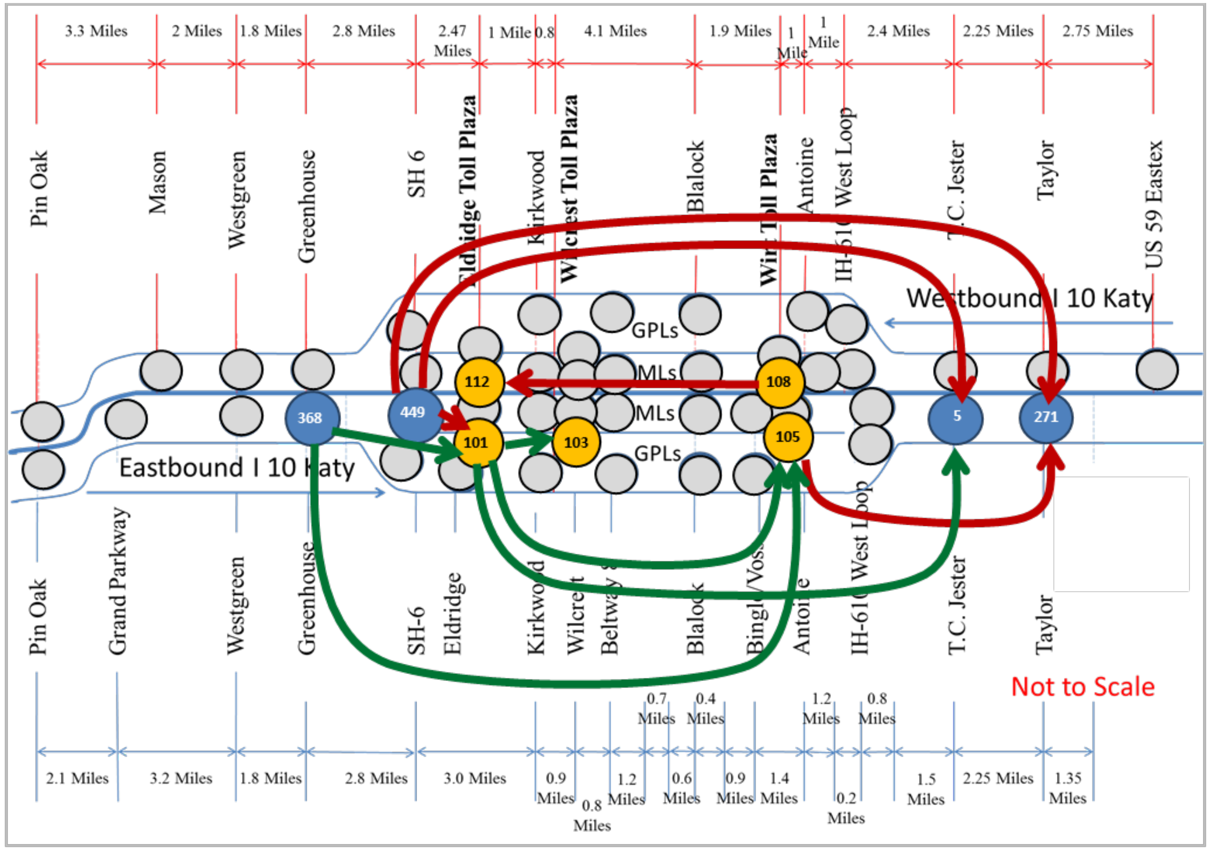
Higher Likelihood of Negative Travel Time Saving or Uneconomical ML Trip



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