9th INTERNATIONAL CONGRESS ON TRANSPORTATION RESEARCH

ICTR 2019

October 24-25
Athens, Greece

Identification of Critical Driving Parameters Affecting Speeding Using Data from Smartphones

Armira Kontaxi

Transportation Engineer, Research Assistant

Together with:
Aristotelis Kokkinakis, Dimitrios Tselentis, George Yannis

National Technical University of Athens
Department of Transportation Planning and Engineering

Scope

- Identify the critical driving parameters that affect speeding using data from:
  - Smartphone devices
  - Naturalistic driving experiments

- Examine whether driving characteristics recorded by smartphone affect and can therefore predict the percentage of speeding time during driving.

Background (1/2)

- Driving behaviour characteristics
  - Speeding
  - Harsh braking/ acceleration/ cornering
  - Seatbelt use
  - Mobile phone use

- Travel behaviour characteristics
  - Total distance
  - Road network type
  - Risky hours driving
  - Trip frequency
  - Vehicle type
  - Weather conditions

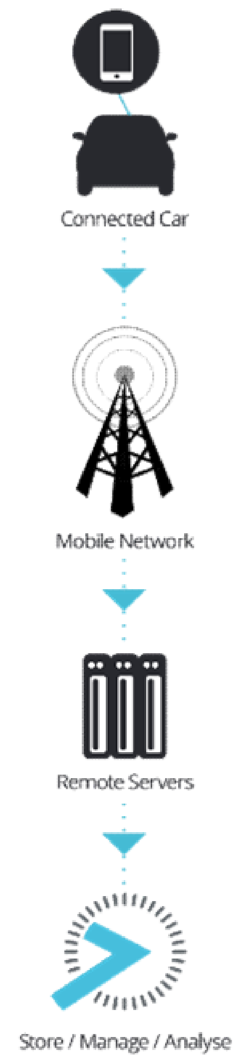
Background (2/2)

- Data collection schemes
  - Smartphones
  - In-vehicle devices
  - On-board diagnostic devices (e.g. OBD-II)

- Data sources
  - Naturalistic driving experiments
  - Driving simulator experiments
  - In-depth accident investigation

Smartphone data collection (1/2)

- A mobile application to record user's driving behaviour (automatic start / stop)
- A variety of APIs is used to read mobile phone sensor data
- Data is transmitted from the mobile App to the central database
- Data are stored in a sophisticated database where they are managed and processed

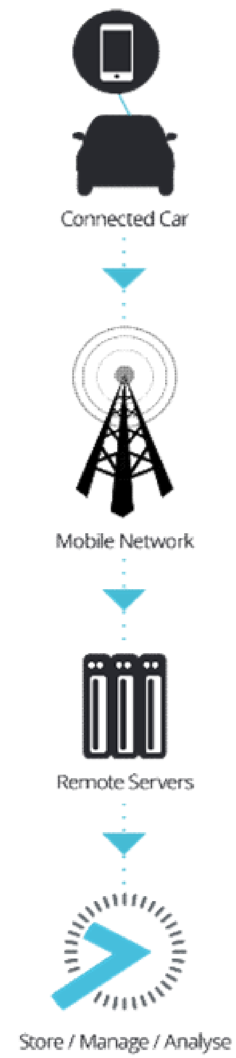


Smartphone data collection (2/2)

- Indicators are designed using:
  - machine learning algorithms
  - big data mining techniques

- Data collected for 21,610 trips from 68 drivers between August 2016 to October 2017 (15 months)

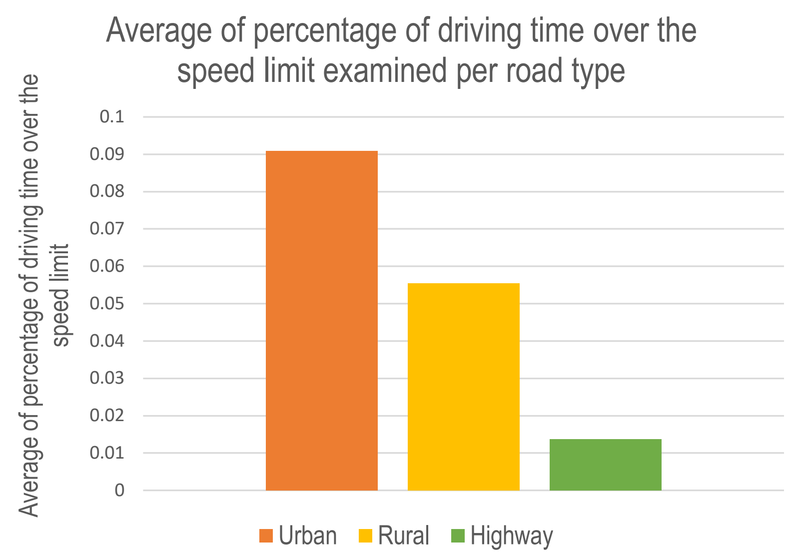
- The database analyzed was in .csv format
  - Drivers' trips are stored per row, the characteristics of which are stored in each column's variables

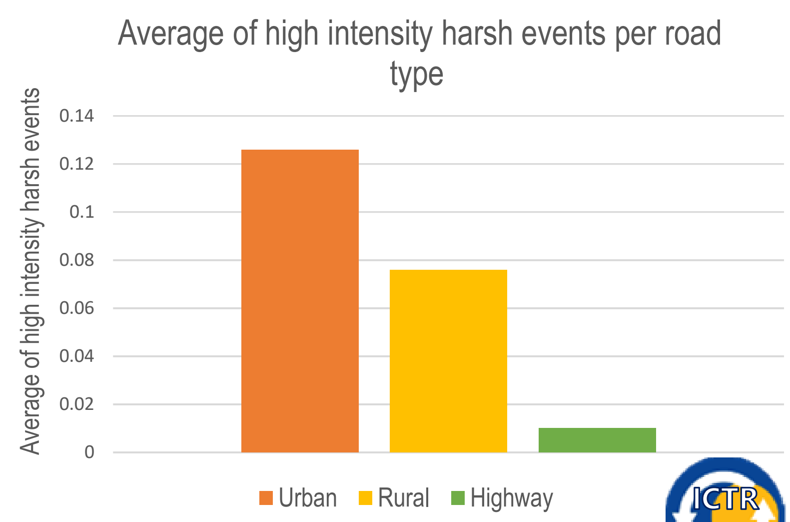


Descriptive statistics

- The highest and lowest percentage of speeding is found in the urban environment and highways, respectively.
  - This is probably because the speed limits on highway are already high enough

- The largest number of high intensity harsh events takes place in the urban environment mainly due to its characteristics e.g.:
  - high traffic conditions
  - lower number of road lanes
  - overall parameters that lead to a more nervous driving performance

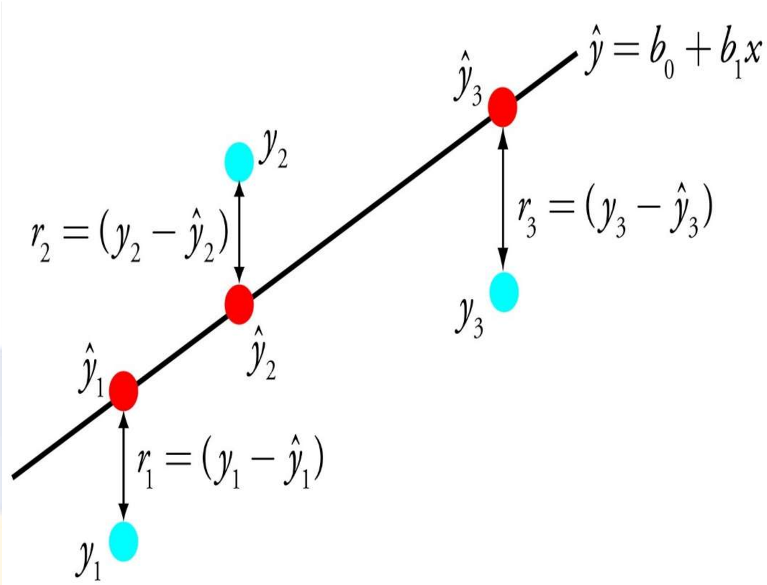




ICTR 2019

Methodology (1/2)

- When a variable Y is linearly depended on more than one variables X (X_1, X_2, X_3,.., X_K), multiple linear regression is used. The relationship between the dependent and the independent variables is given by the following formula:
  - $y_i = \beta_o + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \beta_3 * x_{3i} + \beta_k * x_{ki} + \varepsilon_i$



Methodology (2/2)

- Four linear regression models forecasting the percentage of driving duration of driving above the speed limit were developed: one overall model and three models for each different road type (urban, rural, highway):
  - Model 1: Predicting the percentage of speeding – overall model
  - Model 2: Predicting the percentage of speeding on urban road
  - Model 3: Predicting the percentage of speeding on rural road
  - Model 4: Predicting the percentage of speeding on highway

Results (1/3)

Independent Variables	Model 1 (overall model)				Model 2 (urban road)				Model 3 (rural road)				Model 4 (highway)			
	b_i	t	e_i	e_i*	b_i	t	e_i	e_i*	b_i	t	e_i	e_i*	b_i	t	e_i	e_i*
Constant	0,021	16,182			-10.466	15,990			0.014	10.234			0.002	3.382		
totaldist	0,002	63,679	0,626	9,468	0.014	60.955	1,629	14,808	0.006	63.515	1,385	9,819	0.002	57.137	1,516	7,664
ha_intensity_high	0,031	20,179	0,809	12,235	0.014	5.451	0,297	2,700	0.026	7.441	0,630	4,465	0.031	5.712	0,755	3,819
hb_intensity_high	0,054	19,794	1,006	15,219	0,070	14.091	0,963	8,755	0.032	6.145	0,539	3,818	0.096	12.028	0,831	4,203
hc	-0,002	-6,677	-0,167	-2,519	-0.009	-10.432	-0,406	-3,688	-0.017	-22.407	-1,016	-7,201	-0.033	-5.283	-0,812	-4,103
avdecel	-0,016	-8,674	0,676	10,218	-0.041	-22.659	1,426	12,963	-0.017	-15.455	0,500	3,541	-0.057	-28.715	0,664	3,358
mobileusage	0,014	3,709	0,066	1,000	0.026	6.159	0,110	1,000	0,036	6.666	0,141	1,000	0.104	13.894	0,198	1,000
R^2	0,239				0,193				0,200				0,296			

- Higher effect of the number of high intensity harsh braking events compared to the rest of the variables.
- The percentage of mobile phone usage has the lowest influence for all models.

Results (2/3)

- Based on the elasticity values of the overall model:
  - the influence of the hb_intensity_high variable is the highest among all independent variables.
  - variable "mobileusage" shows the least influence on the model. Specifically, with respect to the most influential variable, it affects the model 15,2 times less.
  - variables "ha_intensity_high" and "avdecel" have the second highest impact on the model, after the most influential variable (0,809 and 0.676 for "ha_intensity_high" and "avdecel", respectively).
  - the impact of variables “totaldist” and “hc” are 9,4 and 2,5 times higher than the variable “mobileusage”.

Results (3/3)

- Regarding the models developed per road type, total distance is found to be the most significant predictor of speeding.

- The number of high intensity harsh acceleration and braking events are the most significant predictors in the overall model.

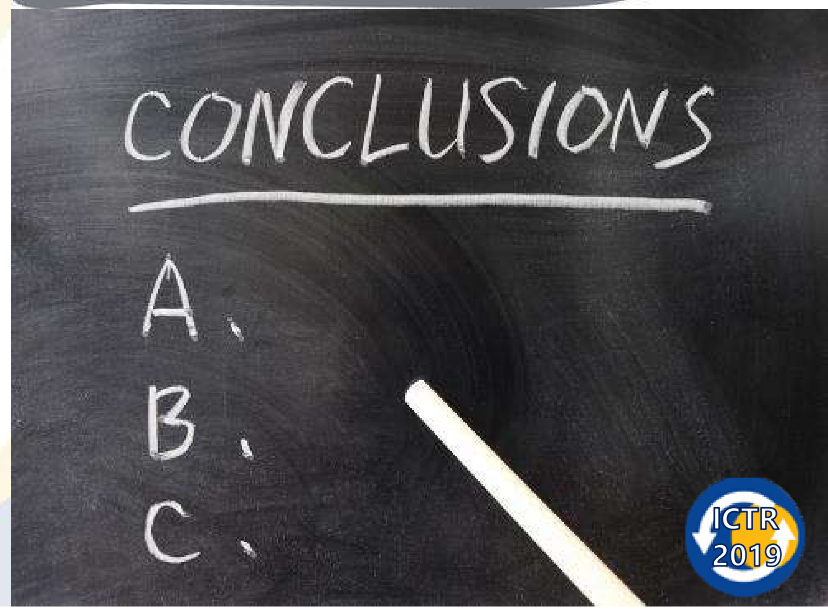
- The relatively low R^2 value indicates that the examined independent variables can partially predict the dependent
  - they can be further improved by examining additional independent variables

Conclusions (1/2)

- Predicting the percentage of speeding time is more accurate on motorways than on other road types. This is probably because traffic and speed are more normalized.

- Distance traveled in each type of road is crucial for predicting speeding duration, as this is the most influential variable. The longer the distance traveled, the faster the speed limits are exceeded.

- The number of high-intensity harsh decelerations affects the predictability of speeding for all road types. This is probably because these characteristics are directly associated with aggressive driving.



Conclusions (2/2)

- As the number of harsh cornering events increases, speeding time decreases. This is rational, as drivers instinctively reduce their speed to make a safe turn.

- High-intensity harsh acceleration and deceleration are associated with aggressive driving and speed limits exceedance. The higher the number of harsh events, the higher the speeding time.

- As the mobile phone usage time increases, speeding time increases. This is probably due to the driving distraction.

Further research

- Apply methodology on a larger drivers sample.
- Collect and analyze more information e.g. gender, age etc. and under different traffic and weather conditions.
- Examine the correlation between speeding and several factors e.g. presence of passengers, seatbelt usage, alcohol usage etc.
- Investigate speeding behaviour of motorcyclists.

9th INTERNATIONAL CONGRESS ON TRANSPORTATION RESEARCH

ICTR 2019

October 24-25
Athens, Greece

Identification of Critical Driving Parameters Affecting Speeding Using Data from Smartphones

Armira Kontaxi

Transportation Engineer, Research Assistant

Together with:
Aristotelis Kokkinakis, Dimitrios Tselentis, George Yannis

National Technical University of Athens
Department of Transportation Planning and Engineering