

Context and Location Awareness in Eco-Driving Recommendations

André Campolina^{1,2}, Azzedine Boukerche¹, and Antonio A. F. Loureiro²

¹PARADISE Research Laboratory, EECS - University of Ottawa, Canada

²Department of Computer Science, Federal University of Minas Gerais, Brazil
{andre.campolina, loureiro}@dcc.ufmg.br, boukerch@site.uottawa.ca

Abstract—Eco-driving techniques are methods that drivers can take in order to improve their vehicles' fuel efficiency. One way of implementing such methods is to recommend changes in driving habits focusing on saving fuel. Recommendation systems in the context of efficient driving may take numerous data sources as input and serve multiple purposes. In this work, we develop a recommendation system that suggests changes in engine revolutions per minute (RPM) that reduce fuel consumption. We simulate the effects of this system using a vehicular sensor dataset that contains location, speed, RPM, and consumption. We also investigate the effects of including location data in the recommendations as a way to leverage local habits and behaviors. Both methods reduced fuel consumption in all recorded trips. Moreover, using only local data yielded a mean fuel reduction of 43%, whereas using the entire dataset reduced the fuel consumption in 56% on average. Upon analyzing the resulting changes, we noted that such a difference in fuel consumption is due to the local recommendations not having access to global optimal data points and accounting for local behavior that is affected by aspects such as terrain.

I. INTRODUCTION

Millions of vehicles enable transportation networks across the world, moving people and cargo, and produce vast amounts of data doing so. These data sources comprise what is called vehicular data, i.e., all data that can be used to describe the state of a vehicle, its occupants, transported goods, and even its surroundings at any given moment. Multiple devices, like smartphones, can collect vehicular data besides the vehicle itself. Vehicular datasets can contain data from multiple vehicles, which enables more precise and detailed representations and models.

Vehicular data benefits multiple areas of interest, one of which is the improvement of fuel efficiency, which means consuming less fuel and emitting less greenhouse effect gases. By conducting analyses on vehicular data, it is possible to identify a set of events and behaviors that decrease such efficiency and act on their causes to avoid them. Eco-driving [1] comprises the set of recommendations and suggestions that aim at reducing the impact of traffic in the environment, such as fuel efficiency, which is the target of this work.

Eco-driving techniques and systems may rely on numerous data sources to reduce fuel consumption. For instance, terrain data [2] coupled with sensor readings can improve engine usage recommendations, while weather and traffic information can improve routes by avoiding rush hours or traffic jams

caused by storms, for instance. Furthermore, these two solutions can be combined to ensure that drivers are not only following optimal routes but also driving efficiently while doing so. In this work, we use a vehicular sensor dataset to provide recommendations about engine revolutions per minute (RPM) to improve the fuel efficiency of trips.

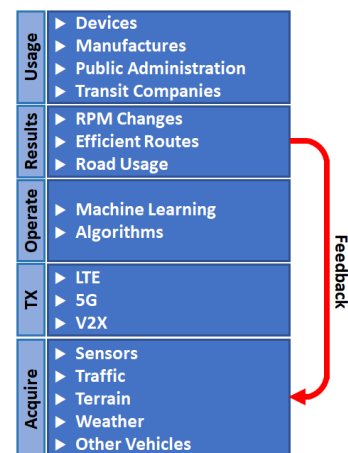


Fig. 1. Recommendation Process Framework

A recommendation process is comprised of different steps, as illustrated in Figure 1, which shows a generic eco-driving suggestion system that might contain other aspects not included there, depending on the particularities of the desired usage. It represents the demand for data and information in a stacked form, that is, the top layer contains some of the entities that may request recommendations or transportation data in general, while the bottommost contains the sources of raw data that will feed the systems. In this work, devices such as smartphones and vehicles' internal processors use the RPM and gear value changes to reduce fuel consumption. The gear value changes are a result of an algorithm that uses machine learning techniques and needs communication methods such as LTE to gather data it requires, which was acquired from sensors in multiple cars. The figure also points that data generated by algorithms may eventually be fed back into the system, as more drivers use the recommendations.

Our approach uses data from multiple trips recorded by a single vehicle model to determine the best eco-driving behavior, given a driving context. This is mainly because

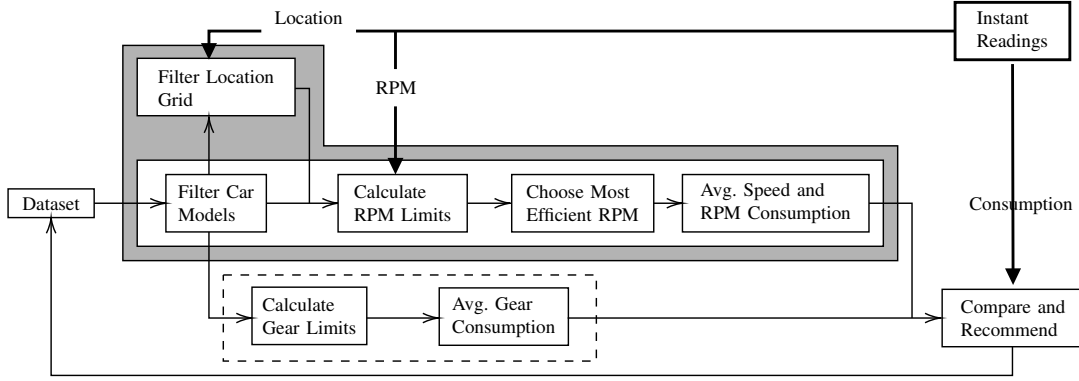


Fig. 2. Framework for different eco-driving recommendation strategies

the data source we use in this work comes from individuals who collected and shared data from their vehicles using a smartphone app (see Section IV). As already pointed out, other datasets could be used as part of this recommendation system. We understand that the driving context encompasses all descriptors of a given moment from a vehicle standpoint. For example, current speed, RPM, neighboring vehicles, terrain, and traffic signs describe the driving context, and, in this work, we focus on the first two descriptors. We base our recommendation on finding the RPM value within similar context that minimizes fuel consumption. In doing that, we search and recommend changes in the driving behavior that is relevant to drivers' contexts.

Figure 2 illustrates a framework with different fuel-saving strategies that we compare in this work. We propose two novel approaches, delimited by the gray and white areas, and we compare them to a strategy proposed by Rettore et al. [3], highlighted by the dashed rectangle. Our threshold-based approach, delimited by the white area, consists of selecting only sensor readings that have RPM values within a specific range from the instant reading. Among these possible values, we identify the ones that historically reduce fuel consumption, given the instant speed, to be the recommended value. We also propose a location-based recommendation, which adds a filter to the threshold-based process. After selecting all readings from the same model, we also select only the observations made within the same spatial grid cell as the current readings. With this step, we leverage location to provide contextually-appropriate recommendations to drivers and avoid recommending driving behaviors unsuited for regional conditions.

The remainder of this work is organized as follows: Section II presents the related work that targets at reducing fuel consumption and eco-driving practices; Section III describes the dataset we used in our investigations and how it was collected; Section IV describes a baseline recommendation approach as well as our proposed strategies, one that uses all observations of a vehicle model and other that uses observations of that model grouped by region; Section V presents our simulation process and the results yielded by the recommendation strategies on it, and finally, Section VI

presents our conclusions and discusses some possible future extensions of this work.

II. RELATED WORK

Collecting sensor data from multiple vehicles in urban areas is a fundamental task in order to gain knowledge about drivers' behavior and its relationship with other aspects of driving, such as fuel consumption. Traditionally, fuel consumption is measured and regulated using driving cycles, which are routines that recreate common traffic situations [4]. With the introduction of new communication and sensing capabilities to more vehicles, it is also possible to acquire sensor data from multiple cars in near real-time, which paves the way to new consumption models and studies based on real-world fuel data.

Some studies found in the literature focus on reducing fuel consumption and CO₂ emissions but do not deal with these variables directly. Instead, researchers pay attention to sensors that monitor the interaction between drivers and vehicles. Zhao et al. [5] used speed sensor data to avoid drivers stopping at red traffic lights. The authors calculated time windows in which vehicles would not be stopped and suggested adjustments in cruise speed to match such windows. Schröder and Cabral [2] used different maps to infer elevations in roads to estimate fuel consumption in a city road network and trace routes that, instead of optimizing traveled distance, optimize fuel consumption.

There are different ways of approaching the task of tracing routes optimized for minimal fuel consumption, depending on the type of data and driver's interaction. Magaña and Muñoz-Organero [6] developed a driving assistant that uses artificial intelligence techniques and sensor readings to issue recommendations that reduce consumption. Saremi et al. [7] reported the usage of GreenGPS, an application that takes sensor readings to infer fuel consumption in streets, and uses this information to trace routes with the smallest fuel consumption. Xiong, Cao, and Yu [8] focused on improving the efficiency of hybrid vehicles by using driving cycles to determine the best battery power usage profiles during a trip.

Eco-driving techniques include a set of behavioral suggestions that vary from maintenance routines to driving habits [1]. Ayyildiz et al. [9] presented a case study of training freight

drivers in eco-driving behaviors that resulted in lower fuel consumption. Their study identified behaviors such as early gear changes and soft acceleration patterns as causes of efficiency loss. Gear changes are a point investigated by Rettore et al. [3]. The authors identified a relationship between instant fuel consumption and current transmission gear, and then investigated the potential impact of a gear recommendation service that suggests changes based on optimal fuel efficiency.

This work is mostly related to the study conducted by Rettore et al. [3]. Our work uses sensor data to recommend changes in engine demand, which, in turn, may result in gear changes that lead to better efficiency. On the other hand, that work recommends gear changes based on pre-established torque requirements, not taking into account current driving states, nor other observations made in the same region.

III. DATASET DESCRIPTION

The data used in this work was collected from the enviroCar project [10], a project that allows users to collect and share data from their vehicles using a smartphone app. Currently, there are over 800 registered drivers using the app, who logged over 17,000 trips, comprised of more than 5.5 million individual observations. Each observation contains location and time data, as well as information about the used vehicle and readings from up to 24 vehicular sensors. However, it is important to mention that not all vehicles will collect data from all sensors, as different models can have different interface projects. In fact, some sensors do not have recorded readings at all.

EnviroCar users collect vehicular data around the world, but the majority of its recorded trips concentrate in Germany, where 94% of the trips pass through German territory. Trips are mainly concentrated in the city of Mönchengladbach (46% of the total), which had a partnership between its administration and the application that encouraged drivers to record their trips in specific days. Because of this concentration, we chose to investigate trips of a particular model of vehicle that has the most observations recorded in the city area. In total, there are 234 trips recorded with this model, accounting for 327 thousand observations

IV. ECO-DRIVING RECOMMENDATIONS

Eco-driving recommendations target improving environmental efficiency by suggesting new behaviors and habits to drivers. For example, these changes can be related to maintenance tasks like aligning the wheels or driving habits like accelerating and braking softly. In this work, we focus on the interaction between drivers and vehicles and on one of the ways it can be measured: gear changes.

In vehicles, gears transmit the power generated on engine revolutions to the wheels, and, in doing that, they pose different relations between the engine's fuel consumption and its power delivery. By pressing harder on the acceleration pedal, drivers demand more power from the engine, causing it to increase its revolutions per minute (RPM), and thus, consume more fuel. In manual transmission vehicles, drivers can choose

to change gears at any point in time. Higher gears bear higher speeds, at the cost of delivering less torque to the wheels. On the other hand, lower gears deliver more torque but can not tolerate high speeds. Figure 3 illustrates the relation between speed and RPM, which presents groupings that coincide with the different gear ratios. This figure uses real vehicular data, grouped using the k-means clustering method.

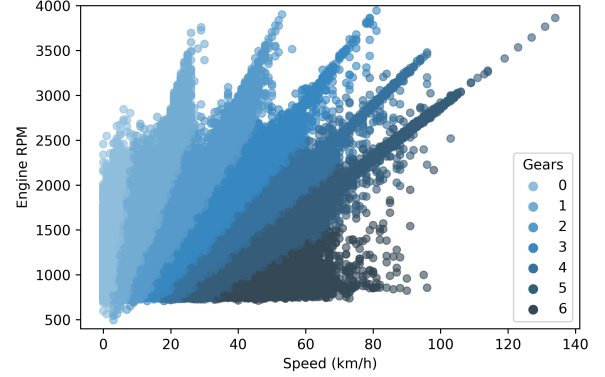


Fig. 3. The relationship between speed and RPM commanded by gears

A. Baseline Approach

The first step in our work is to establish a baseline recommendation system that also targets gear changes in order to improve fuel efficiency. Rettore et al. [3] presented a method that analyzes vehicular data to determine ranges of speeds in which each of the gears is the best option to reduce consumption while keeping engine torque at a minimum determined level. The authors initially cluster RPM and speed values in a similarly to Figure 3. In order to simulate gear changes according to their recommendation schema, they defined speed ranges for each gear, such that the new gear g' assigned to a speed value s is defined by $g'(s)$ as follows:

$$g'(s) = \begin{cases} i & \text{if } \lim(i) \leq s < \lim(i+1), i \in [1, \dots, |G| - 1] \\ 0 & \text{if } s < \lim(1) \\ |G| & \text{if } s \geq \lim(|G|) \end{cases} \quad (1)$$

where G refers to the set of gear values determined by the transmission system and $\lim(g)$ is defined as

$$\lim(g) = \underset{s}{\operatorname{argmin}}(\operatorname{mean}(t|s, g) \geq t_{\min}) \quad (2)$$

where t represents torque values measured by the engine load sensor. Equation 2 serves the purpose of determining the minimum speed at which the mean observed torque value is greater than a defined threshold. By applying this equation, it is possible to determine a speed beyond which changing gears would not result in low torque outputs.

In order to simulate the fuel consumption yielded by adopting the recommendations, the authors assume that the consumption takes the mean consumption value of all observations that share the same speed value. This recommendation does

not always yield a reduction in fuel consumption, given that a trip may contain a new minimum consumption value. To overcome this fact, the authors take the minimum between the calculated and observed consumption values for each observation.

B. Threshold-Based Approach

In this work, we present a new recommendation schema based on this baseline. We consider that grouping all readings for a given speed value in the same gear may result in abnormal or impossible driving scenarios. For instance, there are speeds for which there are more than four possible gears, and changing from the lowest to the highest value or vice-versa would not be feasible in real scenarios. Furthermore, our approach does not consider torque values. Instead, we focus on the RPM readings that also reflect the torque outputted by the engine.

Like our base approach, we consider that the speed value s can not be changed, but other variables, such as rpm r can receive new values, provided that the pair (r, s) has been observed before. We also consider the current torque demand by limiting the range at which r values are considered using a threshold. With that, given an observation with speed s and rpm r , we define the set of candidate state changes C as

$$C = \{o \in O | o_s = s, 0.8r < o_r < 1.2r\} \quad (3)$$

which denotes the set of observations o contained in the dataset O that have the same speed value o_s as the value s being evaluated and whose RPM value o_r is within an arbitrary 20% variation threshold of the original r value.

By considering only those readings that registered the same speed and RPM values within a specific limit, we addressed one of the shortcomings of the baseline recommendation method that is its abnormal gear changes. In order to evaluate this, we ensure that gears are correctly assigned to readings in the case of an RPM change recommendation. We fitted a k-means classifier to assign colors and gear values to the observations in Figure 3, and in our experiments, we use the same readings. The trained classifier has information on the gear ratios that determine gear values. Thus, assigning new gears to observations that received new RPM values is a matter of feeding them to the same k-means classifier.

Equation 3 defines the set of candidate RPM values, and with that, we choose the best candidate according to their mean fuel consumption. Choosing the RPM that results in the highest efficiency does not mean that action will be taken, given that the current vehicle state may be the optimal behavior, considering all data gathered so far, and thus, changing it would result in an increase in consumption. As a result, trips become a mix between recommended new data, in the cases of improvements in efficiency, and original readings, which have the optimal observed behavior.

C. Spatial Clustering and Location-Based Recommendation

Given that there are aspects of the driving activity associated with location, taking them into account is expected to

yield more realistic results. The recommendation policy we developed assumes that instant speed can not be changed. Otherwise, trips' structures would suffer alterations that would require generating new traces. With the targeted variable being engine RPM, local variations capture aspects such as the way drivers interact with their vehicles and the environment.

Furthermore, the approach of using observations from the same regions takes into account environmental variables, such as terrain. Ascending slopes naturally require more power from vehicles climbing them, as compared to a flat or descending stretch. In order to output more power, drivers use lower gears or push the accelerator pedal harder, both of which cause increased fuel consumption. These two changes can also improve efficiency in reverse situations, such as downhill, by using lighter gears and requiring less power from the engine. By taking into account the region in the recommendation task, our approach considers these local variations induced by the terrain and finds the optimal observed actions.

To consider observations' location in the recommendation service, we essentially change how we select the candidates in our proposed recommendation strategy, defined in Equation 3. Our location-based recommendation approach selects candidates based on their current speed and RPM values, given an RPM threshold, but conditioned to the fact that all candidate observations belong to the same spatial cluster, represented by grids that comprise an area of approximately 1 km². With that, we define the set of candidate observations C_{grid} for the original setup comprised of speed s , RPM r , and grid group g as

$$C_{\text{grid}} = \{o \in O | o_g = g, o_s = s, 0.8r < o_r < 1.2r\} \quad (4)$$

where O denotes the set of all observations, and the threshold to consider RPM states as candidates is a 20% variation over the original value.

V. PERFORMANCE EVALUATION

We evaluate the performance of the proposed recommendation algorithms by simulating the recommendation systems' behavior on the enviroCar dataset's sensor readings. We believe that, with access to the required resources, it would be possible to deploy all three systems on small processing devices and connect them to cars' OBD ports to gather real-time readings and give drivers feedback on their consumption. In our simulations, we compare the performance of the proposed recommendation algorithms by comparing them to the original dataset and the recommendations generated by an implementation of the baseline approach by Rettore et al. [3]. We not only evaluate the differences in fuel consumption but also how recommendations would affect drivers' behavior. To evaluate the latter, we look into the gear and RPM values assigned by the methods.

The most significant impact of the eco-driving recommendation strategy we developed is to reduce the fuel consumption, which is related to CO₂ emissions. Reducing these two variables is one of the missions of eco-driving techniques, which aim at making driving more environmentally efficient.

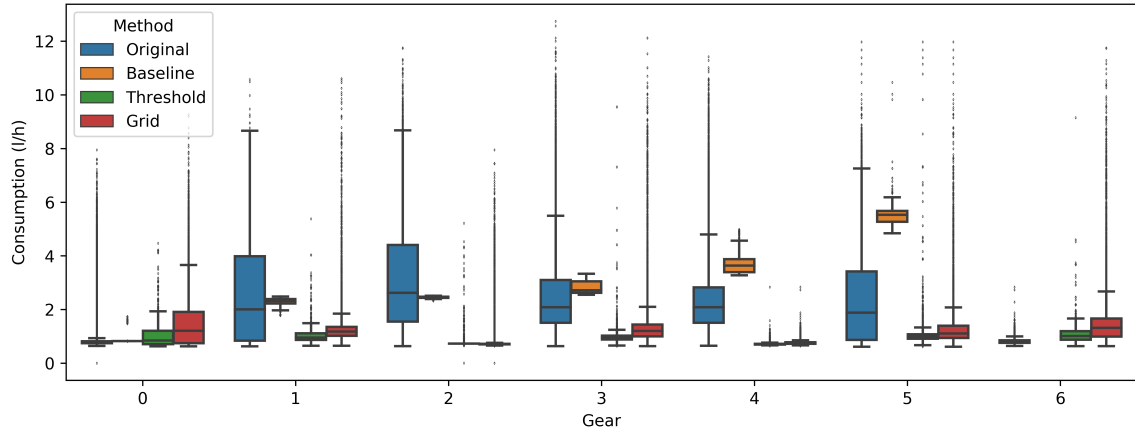


Fig. 4. Fuel consumption distribution per gear

Compared to the original dataset, our baseline approach reduced fuel consumption 23% on average, while our threshold approaches achieved a 43% mean reduction using the spatial clusters, which improved to a 59% reduction when considering the entire dataset. We credit the difference between the baseline and both threshold methods to the grouping of observations made by the baseline algorithm. That approach groups all observations in the same speed value into the same gear, including data points with inadequate original gear and RPM values in its recommendation. On the other hand, our approaches limit their choices of recommended states to a defined limit based on the original values, which limit them to pertinent changes only.

Compared to the original dataset, applying all three methods resulted in less fuel consumption in all recorded trips. Both methods we proposed in this work reduced total fuel consumed on all trips, compared with the baseline method. We present such reductions in Figure 4, by illustrating the consumption distribution divided into gears.

That figure shows that the most significant improvements in fuel consumption were in gears 1 to 5, which correspond to the gears with a well-defined linear relationship between speed and RPM, as seen in Figure 3. Gears 0 and 6, in which the threshold and grid methods presented fuel consumption above the original dataset, represent unusual operation states, such as takeoffs and gear changes. In fact, the increase in fuel consumption on gear 0 can also be a result of the migration of some of its original observations to other gear, as a result of suggested RPM values.

Comparing the two threshold approaches, one using the entire dataset, and the other using grids, we note that the median value using grids is always higher than considering all observations. The reason for that difference lies in the selection of local and global minimum values. The grid method takes into account a subset of the dataset, looking for the local behavior that leads to minimum fuel consumption, while the threshold method can evaluate all observations and find the

global minimum values. This difference is also the reason why the number of outlying points is smaller with the threshold method; that is, it can fit abnormal situations into more common states by recommending behaviors observed in other regions.

We now evaluate the impact of these methods in the resulting gear values, as compared with the original dataset's values. Figure 5 presents the changes in gear values caused by the three methods as compared with the original dataset, as well as the original gear value distribution, and the resulting distributions after each of the recommendation methods. Each of the three matrices presents the joint distribution of gear values, illustrating the changes in values caused by them. Each cell, defined by (o, g) contains the ratio of observations whose original gear value was o and changed to g after each method's recommendations. For example, in the baseline approach, 23% of the observations had their original gear value $o = 0$, and the algorithm kept such values. On the other hand, the same observations with original gear value $o = 0$ were assigned new gears $g = 3$ by the threshold and grid methods.

Regarding the two new proposed methods, we note that they tend to change the original gears. While the baseline method kept 27% of the original gear values, the threshold and grid approaches kept 17% and 24%, respectively. It is also possible to note that the value change patterns of both methods are similar, meaning that the majority of the original gears tend to be changed into the same new values, only differing in the proportion of such values. For instance, the observations initially assigned to gear 4 changed into gears 2 and 5, the first more frequently than the last in the grid method, while we observed the opposite in the threshold method. The similar gear change patterns also make the resulting distributions is also observed in the resulting distributions.

VI. CONCLUSION AND FUTURE WORK

In this work, we presented two new eco-driving recommendation tools. Among the various aspects of driving that eco-driving techniques focus on, we drew our attention to gear

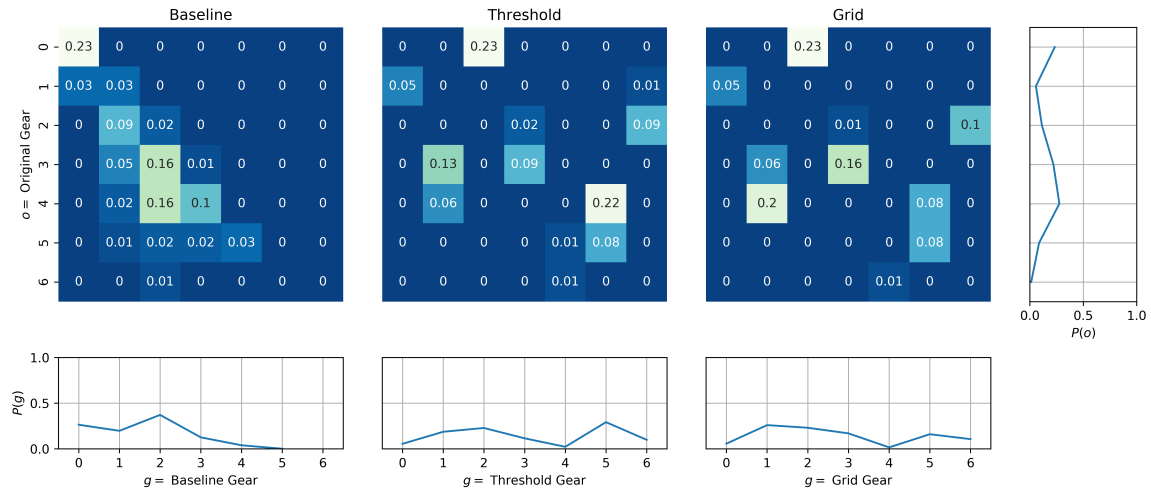


Fig. 5. Comparison of gear transitions from each approach

changes and engine RPM. We proposed a recommendation tool and tested it using two datasets: the entire collection of observations from a single car model, and the observations of the same model grouped into spatial grids. Our recommendation algorithm evaluates the effects of changing RPM values while keeping the same speed on fuel consumption. These changes can result in gear changes as well, and we also simulated this effect.

We developed two new models, one that uses the entire dataset and others that only considers observations made in the same region. In our simulations, both models improved fuel efficiency, both considering the original dataset and the baseline method. However, the model that uses the entire dataset performed better than the regional model, given that the first has access to global optimal observations, instead of local optimal readings. Moreover, the similar changes in gear values stress the fact that both recommendation strategies yield comparable results, although one only uses a subset of the dataset. Nevertheless, our results show that even considering local observations, our recommendation method had a positive impact on fuel consumption.

ACKNOWLEDGMENTS

This work was supported by NSERC, Canada Research Chairs Program, NSERC CREATE - TRANSIT Program,

We consider that such recommendation systems could be deployed and take advantage of vehicular communications, by aggregating the best sensor readings through geocaching policies, and transmitting this knowledge to vehicles as they enter new regions. As future extensions of this work, we plan on expanding the analysis of our methodology to other vehicle models, eventually using the entire dataset and evaluating cross-model recommendations. Moreover, we plan on creating a more powerful simulation that considers the impacts on speed after taking RPM recommendations. This new simulation would result in generating new artificial traces based on the original observations.

CNPq, CAPES, grants #15/24494-8 & #18/23064-8, São Paulo Research Foundation (FAPESP).

REFERENCES

- [1] A. Sanguinetti, K. Kurani, and J. Davies, "The many reasons your mileage may vary: Toward a unifying typology of eco-driving behaviors," *Transportation Research Part D: Transport and Environment*, vol. 52, pp. 73–84, 2017.
- [2] M. Schröder and P. Cabral, "Eco-friendly 3D-Routing : A GIS based 3D-Routing-Model to estimate and reduce CO₂ -emissions of distribution transports," *Computers, Environment and Urban Systems*, vol. 73, no. July 2018, pp. 40–55, 2019.
- [3] P. H. L. Rettore, A. B. Campolina, L. A. Villas, and A. A. F. Loureiro, "A method of eco-driving based on intra-vehicular sensor data," in *2017 IEEE Symposium on Computers and Communications (ISCC)*. Heraklion, Greece: IEEE, Jul. 2017, pp. 1122–1127.
- [4] S. R. Kancharla and G. Ramadurai, "Incorporating driving cycle based fuel consumption estimation in green vehicle routing problems," *Sustainable Cities and Society*, vol. 40, no. September 2017, pp. 214–221, Jul. 2018.
- [5] Y. Zhao, S. Li, S. Hu, L. Su, S. Yao, H. Shao, H. Wang, and T. Abdelzaher, "GreenDrive," in *Proceedings of the 8th International Conference on Cyber-Physical Systems - ICCPS '17*. New York, New York, USA: ACM Press, 2017, pp. 229–238.
- [6] V. C. Magana and M. Munoz-Organero, "Artemisa: A Personal Driving Assistant for Fuel Saving," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2437–2451, 2016.
- [7] F. Saremi, O. Fatemeh, H. Ahmadi, H. Wang, T. Abdelzaher, R. Ganti, H. Liu, S. Hu, S. Li, and L. Su, "Experiences with GreenGPSFuel-Efficient Navigation Using Participatory Sensing," *IEEE Transactions on Mobile Computing*, vol. 15, no. 3, pp. 672–689, Mar. 2016.
- [8] R. Xiong, J. Cao, and Q. Yu, "Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle," *Applied Energy*, vol. 211, no. 5, pp. 538–548, Feb. 2018.
- [9] K. Ayyildiz, F. Cavallaro, S. Nocera, and R. Willenbrock, "Reducing fuel consumption and carbon emissions through eco-drive training," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 46, pp. 96–110, Apr. 2017.
- [10] A. Bröring, A. Remke, C. Stasch, C. Autermann, M. Rieke, and J. Möllers, "enviroCar: A Citizen Science Platform for Analyzing and Mapping Crowd-Sourced Car Sensor Data," *Transactions in GIS*, vol. 19, no. 3, pp. 362–376, 2015.