MOBILE SENSING AND MACHINE LEARNING FOR IDENTIFYING DRIVING SAFETY PROFILES

Eleni G. Mantouka, Ph.D. Candidate*
National Technical University of Athens
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1515; Fax: +30 210 772 1454; e-mail: elmant@central.ntua.gr

Emmanouil N. Barmpounakis, Ph.D.
National Technical University of Athens
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1515; Fax: +30 210 772 1454; e-mail: manosbar@central.ntua.gr

Eleni I. Vlahogianni, Ph.D.
Assistant Professor
National Technical University of Athens
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1369; Fax: +30 210 772 1454; e-mail: elenivl@central.ntua.gr

ACKNOWLEDGEMENTS
This research has exploited data provided by OSeven Telematics, London, UK.

**corresponding author*

Word count: 1450 words text (1 figure + 2 tables)

Submission date: 8/11/2017

INTRODUCTION

The world's number of cars on the road is expected to double by 2040, with the absolute value reaching 2 billion cars (*1*). This will cause an excessive increase in the interactions between drivers, which have a significant impact on road safety, fuel economy and congestion. The management of these challenges relies on the ability to forecast future demand, optimize urban mobility and improve driver efficiency (*2*, *3*). Overcoming such problems has provoked the increased interest of researchers to collect driving data, aiming to identify risky driving and improve driver efficiency.

Different drivers vary in relation to the manner they alter their longitudinal (accelerate, decelerate) and lateral position (steering), how much distance they keep to follow a vehicle safely and comfortably and how much time they drive above the speed limit (speeding) (*4*).

The rapid advances of new Information and Communication Technologies (ICT), such as GPS devices, mobile phones, Bluetooth, etc., offer the capability of monitoring driving behavior and collect driving and travel data in a cost-effective manner. The always increasing processing capacity of the smartphone, coupled with its wireless communications features, allow continuous, inexpensive and fast data collection (*5*). Smartphones are programmable and come with a growing set of powerful embedded sensors, such as accelerometer, digital compass, gyroscope, GPS, microphone, and camera, which are enabling the emergence of sensing applications, even without the engagement of the users.

The identification of unsafe driving styles coupled with the most relevant incentives to avoid them, is key for the achievement of safe and efficient driving. Previous research (*6–12*) has led to the development of systems that investigate driving behavior, but none of them results in a universal system which can detect not only aggressive driving profiles, but also distraction from driving as well as risk taking in the sense of driving over the speed limits.

Taking into account all the above, the research questions that arise from the literature review are the following:

- Is it possible to produce the most appropriate characterization for every trip, concerning all driving parameters, such as aggressiveness, speeding and driver's distraction from driving task, under differing driving conditions?
- Is it possible to provide relevant characterizations concerning each driver's profile?

In this paper, we aim to identify aggressive and dangerous driving profiles using data collected from smartphone sensors. More specifically, a two stage clustering approach is implemented in order to, first, distinguish aggressive from non–aggressive driving and, then, reveal additional unsafe behaviors, namely distraction (mobile phone usage) and risk taking (perform speed limit violations) in both the aggressive and non-aggressive states. Finally, a discussion on the average behavior of each driver is provided, underlying whether drivers improve or deteriorate their driving style.

METHODOLOGY

Unsafe driving profiles include driving in an aggressive manner, being distracted from the driving task (use of mobile phone) as well as developing risk taking behaviors (speeding) while driving. Although there is a high correlation between aggressiveness and other unsafe driving behavior, it is possible to drive aggressively without driving unsafely in another manner. To this end, in this paper a two-level K – Means clustering is implemented to provide appropriate characterizations for each trip. The optimal number of clusters emerged from the Calinski-Harabasz criterion and clustering is implemented on Euclidean distance matrix. The first level of clustering aims to separate aggressive from non–aggressive trips. The second level of clustering addresses the variability of behavior in both the aggressive and non-aggressive states based on whether the driver is i) driving over the speed limit, and ii) uses his/her mobile phone. Based on the aforementioned methodology, each trip may be classified in either one of the following 6 categories ranked by importance to driving safety:

- Safe behavior
- Aggressive behavior
- Risky behavior
- Distracted Behavior
- Aggressive/Risky behavior
- Aggressive/Distracted behavior

Data Collection

Data is collected from an already developed smartphone application for both iOS and Android devices. After removing data noise, the final database includes 10212 trips made from 129 unique drivers from June 2016 to April 2017 in urban and rural road networks. For each trip, numerous variables are extracted to be included as input features in the clustering (TABLE 1). Statistical measurements of acceleration and deceleration during a trip are included that describe how smoothly the driver changes his/her longitudinal position. In addition, speeding measurements are collected that describe smoothly and with speed excess driving, as well as mobile usage indicators are estimated that describe how cautious the driver is.

TABLE 1 Description of Variables Used in Clustering

Variable	Description
Harsh acceleration per km	The number of harsh accelerations per km traveled.
Harsh brakes per km	The number of harsh brakes per km traveled.
Smoothness indicator	The sum of differences of squares of final and initial speed, divided by trip distance.
Standard deviation of acceleration	The standard deviation of acceleration performed in each trip.
Percent of mobile usage	The percentage of trip duration when the driver uses his/her mobile phone.
Percent of speeding	The percentage of trip duration when the driver drives over the speed limit.

FINDINGS

1st Level Clustering

In order to detect aggressive driving and provide the relevant characterization for each trip, K-means clustering is performed. The number of clusters is set to $k = 2$. The variables that are used describe the number of harsh alterations of the longitudinal position of the vehicle (acceleration and braking), while the rest of them are essentially indices of the average acceleration of the trip (smoothness indicator and standard deviation of acceleration). Results indicate that almost 74% of the trips are not featured by aggressive driving characteristics. However, the remaining 26% of the sample reveals an aggressive behavior while driving.

Moreover, findings revealed that average acceleration is roughly equal for both categories of trips, although in the case of aggressive trips the number of harsh events that are detected are much more.

2nd Level Clustering

In order to detect additional unsafe behaviors while driving, a further categorization of the trips was made. A second level clustering was performed for both aggressive and non–aggressive trips. In this case the number of clusters is set to $k = 3$. Results have shown that drivers who drive over the speed limit do not intend to use their mobile phones and the corresponding trips were characterized as "risky trips". Moreover, 69% of the trips featured by aggressiveness do not have any additional unsafe driving characteristics, concerning distraction and speeding. Such trips are characterized as "aggressive trips". Finally, those trips where the driver was distracted from the driving task are characterized as "distracted trips". TABLE 2 presents the cluster centers for both first and second level clustering and the corresponding number of trips of each cluster.

TABLE 2 Cluster centers of Two-level clustering

1st level clustering					
Variable/Cluster	**Harsh acceleration/km**	**Harsh brake/km**	**Smoothness indicator**	**Std deviation of acceleration**	**Number of trips**
Non – aggressive	0.13	0.07	0.32	1.15	7534
Aggressive	0.72	0.19	0.45	1.62	2678
2nd level clustering					
Variable/Cluster	**Percent of mobile usage**		**Percent of speeding**		**Number of trips**
Aggressive trips					
Aggressive trips	0.03		0.07		1837
Distracted trips	0.44		0.11		226
Risky trips	0.02		0.41		615
Non – aggressive trips					
Safe trips	0.02		0.07		5183
Distracted trips	0.52		0.06		566
Risky trips	0.02		0.29		1785

Drivers' Profiles

Although the majority of trips are characterized as safe trips, a significant number of trips are distinguished by aggressive driving characteristics as well as exceed speeding (FIGURE 1). As expected, it is observed that each driver appears to behave differently in every trip, developing each one of the unsafe driving behaviors, which are identified. Thus, with a different frequency and in different circumstances, each driver may drive aggressively, may use the mobile phone during driving and/or take on risky behavior by speeding. Therefore, it is not possible to assign a specific driving profile to each driver. Nevertheless, the level of behavioral variability between trips can be measured and whether or not the driver improves his / her behavior between successive trips.

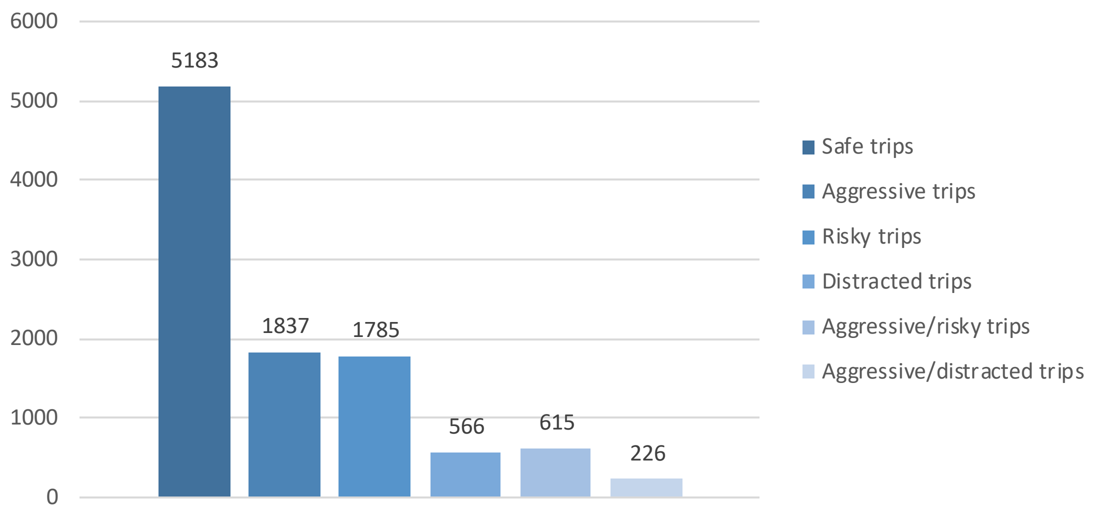

FIGURE 1 Distribution of the trips between the six identified driving states

CONCLUSION

In this paper, we aimed to identify unsafe driving behaviors and provide relevant trip characterizations. For this purpose, two level K- means clustering was implemented on data collected through smartphone sensors. Results indicate that drivers behave differently every time, performing trips that fall within each one of the recognized categories of safe and unsafe driving style. Moreover, the large values of the estimated volatility measure reveal that drivers do not have a stable driving profile, but instead they change the way they drive on every trip. Further research could highlight the circumstances in which each driver performs a certain behavior, such as time of the day, weather as well as road and network conditions.

The results described above can provide a complete view on different driving styles, as well as an efficient way to identify unsafe driving behavior. Furthermore, these results could be useful for organizations, such as insurance companies, who are interested in identify driving style of their customers in order to offer them personalized and efficient services. Finally, findings could be exploited in real-time recommendation systems, which aim to improve driving behavior in relation to safe, efficient and sustainable driving.

REFERENCES

1. Smith, M. N. The Number of Cars Will Double Worldwide by 2040. *Business Insider UK*. http://uk.businessinsider.com/global-transport-use-will-double-by-2040-as-china-and-india-gdp-balloon?utm_source=feedly&utm_medium=webfeeds.

2. Tulusan, J., T. Staake, and E. Fleisch. Providing Eco-Driving Feedback to Corporate Car Drivers: What Impact Does a Smartphone Application Have on Their Fuel Efficiency? *Proc. UbiComp '12*, 2012, pp. 212–215. https://doi.org/10.1145/2370216.2370250.

3. Castignani, G., T. Derrmann, R. Frank, and T. Engel. Driver Behavior Profiling Using Smartphones: A Low-Cost Platform for Driver Monitoring. *IEEE Intelligent Transportation Systems Magazine*, Vol. 7, No. 1, 2015, pp. 91–102. https://doi.org/10.1109/MITS.2014.2328673.

4. Miyajima, B. C., Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, K. Takeda, and F. Itakura. Driver Modeling Based on Driving Behavior and Its Evaluation in Driver Identification. No. 2, 2007, pp. 427–237.

5. Araújo, R., Â. Igreja, R. De Castro, and R. E. Araújo. Driving Coach: A Smartphone Application to Evaluate Driving Efficient Patterns. *IEEE Intelligent Vehicles Symposium, Proceedings*, Vol. 1, No. 1, 2012, pp. 1005–1010. https://doi.org/10.1109/IVS.2012.6232304.

6. Ma, C., X. Dai, J. Zhu, N. Liu, H. Sun, and M. Liu. DrivingSense : Dangerous Driving Behavior Identification Based on Smartphone. Vol. 2017, 2017, pp. 1–11.

7. Ghose, A., A. Chowdhury, V. Chandel, T. Banerjee, and T. Chakravarty. An Enhanced Automated System for Evaluating Harsh Driving Using Smartphone Sensors. *Proceedings of the 17th International Conference on Distributed Computing and Networking - ICDCN '16*, 2016, pp. 1–6. https://doi.org/10.1145/2833312.2849555.

8. Liu, T., Y. Yang, G. Bin Huang, Y. K. Yeo, and Z. Lin. Driver Distraction Detection Using Semi-Supervised Machine Learning. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 4, 2016, pp. 1108–1120. https://doi.org/10.1109/TITS.2015.2496157.

9. Quek, Z. F., and E. Ng. Driver Identification by Driving Style. 2013, pp. 1–4.

10. Meseguer, J. E., C. T. Calafate, J. C. Cano, and P. Manzoni. DrivingStyles: A Smartphone Application to Assess Driver Behavior. *Proceedings - International Symposium on Computers and Communications*, 2013, pp. 535–540. https://doi.org/10.1109/ISCC.2013.6755001.

11. Aljaafreh, A., N. Alshabatat, and M. S. Najim Al-Din. Driving Style Recognition Using Fuzzy Logic. *2012 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2012*, 2012, pp. 460–463. https://doi.org/10.1109/ICVES.2012.6294318.

12. Paefgen, J., F. Kehr, Y. Zhai, and F. Michahelles. Driving Behavior Analysis with Smartphones: Insights from a Controlled Field Study. *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, No. January, 2012, p. 36:1–36:8. https://doi.org/10.1145/2406367.2406412.