

# CMetric: A Driving Behavior Measure using Centrality Functions

Rohan Chandra, Uttaran Bhattacharya, Trisha Mittal, Aniket Bera, and Dinesh Manocha

University of Maryland, College Park

Supplemental version including Code, Video, Datasets at <https://gamma.umd.edu/cmetric/>

**Abstract**—We present a new measure, CMetric, to classify driver behaviors using centrality functions. Our formulation combines concepts from computational graph theory and social traffic psychology to quantify and classify the behavior of human drivers. CMetric is used to compute the probability of a vehicle executing a driving style, as well as the intensity used to execute the style. Our approach is designed for realtime autonomous driving applications, where the trajectory of each vehicle or road-agent is extracted from a video. We compute a dynamic geometric graph (DGG) based on the positions and proximity of the road-agents and centrality functions corresponding to closeness and degree. These functions are used to compute the CMetric based on style likelihood and style intensity estimates. Our approach is general and makes no assumption about traffic density, heterogeneity, or how driving behaviors change over time. We present efficient techniques to compute CMetric and demonstrate its performance on well-known autonomous driving datasets. We evaluate the accuracy of CMetric and compare with ground truth behavior labels and with that of a human observer by performing a user study over a long vehicle trajectory.

## I. INTRODUCTION

Autonomous driving is an active area of research with significant developments in perception, planning, and control, along with the integration of different methods and evaluation [2]. Recent developments in perception technologies have resulted in good techniques for object recognition and tracking the positions of vehicles and road-agents using commodity visual sensors (e.g., cameras and lidars). One of the major challenges is to develop robust techniques for planning and decision-making, that can be used to compute collision-free and socially-acceptable trajectories for autonomous vehicles. This problem gets more challenging in urban environments, where the ego-vehicle is driving in close proximity to human drivers of vehicles, buses, trucks, bicycles, as well as pedestrians. A key issue in autonomous driving is safety, and the most important criteria is to compute safe and collision-free trajectories of the ego-vehicles. A recent trend is to classify and account for the driving behavior of other road agents and take them into account to perform behavior-aware planning [2], [3], [4], [5].

There is considerable work in social psychology and traffic modeling on modeling and analyzing driver behaviors. Many studies from social traffic psychology [6], [7] conclude that driving behavior falls into three broad categories—aggressive, neutral, and conservative. However, the exact definitions of these categories vary across the studies. Sagberg et al. [8] summarized these studies and developed a uniform definition such that each behavioral category can be determined in terms of specific *styles* (See Table I). For example, aggressive driving may be manifested in styles such as overspeeding, overtaking, sudden lane-changes, etc. A style refers to a specific maneuver that a driver may perform and can be related to maneuver-based road-agent behavior [9], [4].

In this paper, we mainly focus on behavior prediction from realtime traffic videos. Some of the earlier work is based on

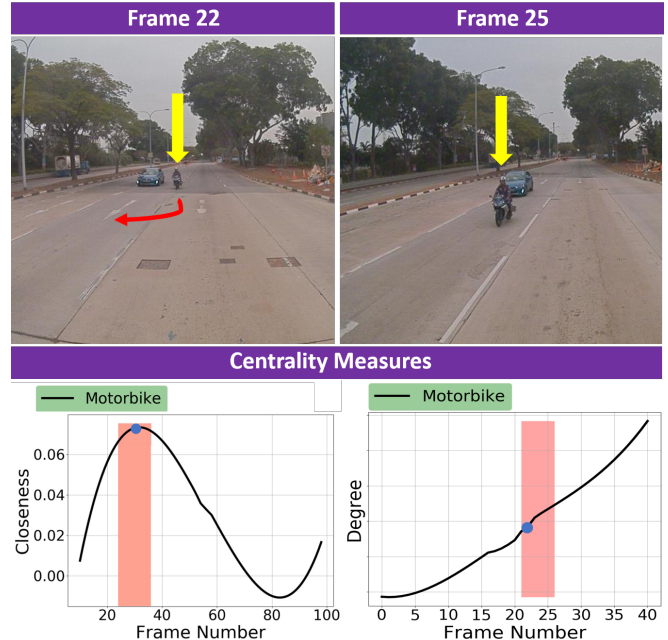


Fig. 1: **Characterization of weaving motorbike using CMetric:**

We use two centrality functions to measure different forms of aggressive styles such as weaving, overtaking, and overspeeding. For instance, the *closeness* centrality function characterizes weaving by checking for extreme points of the function. In this example, the extreme point (blue circle) of the closeness centrality lies between the 20<sup>th</sup> and 25<sup>th</sup> frames (red region), which is the duration during which the motorbike weaves past the blue car.

data-driven or machine learning methods that rely on large datasets of traffic videos with behavior labels [10], [9]. While there are a lot of recent datasets for autonomous driving [11], [12], [9], they are mostly used for scene segmentation, object recognition, or vehicle trajectories, and do not contain behavior labels. Some other methods have been proposed for automatically classifying behaviors from trajectories based on spectral analysis [13], neural networks [14], and game theory [15]. Some of these methods assume that the driver behavior does not change over a long trajectory, or they do not account for all the driving styles [8].

**Main Contributions:** We present a realtime and general metric called CMetric for characterizing driving behavior based only on the positions of vehicles. We use graph theory to model the spatial interactions between the drivers through weighted dynamic geometric graphs. Our key insight for classifying driving behavior is based on distinguishing specific styles of drivers through vertex centrality measures [16]. In graph theory and network analysis, centrality measures are real-valued functions on the vertices of a graph. We show that centrality functions have nice analytical properties that can be exploited to measure the likelihood and intensity of the driving styles. The CMetric measure is proposed as a

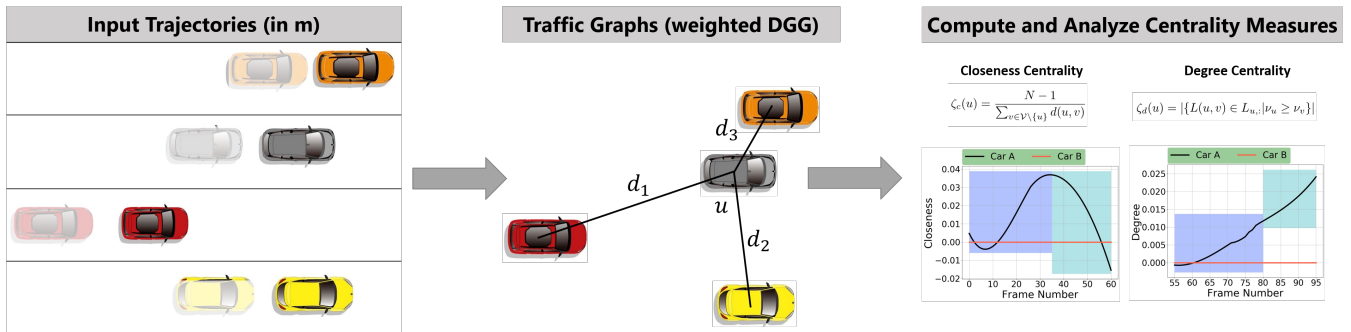


Fig. 2: **Overview of CMetric:** (left:) The sensors on autonomous vehicle observe the positions of other vehicles or road-agents; (middle:) The positions and corresponding spatial distances between vehicles are represented through a graph, DGG; (right:) Our CMetric uses the closeness and degree centrality functions to measure the style of each vehicle. These styles are used to classify a global driving behavior (such as aggressive or conservative) for each vehicle.

combination of the centrality functions and we use to classify global behaviors and different styles.

We have evaluated the performance of our approach on multiple autonomous driving urban datasets. We computed the accuracy of CMetric and compared the behavior classification results with ground truth, corresponding to human-annotated labels. We observed high classification accuracy of more than 90%. We also performed a user-study to compare the behavior classification results with that of a human observing the same video of the vehicle moving over a long time period. We notice a time difference of  $0.23s - 1.28s$  between the times instances when a human observer and our CMetric algorithm observe certain aggressive or conservative global behaviors.

As compared to prior methods, our CMetric measure offers the following benefits. It is a realtime algorithm that automatically operates on the data observed through cames, and does not require any parameters to be adjusted manually. Moreover, the only inputs that are required are the positions of the vehicles in the world coordinate frame. Furthermore, our model does not assume knowledge about driver-related factors and is designed to identify driving behavior similar to how humans intuitively infer driving behavior by simply “observing the vehicle trajectories”.

## II. RELATED WORK

We give a brief overview of prior work in classifying driver behaviors, intent prediction, and graph-based traffic networks.

### A. Driver Behaviors

At a broad level, prior approaches to analyze driver behaviors can be classified into three categories. The first category of methods classify driver behavior based on the characteristics of drivers such as age, gender, blood pressure, personality, occupation, hearing, etc. Dahlen et al. [17] explored driver personalities and how they connected to aggressive driving. Rong et al. [18] investigated the causes of tailgating and determined that indicative features include blood pressure, hearing, and driving experience. Social Psychology studies [19] have found that aggressiveness may also be correlated to the background of the driver, including age, gender, occupation, etc.

The second category of methods is based on environmental factors such as weather or traffic conditions [20], [21]. The study conducted in [21] was designed to investigate the

effects of weather-controlled speed limits and signs for slippery road conditions on driver behavior. Other studies [20] have correlated changes in traffic density with varying driver behavior.

The third category refers to psychological aspects that affect driving styles. Psychological aspects include drunk driving, driving under influence, and state of fatigue. It is shown [22] that driving under influence induces delayed responses in acceleration and deceleration. Jackson et al. [23] show that a state of fatigue manifests the same characteristics as driving under the influence, but without the effect of substance intoxication. Other techniques evaluate the impact of mobile phone operation on driver behaviors [24]. Our approach is complementary and can be combined with these methods.

### B. Behavior Prediction

Apart from behavior modeling, some methods have been proposed for behavior prediction and analysis of human pedestrians or traffic agents [25], [26]. Some recent approaches use reinforcement learning and imitation learning techniques to recognize various driver intentions [27], [1], while others have used RNN and LSTM-based networks for data-driven driver intent prediction [5].

### C. Graph-Based Traffic Networks

Traffic networks have been studied extensively to predict traffic flow [28] or traffic density [29], [30] at a macroscopic scale. These techniques have been used for Predicting traffic flow is important for applications such as congestion management and vehicle routing. Graphs have also been used for trajectory prediction [14], [31], [32]. Chandra et al. [13] model traffic entities using dynamic weighted graphs. They predict the driving behavior by training a neural network on the eigenvectors of the traffic-graphs. However, their approach requires a large amount of training data with behavior label annotations, which can be time-consuming and expensive to collect. Further, they assume that a driver’s behavior is constant and does not change with time. Our centrality-based approach is more general and overcomes these limitations.

## III. BACKGROUND AND OVERVIEW

In this section, we give an overview of traffic driving behaviors and centrality functions.

### A. Categorizing Driving Behavior

Several criteria have been proposed in psychology and robotics literature to characterize driving behaviors. These include explicit formulations or scales to measure aggressive driving. The Driving Anger Scale (DAS) [33] consists of 14 scenarios rated on a 5-point Likert scale (1 =not at all; 5 =very much) measuring the amount of anger experienced during an offensive situation (e.g., aggressive overtaking). The DAS assesses the propensity to become angry while driving and higher scores reflect greater driving anger. The DAS was extended to DAX (Driving Anger Expression) that identifies four ways people express their anger when driving, and they can be combined to form a Total Aggressive Expression Index.

More generally, other scales consider behaviors based on question-answer based analysis. For example, the Multidimensional Driving Style Inventory (MDSI) [34] developed a scale measuring eight factors, each one representing a specific driving style— dissociative, anxious, risky, angry, high-velocity, distress reduction, patient, and careful. Another scale, the Driving Behaviour Inventory (DBI), was developed to study dimensions of driver stress [35], including driving aggression, dislike of driving, tension and frustration, and irritation. Similarly, the Driving Style Questionnaire (DSQ) [36] is composed of six independent dimensions of driving style that are labeled— speed, calmness, social resistance, focus, planning, and deviance.

These scales measure global driving behaviors from different perspectives and therefore interpret the meaning of the behaviors in different ways. Consequently, behavior labels seem to represent somewhat different concepts that are hard to summarize. Therefore, we follow the classification principle [8], which proposes that each global behavior is a function of specific driving maneuvers or *styles*. This principle suggests that, rather than attempting to classify a vehicle’s behavior according to global labels that are interpreted differently by different scales, it is more useful to predict the specific styles that constitute the global behavior. We use these principles proposed by [8] in the following definition:

**Definition III.1.** *Driving behavior refers to the high-level global behavior, such as aggressive or conservative driving. Each global behavior consists of one or more underlying specific styles. For example, an aggressive driver (global behavior) may frequently overspeed or overtake (specific styles).*

We summarize the global behaviors and their constituent specific styles in Table I. However, while Definition III.1 suggests a definitive and uniform way to classifying driving behaviors, current scales or metrics in psychology or robotics for measuring driving behavior cannot classify the specific styles in Table I, or do not provide realtime performance. Our goal in this work is to develop a computational metric that measures the following specific styles— *Overtaking, overspeeding, sudden lane-changes, and weaving* from the trajectories. We state :

**Problem III.1.** *In a traffic video with  $N$  vehicles during any time-period  $\Delta t$ , given the spatial coordinates in the world coordinate frame of all vehicles, our overall objective is to classify the specific styles for all drivers during  $\Delta t$  based on the styles described in Table I.*

### B. Centrality Functions

Centrality functions [16] are real-valued functions that characterize the behavior of vertices in a graph. Such functions can be defined as  $\zeta : \mathcal{V} \rightarrow \mathbb{R}$ , where  $\mathcal{V}$  denotes the set of vertices and  $\mathbb{R}$  denotes a real number. The particular characteristic that is modeled through centrality functions depends on the type of graph. A few examples of centrality functions include closeness centrality, degree centrality, and eigenvector centrality. These functions have been used to identify influential personalities in social media networks, identify key infrastructure nodes on the internet, rank web-pages in search engines, and to discover the origin of epidemics. In our approach, we use the following characteristics of centrality functions.

**Definition III.2. Closeness Centrality:** *For a given vertex,  $u$ , in a connected DGG, the closeness centrality function,  $\zeta_c(u)$ , is defined as,*

$$\zeta_c(u) = \frac{N - 1}{\sum_{v \in \mathcal{V} \setminus \{u\}} d(u, v)}, \quad (1)$$

where  $v \in \mathcal{V} \setminus \{u\}$  denote all vertices in the connected DGG other than  $u$ .

The closeness centrality for a given vertex computes the reciprocal of the sum of the edge lengths of the shortest paths between the given vertex and all other vertices in the connected DGG. By definition, the higher the closeness centrality value, the more centrally the vertex is placed.

**Definition III.3. Degree Centrality:** *For a given vertex,  $u$ , in a connected DGG, the degree centrality function is defined as,*

$$\zeta_d(u) = |\{L(u, v) \in L_{u,:} | \nu_u \geq \nu_v\}|, \quad (2)$$

where  $|\cdot|$  computes the number of non-zero entries of a vector or matrix and  $L_{u,:}$  denotes the  $u^{\text{th}}$  row of  $L$ .

The degree centrality computes the number of edges between the given vertex and connected vertices in the graph.

We exploit the analytical properties, including the derivatives and extreme values of these centrality functions, to develop our new metric in the following section. This metric computes the likelihood and intensity of different driving styles.

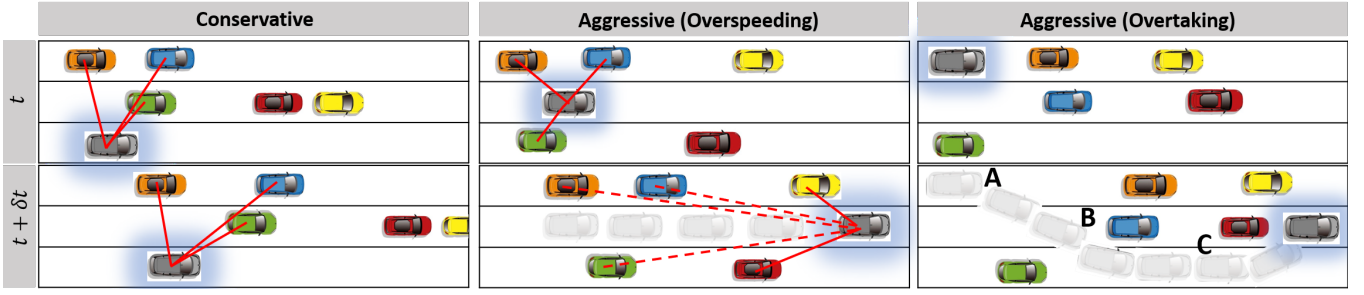
## IV. DRIVING BEHAVIOR CLASSIFICATION USING CMETRIC

We use centrality functions to develop a novel metric which answers two questions,

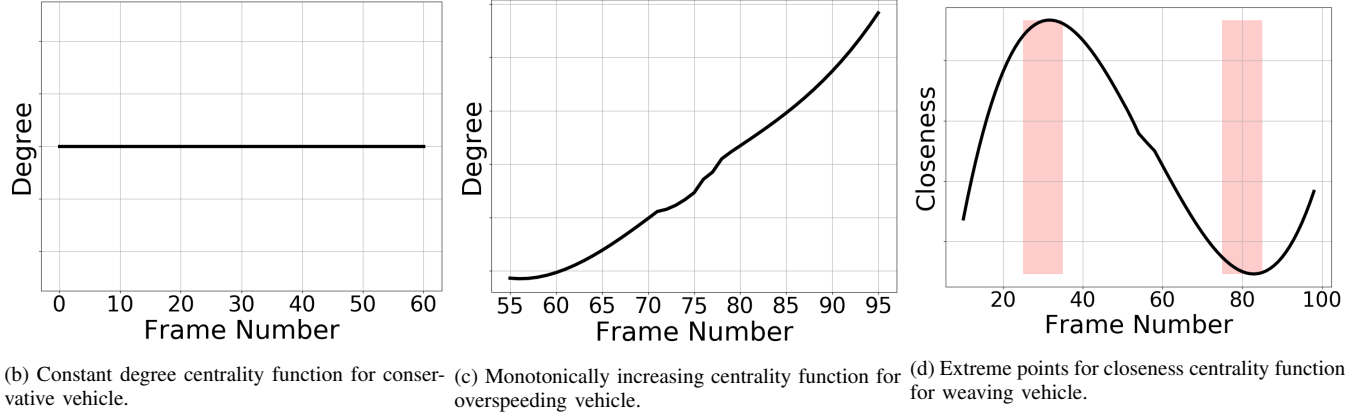
- (Likelihood) At time  $t$ , how likely is it that a vehicle executes a specific style?
- (Intensity) At time  $t$ , what is the intensity with which a vehicle executes a specific style?

The *specific styles* can then be used to assign *global behaviors* according to Table I. Our model (Figure 2) is a computational model for behavior classification and consists of the following steps:

- 1) Obtain the positions of all vehicles using sensors deployed on the autonomous vehicle and form Dynamic Geometric Graphs (Section IV-A).
- 2) Compute the closeness and degree centrality function values using the definitions in Section III-B.



(a) In all three scenarios, the ego-vehicle is a gray vehicle marked with blue glow outline. (left) A conservative vehicle, (middle) overspeeding vehicle in the same lane, and (right) weaving and overtaking vehicle.



(b) Constant degree centrality function for conservative vehicle. (c) Monotonically increasing centrality function for overspeeding vehicle. (d) Extreme points for closeness centrality function for weaving vehicle.

**Fig. 3: Measuring the Likelihood of Specific Styles with CMetric:** CMetric measures (degree and closeness centrality) the likelihood of the specific style of the ego-vehicle (grey with a blue glow) by computing the magnitude of the derivative of the centrality functions as well as the functions' extreme points. In Figure 3b, the derivative of the degree centrality is 0 because the ego-vehicle does not observe any additional new neighbors (See Section IV-B) so the degree centrality is a constant function; therefore the vehicle is conservative. In Figure 3c, the vehicle overspeeds, and consequently, the rate of observing new neighbors is high, which is reflected in the magnitude of the derivative of the degree centrality being positive. Finally, in Figure 3d, the ego-vehicle demonstrates overtaking/sudden lane-changes and weaves through traffic. This is reflected in the magnitude of the slope and the location of extreme points, respectively, of the closeness centrality function. Our approach can classify these behaviors based on CMetric.

- 3) Use the CMetric value to measure the likelihood and intensity of specific driving styles listed in Table I (Section IV-B).

We explain these steps in detail in the remainder of this section.

#### A. Traffic Representation Using Dynamic Geometric Graphs (DGGs)

In this section, we describe the representation of traffic through Dynamic Geometric Graphs (DGGs). We assume that the trajectories of all the vehicles in the video are extracted and given to our algorithm as an input. Given this input, we first construct a DGG [37] at each time-step. We define a dynamic geometric graph as follows:

**Definition IV.1.** A Geometric Graph is an undirected graph with a set of vertices  $\mathcal{V}$  and a set of edges  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  defined in the 2-D Euclidean metric space with metric function  $f(x, y) = \|x - y\|^2$ . Two vertices  $v_i, v_j \in \mathcal{V}$  are connected if, and only if, their  $f(v_i, v_j) < r$  for some constant  $r$ .

A Dynamic Geometric Graph (DGG) is a geometric graph with a set of vertices  $\mathcal{V}(t)$  and a set of edges  $\mathcal{E}(t)$ , where  $\mathcal{V}(t)$  and  $\mathcal{E}(t)$  are the sets of vertices and edges as functions of time.

We represent traffic at each time instance with  $N$  road-agents using a DGG, where the positions of vehicles, including motorbikes and scooters, represent the vertices. In particular, we represent a vehicle position as a point in  $\mathbb{R}^2$ .

Thus,  $v_i \leftarrow [x_i, y_i]^\top$ , where  $[x_i, y_i]^\top$  is the 2-D spatial coordinates (e.g., in meters) of the  $i^{\text{th}}$  vehicle in the global coordinate frame. For a DGG,  $\mathcal{G}$ , the adjacency matrix,  $A \in \mathbb{R}^{N \times N}$  is given by,

$$A(i, j) = \begin{cases} d(v_i, v_j) & \text{if } d(v_i, v_j) < \mu, i \neq j, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

where  $d(v_i, v_j)$  denotes the Euclidean distance between the  $i^{\text{th}}$  and  $j^{\text{th}}$  vehicles, and  $\mu$  is a distance threshold parameter. We discuss its implementation in Section V.

For an adjacency matrix  $A$  at each time instance, the corresponding degree matrix  $D \in \mathbb{R}^{N \times N}$  is defined as a diagonal matrix with main diagonal  $D(i, i) = \sum_{j=1}^N A(i, j)$  and 0 otherwise. Further, the symmetric Laplacian matrix can be obtained by subtracting  $A$  from  $D$ ,

$$L(i, j) = \begin{cases} D(i, i) & \text{if } i = j, \\ -e^{-d(v_i, v_j)} & \text{if } d(v_i, v_j) < \mu, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The Laplacian matrix for each time-step is correlated with the Laplacian matrices for all previous time-steps. Let the Laplacian matrix at a time instance  $t$  be denoted as  $L_t$ . Then, the laplacian matrix for the next time-step,  $L_{t+1}$  is given by the following update,

$$L_{t+1} = \begin{bmatrix} L_t & 0 \\ 0 & 1 \end{bmatrix} + \delta\delta^\top, \quad (5)$$

where  $\delta \in \mathbb{R}^{d \times 2}$  is a sparse matrix with  $\|\delta\|_0 \ll n$ . The update rule in Equation 5 enforces a vehicle to add edges connections to new vehicles while retaining edges with previously seen vehicles. The presence of a non-zero value in the  $j^{\text{th}}$  row of  $\delta$  indicates that the  $j^{\text{th}}$  road-agent has formed an edge connection with a new vehicle, that has been added to the current DGG. The size of  $L_t$  is fixed for all time  $t$  and is initialized as a zero matrix of size  $N \times N$ , where  $N$  is max number of agents.  $L_t$  is updated in-place with time and is reset to a zero matrix once the number of vehicles crosses  $N$ . The diagonal elements of the Laplacian matrix in Equation 5 is used by the degree centrality to characterize overspeeding IV-B.

---

**Algorithm 1:** CMetric Measure outputs the Style Likelihood Estimate (SLE) and Style Intensity Estimate (SIE) for a vehicle,  $u$ , in a given time-period  $\Delta t$ .

---

**Input :**  $u = v_i \leftarrow [x_i, y_i]^\top \forall v_i \in \mathcal{V}(t)$

**Output:** SLE $_k(t)$ , SIE $_k(t)$

```

1  $t = 0$ 
2 for each  $u \in \mathcal{V}(t)$  do
3   while  $t \leq T$  do
4     // Compute Centrality //
5      $\zeta_c(u) = \frac{N-1}{\sum_{v \in \mathcal{V} \setminus \{u\}} d(u,v)}$ 
6      $\zeta_d(u) = |\{L(u,v) \in L_{u,:} | \nu_u \geq \nu_v\}|$ 
7      $t \leftarrow t + 1$ 
8   end
9   // Compute CMetric //
10  Form  $\mathcal{M}_{\Delta t}(u)$  using Definition IV.2
11  for  $k = 0, 1$  do
12    // Compute Likelihood and Intensity //
13    SLE $_k(t) = \left| \frac{\partial \mathcal{M}_{\Delta t}(u)[k,:]}{\partial t} \right|$ 
14    SIE $_k(t) = \left| \frac{\partial^2 \mathcal{M}_{\Delta t}(u)[k,:]}{\partial t^2} \right|$ 
15  end
16 end

```

---

### B. CMetric Measure

In a given time-period  $\Delta t$ , we use the notion of CMetric to measure various driving styles. The CMetric is defined as:

**Definition IV.2. CMetric ( $\mathcal{M}_{\Delta t}(u)$ ):** Given a time period  $\Delta t$ , the CMetric value for a particular vehicle,  $u$ , is a matrix

$$\mathcal{M}_{\Delta t}(u) \in \mathbb{R}^{3 \times \Delta t}, \text{ where } \mathcal{M}_{\Delta t}(u) = \begin{bmatrix} \zeta_c(u) \\ \zeta_d(u) \end{bmatrix}.$$

where  $\zeta_c(u)$  and  $\zeta_d(u)$  are the closeness and degree centrality functions defined in Section III-B. We use the CMetric to measure the Style Likelihood Estimate (SLE) and Style Intensity Estimate (SIE) of driving behavior styles.

1) The Style Likelihood Estimate (SLE) of a specific driving style is the probability of its occurrence and

is measured by computing the magnitude of the row derivatives of  $\mathcal{M}_{\Delta t}(u)$  with respect to time. Higher magnitudes of the row derivatives and the existence of local extreme points indicate higher likelihood. The SLE formula is given by,

$$\text{SLE}_k(t) = \left| \frac{\partial \mathcal{M}_{\Delta t}(u)[k,:]}{\partial t} \right| \quad (6)$$

2) The Style Intensity Estimate (SIE) of a specific driving style is the severity with which the style is executed by a vehicle. Specific styles executed over shorter time-frames are considered more intense than styles executed over longer time-frames. SIE is computed by taking the second row derivatives of  $\mathcal{M}_{\Delta t}(u)$  with respect to time and measuring the  $\varepsilon$ -sharpness of local extreme points. The SIE formula is given by,

$$\text{SIE}_k(t) = \left| \frac{\partial^2 \mathcal{M}_{\Delta t}(u)[k,:]}{\partial t^2} \right| \quad (7)$$

The row index  $k$  indicates the centrality function that is used.  $k = 0, 1$  corresponds to the closeness and degree centrality, respectively. We can compute the time  $t_{\text{SLE}}$  of maximum likelihood using

$$t_{\text{SLE}} = \arg \min_{t \in \Delta t} \text{SLE}_k(t)$$

1) *Overtaking/Sudden Lane-Changes:* Overtaking is the act of one vehicle going past another vehicle, traveling in the same or adjacent lane, in the same direction. From Definition 1 in Section III-B, the value of the closeness centrality increases a vehicle moves towards the center and decreases as it moves away from the center. Therefore, the SLE of overtaking can be computed by measuring the rate of change of the closeness centrality. The maximum likelihood SLE $_{\text{max}}$  is,

$$\text{SLE}_{0,\text{max}} = \max_{t \in \Delta t} \text{SLE}_0(t).$$

The SIE of overtaking is computed by simply measuring SIE $_0$ .

2) *Weaving:* Weaving is the act of a vehicle shifting its position from a side lane towards the center, and vice-versa [38]. In such a scenario, the closeness centrality function values oscillates between low values on the side lanes and high values towards the center. Mathematically, these oscillations in the closeness centrality values can be detected by finding the extreme values (points at which function has a local minimum or maximum) of the closeness centrality function. Mathematically, the extreme values, or points of local maximum or minimum can be found at those time instances when SLE $_0(t) = 0$ . To differentiate from constant functions, we impose the condition that the  $\varepsilon$ -sharpness [39] of the closeness centrality be non-zero:

$$\max_{t \in \mathcal{B}_\varepsilon(t^*)} \text{SLE}_0(t) - \text{SLE}_0^* \neq 0,$$

where  $\mathcal{B}_\varepsilon(y) \in \mathbb{R}^d$  is the unit ball centered around  $y$  WHAT IS Y with radius  $\varepsilon$  in  $d$  dimensions, and  $t^* = \arg \min_t \text{SLE}_0$ . The SIE $_0(t)$  is computed by measuring the sharpness of the local minimum or maximum which is expressed using the  $\varepsilon$ -sharpness value.

TABLE I: Definition and categorization of driving behaviors [8] used by our algorithm. Our approach models the specific styles through various centrality functions highlighted on the right. We present a novel mapping between different styles/behaviors with our CMetric measures.

Global Behaviors	Specific Styles	Centrality	Our CMetric Measures	
			Style Intensity Estimate	Style Likelihood Estimate
Aggressive	Overspeeding	Degree ( $\zeta_d$ )	Magnitude of 2 <sup>nd</sup> Derivative	Magnitude of Derivative
	Overtaking / Sudden Lane-Change	Closeness ( $\zeta_c$ )	Magnitude of 2 <sup>nd</sup> Derivative	Magnitude of Derivative
	Weaving	Closeness ( $\zeta_c$ )	$\varepsilon$ -sharpness of Local Extreme Points	Local Extreme Points
Conservative	Driving Slowly or uniformly	Degree ( $\zeta_d$ )	Magnitude of 2 <sup>nd</sup> Derivative	Magnitude of Derivative
	No Lane-change	Closeness ( $\zeta_c$ )	Magnitude of 2 <sup>nd</sup> Derivative	Magnitude of Derivative

3) *Overspeeding*: We use the degree centrality to classify overspeeding. The degree of  $i^{\text{th}}$  vehicle, ( $\theta_i \leq n$ ), at time  $t$ , can be computed from the diagonal elements of the Laplacian matrix  $L_t$  (Section IV-A). As  $L_t$  is formed by adding rows and columns to  $L_{t-1}$ , the degree of each vehicle monotonically increases. Intuitively, an aggressively overspeeding vehicle will observe new neighbors (increasing degree) at a higher rate than neutral or conservative vehicles. Therefore, the likelihood of overspeeding can be measured by computing  $\text{SLE}_1(t)$ . Similar to overtaking, the maximum likelihood estimate is given by

$$\text{SLE}_{1,\max} = \max_{t \in \Delta t} \text{SLE}_1(t)$$

Figures 3c and 3b visualizes how the degree centrality can distinguish between an overspeeding vehicle and a vehicle driving at a uniform speed.

4) *Conservative Vehicles*: Conservative vehicles, on the other hand, conform to a single lane [40] as much as possible, and drive at a uniform speed [8], typically at or below the speed limit. The values of the closeness and degree centrality functions in the case of conservative vehicles thus, remain constant. More formally, the likelihood that a vehicle stays in a single lane during time-period  $\Delta t$  is higher when,

$$\text{SLE}_0(t) = 0 \text{ and } \max_{t \in \mathcal{B}_\varepsilon(t^*)} \text{SLE}_0(t) - \text{SLE}_0^* = 0,$$

and the likelihood that a vehicle will prefer to drive at uniform speed is higher when,

$$\text{SLE}_1(t) = 0 \text{ and } \max_{t \in \mathcal{B}_\varepsilon(t^*)} \text{SLE}_1(t) - \text{SLE}_1^* = 0.$$

The more conservative a driver is, the higher their intensity value and can be similarly measured by computing the  $\text{SIE}_k(t)$ ,  $k = 0, 1$ .

In summary, the specific driving styles listed in Table I can be characterized by CMetric by exploiting its analytical properties, including the derivative, second derivative, extreme values, and the notion of  $\varepsilon$ -sharpness of the local minimum (or maximum) regions.

## V. IMPLEMENTATION AND EVALUATION

We first highlight the datasets used in our work, followed by the evaluation protocol. We report the results for metrics in Section V-C. We conclude the section with an analysis of the CMetric performance.

### A. Dataset

One of the main issues in driving behavior research is the availability of large-scale open-source datasets for driving behaviors that specifically contain labels for aggressive

TABLE II: We report the Time Deviation Error (TDE) (in # frames) for the following driving styles: Overspeeding (OS), Overtaking (OT), Sudden Lane-Changes (SLC), and Weaving (W). The TDE indicates the absolute difference between the times taken by a human and our proposed CMetric to identify a driving style. Lower is better. – indicates that the particular style was not observed in the ground truth.

Dataset	Styles			
	OS	OT	SLC	W
Argoverse [11]	0.25s	–	0.23s	–
SG	0.54s	0.875s	1.21s	1.28s

and conservative vehicles. In light of these limitations, we create such a dataset consisting of 11 traffic videos called SG dataset, by modifying existing large-scale autonomous driving datasets originally intended for trajectory prediction and tracking [11], [12], [41]. Our new dataset consists of videos of dense urban traffic with annotated trajectories from different geographic regions. The behavior labels are labeled and verified via crowd-sourced annotators. We show samples of the dataset in our supplementary video. We also present an analysis for the CMetric in Section V-D. Additionally, we also test CMetric on the Argoverse dataset.

### B. Evaluation Protocol

The goal of CMetric is to observe vehicles in realtime traffic and classify their driving styles at the same time that a human driver would take to perform the same task, and with similar accuracy. To obtain ground truth, we design a user study where we asked ten participants to mark the time-periods in the videos where the vehicle is perceived to be exhibiting a particular style. We then evaluate the accuracy of our metric by measuring the overlap between the time-period marked by the participants and the time-period suggested by CMetric.

More precisely, we first compute a probability mass function  $\mathcal{P}(T)$ , where  $T$  is a random variable that denotes time.  $\mathcal{P}(T = t)$  denotes the probability that a vehicle is exhibiting one of the four behavioral styles at time  $t$ . We report the Time Deviation Error (TDE) that is computed as,

$$\text{TDE}_{\text{style}} = \left| \frac{t_{\text{SLE}} - \mathbb{E}[T]}{f} \right| \quad (8)$$

where  $f$  is the frame rate of the video.  $f = 2$  Hz for the SG dataset and  $f = 10$  Hz for the Argoverse dataset. In other words, the  $\text{TDE}_{\text{style}}$  computes the difference between the mean frame in the user study and the frame with a maximum likelihood for the driving style.

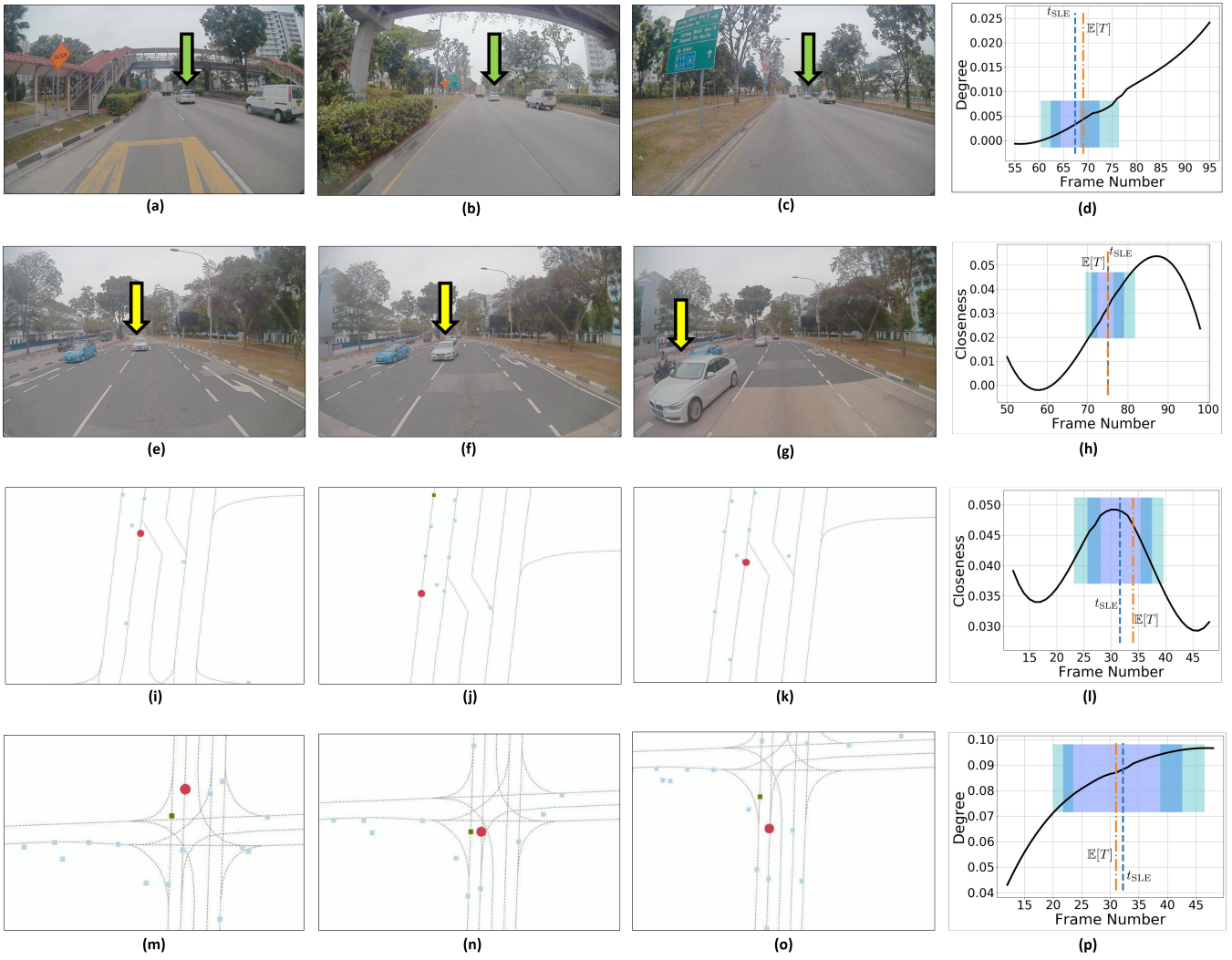


Fig. 4: **Qualitative Analysis:** First and second rows are videos from the SG dataset, while the third row contains a video from the Argoverse dataset. Figures (a), (b), and (c) show the trajectory of an overspeeding vehicle during the 60<sup>th</sup>, 64<sup>th</sup>, and 69<sup>th</sup> frames. In (d), we show a monotonically rising curve for the degree centrality with  $SLE_1(60) < SLE_1(64) < SLE_1(69)$  indicating overspeeding. The shaded colored regions overlaid on each graph are color heat maps that correspond to  $\mathcal{P}(T)$ . Similarly, Figures (e), (f), and (g) show the trajectory of an overtaking vehicle during the 73<sup>rd</sup>, 75<sup>th</sup>, and 78<sup>th</sup> frames. We show a monotonically rising curve for the closeness function between the 70<sup>th</sup>, and 80<sup>th</sup> frames. We show corresponding results for the Argoverse dataset. The third row shows a lane-changing vehicle, while the fourth row demonstrates an overspeeding vehicle.

### C. Results

We report the Time Deviation Error (TDE), in seconds(s), for the following driving styles: Overspeeding (OS), Overtaking (OT), Sudden Lane-Changes (SLC), and Weaving (W), respectively, in Table II. We report the average TDE observed for each style across all videos in a dataset. Additionally, we also compute the accuracy by checking if  $t_{SLE}$  belongs to the range marked by the participants, *i.e.*, we consider our estimation to have accuracy 1 if it lies within the participant-makred range (including the boundaries) and 0 otherwise. For example, on the SG dataset, CMetric lies outside the range for only one out of the 11 traffic videos. Therefore, we count this as a “failed” result and measure the mean accuracy on the SG dataset to be  $10/11 = 0.91$ .

### D. Analysis of CMetric

In Figure 4, we show two sequences from the SG dataset where we use CMetric to characterize overspeeding and overtaking using the degree and closeness centrality functions,

respectively. We also show one sequence from the Argoverse dataset where CMetric characterizes sudden lane-change through the closeness centrality. Figures (a), (b), and (c) show the trajectory of an overspeeding vehicle during the 60<sup>th</sup>, 64<sup>th</sup>, and 69<sup>th</sup> frames. In Figure (d), we show a monotonically rising curve for the degree function. From Section IV-B, the likelihood of overspeeding increases with increasing values of  $SLE_1(t)$ .  $SLE_1(60) < SLE_1(64) < SLE_1(69)$  verifies the CMetric measure for overspeeding. Similarly, Figures (e), (f), and (g) show the trajectory of an overtaking vehicle during the 73<sup>rd</sup>, 75<sup>th</sup>, and 78<sup>th</sup> frames. From Section IV-B, the likelihood of overtaking increases with increasing values of  $SLE_0(t)$ . In Figure (h), we show a monotonically rising curve for the closeness function between the 70<sup>th</sup>, and 80<sup>th</sup> frames. In fact, we also observe extreme values of the closeness centrality function at the 57<sup>th</sup>, and 89<sup>th</sup> frames. The CMetric measure for weaving (Section IV-B) indicates that the vehicle is also weaving through traffic at these time-instances.

## VI. CONCLUSION, LIMITATIONS, AND FUTURE WORK

We present a new measure (CMetric) to classify driver behaviors using centrality functions. CMetric computes the likelihood of a vehicle executing a driving style along with the intensity used to execute the style. Our approach is designed for realtime autonomous driving applications, where the trajectory of each vehicle or road-agent is extracted from a video. We represent traffic using dynamic geometric graph (DGG) and centrality functions, corresponding to closeness and degree, that are used to compute the CMetric.

Our work has some limitations. Currently, our CMetric measure only works for videos that contain more than one vehicle (other than the ego-vehicle). Additionally, we are currently limited to four driving styles, and we plan to investigate additional driving styles such as tailgating. As part of future work, we plan to use the CMetric measure to classify behaviors for improving realtime navigation and trajectory prediction. We would like to evaluate the performance in scenarios with VRUs (vulnerable road users).

## REFERENCES

- [1] F. Codevilla, M. Müller, A. López, V. Koltun, and A. Dosovitskiy, "End-to-end driving via conditional imitation learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–9.
- [2] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and decision-making for autonomous vehicles," *Annual Review of Control, Robotics, and Autonomous Systems*, 2018.
- [3] B. Paden, M. Čáp, S. Z. Yong, D. Yershov, and E. Frazzoli, "A survey of motion planning and control techniques for self-driving urban vehicles," *IEEE Transactions on intelligent vehicles*, vol. 1, no. 1, pp. 33–55, 2016.
- [4] L. Sun, W. Zhan, C.-Y. Chan, and M. Tomizuka, "Behavior planning of autonomous cars with social perception," *arXiv preprint arXiv:1905.00988*, 2019.
- [5] A. Zyner, S. Worrall, and E. Nebot, "A recurrent neural network solution for predicting driver intention at unsignalized intersections," *IEEE RAL*, vol. 3, no. 3, pp. 1759–1764, 2018.
- [6] A. Aljaafreh, N. Alshabat, and M. S. N. Al-Din, "Driving style recognition using fuzzy logic," in *2012 IEEE International Conference on Vehicular Electronics and Safety (ICVES 2012)*. IEEE, 2012.
- [7] W. Wang, J. Xi, A. Chong, and L. Li, "Driving style classification using a semisupervised support vector machine," *IEEE Transactions on Human-Machine Systems*, vol. 47, pp. 650–660, 2017.
- [8] F. Sagberg, Selpi, G. F. Bianchi Piccinini, and J. Engström, "A review of research on driving styles and road safety," *Human factors*, vol. 57, no. 7, pp. 1248–1275, 2015.
- [9] V. Ramanishka, Y.-T. Chen, T. Misu, and K. Saenko, "Toward driving scene understanding: A dataset for learning driver behavior and causal reasoning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7699–7707.
- [10] E. Cheung, A. Bera, E. Kubin, K. Gray, and D. Manocha, "Identifying driver behaviors using trajectory features for vehicle navigation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 3445–3452.
- [11] M.-F. Chang, J. W. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, and J. Hays, "Argoverse: 3d tracking and forecasting with rich maps," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.
- [12] R. Kesten, M. Usman, J. Houston, T. Pandya, K. Nadhamuni, A. Ferreira, M. Yuan, B. Low, A. Jain, P. Ondruska, S. Omari, S. Shah, A. Kulkarni, A. Kazakova, C. Tao, L. Platinsky, W. Jiang, and V. Shet, "Lyft level 5 av dataset 2019," <https://level5.lyft.com/dataset/>, 2019.
- [13] R. Chandra, U. Bhattacharya, T. Mittal, X. Li, A. Bera, and D. Manocha, "Graphrqi: Classifying driver behaviors using graph spectrums," *arXiv preprint arXiv:1910.00049*. To appear in *Proceedings of ICRA*, 2019.
- [14] R. Chandra, T. Guan, S. Panuganti, T. Mittal, U. Bhattacharya, A. Bera, and D. Manocha, "Forecasting trajectory and behavior of road-agents using spectral clustering in graph-lstms," *arXiv preprint arXiv:1912.01118*, 2019.
- [15] Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, and D. Rus, "Social behavior for autonomous vehicles," *Proceedings of the National Academy of Sciences*, vol. 116, no. 50, pp. 24972–24978, 2019.
- [16] F. A. Rodrigues, "Network centrality: An introduction," *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*, p. 177, 2019.
- [17] E. R. Dahlen, B. D. Edwards, T. Tubré, M. J. Zyphur, and C. R. Warren, "Taking a look behind the wheel: An investigation into the personality predictors of aggressive driving," *Accident Analysis & Prevention*, vol. 45, pp. 1–9, 2012.
- [18] J. Rong, K. Mao, and J. Ma, "Effects of individual differences on driving behavior and traffic flow characteristics," *Transportation research record*, vol. 2248, no. 1, pp. 1–9, 2011.
- [19] K. H. Beck, B. Ali, and S. B. Daughters, "Distress tolerance as a predictor of risky and aggressive driving," *Traffic injury prevention*, vol. 15 4, pp. 349–54, 2014.
- [20] A. H. Jamson, N. Merat, O. M. Carsten, and F. C. Lai, "Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions," *Transportation research part C: emerging technologies*, vol. 30, pp. 116–125, 2013.
- [21] P. Rämä, "Effects of weather-controlled variable speed limits and warning signs on driver behavior," *Transportation Research Record*, vol. 1689, no. 1, pp. 53–59, 1999.
- [22] J. Dai, J. Teng, X. Bai, Z. Shen, and D. Xuan, "Mobile phone based drunk driving detection," in *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*. IEEE, 2010, pp. 1–8.
- [23] P. Jackson, C. Hilditch, A. Holmes, N. Reed, N. Merat, and L. Smith, "Fatigue and road safety: a critical analysis of recent evidence," *Department for Transport, Road Safety Web Publication*, vol. 21, 2011.
- [24] M. Saifuzzaman, M. M. Haque, Z. Zheng, and S. Washington, "Impact of mobile phone use on car-following behaviour of young drivers," *Accident Analysis & Prevention*, 2015.
- [25] A. Bera, S. Kim, T. Randhavane, S. Pratapa, and D. Manocha, "Glmp-realtime pedestrian path prediction using global and local movement patterns," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 5528–5535.
- [26] S. J. Guy, J. Van Den Berg, W. Liu, R. Lau, M. C. Lin, and D. Manocha, "A statistical similarity measure for aggregate crowd dynamics," *ACM Transactions on Graphics (TOG)*, 2012.
- [27] S. Qi and S.-C. Zhu, "Intent-aware multi-agent reinforcement learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 7533–7540.
- [28] M. Castro-Neto, Y.-S. Jeong, M.-K. Jeong, and L. D. Han, "Online-svr for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Systems with Applications*, vol. 36, no. 3, Part 2, pp. 6164 – 6173, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0957417408004740>
- [29] A. Ermagun and D. Levinson, "Spatiotemporal traffic forecasting: review and proposed directions," *Transport Reviews*, 2018.
- [30] X. Cheng, R. Zhang, J. Zhou, and W. Xu, "Deeptransport: Learning spatial-temporal dependency for traffic condition forecasting," in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–8.
- [31] X. Li, X. Ying, and M. C. Chuah, "Grip: Graph-based interaction-aware trajectory prediction," *arXiv preprint arXiv:1907.07792*, 2019.
- [32] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," *ArXiv*, vol. abs/1709.04875, 2018.
- [33] J. L. Deffenbacher, E. R. Oetting, and R. S. Lynch, "Development of a driving anger scale," *Psychological reports*, 1994.
- [34] O. Taubman-Ben-Ari, M. Mikulincer, and O. Gillath, "The multidimensional driving style inventory—scale construct and validation," *Accident Analysis & Prevention*, vol. 36, no. 3, pp. 323–332, 2004.
- [35] E. Gulian, G. Matthews, A. I. Glendon, D. Davies, and L. Debney, "Dimensions of driver stress," *Ergonomics*, 1989.
- [36] D. J. French, R. J. West, J. Elander, and J. M. WILDING, "Decision-making style, driving style, and self-reported involvement in road traffic accidents," *Ergonomics*, vol. 36, no. 6, pp. 627–644, 1993.
- [37] B. M. Waxman, "Routing of multipoint connections," *IEEE journal on selected areas in communications*, 1988.
- [38] H. Farah, S. Bekhor, A. Polus, and T. Toledo, "A passing gap acceptance model for two-lane rural highways," *Transportmetrica*, vol. 5, no. 3, pp. 159–172, 2009.
- [39] L. Dinh, R. Pascanu, S. Bengio, and Y. Bengio, "Sharp minima can generalize for deep nets," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 2017, pp. 1019–1028.
- [40] K. I. Ahmed, "Modeling drivers' acceleration and lane changing behavior," Ph.D. dissertation, MIT, 1999.
- [41] Y. Ma, X. Zhu, S. Zhang, R. Yang, W. Wang, and D. Manocha, "Trafficpredict: Trajectory prediction for heterogeneous traffic-agents," *arXiv preprint arXiv:1811.02146*, 2018.