Vehicular Data Analytics for Smart Driving

Dilum Bandara, PhD
Dept. of Computer Science & Engineering
University of Moratuwa
Dilum.Bandara@uom.lk

International Symposium on Applied Analytics for a Smart Society – 2016
Why?

- Reckless Driving
- Fault Detection & Prediction
- Fuel Fraud
- IoT
- Data Analytics

Driving Anomaly Detection

Smart Driving

Tracking & Surveillance
Solution Architecture

- Real-time analysis
  - Driving anomaly detection
  - Fuel fraud
  - Geo fencing
  - Vehicle fault detection

- Historical analysis
  - Driver profiling
  - Driver coaching
  - Predicting sensor failure
  - Case analysis
Dashboard

- **Number of Vehicles**: 3 vehicles
  - 40% from last month

- **Distance Traveled**: 12,309 km
  - 4% from last month

- **Average Speed**: 46 km/h
  - 5% from last month

- **Aggressive Driving**: 76 events per 100km
  - 34% from last month

- **Idle Time**: 3h 15m
  - 72% from last month

- **Speeding Time**: 4h 39m
  - 34% from last month

**Days of Driving**
- Monday: 15%
- Tuesday: 20%
- Wednesday: 30%
- Thursday: 10%
- Friday: 30%
- Saturday: 25%
- Sunday: 10%

**Speed Levels**
- 0 - 30 km/h: 56%
- 30 - 60 km/h: 22%
- 60 - 90 km/h: 13%
- 90 - 120 km/h: 8%
- 120 km/h above: 1%

**Comfort Level**
- 57%
  - 4% from last month

**Idle vs Over Speed**

**Driver Map**
OBD2 Based Analysis

- OBD – On Board Diagnostics
  - Available in many vehicles since 1996
  - OBD2 – In most vehicles since 2005
  - Speed, RPM, Odometer, Coolant Temperature, Padle Position, Oxygen, Mass Air Flow, etc.
Solution Architecture

- Jaggery + Caramel (MVC), WSO2 UES
- WSO2 BAM & CEP
- REST API
- Android
- Bluetooth
App-Level Processing – Real-Time Dashboard
Fuel Economy & Coolant Temperature Monitoring

- Implemented using Siddhi CEP on smartphone
- Minimum impact on battery level
  - Bandwidth saving due to local processing → Reduce energy consumption
Trip Logs

- Standard car → High-end Car
Backend Processing – Reckless Driving

- Hard accelerations & deceleration count above a threshold
  - Per 100 Km
  - Per 1 Hour
- Count depends on average speed of vehicle in last $t$ seconds
- Implemented using Siddhi CEP
- Computed values stored in RDBMS
Driver Profiling

- Detection of anomalies
  - Hidden Markov Model based on acceleration profile
  - Model implemented in BAM
  - Validator implemented in CEP
Sensor Failure Prediction

- Mass Air Flow (MAF) sensor value has a linear relationship with engine RPM.
- When sensor fails, gradient between MAF & RPM reduces with time.
- Rate of change of gradient can predict date of failure.
Fuel Consumption Prediction

- Long-distance bus fitted with a GPS unit & high-precision fuel sensor
- Could you
  - explain variability in fuel consumption
  - predict fuel consumption of a journey
  - give tips to improve fuel consumption
Dataset

- From 13 May 2015 – 31 August 2015
- Parameters
  - Timestamp (date and time)
  - Longitude (Min: 5.918611° N, Max: 9.835556° N)
  - Latitude (Min: 79.516667° E, Max: 81.879167° E)
  - Bearing (0° to 360°)
  - Elevation (Min: 0m, Max: 2,524m)
  - Distance traveled (km) – between two samples
  - Speed (kmh-1)
  - Acceleration (kmh-2)
  - Ignition status (1 – Ignition On or 0 – Ignition Off)
  - Current battery voltage (Min: 0v, Max: 29v)
  - Fuel level (Min: 0L, Max: 218L)
  - Fuel consumption (L)
Bus Route
Fuel Usage
Factors Contributing to Fuel Usage

Total Fuel Consumption Summary Vs Day

Fuel Consumption and Elevation
Factors Contributing to Fuel Usage (Cont.)
Factors Contributing to Fuel Usage (Cont.)

![Graph showing Total Fuel Usage VS No of Harsh Acc/Deacc]
Factors Contributing to Fuel Usage (Cont.)
Variable Importance

- Latitude
- Distance
- FuelLevel
- Longitude
- Speed
- Date
- Bearing
- Elevation
- IgnitionStatus
- Acceleration
- CurrentBatteryVoltage

- Distance
- FuelLevel
- Speed
- Longitude
- Latitude
- Bearing
- Acceleration
- Elevation
- Date
- CurrentBatteryVoltage
- IgnitionStatus
Predicting Fuel Consumption – Random Forrest

Actual Fuel Consumption: 84.08L
Predicted Fuel Consumption: 91.77L
Error: 9.1%
Predicting Fuel Consumption – Gradient Boosting & Neural Network

Fuel Usage Prediction

Fuel Usage Prediction

Fuel Usage Prediction

Fuel Usage Prediction
# Predicting Fuel Consumption (Cont.)

<table>
<thead>
<tr>
<th>Model</th>
<th>Nash-Sutcliffe Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>0.26189</td>
</tr>
<tr>
<td>GB</td>
<td>-0.00240</td>
</tr>
<tr>
<td>ANN</td>
<td>-0.01304</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Error Statistics</th>
<th>RF</th>
<th>GB</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIAS</td>
<td>0.00477</td>
<td>0.00045</td>
<td>0.00274</td>
</tr>
<tr>
<td>MAE</td>
<td>0.02296</td>
<td>0.02585</td>
<td>0.02756</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.04046</td>
<td>0.04715</td>
<td>0.04740</td>
</tr>
</tbody>
</table>
On Going Work

- Dashboard design
- Driver profiling
  - Beyond acceleration profile
  - Correlating with location, time, traffic, & weather
  - Usage-Based Insurance (UBI)
- Quantifying passenger comfort
- Case analysis
  - Traffic, weather
- Driver feedback
  - Real-time & long-term
- Process re-engineering
Research Challenges

- Lack of annotated data
  - Events, GIS, weather
- Optimum sampling frequency – 4Vs of big data
  - GPS 10 Hz, practically < 0.2 Hz
  - OBD2 ~10 Hz per PID
- Enhancing accuracy of detected events
- Correlating with location, time, traffic, & weather
  - Lack of (real-time) data
- Relating numbers to physical events such that drivers could understand
Publications

Acknowledgement

- **Students**
  - Sandareka Wickramanayake (MSc)
  - Shashika Muramudalige (MSc, BSc)
  - Asiri Liyana Arachchi (BSc)
  - Malintha Amarasinghe (BSc)
  - Sasikala Kottegoda (BSc)

- **Research partners**
  - Mr. Nishal Samarasekera (Dept. of TLM)

- **Data**
  - Nimbus Venture (Pvt) Ltd.
  - Many other drivers who help us collected data
Q&A

Dilum.Bandara@uom.lk