Characterising Eco driving with GPS and CANBus data

Karsten Jakobsen
Institute of Computer Science
Aalborg University
Denmark
karstenjjakobsen@gmail.com

Sabrine C. H. Mouritsen
Institute of Computer Science
Aalborg University
Denmark
sabrinechm@gmail.com

1. DATA DESCRIPTION

Name	Data type	Source	Description
vehicleid	integer	ID	Unique identifier for vehicles
timestamp	timestamp	GPS	Date and time of the record
longitude	float	GPS	Longitude coordinate of the vehicle
latitude	float	GPS	Latitude coordinate of the vehicle
speed	float	GPS	Driving speed in km/h
direction	integer	GPS	Direction of the vehicle in degrees. North (0), east(90), south(180) and west(270)
satellites	integer	GPS	Number of visible satellites
rpm	integer	CANBus	The engines rounds per minute
kmcounter	float	CANBus	Mileage record
temperature	float	CANBus	Temperature of the engine in degrees Celsius
throttlepos	float	CANBus	Position of the throttle. No data available
fuellevel	float	CANBus	Fuel level in the tank where 100 is full tank.
totalconsumed	float	CANBus	Total amount of fuel consumed by the vehicle in liters
actualconsumed	float	CANBus	Instantaneous amount of fuel consumption in liters. Undependable
actual_km_l	float	CANBus	Instantaneous km/l. Undependable
acceleration	float	CANBus	Acceleration ???
make	integer	Haulier	The make of the vehicle. No data available
model	integer	Haulier	The model of the vehicle. No data available
capacity	float	Haulier	Number of possible passengers. No data available
weight	float	Haulier	The weight of the vehicle. No data available

Table 1: GPS and CANBus data types

A number of data sources can be used to gather data. GPS data can easily be collected from any vehicle equiped with a GPS system, through which for example a timestamp, longitude, latitude and speed are available. CANBus data provides information about the state of the vehicle. Retriving CANBus data is not as easy as GPS data, and little data is therefore avialable. CANBus data is for example the engines rounds per minute, fuel consumption and temperature. The hauliers may also provide additional information such as vehicle model, capacity and weight.

Table 1 lists the data columns in the provided data set. The data set contains 10,224,846 records collected from 6 minibusses. All vehicles are assumed to be compareable.

Of the four columns for fuel comsumption, **totalconsumed** is the most accurate measure. The instantaneous values, i.e. **actualconsumed** and **actual_km_l**, is said to be undependable by the data provider and should not be used. **fuellevel** is fuel level in the tank in percentage (100 % is full and 0 % is empty) and **totalconsumed** is the amount of fuel consumed. **totalconsumed** does only have a granularity of half a liter, which means it can only be used at an agreegated level of at least a trip in order to get usable results.

Non of the data provided by the haulier is available.

2. DEFINTIONS

2.1 Trips

The provided data set is annotated with a trip identifier, *tid*. A trip is defined as a collection of records with the same vehicle identifier, *vid*, and where each two consecutive records are within 100 seconds of each other. A trip is hence defined from when the vehicle is turn on and not by when the vehicle is moving when it is possible to map-match GPS coordinates to road segments. This defintion is chosen as idle time, i.e. when the engine is running but the vehicle is not moving, is persumed to be an important factor for fuel consumption.

The data set contains 13606 distinct trips containing between 1 and 78511 records. Many of these trips have less than 100 records and are therefore not useful. We remove these trips which results in 2868 distinct trips. The longest trip is 231 km and 278 trips drive less than 1 km.

2.2 km per liter

Km per liter is the total number of kilometers driven in the trip divided by the amount of fuel used. The fuel consumption is extracted from **totalconsumed** by substracting the lowest value from the highest.

All six vehicles in the data set have similar km/l accross their trips (See Figure 1). All trips a plotted along the x-axis with the km/l as the y-value. Most trips have been driven with between 5 and 10 km/l and 278 trips with 0 km/l. The latter is because they do not drive any where.

Km per liter can be used as a measure for how fuel efficient the vehicles are. We classify the trips into three classes based on their km/l. Class 'low' contains trips with between 0 and

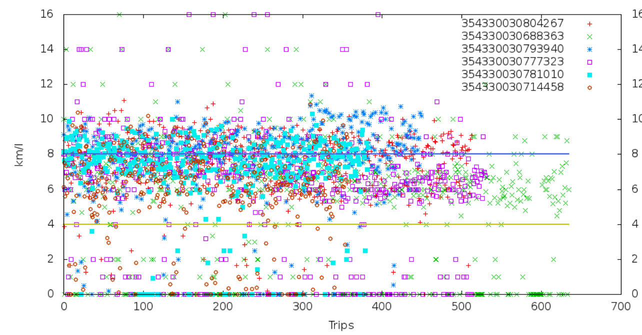

Figure 1: km/l for all trips

4 km/l, being all those that fall below the normal values and drive very fuel inefficient. Class 'medium' and 'high' splits the main cluser into two in order to distinguish the better trips. Class 'medium' contains trips with between 4 and 8 km/l and class 'high' contains the remaning trips with 8 or more km/l. Having only two classes, e.g. 0 to 4 and above 4, will allow us to identify very bad trips but not average versus good. One could also have a class containing trips with more than 10 km/l but only 94 trips will be in this class.

2.3 Idle

We say a vehicle is idle when its speed is zero while the engine is running, i.e. the engines rounds per minut (RPM) is above zero. In these situations the driver might as well have turned the engine off. The data set has 2859110 records with zero speed and non-zero RPM, with a minimal RPM of 41, a maximum RPM of 8192 and an average RPM of 910.4. Let $idle_t$ be the percentage of the trip t where the car is in an idle state.

Figure 2 plots $idle_t$ for each trip t versus the km/l of that trip. The size of the plot indicates the amount of fuel used in the trip and the color signals what class it has been asigned to. It is clear that most of the trips in class 'low' often idles and that the trips in class 'high' clusters around 0 to 4 %

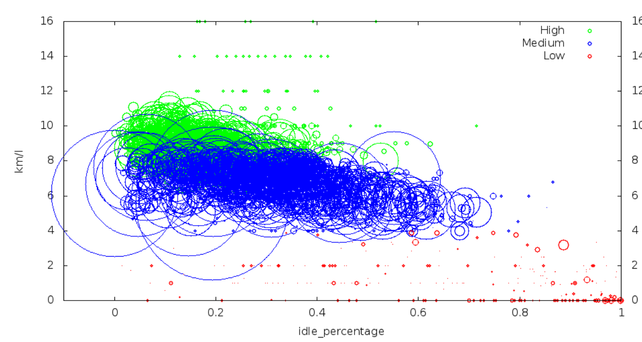

Figure 2: Idle time

2.4 Cruise Control

[] establishes that driving at a constant speed is more fuel echonomic. From observing the data in periods where we think the driver either drivers with cruise control or drives as if he did, we see that the speed varies with ± 1 km/h. Most drivers can drive with a constant speed for a short period. Experienced drivers might be able to drive at a constant speed over a longer time period without cruise control but the speed will generaly vary more. Hence, we say that a driver is using cruise control, or driving as if he did, if the speed does not changes more than ± 1 km/h in at least 40 seconds. Let $cruise_t$ be the percentage of the trip t where the car is in an cruise state.

Figure 3 plots $cruise_t$ for each trip t versus the km/l of that trip. The size of the plot indicates the amount of fuel used in the trip and the color signals what class it has been asigned to. We see that all trips in the 'low' class and many of the trips in the 'medium' class never cruises. Generally, the trips in the 'high' has tends to be cruising more often.

2.5 acckm

Acckm capture the sum of accelerration a vehicle perform on a trip. [] establishes that accelerrating consume extra fuel. A trip with a low accelerration should use less fuel. To normalise acckm we divide it by the lenght of the trip. Acckm is caculated in Algorithm 1. A buffer (dotted line) as shown in Figure 4 is implemented to prevent small variaction in speed (solid line) to effect the acckm

Algorithm 1 *acckm*

$temp = 0$
$counter = 0$
$buffer = 5$
while $i < n$ **do**
  if $v_i - v_{i-1} > buffer$ **then**
    $counter += (v_i - v_{i-1}) - buffer$
    $temp = v_i - buffer$
  else if $v_{i-1} - v_i > buffer$ **then**
    $temp = v_i + buffer$
  end if
  $i += 1$
end while
return $counter/trip_{length}$

Figure 4 show the number of trips on the y axis having a acckm show on the x axis. Trips with a high km/l are mostly clustered with a low acckm and vice versa.

2.6 Stop and go

Stop and go behaviour is not fuel efficient as more fuel is consumed when accelerating that when driving at a constant

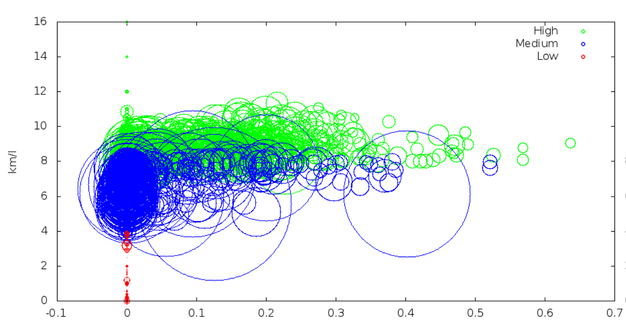

Figure 3: Cruise control

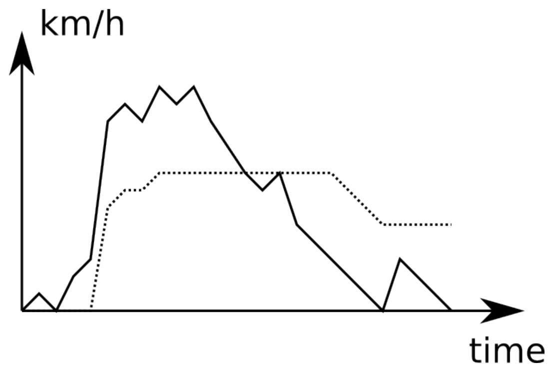

Figure 4: Acckm

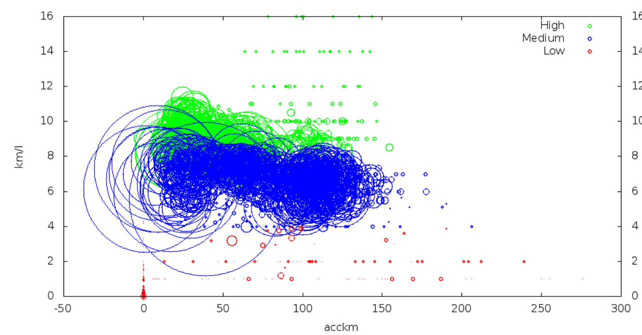

Figure 5: Acceleration kilometers

speed. It is therefore interesting to see if the number of times a vehicle stops and drives on is correlated with km/l. We say a vehicle is stop when it decelerations below 10 km/h and that it drives again when it accelerates up above 15 km/h.

Figure 6 shows the number of stop-and-goes of each trip normalised with the length of that trip versus the km/l of the trip. The size of the plot indicates the amount of fuel used in the trip and the color signals what class it has been asigned to. We see that the trips in the 'high' class have few stop and goes and about half of the trips in the 'medium' class have more stop and goes.

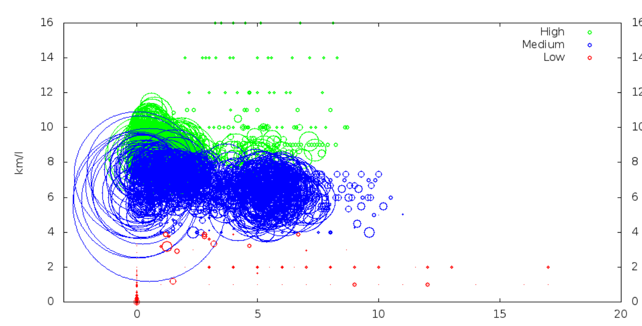

Figure 6: Number of stop and goes