Modeling Drivers' Route Choices and Route Compliance when Interacting with An Eco-Routing Navigation System

Ocobter 1st, 2020

Shan Bao, Ph.D.
University of Michigan-Dearborn
University of Michigan Transportation Research Institute

Content

1. Introduction
2. Methodology: eco-routing application and naturalistic driving experiments
3. Results: modelling driver’s route choice
4. Conclusion

Team Members

- Dr. Bo Yu
- Dillon Funkhouser, Ethan Bezzina, Will Wang
- Zeyang Li, Huajian Zhou, Zeyang Cheng, Haibo Wang

Introduction

- It has been reported that the transportation sector made up the largest proportion (approximately 29%) of the total U.S. greenhouse gas emissions in 2017, of which 41% was represented by passenger cars (EPA, 2019).
- The eco-routing navigation system becomes a potential application to reduce fuel consumption, which optimizes the route based on the most energy efficiency instead of minimizing the travel time or distance (Zeng et al., 2017; Boriboonsomsin and Barth, 2014).
- Route choice models can provide a better understanding of drivers' route choice preferences and their influencing factors while limited studies have been conducted related to eco-routing choices.

Objectives

Eco-routing navigation systems can greatly contribute to the reduction of fuel consumption and greenhouse gas emissions, if drivers are willing to choose and comply with the eco routes provided by these systems.

However, current studies paid limited efforts to drivers' route choices and compliance behavior when interacting with eco-routing navigation systems.

There are two main objectives of this study:

(1) Investigating and predicting preferred routes in drivers from recommendations;

(2) Exploring drivers' decision making on route selection.

Methodology

The methodology section contains five key parts:

(1) Cellphone based eco-routing navigation application development;

(2) Naturalistic driving experiments were designed and conducted;

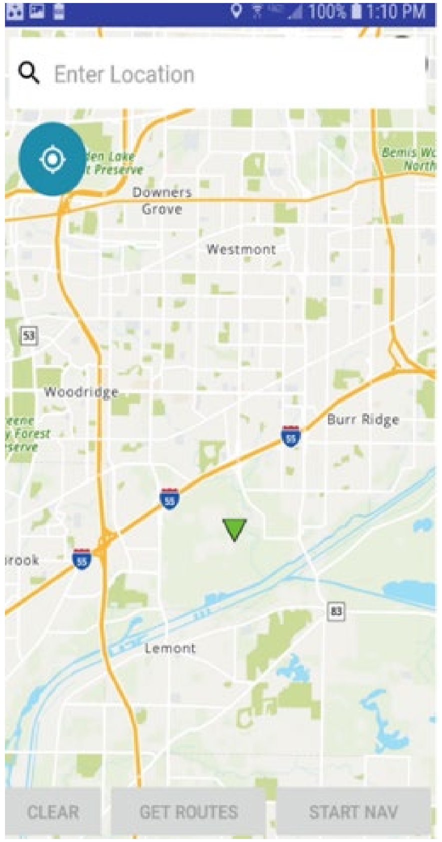
(3) Google API application to correct cellphone GPS data;

(4) Mixed model analyses were applied to interpret drivers' route choices and route compliance;

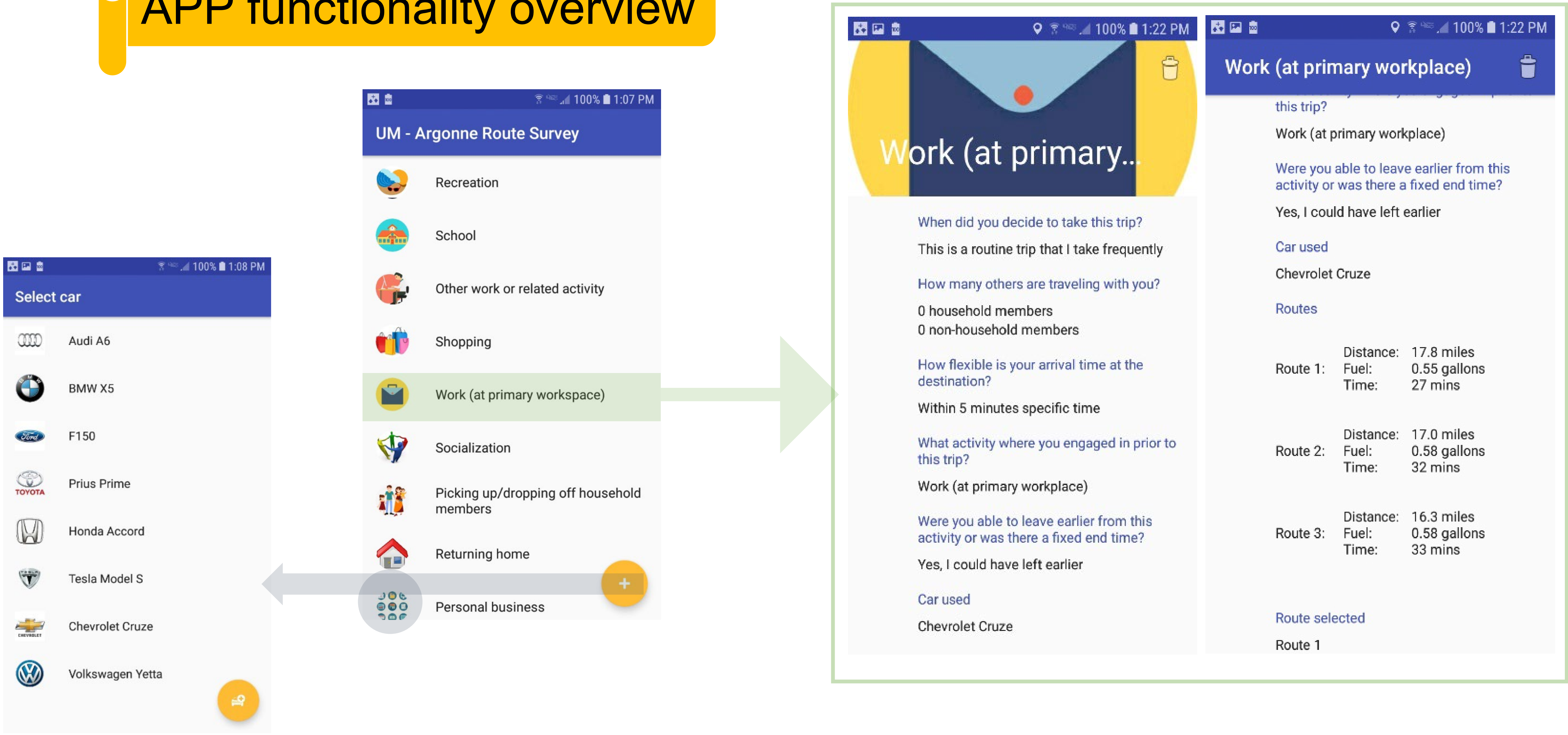
(5) The Multi-label random forests algorithm was used to predict drivers' route choice behavior.

Eco-routing application development

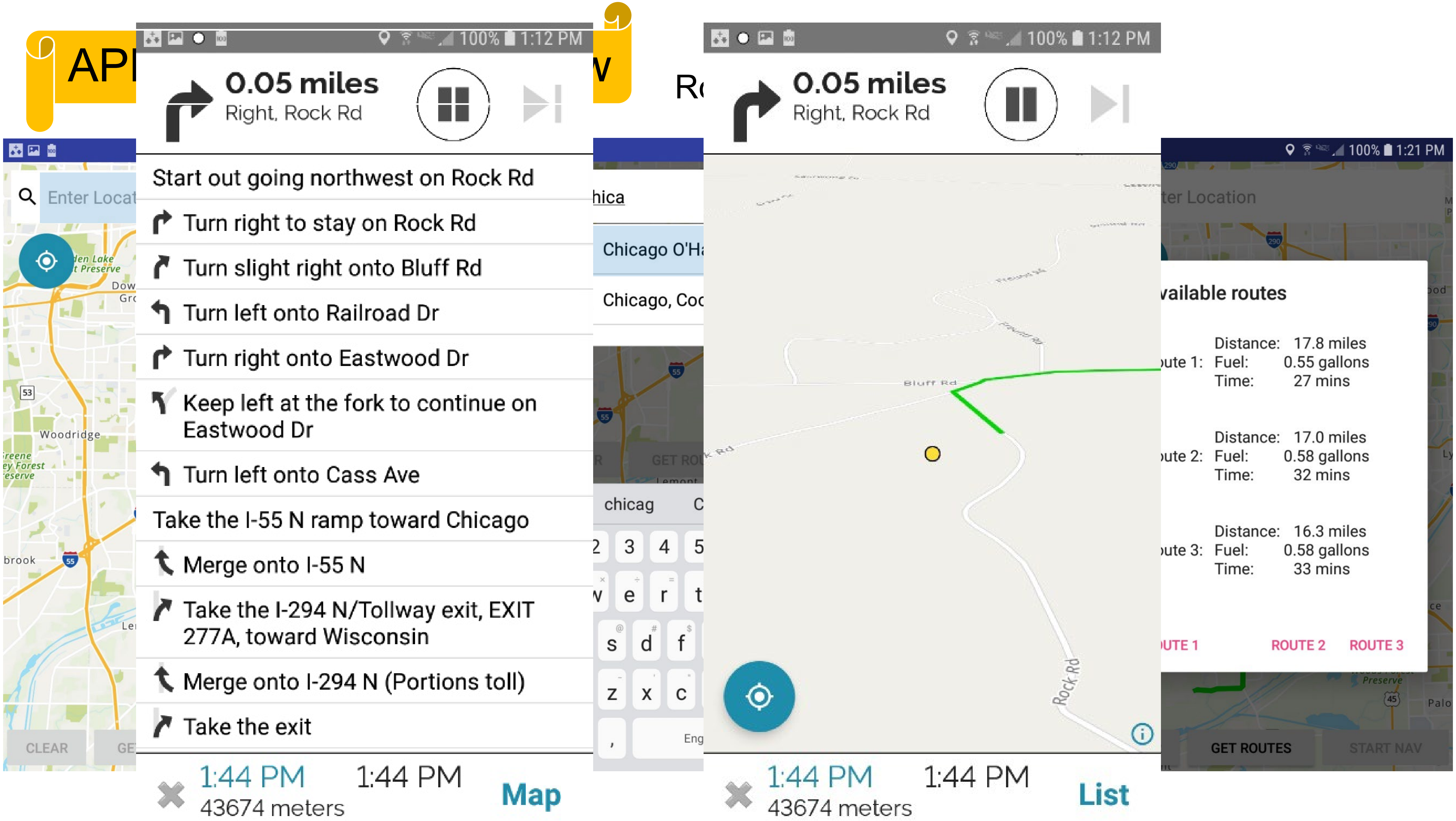
- Collecting basic trip information prior to each trip;
- Providing 1-3 routes to driver with varying estimated time/fuel consumption (eco/fast/balanced);
- Providing turn-by-turn guidance for a selected route;
- Recording trip information, driver route decisions, and actual route taken (GPS co-ord. during trip).



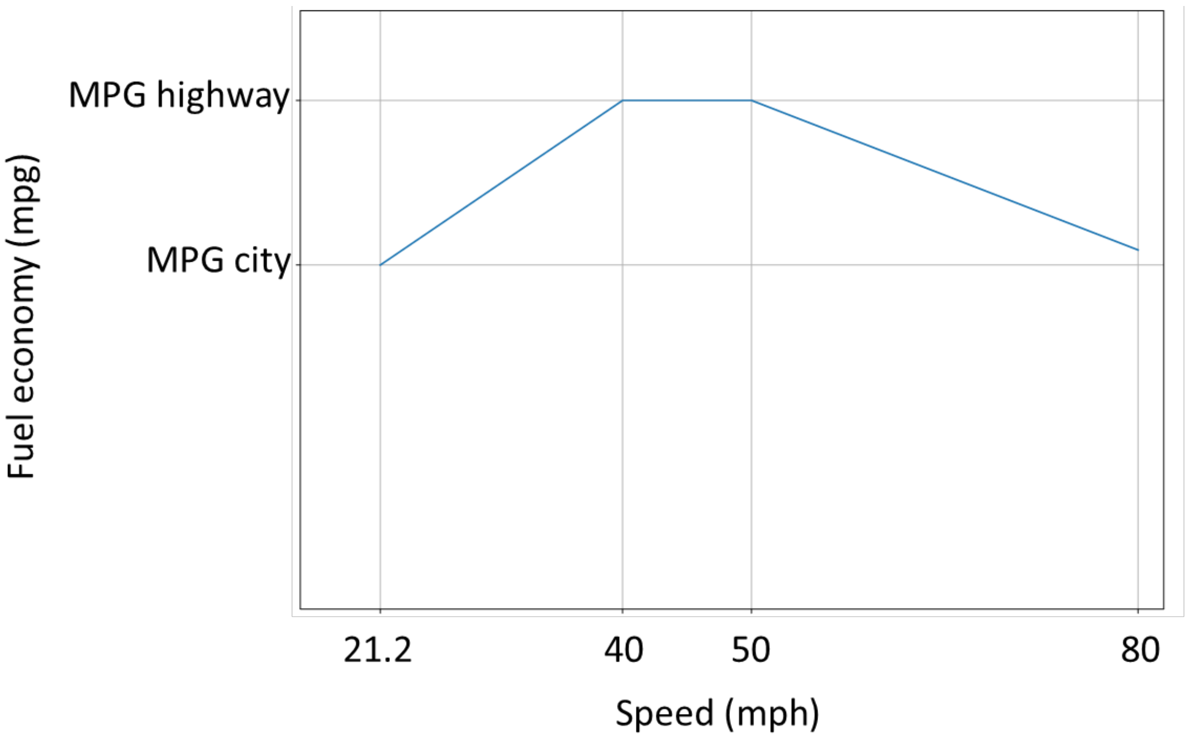
APP functionality overview



List of trips, trip details and select car


0.05 miles
Right, Rock Rd
Start out going northwest on Rock Rd
Turn right to stay on Rock Rd
Turn slight right onto Bluff Rd
Turn left onto Railroad Dr
Turn right onto Eastwood Dr
Keep left at the fork to continue on Eastwood Dr
Turn left onto Cass Ave
Take the I-55 N ramp toward Chicago
Merge onto I-55 N
Take the I-294 N/Tollway exit, EXIT 277A, toward Wisconsin
Merge onto I-294 N (Portions toll)
Take the exit
1:44 PM
1:44 PM
43674 meters
Map
0.05 miles
Right, Rock Rd
Bluff Rd
Rock Rd
1:44 PM
1:44 PM
43674 meters
List
Enter Locat
CLEAR
Woodridge
Chicago O'Ha
Chicago, Coo
chicag
vailable routes
Distance: 17.8 miles
Fuel: 0.55 gallons
Time: 27 mins
Distance: 17.0 miles
Fuel: 0.58 gallons
Time: 32 mins
Distance: 16.3 miles
Fuel: 0.58 gallons
Time: 33 mins
ROUTE 2
ROUTE 3
GET ROUTES
START NAV

APP functionality overview



The effect of speed on the MPG during a highway driving

- Three recommended routes (eco vs. fast vs. balanced) displayed an estimation of the fuel consumption that was calculated considering three main effects: the average velocity, the traffic and the car used.
  - Specifically, the effects of the velocity and car used were modeled by inputting the MPG (Miles per gallon) values in the Urban Dynamometer Driving Schedule (UDDS) cycle (average velocity of 21.2 mph) and in the highway cycle (average velocity of 48.3 mph)
  - Values on U.S. EPA (Environmental Protection Agency) website
  - The effect of traffic was estimated using the data provided by MapQuest API

$$MPG_{global} = \frac{\sum Distances_{FREE\_FLOW} * MPG(Average\ speed) + \sum Distances_{NON\_FREE\_FLOW} * MPG\_city}{Total\ distance}$$

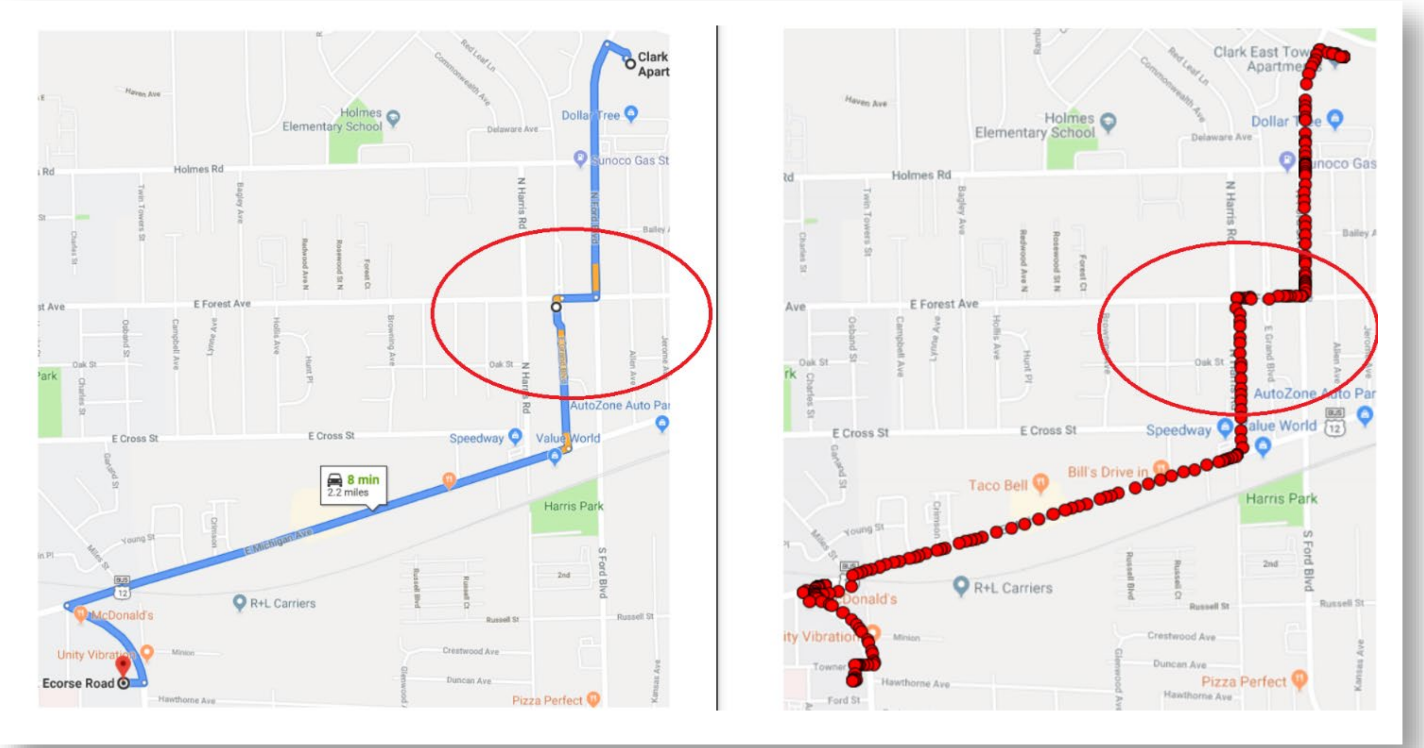
Data collection and analysis

Participant recruitment and data collection procedure

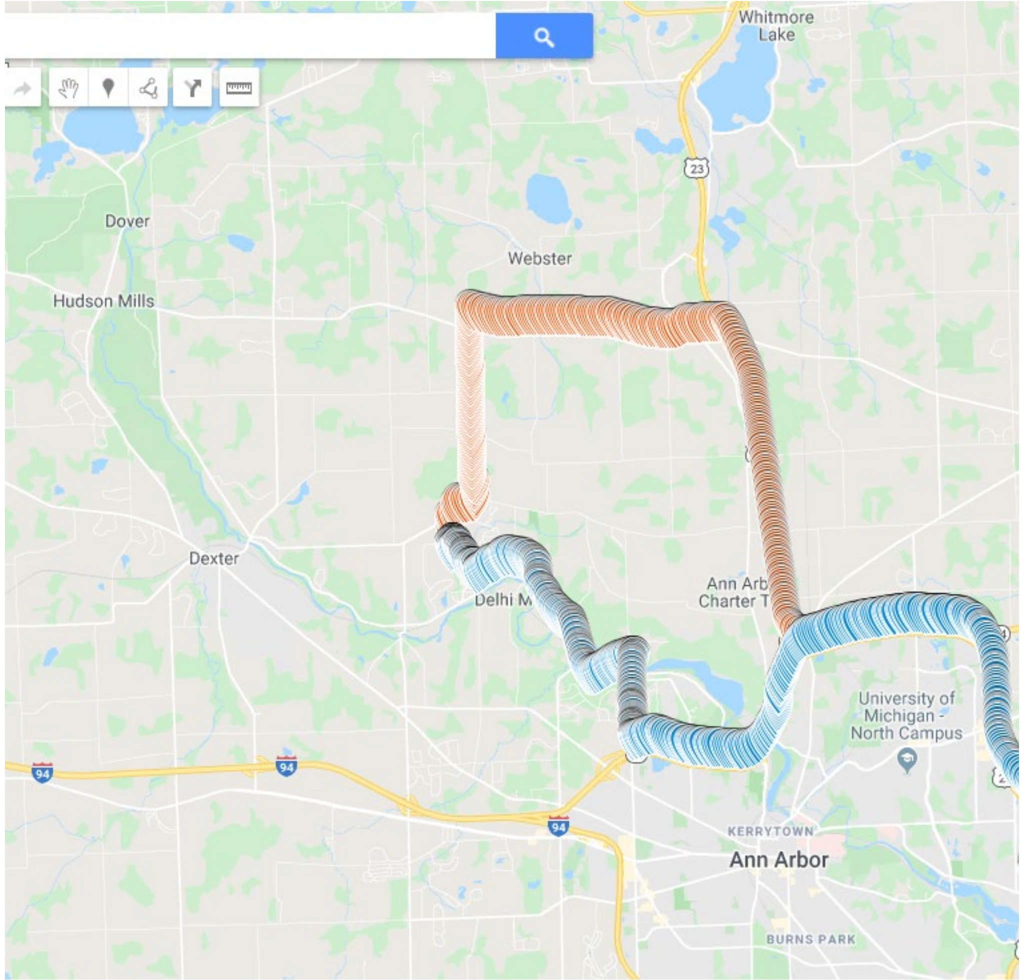
- 43 participants were given a cell phone handset for two weeks
  - Participants who were participating in UMTRI's Ann Arbor Connected Vehicle Test Environment (AACVTE) were recruited to receive the cellular device with the custom-designed, Eco-routing software application (Eco-routing device).
  - This interaction included completing a route-choice survey and selecting from recommended driving routes prior to beginning of some of their driving trips over a 2-week period.
  - Participants were expected to record data for 20 trips and paid to $100 for their time.
  - UM IRB approved.

Data reduction and preparation

- Valid trips were first identified as having a completed survey and a fairly complete GPS dataset.
- A total of 738 valid trips from 1,024 trips
- 39 participants (22 female and 17 male drivers), aged from 20 to 72 years old (Mean=47.3, S.D.=15.3)
- Ovitalmaps software was used to create the trace maps of the Recommended Route.



GPS Data reduction and preparation



1: Correct GPS drift

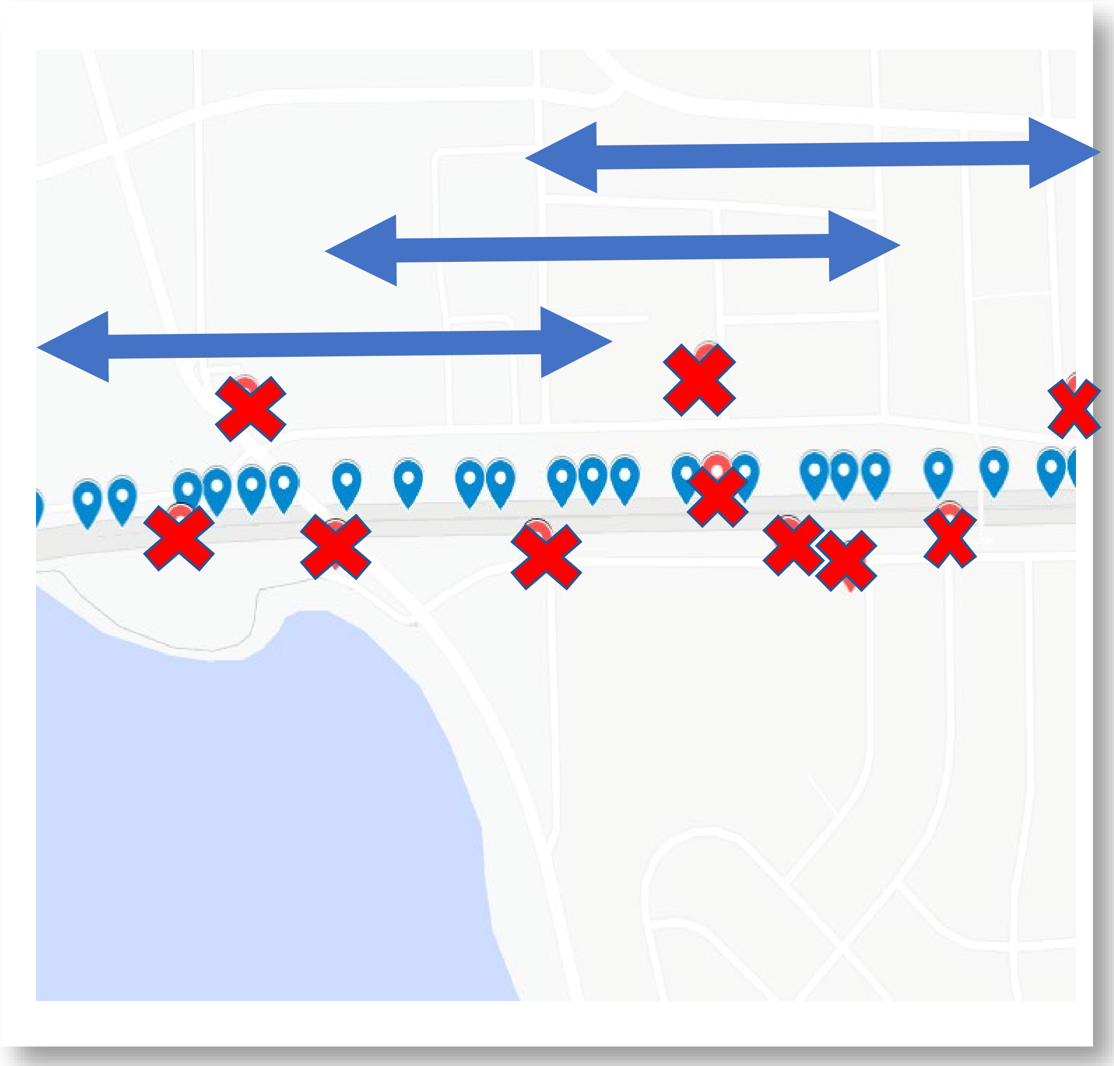
2: Reverse geocoding

3: Identify overlap road segments between real driving and recommended routes

4: Calculate overlap percentage

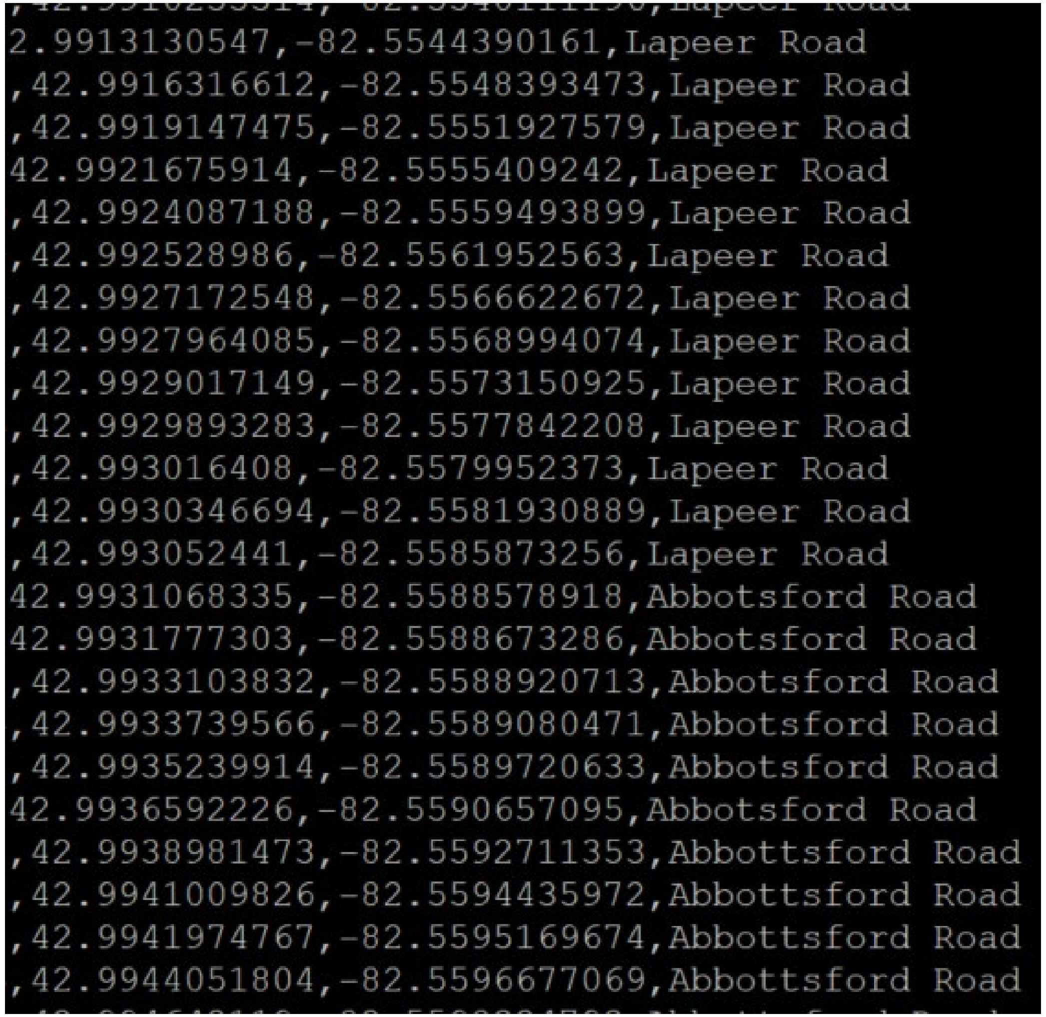
Google API

GPS signal drift issue



- Trip-based correction
- Google Maps API : “Snap to Roads”
- Input 100 GPS data points at each step
- The Roads API returns 100 new/corrected GPS data points

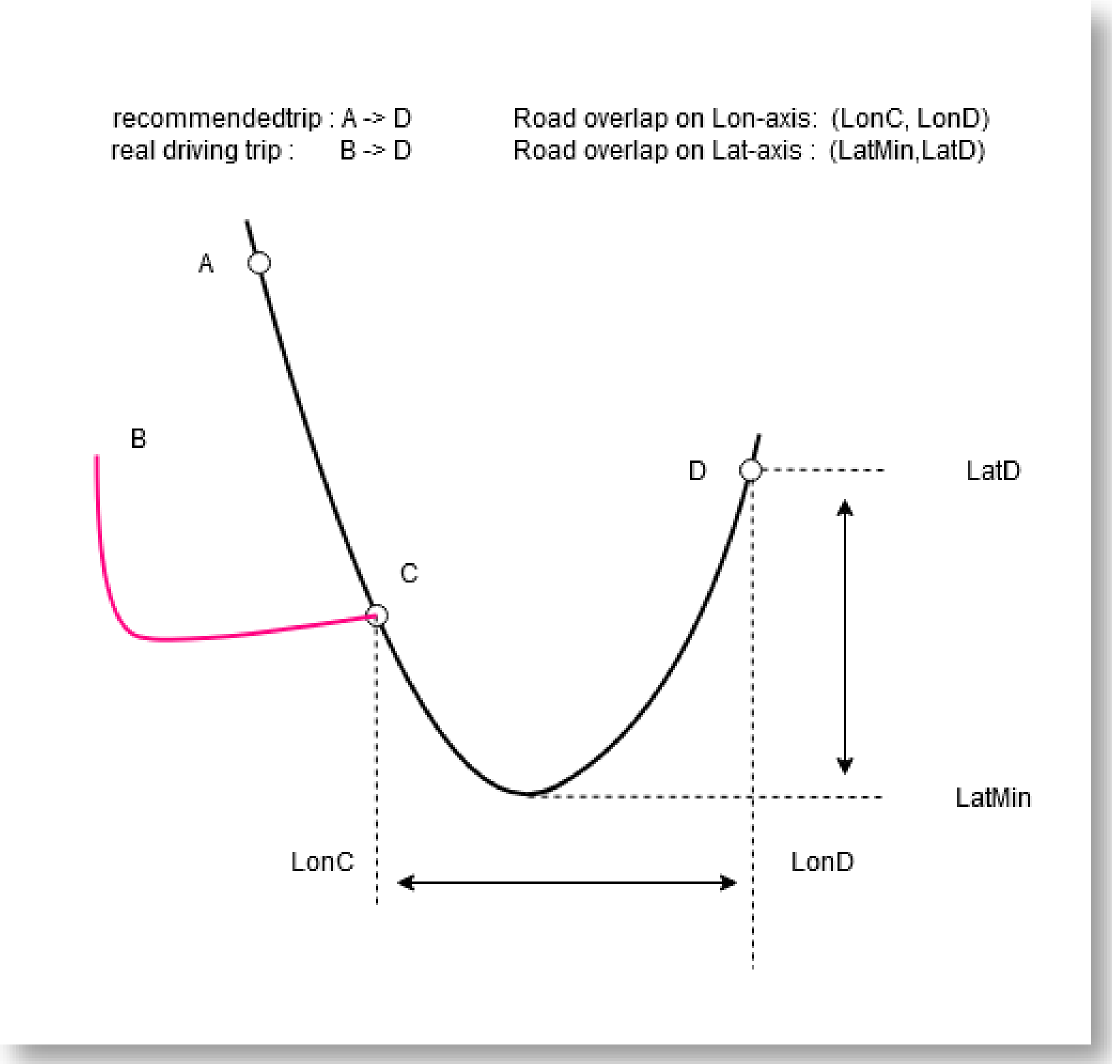
Mapping Location: Road Names



```
2.9913130547,-82.5544390161,Lapeer Road
,42.9916316612,-82.5548393473,Lapeer Road
,42.9919147475,-82.5551927579,Lapeer Road
42.9921675914,-82.5555409242,Lapeer Road
,42.9924087188,-82.5559493899,Lapeer Road
,42.992528986,-82.5561952563,Lapeer Road
,42.9927172548,-82.5566622672,Lapeer Road
,42.9927964085,-82.5568994074,Lapeer Road
,42.9929017149,-82.5573150925,Lapeer Road
,42.9929893283,-82.5577842208,Lapeer Road
,42.993016408,-82.5579952373,Lapeer Road
,42.9930346694,-82.5581930889,Lapeer Road
,42.993052441,-82.5585873256,Lapeer Road
42.9931068335,-82.5588578918,Abbotsford Road
42.9931777303,-82.5588673286,Abbotsford Road
,42.9933103832,-82.5588920713,Abbotsford Road
,42.9933739566,-82.5589080471,Abbotsford Road
,42.9935239914,-82.5589720633,Abbotsford Road
42.9936592226,-82.5590657095,Abbotsford Road
,42.9938981473,-82.5592711353,Abbottsford Road
,42.9941009826,-82.5594435972,Abbottsford Road
,42.9941974767,-82.5595169674,Abbottsford Road
,42.9944051804,-82.5596677069,Abbottsford Road
```

- Reverse geocoding function
- Reverse geocoding converts the geographic coordinates (latitude, longitude) to addresses (like "2901 Baxter Rd, Ann Arbor, MI 48109")
- Only need road names
- Break each trip into road segments

Calculating overlap road segments



- Road Overlap Algorithm:

On the same road, recommended trip segment is (startA, endB); real driving trip segment is (startC, endD)

Define:

$minLon(AB)$ is the minimum Lon value from startA to endB

$maxLon(AB)$ is the maximum Lon value from startA to endB

$minLat(AB)$ is the minimum Lat value from startA to endB

$maxLat(AB)$ is the maximum Lat value from startA to endB

$minLon(CD)$ is the minimum Lon value from startC to endD

$maxLon(CD)$ is the maximum Lon value from startC to endD

$minLat(CD)$ is the minimum Lat value from startC to endD

$maxLat(CD)$ is the maximum Lat value from startC to endD

Project segment AB and segment CD to Lon axis and Lat axis:

Seg AB proj to Lon:

(min(LonA, LonB, $minLon(AB)$),max(LonA, LonB, $maxLon(AB)$))

Seg AB proj to Lat:

(min(LatA, LatB, $minLat(AB)$),max(LatA, LatB, $maxLat(AB)$))

Seg CD proj to Lon:

(min(LonC, LonD, $minLon(CD)$),max(LonC, LonD, $maxLon(CD)$))

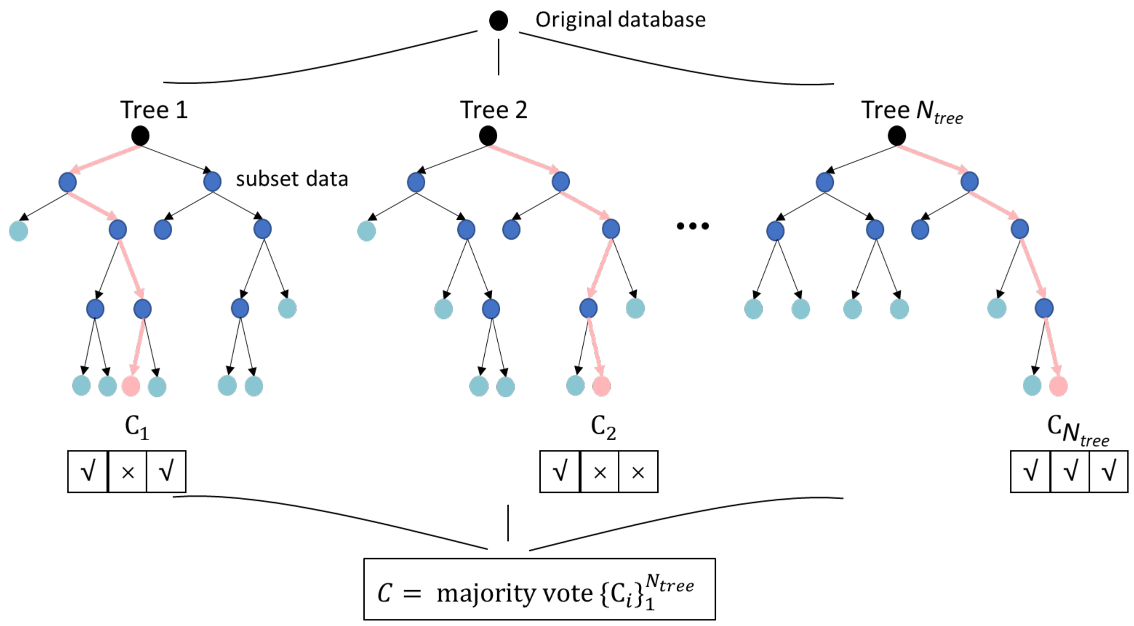
Seg CD proj to Lat:

(min(LatC, LatD, $minLat(CD)$),max(LatC, LatD, $maxLat(CD)$))

Mixed model analyses and Multi-label random forests

Mixed model analyses

Mixed model analyses can be used to model **random and fixed effect** data and has the benefit of well-handling data with **heterogeneous variances** and **auto-correlated observations**. This procedure has been applied in many naturalistic driving study data analyses .



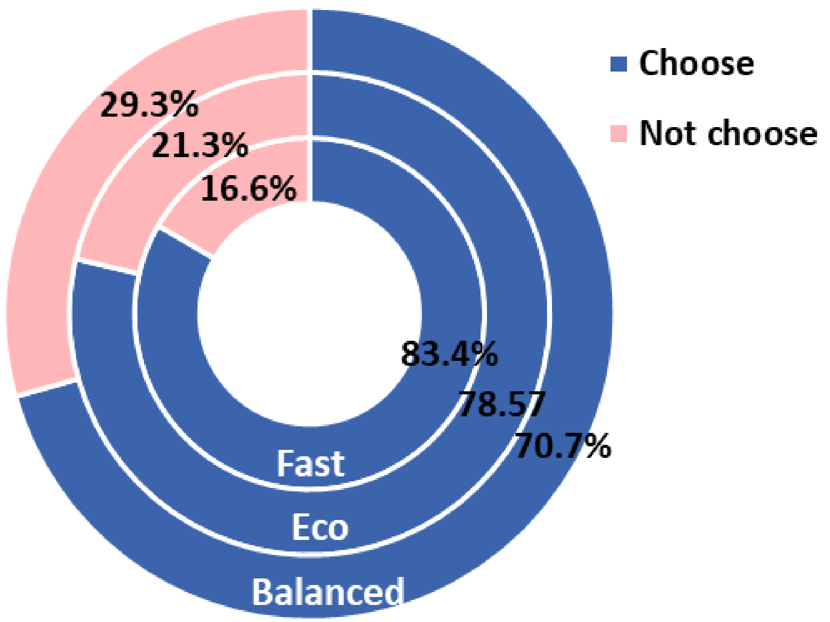
Multi-label random forests (MLRF)

Different from classical multi-class classification problems, rational **multi-label classification** models are capable to take underlying correlations among different labels into consideration.

To extend Random forests to be able to perform multi-label classification, multi-label decision trees that can generalize the entropy of the sample set to adapt to multi-label data sets, are employed to develop the MLRF method.

Results: Modelling driver's route choice

- The driver's route choice is a **multi-label** problem, since each recommended route may have several different features at the same time.
- In general, drivers were more likely to choose the **fast** route, having the highest average probability **83.4%.**
- The next was the **eco** route with a selection probability around **78.57%**.
- The routes with the **balanced** feature had the least likelihood to be selected, averagely **70.7%** for each driver.



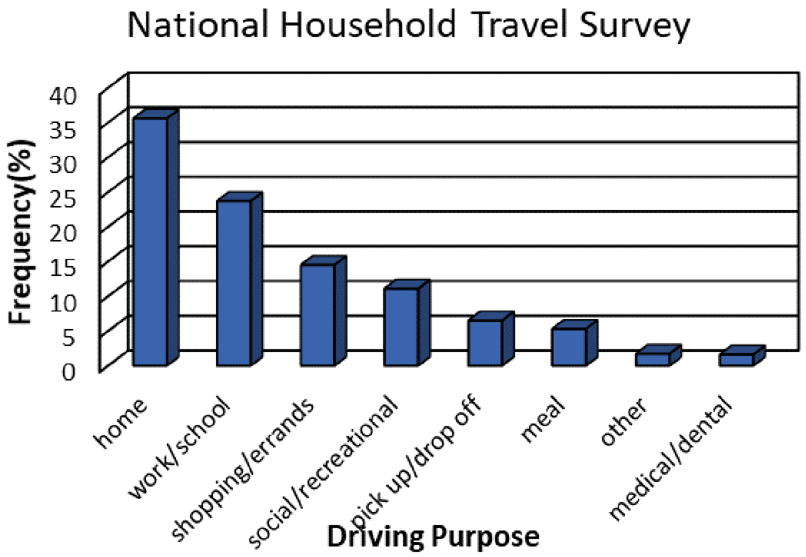
The average probability of route choice

Impacting factors on the eco-route choice

Independent variables

- **Route information**: distance, distance saving, average gas consumption, recommendation sequence, number of routes;
- **Driver characteristics**: age and gender;
- **Subjective data**: purpose, decision time, household passenger, non-household passenger, flexibility, prior activity, and leave earlier.

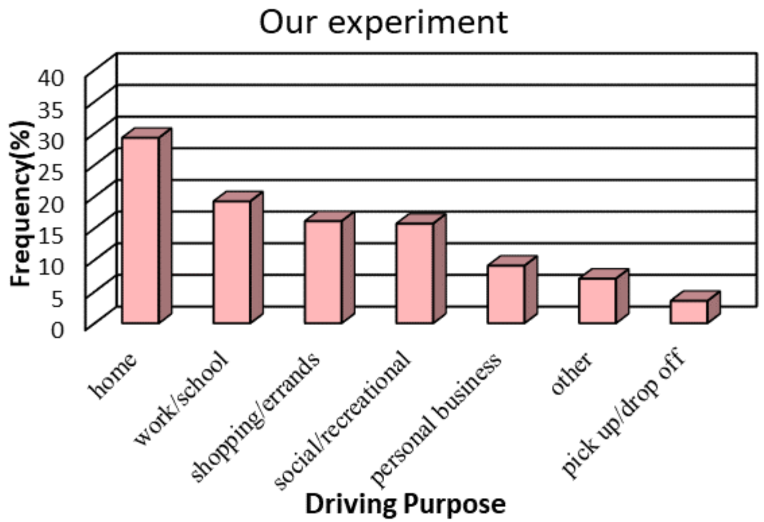
Dependent variable: whether choose the eco route or not



Mixed model analyses

- Drivers were more likely to select the eco route when its **distance was shorter** and **gas consumption per mile was higher.**
- **Giving priority to recommend the eco route** could guide drivers to choose the eco way.

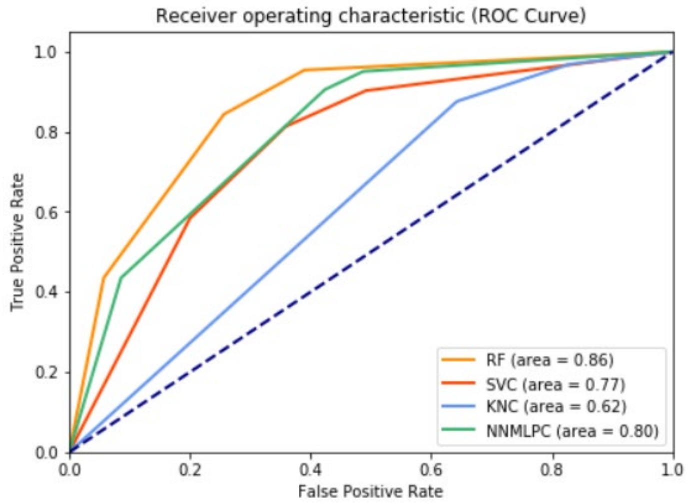
Effect	Estimate	Standard error	DF	t Value	Pr>|t|
Intercept	1.139	0.715	38	1.59	0.119
Distance	-0.056	0.010	643	-5.56	0.019
Average gas consumption	40.868	17.395	643	2.35	<0.001
Sequence					
1st*	0				
2nd	-1.925	0.316	52	-6.10	<0.001
3rd	-2.518	0.491	52	-5.13	<0.001



Predicting driver's route choice behavior

The multi-label Random Forests (RF) classification model

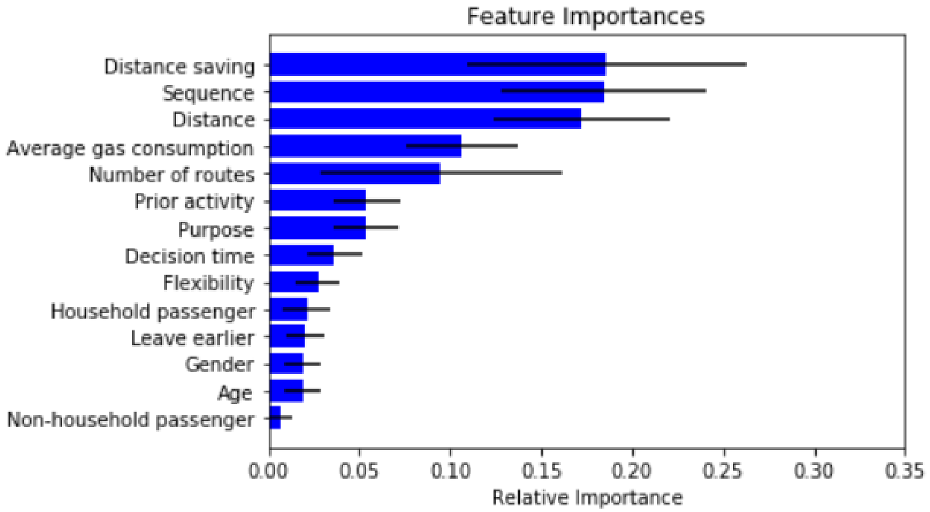
- 14 Independent variables ;
- A multi-label dependent variable: eco, fast, and balanced
- All 737 samples (training: testing=7:3)
- Overall accuracy: training OOB(out-of-bag) accuracy 87.0%, testing 79.3%
- Some other methods: K-neighbors classifier (KNC), Support vector classification (SVC); Neural network multi-layer perceptron classifier (NNMLPC)



Label	Precision	Recall	f1-score
Eco	0.86	0.85	0.86
Fast	0.90	0.95	0.92
Balanced	0.81	0.77	0.78
Overall accuracy	Training (OOB): 0.870; Testing: 0.793		

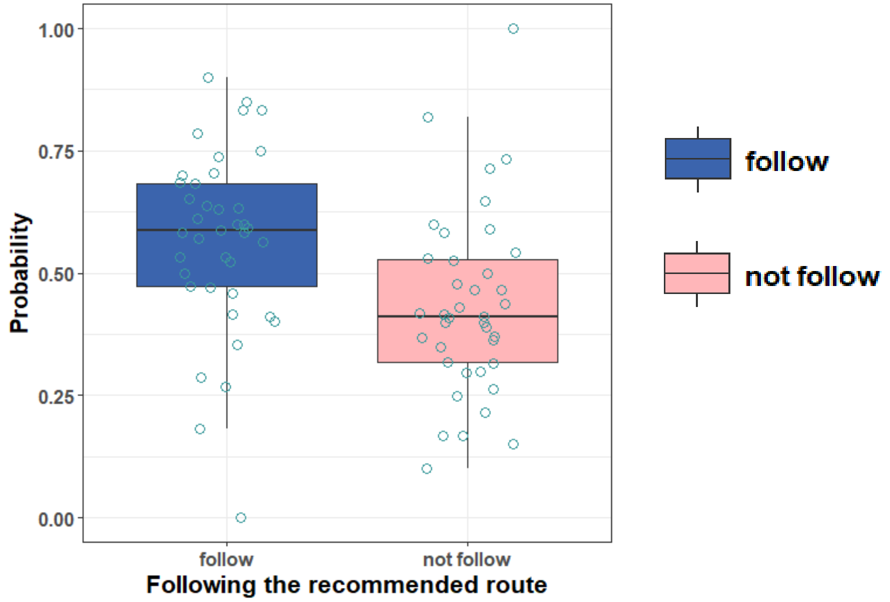
Variable Importance

- Variables from **route information** showed the largest impacts
- The following were subjective data
- No obvious relationships were found in demographic data



Impacting factors on following the recommended route

The average probability that drivers would actually follow the route after they chose from the recommended options was **56.7%**.
eco (61.6%) fast (61.1%) balanced (59.9%)



The average probability of following the recommended route

Independent variables (12)

- **Route information**: eco, fast, and balanced;
- **Driver characteristics**: age and gender;
- **Subjective data**: purpose, decision time, household passenger, non-household passenger, flexibility, prior activity, and leave earlier.

Dependent variable: whether driver's following the selected route

- If driver choose the **eco** or **fast** routes, they are more likely to fully drive along the recommended route.
- Compared with driving alone or with only one household passenger, drivers will comply with the recommended route when there were **three or more household passengers**.

Mixed model results for the recommended route following

Effect	Estimate	Standard error	DF	t Value	Pr>|t|
Intercept	-1.199	0.257	22	-4.67	<0.001
Eco route	0.776	0.215	32	3.61	0.001
Fast route	1.071	0.229	31	4.68	<0.001
Household member					
0	-2.394	0.852	22	-2.81	0.010
1	-2.635	0.893	22	-2.95	0.007
2	-2.003	0.989	22	-2.02	0.055
3 or more*	0				

Conclusions

- Our study data are consistent with national household survey data on trip purposes.
- Drivers were more likely to select the eco route when its distance was shorter and gas consumption per mile was higher.
- Giving priority to recommend the eco route could guide drivers to choose the eco way.
- If driver choose the eco or fast routes, they are more likely to fully drive along the recommended route.
- Compared with driving alone or with only one household passenger, drivers will comply with the recommended route when there were three or more household passengers.

References

Boriboonsomsin, K., Dean, J., & Barth, M. (2014). Examination of attributes and value of ecologically friendly route choices. Transportation Research Record, 2427(1), 13-25.

EPA. (2019). U.S. Transportation Sector Greenhouse Gas Emissions: 1990–2017. Office of Transportation and Air Quality EPA-420-F-19-047.

U.S. Department of Transportation Federal Highway Administration (2020). National Household Travel Survey in 2017. Retrieved January 16, 2020, from https://nhts.ornl.gov/vehicle-trips.

Zeng, W., Miwa, T., & Morikawa, T. (2017). Application of the support vector machine and heuristic k-shortest path algorithm to determine the most eco-friendly path with a travel time constraint. Transportation Research Part D: Transport and Environment, 57, 458-473.

Acknowledgment

- Center for Connected and Automated Transportation (CCAT)
- Department of Energy (DOE)
- Argonne National Laboratory (ANL)

Thank you! Questions?

Shanbao@umich.edu