Demystifying Interactions Between Driving Behaviors and Styles Through Self-clustering Algorithms

Yu Zhang[1,2], Wangkai Jin[1], Zeyu Xiong[1], Zhihao Li[1], Yuyang Liu[1,3P], and Xiangjun Peng[1,4P](✉)

[1] User-Centric Computing Group, University of Nottingham, Ningbo, China
xjpeng@cse.cuhk.edu.hk
[2] School of Computer Science, University of Nottingham, Nottingham, UK
[3] Department of Computer Science, University of Machester, Machester, UK
[4] Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong

Abstract. We argue that driving styles demand adaptive classifications, and such mechanisms are essential for adaptive and personalized Human-Vehicle Interaction systems. To this end, we conduct an in-depth study to demystify complicated interactions between driving behaviors and styles. The key idea behind this study is to enable different numbers of clusters on the fly, when classifying driving behaviors. We achieve so by applying Self-Clustering algorithms (i.e. DBSCAN) over a state-of-the-art open-sourced dataset of Human-Vehicle Interactions. Our results derive 8 key findings, which showcases the complicated interactions between driving behaviors and driving styles. Hence, we conjecture that future Human-Vehicle Interactions systems demand similar approaches for the characterizations of drivers, to enable more adaptive and personalized Human-Vehicle Interaction systems. We believe our findings can stimulate and benefit more future research as well.

Keywords: Driving behaviors · Driving styles · Adaptive & personalized human-vehicle interactions

1 Introduction

Responses from Driver-Vehicle Interactions, whether they satisfy drivers' expectations or not, have significant impacts on users' trusts in terms of Autonomous Driving. To deliver user-expected interactions, detailed insights from Driving statistics are the most critical parts of modern Human-Vehicle Interaction systems. For instance, users' trust in Autonomous Vehicles are highly dependent with such responses, which rely on the detailed insights from driving statistics [16,30,37]. Recent efforts characterize driving behaviors empirically, and further classify them into multiple driving styles in static partitions. However, with the growing popularity of Autonomous Vehicles, computational methods, rather

H. Krömker (Ed.): HCII 2021, LNCS 12791, pp. 335–350, 2021.
https://doi.org/10.1007/978-3-030-78358-7_23

than empirical methods, can potentially fit better within personalized Human-Vehicle Interactions, in practice.

With such a mindset, we argue that **conventional classifications of driving styles are not suitable for adaptive and personalized Human-Vehicle Interaction systems**. We disagree with the conventional approach from the following two aspects. First, static classifications of driving styles are not adaptive during the driving procedures; and second, driving styles, derived from empirical studies, are insufficient to contribute to personalized Human-Vehicle Interaction techniques. We believe that the root causes of the above issues are because complicated interactions between driving styles and behaviors remain under-studied, in terms of both mechanisms and findings.

Our goal is to demystify complicated interactions between driving behaviors and driving styles, to reveal the opportunities for adaptive and personalized Human-Vehicle Interactions. We make the **key observation** that the conventional classifications of driving styles rely on static partitions of driving behaviors, obtained from empirical studies. In other words, the problem is abstracted as clustering techniques, with the pre-determined number of clusters. To this end, our key idea is to apply computational techniques to eliminate the needs for pre-determined number of clusters. Hence, we utilize self-clustering algorithms, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), for adaptive classifications of driving styles and hidden patterns of driving behaviors.

We perform our studies over BROOK, a state-of-the-art and open-sourced dataset for Human-Vehicle Interactions. Our studies have included 34 drivers in 11 dimensions of driving statistics [23]. In total, we make 8 key findings through our studies. Our studies start with rigorous examinations of the impacts from different DBSCAN configurations, representative driver groups, time-series variations, road conditions and etc. Furthermore, we characterize in-depth characteristics of driving styles, by breaking down detailed features and analyzing the overlap across different styles. Based on the above findings, we confirm that the interactions between driving behaviors are more complicated, and our DBSCAN-based approach is more applicable in this context, compared with conventional partitions of driving styles.

We make the following **key contributions** in this paper:

– We address the problem that conventional classifications of driving styles overlook the opportunities for adaptive and personalized Human-Vehicle Interactions, and identify that the static partitions of driving styles are the key limitation in conventional approaches.
– To the best of our knowledge, we are the first to propose and utilize Density-Based Spatial Clustering of Applications with Noise (DBSCAN), for adaptive classifications of driving styles and hidden patterns of driving behaviors.
– We experimentally characterize and examine the effects of our DBSCAN-based approach over BROOK, a state-of-the-art and open-sourced dataset for Human-Vehicle Interactions.

– We retrieve 8 key findings from the above studies, by rigorously changing different configurations. These observations can serve as starting guidelines for adaptive and personalized Human-Vehicle Interactions systems in the future, for both research and industrial communities.

The rest of this paper would be organized as follow. Section 2 introduces the related works in studying driving behaviors and driving styles. Section 3 gives details about the experiment methodology. Results are shown in Sect. 4. Section 5 reports an discussion in order to inspire potential principles. Section 6 presents the conclusions and future work.

2 Background and Motivation

Modern Methodology, to evaluate driving style, can be divided into Subjective Evaluation and Objective Evaluation. Subjective Evaluation is carried out through quesionnaires and surveys, to obtain empirical results. For instance, [24] first proposed Driver Behavior Questionnaire (DBQ), to reflect bad driving behavior by self-reporting. Since then, follow-up efforts, based on DBQ, investigates on the impacts of regions, cultures, ages and genders, through the variations of drivers' behaviors. For instance, [13] developed Driver Style Questionnaire (DSQ), to study correlations between Traffic Accidents and Driving Behaviors (e.g. Speed and Distance of Vehicles). Another example is Multidimensional Driving Style Inventory (MDSI) [31] enables the capability to evaluate driving styles from multiple dimensions. More specifically, it defines the structure of driving styles and explicitly classifies them into four categories.

However, these methods are subjective and could be influenced by some external factors. Hence, Subjective Evaluations demand high standards of the effectiveness of the driving questionnaire and experts' experience. To this end, Objective Evaluation is proposed to complement this method. Objective Evaluations analyze driving styles through driving statistics, which are obtained from a driving simulator [4] or in-field vehicles [15]. In the context of driving styles, [5] proposes a classification and recognition model for driving behavior based on sparse representations. More specifically, the vehicle motion tracks, obtained by vision, are used as input in this model, and the sparse representation approach is used to mine the features of the driving behavior decision. Another example is as follow. [33] proposes a pattern recognition method, which utilizes triaxial accelerometers' statistical data to evaluate normal and aggressive driving styles. They further discuss time-domain feature extraction. Also, [20] divides the following behavior according to the difference of patience and puts forward the view, that the following time can measure the driving style. The Mean and Standard Deviations of relevant indicators are often used to differentiate the driver's driving styles, based on the assumption of a normal distribution for these indicators [17,34]. However, these measures, derived directly from sequential observations, are based on static criterion, which can be inconsistent with established parametric distributions [7].

Previous attempts, to characterize driving styles, focus on the differentiated trends of driving statistics (e.g. driving speeds, headway distance, following time, multi-modal information and etc.) [3,5,6,10,12,20,33]. Such studies overlook the effects of time-series and results in coarse-grained decision-making procedures, for the determinations and classifications of driving styles. Hence, we make the key observation that static partitions of driving styles/behaviors may not be suitable for adaptive and personalized Driver-Vehicle Interactions. To this end, we propose to utilize self-clustering algorithms, and compare auto-generated patterns on-the-fly instead of statically partitioned driving styles.

3 Study Methodology

3.1 Dataset Description

Our study uses a public multi-modal database for Human-Vehicle Interaction – BROOK [23]. BROOK contains 34 drivers' data under four driving scenarios (both manual and automated driving), such as Time, Vehicle Speed, Vehicle Acceleration and Vehicle Coordinates. We utilize representative drivers' driving statistics as input, in terms of time series.

3.2 Dataset Pre-processing

Before the bulk of this study, we first normalize all statistics based on the following insights.

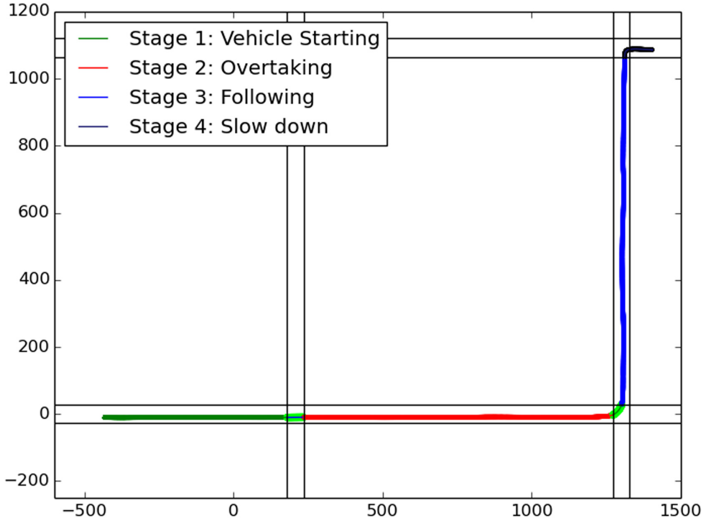


Fig. 1. Movement track of vehicle.

Data Cleanup. The original database records consist of both only stable stages and star-up stages and parking stages. Since the driving styles are characterized within relatively stable stages, we eliminate the unstable stages to ensure our studies are consistent with others. This representative length of each driving scenario is enough for driving behavior data analysis. The route map is fixed and we show it in Fig. 1.

As shown in Fig. 1, the whole driving route can be divided into four stages[1], which are marked in different colors. The statistics reflect that, there are huge gaps between stable stage and unstable stage, in terms of driving behaviors. For instance, driving speed usually fluctuates within a certain range, where unstable stages drift more randomly. Hence, we consider the statistics, without the fixed range, as a noise source, which is groundless for the characterizing driving styles.

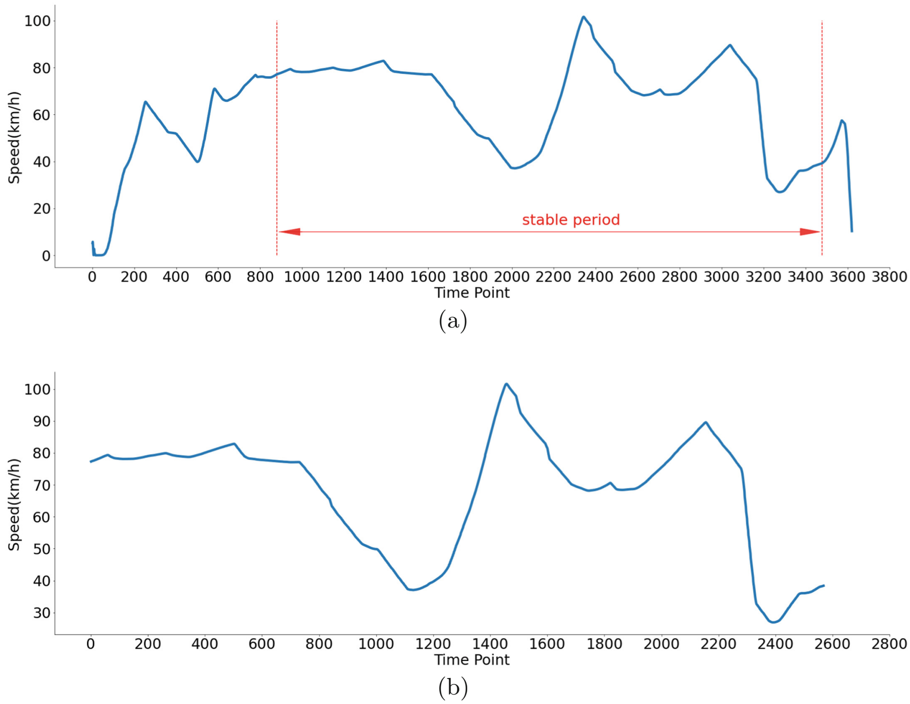


Fig. 2. Filter driving data

Feature Selection. We perform feature selections based on the following insights, where we present backup information in Fig. 2.

Figure 2-(a): **One-way Analysis of Variance (ANOVA).** Feature selection, as a pre-processing stage, aims to select the most discriminative features.

[1] The stages don't take traffic lights into account.

From the perspective of clustering, removing irrelevant features won't negatively impact the accuracy of clustering. This is because irrelevant variables may increase noise and mask the underlying pattern or structure in the dataset, as suggested by [36]. Moreover, such cleanup can reduce the required storage and processing time. Our decisions are to utilize Wrapper Approach [2] on the whole dataset, and use these selected features to construct the clusters. The quality of clustering is an indicator of whether the subset of features is satisfactory, where ANOVA method was used. Table 1 presents Variance Values of corresponding features. To this end, those featuress with small variance values (i.e., a variance value <0.01) are regarded as meaningless features and then removed [9,28]. Therefore, the features, such as Position$_x$, Position$_z$, Rotation$_y$, Rotation$_w$, Speed (km/h), Steering wheel position$_1$ and Headaway time (sec), remain for further analysis.

Table 1. Comparison of different obfuscations in terms of their transformation capabilities

Feature	Variance value	Feature	Variance value
Position$_x$	$4.906680e+04$	Rotation$_x$	$3.873823e-07$
Position$_y$	$8.139139e-29$	Rotation$_y$	$6.590198e-01$
Position$_z$	$1.024064e+05$	Rotation$_z$	$9.440301e-08$
Speed (km/h)	$1.172809e-01$	Rotation$_w$	$1.172809e-01$
Steering Wheel Position$_-$	$8.841402e-04$	Steering Wheel Position$_+$	$9.772212e-02$
Gas Pedal Position	$2.234481e-03$	Brake Pedal Position	$0.000000e+00$
Engine Running	$0.000000e+00$	Distance Ahead (meters)	$0.000000e+00$
Time to Collision (sec)	$0.000000e+00$	Headway Time (sec)	$1.534207e-01$

Figure 2-(b): **Principal Component Analysis (PCA).** PCA is a type of unsupervised method of reducing dimension, which produces latent factors that are known as primary components (PCs). The BROOK database consists of many kinds of data streams (e.g., Vehicle Speed, Vehicle Acceleration), the scales and units of different data are different as well. Since PCA is sensitive to the relative scaling of the original data [14], data normalization needs to be applied before PCA. To transform them into suitable formats for Object Similarity Calculations, we apply Min-Max Normalization to regularize all statistics to facilitate this need [22]. After that, we project the data onto the maximum feature vector to obtain a one-dimensional feature space to find the principal components representing each sample. With each subsequent component explaining less, the first component explains most of the variance in the data.

3.3 Clustering Algorithm

One of the most common clustering strategies is the K-Means Clustering, which requires a pre-determined number of clusters. However, preconditioning the number of clusters are quite challenging since sophisticated knowledge of the domain is required. In our context, we aim to relax such constraints so that we are capable to obtain more insights of the spatial correlations among driving behaviors. To this end, we choose Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [8], for adaptive classifications of driving styles and hidden patterns of driving behaviors. In this way, we aim to demystify the patterns within driving procedures. DBSCAN is a self-clustering algorithm, where the number of clusters is not necessary to be pre-determined. DBSCAN continuously merges two most similar clusters into a new cluster in each iteration until satisfying certain termination criterion (e.g. distance threshold) [19]. Hereby, we elaborate more details of this algorithm and our design choices as follow.

Distance Measurement. When performing data clustering, a basic step is to choose an appropriate distance calculation method that quantifies how similar individuals concern measurements provided in the variables. The most commonly used distance measurement is Euclidean Distance, where we take into account at the first place. For completeness of our study, we also utilize Manhattan Distance and Chebyshev Distance to quantify these effects [1] and the results are reported in Table 2. Hereby, we demonstrate these distances to the DBSCAN algorithm mathematically, as shown in the following equations.

$$D_{\text{Euclidean}} = \sqrt{\sum_{i=1}^{N} (x_{1i} - x_{2i})^2} \tag{1}$$

$$D_{\text{Manhattan}} = \sum^{n} |x_{1i} - x_{2i}| \tag{2}$$

$$D_{\text{Chebyshev}}(x, y) = \max_i (|x_i - y_i|) \tag{3}$$

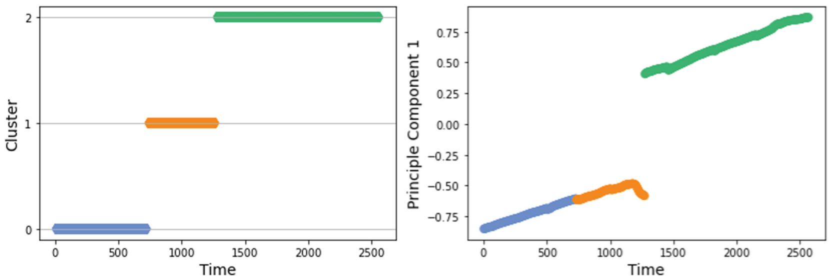
Parameters Setting. DBSCAN also needs to take *minPoints* and *epsilon* as input parameters. The *minPoints* refers to the minimum number of data points within the cluster, and *epsilon* refers to the max radius of the cluster. If the *minPoints* value is too small, more core objects will be generated, leading to too many clusters. On the contrary, two adjacent clusters with higher density may be merged into the same cluster, resulting in fewer clusters. The influence of the value selection of *epsilon* also has similar effects. We use the principle *minPoints* = $2 \cdot dim$ [26] to select an appropriate range of *minPoints*. After intensive rounds of rigorous testing, we identify NINE different setups of both parameters to serve as the representatives of all possible combinations. This is because our goal is not to provide a recommended, near-optimal setup but to demystify the interactions of driving behaviors and styles in detail.

4 Experiment Results

In this section, we present several key results and relevant findings, to showcase the complicated interactions between driving behaviors and driving styles.

4.1 Conventional Classification Against Self-clustering

Finding 1: Heterogeneous Styles can be Generated from DBSCAN. We select the most important dimension (Principle Component 1) after PCA dimension reduction to represent the original driving behavior feature set. The larger the feature variance, the more original data information can be retained. Although the percentage of Principle Component 1's variance in all feature variances is only 0.65990205, it can also retain most of the required information for information visualization. After using DBSCAN algorithm to assign data points, we showcase the clustering results (as shown in Fig. 3-(a)) and the characteristic examples of the change, in terms of microtubule length versus time, are shown in Fig. 3-(b).



(a) Cluster assignment (b) Driving behavior signature

Fig. 3. Driving-style quantification results of Driver Group 3 under road condition 4.

Table 2. Clustering results for different driver groups under the same road conditions.

Distance measurements	Parameters combination identification									
	eps	0.125	0.25	0.5	0.125	0.25	0.5	0.125	0.25	0.5
	minPoints	3	3	3	6	6	6	9	9	9
Eulidean distance	Group 1	6	5	3	6	5	3	7	5	3
Manhattan distance		8	5	3	10	5	3	10	5	3
Eulidean distance		5	3	3	5	3	3	6	3	3
Eulidean distance	Group 2	12	3	3	12	3	3	13	3	3
Manhattan distance		15	5	3	17	6	3	17	6	3
Eulidean distance		10	3	3	10	3	3	10	3	3
Eulidean distance	Group 3	6	3	3	6	4	3	6	5	3
Manhattan distance		6	5	3	6	6	3	6	6	3
Eulidean distance		5	3	3	6	4	3	6	5	3

Different from conventional methods, self-clustering method automatically divides the whole driving stage's data into three categories. We report similar results as presented in Table 2 by adopting the same research method for multiple drivers' data.

Finding 2: Time-Series Variation is Considered by DBSCAN. As displayed in Fig. 3, there are certain continuities in behaviors from a single complete time interval. For instance, the clustering part with green color shows that the drivers exhibits the same driving style over this period. However, behaviors for the whole timeline do not exhibit the same characteristics, while they change dramatically across all driving events. This phenomenon coincides with the different driving stages presented in Fig. 1: the driver entered the following stage after overtaking, which further expounds that drivers will give different driving styles in various driving events. What's more, as is evident from Fig. 3-(b), the overtaking stage also consists of different driving styles.

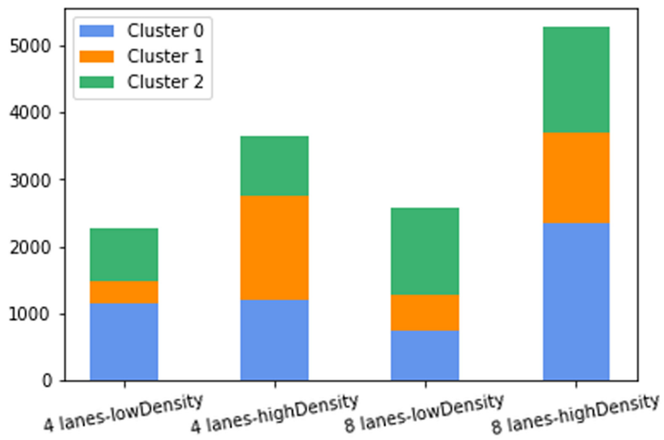


Fig. 4. Clustering results under different road conditions.

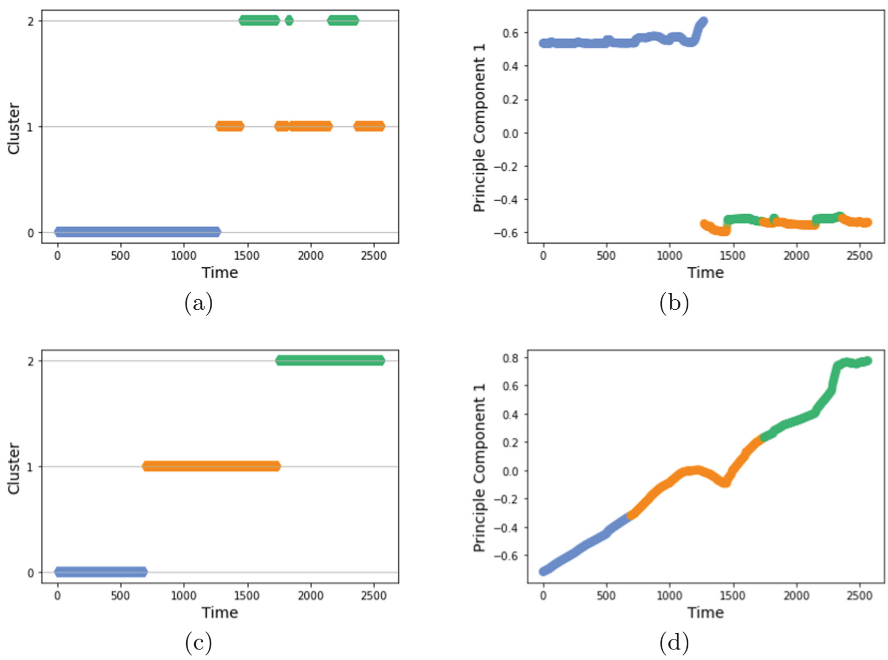
Finding 3: Driving Styles can be Customized for Different Roads. In order to verify whether the clustering results of drivers change under different traffic conditions, we conduct experiments on driving data under different road conditions. We observe that, though the proportion of different driving styles in the whole driving event has changed, the clustering results in a single driving event are still three. Nevertheless, the same driving style has different characterization probability under different road conditions. As shown in Fig. 4, the drivers will show the driving style 1 (cluster 1) with a smaller-time period, under relatively small traffic density.

Finding 4: Degree of Expression of Different Styles. We also observe that different styles have different degrees of representation in the whole driving event. Taking the clustering results, obtained from this setting (i.e. eps = 0.5, minPoints = 9, Distance Measures = Eulidean Distance) as an example, the three driver groups lead to three clustering results as driving styles. But the degrees of different styles are different. Table 3 backs up this finding: Drivers 1 is style 1 56% of the whole time, with the time of style 2 and style 3, 30% and 14% respectively. The pattern for drivers 2 and drivers 3 appear to be reversed. Style 3 is the domain style during this period, with 54% for drivers 2 and 50% for drivers 3. It can be seen that the differences between individuals are significant.

Table 3. Different levels of style representation.

	Style 1	Style 2	Style 3
Drivers 1	56%	30%	14%
Drivers 2	18%	28%	54%
Drivers 3	29%	21%	50%

4.2 In-Depth Driving Style Analysis



(a) (b) (c) (d)

Fig. 5. Quantification results of driving styles after breakdown analysis.

Finding 5: Isolated Features can Greatly Impact the Classifications. We perform breakdown analysis by removing key features. Hereby, we isolate

Position-related features or Rotation-related features. Figure 5-(a) and Fig. 5-(b) report the results, where we remove Position-related features; and Fig. 5-(c) and Fig. 5-(d) report the results, where we remove Rotation-related features. We observe that, though the classification results change, the clustering results are still complicated. After comparing Fig. 5-(a)/(c) with Fig. 5-(b)/(d), though the driving styles are still classified into three clusters, removing features have significantly impacted the robustness of classifications, in terms of timeline, due to greatly-impacted features.

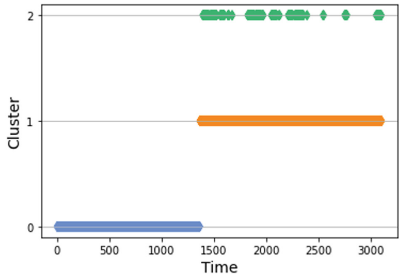
Finding 6: Transient Effects form Mutable Driving Styles. [21] considers driving styles are transient, and explain that a driver can be aggressive at one period but normal for others. Our experiments also back up that, it's possible for the mutation of driving style, from computational perspectives. As shown Fig. 5-(a), two different driving styles alternately characterize the driver's behavior in the second half part of the whole procedure. This can be explained as the driver's step on the brake and the oil port in congestion, showing a particular different style compared with drivers with stable driving speed on highway. To this end, our mechanism are more adaptive and robust compared with conventional methods.

4.3 Driving Styles Overlap

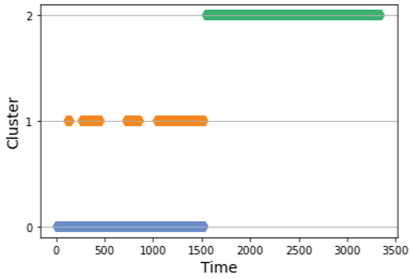
Finding 7: Clusters That are More Likely to Overlap. We extend our studies by changing the time window to re-cluster the statistics, and combine the results after using different sizes of time window. We observe that the same driving behavior is possible to be classified into different styles. The part between the two red lines in Fig. 6 reflects this complex clustering situation, some data points are clustered such that they belong to two different clusters. More specifically, Cluster 1 is more likely to overlap with the other two driving styles.

Finding 8: Scenarios That are More Likely to Occur Overlap. In Fig. 6-(c), the overlap occurs when the vehicles encounters a traffic light. This is because the steering of the car is not completely independent of acceleration or braking under the driving event. Thus, wrongly classifying the steering as acceleration or braking leads to this overlap. While in the situations at Fig. 6-(a) and -(b), there are complicated traffic conditions, like congestion leading the driver to have driving behaviors, that are not routine.

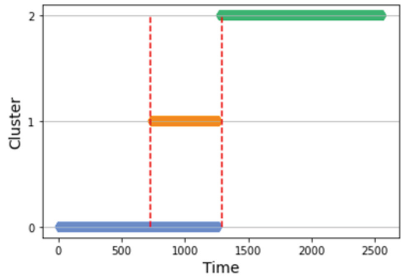
So far, we have compared the clustering results based on multiple drivers and a single time series. We find that not only different drivers will show various styles, but also the same driver will show style differences among driving periods.



(a) Drivers 1



(b) Drivers 2



(c) Drivers 3

Fig. 6. Overlaps between driving styles.

5 Discussion

This work analyze the clustering results of driving behaviors and extensively retrieve multiple findings from computational perspectives. Unlike conventional classifications of driving styles, our findings show that our proposed mechanisms for driving styles reveal more opportunities for adaptive and personalized driving styles' characterizations. Driving styles will migrate under the influence of traffic environments, road conditions, and other environmental factors. We can only conjuncture that drivers are more likely to be in a specific driving style, instead of classifying them into one specific type. Based on our findings, we further relate the key points hereby to stimulate new insights and follow-up investigations.

5.1 Driving Behavior Combination

Driving behaviors can describe combined events, because different driving behaviors are not always distinguished and independent. For example, acceleration is often followed by turning at traffic lights, and deceleration behavior is accompanied by stopping in front of traffic lights. After turning to the other direction, the driver can adjust the speed and lane to better drive experience. This phenomenon makes it difficult to distinguish steering from acceleration or deceleration behavior under certain conditions.

5.2 Traffic Environment

The rising number of vehicles can exacerbate road congestion and render the flow of traffic more complicated. Certain events may be detrimental in some situations. The difference in road conditions will bring difficulties in identifying driving style because driving behavior will change. For instance, the calm type driver will frequently step on the brake and the oil port in the case of congestion, showing a particular aggressive style, while the aggressive type will maintain a relatively stable driving speed on highway. Our studies reveal that, the analysis of driving behaviors in specific road conditions is more critical for style representation.

5.3 Driving Behavior Levels

[32] divides the completion of a driving main task into four main levels: strategic level, mode level, operational level and scene awareness level. Driving style can be reflected on any level. There are: (1) Decision preferences at the strategic level, such as selecting short-distance routes [18]; (2) Driving mode preferences at the mode level, such as frequent lane change, near-following, and far-following [27]; and (3) Operating mode preferences at the operational level, such as uniform acceleration, rapid acceleration, and whether to turn on the turn signal in time [25]. At the level of perceptions, there are recognition preferences such as whether to observe the external area adequately before the lane changes and whether the sight line deviates from the path for a long time [35].

We also vision that our study is complementary to other relevant works as well. [11] provides alternative mechanisms to obtain drivers' multi-modal statistics in a more user-friendly manner. [30] examines the influences of user trust in auto-vehicles, by applying BROOK [23] as the dataset. [29] also ignites the opportunities for more practical infrastructure to enhance the dataset. We believe future works are both essential and promising.

6 Conclusions

We argue that driving styles demand adaptive classifications, and such mechanisms are essential for adaptive and personalized Human-Vehicle Interaction systems. To this end, we conduct an in-depth study to demystify complicated interactions between driving behaviors and styles. The key idea behind this study is to enable different numbers of clusters on the fly, when classifying driving behaviors. We achieve so by applying Self-Clustering algorithms (i.e. DBSCAN) over a state-of-the-art open-sourced dataset of Human-Vehicle Interactions. Our results derive 8 key findings, which showcases the complicated interactions between driving behaviors and driving styles. Hence, we conjecture that future Human-Vehicle Interactions systems demand similar approaches for the characterizations of drivers, to enable more adaptive and personalized Human-Vehicle Interaction systems. We believe our findings can stimulate and benefit more future research as well.

Acknowledgements. We thank for the anonymous reviewers from HCI'21 Regular Paper Track and all members of User-Centric Computing Group for their valuable and insightful feedbacks, especially Mr. Zhentao Huang. This project is a part of the BROOK project from the User-Centric Computing Group in the University of Nottingham Ningbo China [23].

References

1. Aghabozorgi, S., Shirkhorshidi, A.S., Wah, T.Y.: Time-series clustering-a decade review. Inf. Syst. **53**, 16–38 (2015)
2. Alelyani, S., Tang, J., Liu, H.: Feature selection for clustering: a review. Data Clustering: Algorithms Appl. **29**(1), 230 (2013)
3. Augustynowicz, A.: Preliminary classification of driving style with objective rank method. Int. J. Automot. Technol. **10**(5), 607–610 (2009)
4. Chen, S.W., Fang, C.Y., Tien, C.T.: Driving behaviour modelling system based on graph construction. Transp. Res. Part C: Emerging Technol. **26**, 314–330 (2013)
5. Chen, Z.J., et al.: Vehicle behavior learning via sparse reconstruction with $\ell_2 - \ell_p$ minimization and trajectory similarity. IEEE Trans. Intell. Transp. Syst. **18**(2), 236–247 (2016)
6. Constantinescu, Z., Marinoiu, C., Vladoiu, M.: Driving style analysis using data mining techniques. Int. J. Comput. Commun. Control **5**(5), 654–663 (2010)
7. Driggs-Campbell, K., Govindarajan, V., Bajcsy, R.: Integrating intuitive driver models in autonomous planning for interactive maneuvers. IEEE Trans. Intell. Transp. Syst. **18**(12), 3461–3472 (2017)
8. Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise. KDD **96**, 226–231 (1996)
9. Hua, J., Tembe, W.D., Dougherty, E.R.: Performance of feature-selection methods in the classification of high-dimension data. Pattern Recogn. **42**(3), 409–424 (2009)
10. Huang, J., Lin, W.C., Chin, Y.K.: Adaptive vehicle control system with driving style recognition based on headway distance (Oct 2 2012), uS Patent 8,280,560
11. Huang, Z., et al.: Face2multi-modal: In-vehicle multi-modal predictors via facial expressions. In: Adjunct Proceedings of the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2020, Virtual Event, Washington, DC, USA, 21–22 September, 2020, pp. 30–33. ACM (2020). https://doi.org/10.1145/3409251.3411716
12. van Huysduynen, H.H., Terken, J., Eggen, B.: The relation between self-reported driving style and driving behaviour: a simulator study. Transp. Res. Part F: Traffic Psychol. Behav. **56**, 245–255 (2018)
13. Ishibashi, M., Okuwa, M., Doi, S., Akamatsu, M.: Indices for characterizing driving style and their relevance to car following behavior. In: SICE Annual Conference 2007 pp. 1132–1137. IEEE (2007)
14. Janžekovič, F., Novak, T.: PCA-a powerful method for analyze ecological niches. Principal component analysis-multidisciplinary applications, pp. 127–142 (2012)
15. Johnson, D.A., Trivedi, M.M.: Driving style recognition using a smartphone as a sensor platform. In: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 1609–1615. IEEE (2011)
16. Lee, J.G., Kim, K.J., Lee, S., Shin, D.H.: Can autonomous vehicles be safe and trustworthy? effects of appearance and autonomy of unmanned driving systems. Int. J. Hum.-Comput. Inter. **31**(10), 682–691 (2015)

17. Lefevre, S., Carvalho, A., Borrelli, F.: A learning-based framework for velocity control in autonomous driving. IEEE Trans. Autom. Sci. Eng. **13**(1), 32–42 (2015)
18. Li, G., Li, S.E., Cheng, B.: Field operational test of advanced driver assistance systems in typical Chinese road conditions: the influence of driver gender, age and aggression. Int. J. Automot. Technol. **16**(5), 739–750 (2015)
19. Lin, Q., Wang, W., Zhang, Y., Dolan, J.M.: Measuring similarity of interactive driving behaviors using matrix profile. In: 2020 American Control Conference (ACC), pp. 3965–3970. IEEE (2020)
20. MacAdam, C., Bareket, Z., Fancher, P., Ervin, R.: Using neural networks to identify driving style and headway control behavior of drivers. Veh. Syst. Dyn. **29**(S1), 143–160 (1998)
21. Murphey, Y.L., Milton, R., Kiliaris, L.: Driver's style classification using jerk analysis. In: 2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems, pp. 23–28. IEEE (2009)
22. Patro, S., Sahu, K.K.: Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462 (2015)
23. Peng, X., Huang, Z., Sun, X.: Building BROOK: A Multi-modal and Facial Video Database for Human-Vehicle Interaction Research. the 1st Workshop of Speculative Designs for Emergent Personal Data Trails: Signs, Signals and Signifiers, co-located with the 2020 CHI Conference on Human Factors in Computing Systems, (CHI), Honolulu, HI, USA, 25–30 April 2020 pp. 1–9 (2020), https://arxiv.org/abs/2005.08637
24. Reason, J., Manstead, A., Stradling, S., Baxter, J., Campbell, K.: Errors and violations on the roads: a real distinction? Ergonomics **33**(10–11), 1315–1332 (1990)
25. Sagberg, F., Selpi, Bianchi Piccinini, G.F., Engström, J.: A review of research on driving styles and road safety. Human Factors **57**(7), 1248–1275 (2015)
26. Sander, J., Ester, M., Kriegel, H.P., Xu, X.: Density-based clustering in spatial databases: the algorithm gdbscan and its applications. Data Min. Knowl. Disc. **2**(2), 169–194 (1998)
27. Satzoda, R.K., Gunaratne, P., Trivedi, M.M.: Drive quality analysis of lane change maneuvers for naturalistic driving studies. In: 2015 IEEE Intelligent Vehicles Symposium (IV), pp. 654–659. IEEE (2015)
28. Scholz, M., Gatzek, S., Sterling, A., Fiehn, O., Selbig, J.: Metabolite fingerprinting: detecting biological features by independent component analysis. Bioinformatics **20**(15), 2447–2454 (2004)
29. Song, Z., Wang, S., Kong, W., Peng, X., Sun, X.: First attempt to build realistic driving scenes using video-to-video synthesis in opends framework. In: Janssen, C.P., Donker, S.F., Chuang, L.L., Ju, W. (eds.) Adjunct Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019, Utrecht, The Netherlands, 21–25 September 2019, pp. 387–391. ACM (2019). https://doi.org/10.1145/3349263.3351497
30. Sun, X., et al.: Exploring personalised autonomous vehicles to influence user trust. Cogn. Comput. **12**(6), 1170–1186 (2020). https://doi.org/10.1007/s12559-020-09757-x
31. Taubman-Ben-Ari, O., Mikulincer, M., Gillath, O.: The multidimensional driving style inventory–scale construct and validation. Accid. Anal. Prev. **36**(3), 323–332 (2004)
32. Toledo, T., Koutsopoulos, H.N., Ben-Akiva, M.: Integrated driving behavior modeling. Transp. Res. Part C Emerg. Technol. **15**(2), 96–112 (2007)

33. Vaitkus, V., Lengvenis, P., Žylius, G.: Driving style classification using long-term accelerometer information. In: 2014 19th International Conference on Methods and Models in Automation and Robotics (MMAR), pp. 641–644. IEEE (2014)
34. Wang, J., Lu, M., Li, K.: Characterization of longitudinal driving behavior by measurable parameters. Transp. Res. Rec. **2185**(1), 15–23 (2010)
35. Xian, H., Jin, L.: The effects of using in-vehicle computer on driver eye movements and driving performance. Adv. Mech. Eng. **7**(2), (2015)
36. Xie, B., Pan, W., Shen, X.: Penalized model-based clustering with cluster-specific diagonal covariance matrices and grouped variables. Electron. J. Stat. **2**, 168 (2008)
37. Zheng, P., McDonald, M.: Manual vs. adaptive cruise control-can driver's expectation be matched? Transp. Res. Part C: Emerg. Technol. **13**(5–6), 421–431 (2005)