Aggregate driver model to enable predictable behaviour

A Chowdhury[1], T Chakravarty[1]**, T. Banerjee[1] and P. Balamuralidhar[2]**
[1]TCS Innovation Labs, BIPL, Salt Lake, Sector-V, Kolkata, India
[2]TCS Innovation Labs, Abhilash, Plot-96, EPIP, Whitefield, Bangalore

E-mail: tapas.chakravarty@tcs.com

Abstract .   The categorization of driving styles, particularly in terms of aggressiveness and skill is an emerging area of interest under the broader theme of intelligent transportation. There are two possible discriminatory techniques that can be applied for such categorization; a micro-scale (event based) model and a macro-scale (aggregate) model. It is believed that an aggregate model will reveal many interesting aspects of human-machine interaction; for example, we may be able to understand the propensities of individuals to carry out a given task over longer periods of time. A useful driver model may include the adaptive capability of the human driver, aggregated as the individual propensity to control speed/ acceleration. Towards that objective, we carried out experiments by deploying smartphone based application to be used for data collection by a group of drivers. Data is primarily being collected from GPS measurements including position & speed on a second-by-second basis, for a number of trips over a two months period. Analysing the data set, aggregate models for individual drivers were created and their natural aggressiveness were deduced. In this paper, we present the initial results for 12 drivers. It is shown that the higher order moments of the acceleration profile is an important parameter and identifier of journey quality. It is also observed that the Kurtosis of the acceleration profiles stores major information about the driving styles. Such an observation leads to two different ranking systems based on acceleration data. Such driving behaviour models can be integrated with vehicle and road model and used to generate behavioural model for real traffic scenario.

1. Introduction

Modern smartphones have sensors like accelerometer, GPS, gyroscope, magnetometer etc. Numbers of applications are being built that uses smartphone as data collection sensors. Some of these sensors can simply be used to obtain the speed/ acceleration of a moving vehicle. Thus, smartphone based driving style recognition and classification is an important field of ongoing study [1]. Authors [1] have identified that the said work needs to be "extended to differentiate how individual drivers vary in their style of driving from day-to-day". Smartphone based sensing and modelling aggressive driving is also studied through machine learning technique [2]. Some important factors associated with driving pattern are identified in Näätanen & Summala [3]; these factors are important indicators of propensity of accident. Further, driver classification based on extracted features is outlined in [4]. This ve-dyna model allows driver classification and can be used to simulate "behaviour for common driving tasks". It is believed that a useful driver model may include the adaptive capability of the human driver, aggregated as the individual propensity to control speed/ acceleration. It can also be hypothesized that

---

[1]   To whom any correspondence should be addressed.

an aggregate model describing natural driving styles may reveal many more interesting aspects of human-machine interaction. Our investigations for similar purpose also use smartphone as a sensor. We take a statistical approach towards identification of aggression/ skill. The main objective of our study is the parametric identification of driving behaviour from the GPS data where different parameters identifying different driving styles are extracted. This is done in two ways: One is to identify significant parameters which influence signatures relevant to different journeys of a given driver and secondly to identify significant parameters that indicate different driving styles when a number of drivers are compared. The experimentation method and preliminary analysis based on statistical approach is presented in following sections.

2. Experiment and observation

A mobile application created by Tata Consultancy Services Ltd (TCS) was used to collect appropriate data from a number of registered TCS associates. The application residing in each user's smartphone is used to capture GPS data (location & speed) at 1 second interval, when the person is driving. The user assortment consists of both Android & iOS platforms. The data collection pilot was carried out for two months where the registered drivers went about their usual routine. For the preliminary analysis, we utilized 39 observations (trips) for each of the 12- randomly selected drivers. It is understood that the data represents their individual driving behaviour when faced with real life scenario and not a simulated set-up.

2.1 Data pre-processing

For purpose of analysis, we have selected trips of typical duration of approximately 80-90 minutes. The forward/longitudinal acceleration is calculated from the respective GPS speed measurements by numerical differentiation. While analyzing driver behavior, the tail portion of acceleration pdf (representing one completed trip) is considered important since these represent high acceleration/deceleration zones. An acceleration profile with heavy tail i.e. high value of kurtosis is considered as a signature of relatively high risk driving. Hence, kurtosis is an important parameter for these distributions. The kurtosis of forward acceleration is computed for each journey. After this 1st level – trip level of aggregation, we collate kurtosis values for all the journeys completed by the said driver. The analysis is focused on relationship of kurtosis with each driver. Towards that objective, we make an ANOVA (Analysis of Variance) based Design of Experiment to check whether the kurtosis values have dependence on individual driving styles or not. Essentially, we follow a completely randomized single-factor experiment where each driver constitutes *treatment* and the observed kurtosis value for each completed trip is considered a *replicate*.

2.3 Test for kurtosis

Table 1 shows the results of ANOVA test with null hypothesis H_0: Driver is not relevant for kurtosis value of acceleration". ANOVA shows that drivers impact kurtosis of forward acceleration.

Table 1: ANOVA results for kurtosis of 12 drivers

Cause of the variation	df (Degree of freedom)	SS (Sum of Squares)	MS (Mean Square)	F_0
driver	11	60949	5541	2.425
Residuals	456	1034850	2284	

Table 1 shows the result of ANOVA test for 12 drivers' kurtosis data with null hypothesis H_0. We know that the critical value of F-distribution at 1% probability level and degrees of freedom of 11 & 456 respectively, $F_{0.01,11,456} = 2.2867$. Since $F_0 > F_{0.01,11,456}$, we reject H_0 and conclude that the individual driving style significantly influences kurtosis value of the acceleration profile. The analysis is yet to be performed for a larger sized set of drivers. From the above, one can state that the kurtosis of forward acceleration profile is an important indicator of driving behaviour and can potentially serve

as aggression/ skill indicator. Thus, 'risk-propensity ranking' based on these values (collected for a group of drivers in similar spatiotemporal environment) will be an important statistical classification. Figure 1 gives the box plot for these 12 drivers' kurtosis data. Variation of kurtosis values across the drivers is clearly visible. Now to evaluate the performance of drivers, we compute different performance metrics. Based on those metrics we decide the rankings of these drivers. As described earlier, each driver has distinct nature of observed kurtosis values of forward acceleration. Every driver has his/her own propensity defined by his/her of kurtosis data.
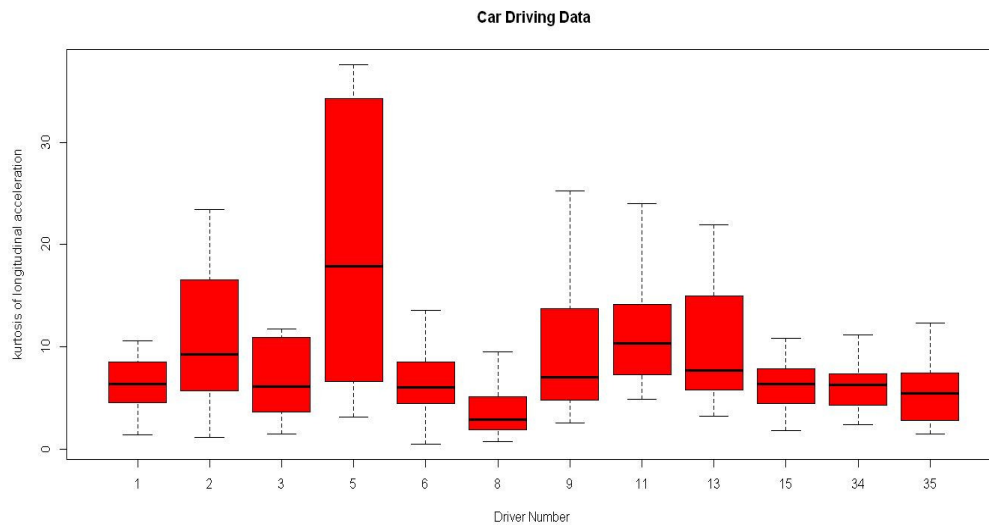

Figure 1: Box plot of kurtosis of longitudinal acceleration for different drivers (with outlier)

3. Results

We compute kurtosis for each completed trip performed by a driver. This is a trip level identifier of the driver behaviour, in terms of aggression and skill; this is a trip level aggregation. Further, for each driver, all the kurtosis values for all the trips completed in 2 month time-window are stored. For our nonparametric descriptive statistical approach, we choose median of all the kurtosis data corresponding to a driver to be representative of that driver's normal driving behaviour. Higher value of kurtosis means heavy tail. This indicates incidents of harsh acceleration or hard brake which, in turn, is usually considered as a signature of aggression. Hence, having higher median will mean more aggressive driving behaviour. Simultaneously, the kurtosis profile for each driver is unique in terms of the kurtosis data range seen for the set of completed journeys. Since each driver is facing a very similar contextual environment (on a daily basis) within the test cycle, one can logically infer that a widely varying acceleration profile signifies lack of driving skill. To quantify such, we implement IQR (Interquartile range) as a robust sample estimator of skill. Thus, low median of kurtosis samples correspond to caution and low IQR for the same corresponds to unskilled driving. Next, we rank the drivers on these two bases and attempt to classify each driver on aggression-skill plane. It is understood that the two said measures, for an individual, do not form orthogonal bases. The final summary of results containing rank is given in table 2.

Table 2: Ranking (Aggression Rank and Skill Rank) based on kurtosis data for 12 drivers

Driver ID	D001	D002	D003	D005	D006	D008	D009	D011	D013	D015	D034	D035
Aggression Rank (median)	6	10	4	12	3	1	8	11	9	7	5	2

Skill Rank (IQR)	4	11	7	12	5	2	9	8	10	3	1	6
Spearman's rank correlation 0.727												

Table 2 shows that for this set of 12 drivers', there exists a positive strong correlation between the two respective rankings. We further use the two ranks (normalized) for classifying the drivers in the aggression-skill plane. This is shown in figure 2. The 12 drivers are placed in one of the four quadrants.
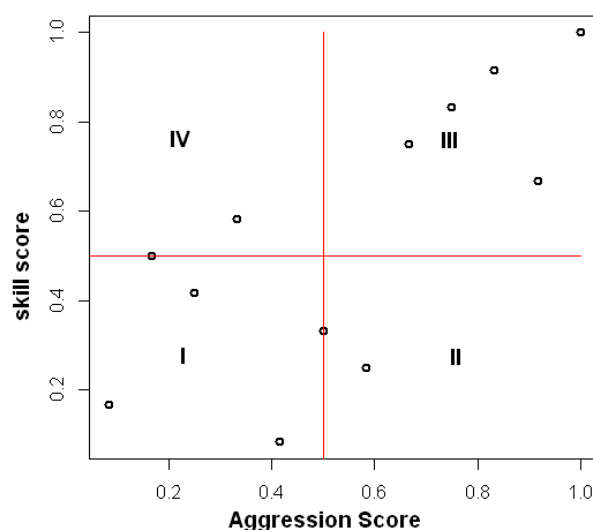

Figure 2: Driver classification: location of the 12 drivers in skill-aggression plane
(I: Cautious; II: skilled & aggressive; III: Unskilled & aggressive; IV. Anxious)

Due to the fact that the aforesaid classification is not dependent on specific parametric threshold value, this classification is considered relative and pertains to this group only.

5. Concluding remarks

In this paper, the methods of analysing the statistical parameters are used to obtain aggregated driver ranking, amongst peer group of drivers. Such rankings signify driver's natural aggression and skill at an aggregated level. Since kurtosis is estimated from entire acceleration profile, handfuls of anomalous events have less influence and the overall driving pattern plays a greater role. These rankings can be used for performance comparison; for cases of demographic & contextual similarity. Also in terms of parametric representation, such analysis will eventually lead to general driver behaviour model based on statistical estimation.

Reference

[1] Van Ly M., Martin S., Trivedi M. M. "Driver classification and driving style recognition using inertial sensors". In Intelligent Vehicles Symposium (IV), June 2013 IEEE, pp. 1040-1045. IEEE.
[2] Hong, J. H., Margines, B., & Dey, A. K. (2014, April). "A smartphone-based sensing platform to model aggressive driving behaviors". In Proceedings of the 32nd annual ACM conference on Human factors in computing systems (pp. 4047-4056). ACM.
[3] Näätänen R. and Summala H. "A model for the role of motivational factors in drivers' decision-making." Accident Analysis & Prevention 6.3 (1974): pp. 243-261.
[4] Irmscher M., and Ehmann M. "Driver classification using ve-dyna advanced driver". No. 2004-01-0451. SAE Technical Paper, 2004.