LARGE AMOUNT OF CONGESTION CAUSED BY TRAFFIC CRUISING

LOOKING FOR PARKING

DEADHEAD VEHICLES FOR HIRE

10 MILES

7-13 MILES

PROBLEM
INTERSECTION WITH SENSOR

INTERSECTION WITHOUT SENSOR

PROBLEM

<table>
<thead>
<tr>
<th>HASHED_MAC</th>
<th>TIME</th>
<th>SENSOR</th>
<th>STRENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>KD98SDK8AH</td>
<td>8:32:01</td>
<td>276105</td>
<td>-52</td>
</tr>
<tr>
<td>8DJSKDL5X0</td>
<td>8:32:01</td>
<td>276102</td>
<td>-55</td>
</tr>
<tr>
<td>439WOA09A</td>
<td>8:32:01</td>
<td>265402</td>
<td>-75</td>
</tr>
<tr>
<td>777AJDKAL8</td>
<td>8:32:05</td>
<td>293010</td>
<td>-50</td>
</tr>
<tr>
<td>QKSJ239A99</td>
<td>8:32:07</td>
<td>251040</td>
<td>-45</td>
</tr>
<tr>
<td>DQWPPOA09</td>
<td>8:32:10</td>
<td>265402</td>
<td>-49</td>
</tr>
<tr>
<td>KD98SDK8AH</td>
<td>8:32:11</td>
<td>265302</td>
<td>-54</td>
</tr>
<tr>
<td>TECHNICAL CHALLENGES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>INCOMPLETE GRID</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensors only cover</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37% of the grid.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BIG DATA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensors produce</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200K observations / hr.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SENSOR DETECTION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensors detect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38% of devices w/ wifi on.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PRIVACY AND DATA GOVERNANCE

DATA IS ANONYMIZED

INEXACT LOCATION

AGGREGATED VIEWS

RAW DATA NOT RETAINED

DEPLOYABLE BEHIND TRANSPORTATION DATA COLLABORATIVE
OBJECTIVES

1. Repurpose sensor network
2. Differentiate types of traffic cruising
3. Visualize with a heat map
4. Scale up and deploy
ANONYMOUS SENSOR READINGS → ESTIMATED PATHS → METADATA → LABELED PATHS → AGGREGATE HEATMAP

ENCRIPTED HASH | TIME | SENSOR
---|---|---
K098SIDKHB8H8X | 8:32:00 | 2761005
809KDLX30WKK | 8:32:01 | 276102
DQWPOAO9065D | 8:32:01 | 265402
K9G5AE9V0X03 | 8:32:01 | 265302
DXFOSPHISHH8H | 8:32:01 | 265301
0186SDVCBNAA | 8:32:01 | 273777
00LPZZZ998A999 | 8:32:02 | 265000
NVMCS90AUKSSS | 8:32:02 | 265003
QAL3ECDHDD000 | 8:32:02 | 265434
DQWPOAO9065D | 8:32:02 | 273099
0186SDVCBNAA | 8:32:02 | 273010
00LPZZZ998A999 | 8:32:02 | 265001
ASKDPPPLQSLAA | 8:32:03 | 265000

APPROACH OVERVIEW
GROUP ANONYMOUS READINGS $\rightarrow$ REMOVE SHORT PATHS $\rightarrow$ REMOVE FALSE DETECTIONS $\rightarrow$ ESTIMATE ROUTE $\rightarrow$ SEGMENT AND LABEL

PIPELINE
ANONYMOUS SENSOR READINGS → ESTIMATED PATHS → METADATA → LABELED PATHS → AGGREGATE HEATMAP

APPROACH OVERVIEW

<table>
<thead>
<tr>
<th>ENCRYPTED HASH</th>
<th>TIME</th>
<th>SENSOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>KD985DQ68A68X</td>
<td>8:32:00</td>
<td>276105</td>
</tr>
<tr>
<td>803K6L3K0W0K</td>
<td>8:32:01</td>
<td>276102</td>
</tr>
<tr>
<td>DQWOPPOG0656</td>
<td>8:32:01</td>
<td>265402</td>
</tr>
<tr>
<td>KD096P6G6X03</td>
<td>8:32:01</td>
<td>265302</td>
</tr>
<tr>
<td>DQK0P66H05H00</td>
<td>8:32:01</td>
<td>265101</td>
</tr>
<tr>
<td>G16D9SVCYBNAA</td>
<td>8:32:01</td>
<td>263777</td>
</tr>
<tr>
<td>02D9PZ22989999</td>
<td>8:32:02</td>
<td>265000</td>
</tr>
<tr>
<td>NVK65K054UK55</td>
<td>8:32:02</td>
<td>265003</td>
</tr>
<tr>
<td>QAL3E9CHD0000</td>
<td>8:32:02</td>
<td>265434</td>
</tr>
<tr>
<td>DQWOPPAO6666</td>
<td>8:32:02</td>
<td>275099</td>
</tr>
<tr>
<td>G16D9SVCYBNAA</td>
<td>8:32:02</td>
<td>273010</td>
</tr>
<tr>
<td>02D9PZ22989999</td>
<td>8:32:02</td>
<td>265001</td>
</tr>
<tr>
<td>ASK8R6R696L6AA</td>
<td>8:32:03</td>
<td>265000</td>
</tr>
</tbody>
</table>

25 MPH
2 STOPS
0.35 DISTANCE RATIO
ETC.
METADATA COLLECTION

PATH

FEATURES

NUMBER OF TIMES PATH CROSSSED
AVERAGE SPEED
MAX SPEED
STANDARD DEVIATION OF SPEED
NUMBER OF LONG STOPS
PERCENTAGE OF TIME DRIVING
PERCENTAGE OF TIME WALKING
ANONYMOUS SENSOR READINGS ➔ ESTIMATED PATHS ➔ METADATA ➔ Labeled Paths ➔ AGGREGATE HEATMAP

APPRAOCH OVERVIEW
DEFINING CRUISING

PROBLEM

1. NO STANDARD DEFINITION OF CRUISING
2. CURRENTLY NO GROUND TRUTH

SOLUTION

1. MULTI-STEP CLASSIFICATION
2. LABEL SUBSET OF EXTREME CASES
3. MACHINE LEARNING TO IDENTIFY CRUISING INDICATORS
USING DISTANCE RATIO

\[ \frac{5}{5} = 1.0 \]

PROBABLY NOT CRUISING

\[ \frac{14}{2} = 7.0 \]

PROBABLY CRUISING
MULTI-STEP CLASSIFICATION

ALL TRIPS
MULTI-STEP CLASSIFICATION

PROBABLY CRUISING

- 12.7%

PROBABLY NOT CRUISING

- 39.5%
- 47.8%
MULTI-STEP CLASSIFICATION

PROBABLY CRUISING

PROBABLY NOT CRUISING

TRAIN ALGORITHM
MULTI-STEP CLASSIFICATION

PROBABLY CRUISING

PROBABLY NOT CRUISING

TRAIN ALGORITHM

TEST ALGORITHM
MULTI-STEP CLASSIFICATION

PROBABLY CRUISING

PROBABLY NOT CRUISING

TRAIN ALGORITHM

TEST ALGORITHM
MULTI-STEP CLASSIFICATION

TESTING ALGORITHM

DECISION TREE
ACCURACY: 96.6%

LOGISTIC REGRESSION
ACCURACY: 91.4%

GRADIENT BOOSTING
ACCURACY: 97.1%
MULTI-STEP CLASSIFICATION

TESTING ALGORITHM

- DECISION TREE
  Accuracy: 96.6%

- LOGISTIC REGRESSION
  Accuracy: 91.4%

- GRADIENT BOOSTING
  Accuracy: 97.1%
LABELING ALL DATA

ALL TRIPS
LABELING ALL DATA

ALL TRIPS

PROBABLY CRUISING

PROBABLY NOT CRUISING
LABELING ALL DATA

UNLABELED

PROBABLY NOT CRUISING

PROBABLY CRUISING
LABELING ALL DATA

UNLABELED

GRADIENT BOOSTING ALGORITHM

PROBABLY NOT CRUISING

PROBABLY CRUISING
LABELING ALL DATA

UNLABELED

PROBABLY NOT CRUISING

PROBABLY CRUISING

GRADIENT BOOSTING ALGORITHM

Labeled Cruising

Labeled Not Cruising
IDENTIFYING FOR-HIRE VEHICLES
IDENTIFYING FOR-HIRE VEHICLES

FOR-HIRE VEHICLE EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

FOR-HIRE VEHICLE EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

FOR-HIRE VEHICLE EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

FOR-HIRE VEHICLE EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

FOR-HIRE VEHICLE EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

FOR-HIRE VEHICLE EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)

UNIQUE SENSORS / TOTAL READS =

22 / 22 = 1.0 [HIGH DISPERSION]
IDENTIFYING FOR-HIRE VEHICLES
IDENTIFYING FOR-HIRE VEHICLES

**BUS EXAMPLE**

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

**BUS EXAMPLE**

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

**BUS EXAMPLE**

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

**BUS** EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)
IDENTIFYING FOR-HIRE VEHICLES

**BUS EXAMPLE**

4 LARGE GAPS IN READ TIMES (5 TRIPS)

**UNIQUE SENSORS / TOTAL READS =**

\[
\frac{5}{25} = 0.2 \quad \text{[LOW DISPERSION]}
\]
IDENTIFYING FOR-HIRE VEHICLES

**FOR-HIRE VEHICLE** EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)

UNIQUE SENSORS / TOTAL READS =

\[ \frac{22}{22} = 1.0 \]  [HIGH DISPERSION]

**BUS** EXAMPLE

4 LARGE GAPS IN READ TIMES (5 TRIPS)

UNIQUE SENSORS / TOTAL READS =

\[ \frac{5}{25} = 0.2 \]  [LOW DISPERSION]
ANONYMOUS SENSOR READINGS ➔ ESTIMATED PATHS ➔ METADATA ➔ LABELED PATHS ➔ AGGREGATE HEATMAP

APPROACH OVERVIEW

<table>
<thead>
<tr>
<th>ENCRYPTED HASH</th>
<th>TIME</th>
<th>SENSOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>KD9850K8A608X</td>
<td>8:32:00</td>
<td>276105</td>
</tr>
<tr>
<td>30950350K80W9K</td>
<td>8:32:01</td>
<td>276102</td>
</tr>
<tr>
<td>D2W2P0OA90C5D</td>
<td>8:32:02</td>
<td>263402</td>
</tr>
<tr>
<td>KD0SPALG9X03</td>
<td>8:32:02</td>
<td>266502</td>
</tr>
<tr>
<td>DEKPOPHS30X1</td>
<td>8:32:02</td>
<td>265101</td>
</tr>
<tr>
<td>01CB20CVBNNAA</td>
<td>8:32:02</td>
<td>273777</td>
</tr>
<tr>
<td>09L2PZZZ986A99</td>
<td>8:32:02</td>
<td>265000</td>
</tr>
<tr>
<td>N9M5C9OAU55S</td>
<td>8:32:02</td>
<td>265101</td>
</tr>
<tr>
<td>QA05B2K4900D</td>
<td>8:32:02</td>
<td>265434</td>
</tr>
<tr>
<td>D2W2P0OA90C5D</td>
<td>8:32:02</td>
<td>273099</td>
</tr>
<tr>
<td>01CB20CVBNNAA</td>
<td>8:32:02</td>
<td>273010</td>
</tr>
<tr>
<td>09L2PZZZ986A99</td>
<td>8:32:02</td>
<td>265001</td>
</tr>
<tr>
<td>A9KDP999QSLAA</td>
<td>8:32:03</td>
<td>265000</td>
</tr>
</tbody>
</table>

25 MPH
2 STOPS
0.35 DISTANCE RATIO
ETC.
AGGREGATION

PATH ONE
PATH TWO
PATH THREE

AGGREGATED COUNTS
CRUISING IN DOWNTOWN SEATTLE

As part of The Data Science for Social Good Program at The University of Washington, the Traffic Cruising Team has produced a heatmap to identify cruising in the downtown Seattle area.

SELECT TYPE OF CRUISING
- PARKING

SELECT DAY
- MON
- TUE
- WED
- THR
- FRI
- SAT
- SUN

SELECT TIME
- ALL DAY
- MORNING
- MIDDAY
- EVENING

NUMBER OF PATHS
- 0 - 59
- 59 - 186
- 186 - 296
- 296 - 458
- 458 - 567
- 567 - 681
- 681 - 855
- 855 - 1055
- 1055 - 1345
- 1345+

NUMBER OF PATHS

TIME
0 10 20 30 40 50 60 70 80 90 10 11 12 13 14 15 16 17 18 19 20
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
CRUISING IN DOWNTOWN SEATTLE

As part of The Data Science for Social Good Program at The University of Washington, the Traffic Cruising Team has produced a heatmap to identify cruising in the downtown Seattle area.

SELECT TYPE OF CRUISING
- PARKING

SELECT DAY
- MON
- TUE
- WED
- THR
- FRI
- SAT
- SUN

SELECT TIME
- ALL DAY
- MORNING
- MIDDAY
- EVENING

NUMBER OF PATHS

<table>
<thead>
<tr>
<th>Number of Paths</th>
<th>Color Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 59</td>
<td>0 - 59</td>
</tr>
<tr>
<td>59 - 186</td>
<td>59 - 186</td>
</tr>
<tr>
<td>186 - 296</td>
<td>186 - 296</td>
</tr>
<tr>
<td>296 - 456</td>
<td>296 - 456</td>
</tr>
<tr>
<td>456 - 567</td>
<td>456 - 567</td>
</tr>
<tr>
<td>567 - 681</td>
<td>567 - 681</td>
</tr>
<tr>
<td>681 - 855</td>
<td>681 - 855</td>
</tr>
<tr>
<td>855 - 1055</td>
<td>855 - 1055</td>
</tr>
<tr>
<td>1055 - 1345</td>
<td>1055 - 1345</td>
</tr>
<tr>
<td>1345+</td>
<td>1345+</td>
</tr>
</tbody>
</table>

NUMBER OF PATHS VS TIME

- Graph showing the number of paths over time from 12 AM to 10 PM.
CONCLUSION

EXISTING TRAFFIC SENSORS CAN BE USED TO IDENTIFY CRUISING

YEAR-ROUND OBSERVATION OF CRUISING PATTERNS

CITY OF SEATTLE BENEFITS

SAME SYSTEM CAN BE APPLIED TO OTHER CITIES
CONCLUSION

EXISTING TRAFFIC SENSORS CAN BE USED TO IDENTIFY CRUISING

YEAR-ROUND OBSERVATION OF CRUISING PATTERNS

CITY OF SEATTLE BENEFITS

SAME SYSTEM CAN BE APPLIED TO OTHER CITIES
EXISTING TRAFFIC SENSORS CAN BE USED TO IDENTIFY CRUISING

- YEAR-ROUND OBSERVATION OF CRUISING PATTERNS
- CITY OF SEATTLE BENEFITS
- SAME SYSTEM CAN BE APPLIED TO OTHER CITIES
CONCLUSION

EXISTING TRAFFIC SENSORS CAN BE USED TO IDENTIFY CRUISING

YEAR-ROUND OBSERVATION OF CRUISING PATTERNS

CITY OF SEATTLE BENEFITS

SAME SYSTEM CAN BE APPLIED TO OTHER CITIES
VALIDATION
<table>
<thead>
<tr>
<th>MODEL</th>
<th>ACCURACY</th>
<th>AUC_ROC</th>
<th>PRECISION</th>
<th>RECALL</th>
<th>F1_SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECISION TREE</td>
<td>0.966 +/- 0.000</td>
<td>0.933 +/- 0.000</td>
<td>0.993 +/- 0.000</td>
<td>0.868 +/- 0.000</td>
<td>0.926 +/- 0.000</td>
</tr>
<tr>
<td>LOGISTIC REGRESSION</td>
<td>0.925 +/- 0.001</td>
<td>0.870 +/- 0.001</td>
<td>0.921 +/- 0.001</td>
<td>0.762 +/- 0.001</td>
<td>0.834 +/- 0.001</td>
</tr>
<tr>
<td>GRADIENT BOOSTING CLASSIFIER</td>
<td>0.967 +/- 0.000</td>
<td>0.935 +/- 0.000</td>
<td>0.994 +/- 0.000</td>
<td>0.871 +/- 0.000</td>
<td>0.928 +/- 0.000</td>
</tr>
<tr>
<td>FEATURE</td>
<td>IMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUMBER OF TIMES PATH CROSSED</td>
<td>0.3442</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVERAGE SPEED</td>
<td>0.1191</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENTAGE OF TIME DRIVING</td>
<td>0.1694</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOUR</td>
<td>0.0101</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAXIMUM SPEED</td>
<td>0.0442</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STD DEV OF SPEED</td>
<td>0.0556</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STOPS 0 - 2 MIN</td>
<td>0.0323</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STOPS 10 - 15 MIN</td>
<td>0.0126</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STOPS 15+ MIN</td>
<td>0.0377</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIME PERCENTAGE CHILLING</td>
<td>0.0906</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIME PERCENTAGE WALKING</td>
<td>0.0351</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SENSOR DETECTION RATE

2ND AVE SENSOR

START
- 58% 2ND • STEWART
- 32% 2ND • PINE
- 34% 2ND • PIKE
- 22% 2ND • UNIVERSITY
- 31% 2ND • SENECA
- 47% 2ND • SPRING
- 36% 2ND • MADISON
- 36% 2ND • MARION
- 35% 2ND • COLUMBIA
- 49% 2ND • CHERRY

END
- 2ND • JAMES
RESULTS
RESULTS
RESULTS

PARKING TUE MORNING

PARKING SAT EVENING

NUMBER OF PATHS

0 - 19
19 - 56
56 - 81
81 - 108
108 - 147
147 - 179
179 - 225
225 - 271
271 - 339
339+

NUMBER OF PATHS

0 - 13
13 - 62
62 - 95
95 - 132
132 - 167
167 - 196
196 - 235
235 - 272
272 - 404
404+