Estimating the effects of the Macroscopic Traffic Parameters on the overall Cooperative Awareness Message generations

Franka van Jaarsveld
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
f.vanjaarsveld@student.utwente.nl

ABSTRACT

Cooperative Awareness Messages (CAMs) are generated by Intelligent Transport Systems when the difference in their speed, direction, or position exceeds a given threshold compared to the previously generated CAM. The frequency of CAM generations is not set, and the channel used in vehicular communication has a limited spectrum. To understand the impact of these messages on the vehicular communication channel, it is important to have an estimate of the generated CAMs in different traffic scenarios. It is relatively easy to deduct traffic (macroscopic) parameters from such scenarios, but CAM generations cannot easily be predicted using this data, as they are generated not by macroscopic but microscopic (vehicle) parameters. This research aims to develop a tool that will analyse a data set of microscopic parameters and create an estimate of the subsequent CAM generations, as well as the corresponding macroscopic parameters, and analyse the resulting data for the influence of the macroscopic parameters on CAM generation.

Keywords

Intelligent Transport Systems, Cooperative Awareness Messages, Macroscopic Traffic Parameters, Network Congestion

1. INTRODUCTION

Autonomous vehicles (AVs), also known as Intelligent Transport Systems (ITS-es), have been the subject of science-fiction as far back as the 1930s [7] and understandably so. Human errors "have been responsible for 90% of road fatalities, such as speeding, alcohol impairment, distractions, and induced fear" [4]. Self-driving vehicles would thus significantly increase road safety. An important part of the development of this technology is the communication between AVs. One of the ways in which autonomous vehicles communicate amongst each other is through Cooperative Awareness Messages, otherwise referred to as CAMs. These messages have everything to do with microscopic parameters, also known as vehicle parameters. These are variables concerning the behaviour of individual vehicles, such as their speed, direction, and acceleration. The change in the value of these parameters determines when a new CAM is broadcast, and these same parameters make up the contents of the broadcast message.

33rd *Twente Student Conference on IT* July 3rd, 2020, Enschede, The Netherlands.

A vehicular communication standard has already been defined for current-day ITS implementations (IEEE 802.11p), but it is unknown whether the vehicular communication channel is capable of handling the load which will be placed upon it by CAMs, as the available bandwidth is limited and the frequency of CAM generations is variable.

Analysing a traffic scenario in terms of the microscopic parameters of the vehicles in such a scenario would lend itself to an estimation of the generated CAMs. However, it is much easier to analyse traffic scenarios in terms of macroscopic parameters such as traffic density, flow, and average speed. Unfortunately, not much is known about the relationship between macroscopic parameters and CAM generations. This research thus aims to create a tool that will process microscopic data and calculate the corresponding macroscopic parameters and CAM generations. This tool will then be applied to an existing microscopic data set to analyse the effect of macroscopic parameters on CAM generations. The possibility to estimate CAM generations in different traffic scenarios may facilitate further research into CAM generations and the capability of the ITS communication network to handle the requirements of future C-ITS development towards autonomous vehicles.

This paper aims to describe the development of the tool and analyse the generated output data. First, some background information is given on Intelligent Transport Systems and Cooperative Awareness Messages, after which the problem that is the focal point of this research is explained. After discussing some related works, the research goal and question is defined. Section 5 then describes the development of the tool in detail. The results of this tool are then presented and analysed, after which the research is concluded and future work is considered.

2. BACKGROUND

In order to clearly define the goal of this research and approach the development of the script as described in the methodology, some background must be given on Intelligent Transport Systems and Cooperative Awareness Messages.

2.1 Intelligent Transport Systems

Intelligent Transport Systems are already a part of modern-day life and can be seen in vehicles in the form of emergency vehicle and traffic jam warnings. The final goal of the development of ITS is fully autonomous vehicles, which has gained quite some media attention, as it could provide many improvements to everyday life. Roads would be safer as autonomous vehicles reduce human errors, driving would be more comfortable and travel time would decrease. Additionally, there could also be environmental benefits, as fuel consumption and emissions are decreased and traffic efficiency increased. However, there is much to be done to reach this final goal. To further the progress towards autonomous vehicles, vehicles must become aware of one another on the road, and communicate with each other. This is where Cooperative Intelligent Transport Systems come into the picture.

2.2 Cooperative Awareness Messages

To enable cooperative awareness within ITS, the European Telecommunication Standard Institute (ETSI) delivered the EN 302 637-2 standard defining Cooperative Awareness Messages (CAMs) [5]. These messages contain data on the position, direction, and speed of a vehicle and are periodically broadcast over the vehicular ad-hoc network for all nearby vehicles to receive. These variables are also called microscopic traffic parameters, as they reflect the movement of an individual car (unlike macroscopic traffic parameters, which reflect the average state of traffic flow [2]). A new CAM is generated when one of the following holds [5]:

- 1000ms has passed since the previous CAM was generated
- At least 100ms has passed since the previous CAM was generated, and at least one of the microscopic parameters has changed enough compared to the contents of the previous CAM:
  - 4 meters of displacement
  - 4 degrees of directional change
  - 0.5m/s change in speed

As a result, the frequency of CAM generations is not fixed, with fewer messages being generated when drivers behave predictably (drive slowly, at a constant speed and in a straight line), and more when they accelerate, decelerate, turn, or drive at high speed [5].

2.3 Problem statement

There already exists a standard for the communication between vehicles, IEEE 802.11p. This standard fits the requirements for the current implementation of ITS communications, but one of the challenges facing the future development of C-ITS is network congestion, as data generated by participating vehicles can be severely high [1]. Not only is the number of participating vehicles unknown and unrestricted, there also exist traffic scenarios where the number of generated CAMs might be very high, leading to network congestion. However, it is not certain what exact traffic scenarios will result in this high generation of CAMs. Furthermore, because of the varying frequency of CAM generations, their impact on the network is difficult to predict. Having said all that, knowing and understanding the relationship between traffic scenarios and CAM generations is needed in order to understand the impact of traffic scenarios on network congestion. Microscopic parameters may be easy to analyse in traffic scenarios with very few vehicles, but these situations are not likely to cause network congestion due to CAM generations. It is thus important to analyse busy traffic scenarios, and traffic parameters are much easier to analyse than vehicle parameters when viewing such a scenario. For example, when viewing a busy intersection, it is easier to analyse the traffic flow (the number of cars passing the intersection during a unit of time) or traffic density (the number of vehicles per distance unit) than the individual speed, acceleration, and direction of each vehicle.

3. RELATED WORK

Campolo et al. [1] analysed the evolution of C-ITS vehicular networks. Different phases in this evolution are discussed, as well as the challenges that lie ahead. One of the challenges discussed in this paper is the limited bandwidth and the possibility of network congestion caused by CAM generations.

Hoogendoorn and Knoop [2] give a formal definition of traffic flow and traffic density, as well as defining the mathematical functions (4) and (5), which show the relationship between macroscopic and microscopic parameters.

Krajewski et al. [3] have collected data from German highways in their data set HighD. This data set is used in testing our script, as well as in the analysis stage of this research.

Lyamin et al. [5] analyse the performance of ETSI EN 302 637-2. In their paper, a clear overview of the standardisation of CAM generations is given, detailing exactly when a new CAM is generated.

4. RESEARCH GOAL

Although there exists research on the relationship between microscopic parameters and CAM generations, and that between microscopic and macroscopic parameters, the relationship between traffic scenarios and CAM generations has unfortunately not yet been thoroughly investigated. As explained in Section 2.3, the easiest way to analyse traffic scenarios is though macroscopic parameters. The goal of this research is therefore to create a tool that will take a microscopic data set as input and create an estimate of the macroscopic parameters and the generated CAM messages. The output of this tool will then be analysed to answer the research question stated below.

4.1 Research Question

What trends can be observed when analysing the relationship between macroscopic traffic parameters and Cooperative Awareness Message generations?

5. METHODOLOGY

In order to approach the research question, a Python script was written [6]. The development of this script is detailed in this section.

5.1 Input handling

For testing and analysis purposes, the HighD data set was used in the development of this tool. HighD is a large set of vehicle trajectory data from German highways available for non-commercial use [3]. Each recording in this data set has three CSV files, each starting with the recording ID (indicated as XX in this section).

- **XX_recordingMeta:** contains general recording data, such as the total number of vehicles in that recording.
- **XX_tracksMeta:** contains data per track, a track being the data concerning a particular vehicle within a recording. An example of this is the number of frames that a vehicle is in the recording.
- **XX_tracks:** contains frame-specific data of vehicles, such as their x- and y-velocity.

Table 1. The locations of the relevant data in the HighD data set, where XX is the recording number

Filename	Variable	Table index
XX_recordingMeta	Recording ID	0
	Framerate	1
	Recording duration	7
	Number of vehicles	10
XX_tracksMeta	Number of frames	5
	Average velocity	11
XX_tracks	Vehicle ID	1
	X-position	2
	Y-position	3
	X-velocity	6
	Y-velocity	7
	Front sight distance	10
	Distance headway	12

The relevant data for this script is defined as global variables in *main.py* (an overview of which can bee seen in Table 1), which collects this data from the input file, runs the functions to calculate the macroscopic parameters and CAM generations and then writes the results to an output CSV file. These variables may be at a different index when using another input data set (this can easily be changed in *main.py*), but this three-file structure must be adhered to in order to run the tool. Once all the necessary data is collected from the input files, it is passed to *estimateCAMs.py* and *macroscopicParams.py*.

5.2 CAM calculations

The estimation of CAM generations is done once per vehicle in the recording and then summed. It is assumed that there was a CAM generation at the start of the recording (unfortunately there is no way of knowing when the last CAM was really generated) so the estimation starts at 1. This does not have a big impact on the results, since the total CAM generations of a recording is divided by the duration of the recording, so the final value is an average of the CAM generations per second. Algorithm 1 details the calculation of CAM generations per recording. The CAM condition check is done according to the conditions detailed in Section 2.2.

Algorithm 1: CAM calculation

```
Result: The number of CAMs for the given recording
Input: fps, frame_amount, positionList, velocityList,
       directionList
prev_cam = 0;
minimum_wait = 100/(1000/fps);
maximum_wait = 1000/(1000/fps);
cams = 1;
index = prev_cam + minimum_wait;
while index < frame_amount do
    Check if the CAM conditions are met;
    if CAM conditions are met OR index >=
     (prev_cam + maximum_wait) then
        cams = cams + 1;
        prev_cam = i;
        index = prev_cam + minimum_wait;
    else
        index = index + 1;
    end
end
```

The change in position and velocity is calculated using Pythagoras' theorem and the known old and new x- and y-values. Unfortunately, the input data set does not contain directional data (see Table 1), so this is created using the following equation:

$$\theta = \tan^{-1} \frac{v_y}{v_x} \tag{1}$$

Where θ is the angular direction, and v_x and v_y the x- and y-velocity, respectively. If both v_y and v_x are 0, a *None* value is placed in the dictionary. Otherwise, the angular direction is converted to degrees and stored in the dictionary. When calculating the change in direction, if both the old and new direction is *None*, it is assumed that there is no directional change. If either the new or old direction is *None*, the algorithm searches for the first non-*None* value in the data points between the new and old directions. If neither values are *None*, their absolute difference is simply calculated. Algorithm 2 shows this in more detail.

Algorithm 2: Calculate the directional change

```
Result: The difference between two directional data
        points, which may be of type None
Input: old_index, new_index, directionList,
       positionList
old_direction = directionList[old_index];
new_direction = directionList[new_index];
if old_direction and new_direction are None then
    return 0;
else if new_direction is None then
    index = new_index − 1;
    while index > old_index do
        if directionList[index] is None then
            index = index − 1;
        else
            direction_new = directionList[index];
            break;
        end
    end
    if index == old_index then
        return 0;
end
else if old_direction is None then
    index = old_index − 1;
    while index >= 0 do
        if directionList[index] is None then
            index = index − 1;
        else
            direction_old = directionList[index];
            break;
        end
    end
    if index == -1 then
        return 0;
end
return abs(direction_old − direction_new);
```

5.3 Macroscopic parameter calculations

The calculation of the macroscopic parameter values is done once per recording. The file *macroscopicParams.py* receives the number of vehicles in the recording, the recording duration, a dictionary containing the headway data of the recorded vehicles, and a list of average vehicle speeds.

Traffic flow q is defined as the "average number of vehicles (n) that pass a cross-section during a unit of time (T)" [2].

Thus we have the following equation:

$$q = \frac{n}{T} \tag{2}$$

Since both variables in this equation are passed to *macroscopicParams.py*, the flow is calculated using this equation.

Traffic density k is defined as the "number of vehicles (m) per distance unit (X)" [2]. Thus we have the following equation:

$$k = \frac{m}{X} \tag{3}$$

Unfortunately, the distance in the recording is unknown, so the equation must be expanded:

$$k = \frac{m}{X} = \frac{m}{\sum_{i=1}^{m} s_i} = \frac{1}{\bar{s}} \tag{4}$$

This shows the relationship between the traffic density k and the average distance headway $\bar{s}$, distance headway being defined as the distance between two consecutive vehicles on a road. Using the headway dictionary, the mean of the mean headway of each vehicle is taken and inverted to get the density. When a vehicle has no other vehicles in front of it, their distance headway is set to 0 in the input data. The use of these 0-values would impact the data quite severely, so whenever the distance headway is 0, the front sight distance is taken instead. This is the distance from the front of the vehicle to the edge of the recording frame.

Finally, we have the average speed, which is the mean of the list of average vehicle speeds. Each vehicle's average speed weighs equally in this calculation, no matter how many recording frames they are in. This does not have a big impact on the results, as the average number of vehicles in a recording is 1842, with the minimum being at 607. The small number of vehicles at the start and end of the recording which are in only a few frames, therefore, do not have a great impact on the final average speed as they do not weigh up against the many more vehicles in the rest of the recording. Furthermore, since the recordings are of highway traffic, the vehicles are likely travelling at a relatively constant speed.

Each value is rounded to 4 decimals when returned to *main.py* so they are easier to analyse in the *results.csv* file.

5.4 Output

To use the tool, one simply runs *main.py*. While it is running, the file paths will be printed as it finishes the calculations for each path. Once the program is finished running, the results will appear in the same folder as *main.py* in a file named *results.csv*. This CSV file can then be opened in a spreadsheet program of your choice. If one were to run the calculations again, the previous contents of *results.csv* will be overwritten.

6. ANALYSIS

Once the script was finished, it was run using the highD data set as input data. The resulting CSV file was used to produce the graphs discussed in this section and can be seen in the GitHub repository *Frankavj/camEstimation* [6].

6.1 CAM generations and traffic flow

Equation (2) can be expanded as follows:

$$q = \frac{n}{T} = \frac{n}{\sum_{i=1}^{n} h_i} = \frac{1}{\bar{h}} \tag{5}$$

This expansion shows the relationship between traffic flow and average time headway, the time headway being the time it takes for a vehicle to reach a certain point after the vehicle directly in front of it passed it. When the average speed of vehicles is higher and their distance headway remains unchanged, the time headway will decrease.

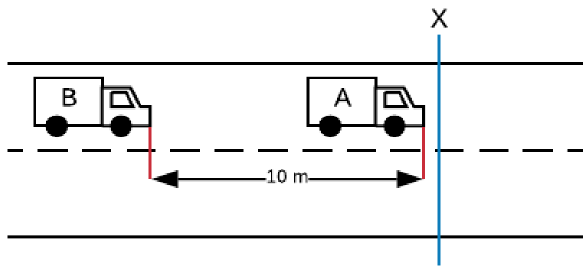

Figure 1. Two vehicles with 10 meters distance headway travelling towards point X.

Take for example two vehicles, A and B, which are travelling at 10 m/s and have a distance headway of 10 meters (figure 1). After vehicle A passes a point X, it will take vehicle B one second to reach X as it is travelling at 10 m/s and is 10 meters behind vehicle A. The time headway of B is thus one second in this situation. Say now that both vehicles increase their speeds to 20 m/s, but do not change their distance headway of 10 meters. In this situation, it would only take vehicle B half a second to pass X after vehicle A passes it, as it still has to cross 10 meters of distance but it is travelling twice as fast. Following this example, one might say that a higher average speed results in a lower time headway and that, following equation (5), the traffic flow must subsequently be higher.

Having said all this, we have assumed that the distance headway remains unchanged. This cannot be assumed though, as drivers are generally taught to keep a larger distance between themselves and the car in front of them when driving at a higher speed. If one assumes that drivers keep the appropriate distance headway, traffic flow may not be influenced by the average speed of vehicles, as the time headway would remain unchanged. Thus, no difference in the frequency of CAM generations is expected to be seen with different traffic flow values based on the corresponding vehicle speeds.

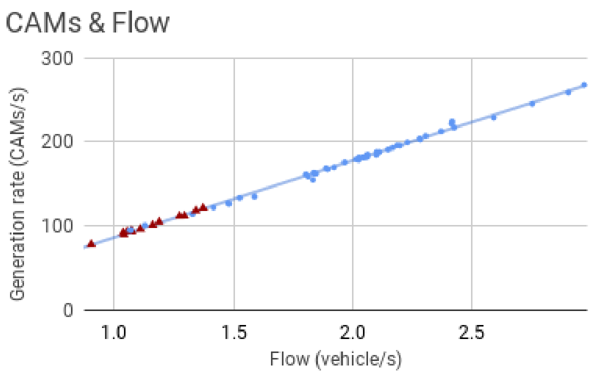

Figure 2. The CAM generations per second plotted against the traffic flow. Each point represents a recording in the data set. The red triangular data points represent a data group discussed in Section 6.2.

However, when plotting the traffic flow in the recordings against the estimated CAM generations per second, a clear positive linear relationship can be observed (see Figure 2). This must, therefore, be caused either by drivers not keeping the preferred distance, or because there are simply more vehicles emitting CAMs.

6.2 CAM generations and traffic density

Equation (4) shows the relationship between traffic density and average distance headway. As explained in the previous section, drivers are generally taught to keep a larger distance from the vehicle in front of them when increasing their speed. Thus, the average distance headway increases when the average speed is higher, and according to equation (4), the density would then be lower. Thus, one could theorise that a traffic scenario with a low density of vehicles would have a higher frequency of CAM generations than if the traffic density were higher. However, low traffic density also means that fewer vehicles emitting CAMs than at a high density. Therefore, it is difficult to formulate a prediction of the relationship between traffic density and CAM generations.

When plotting the traffic density against the CAM generation frequency, again a positive linear relationship is observed (see Figure 3). There is however a group of outliers that have a significantly lower frequency of CAM generations. This group has been plotted as red triangles instead of blue dots in Figure 3. The lower CAM generational value can be explained by looking at these same data points in Figure 2, again plotted using red triangles. These points make up the part of the data with the lowest traffic flow values, explaining the lower value of CAM generations per second in Figure 3.

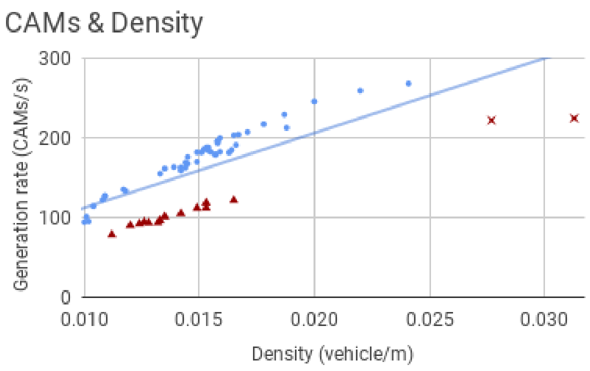

Figure 3. The CAM generations per second plotted against the traffic density. Each point represents a recording in the data set. The different colours and shapes as used to refer to different groups of outlying data in Section 6.2.

Other interesting points in this graph are the two with the highest density values, which have been plotted using a red cross instead of a blue circle. These have a lower CAM generation frequency value than would be expected when one looks at the other data points. This can be explained by taking a look at Figure 5. Here, the same two points are again plotted using red crosses. The two recordings have a significantly lower average speed than the other recordings, which explains the low CAM generations in Figure 3, as vehicles travelling at a lower speed emit fewer CAMs. The high density and low speed suggest that there may have been heavy traffic in these recordings.

6.3 Traffic density and flow

Both the traffic density and traffic flow have been found to have a positive linear relationship with the frequency of CAM generations. However, before any conclusions can be drawn from these graphs, one must also analyse the relationship between traffic density and flow. As can be seen in Figure 4, there is a positive linear relationship between these two macroscopic variables, though not a very steep one. However, one must take this into account, as the positive linear relationships seen in Figure 2 and 3 are amplified by the relationship seen in Figure 4.

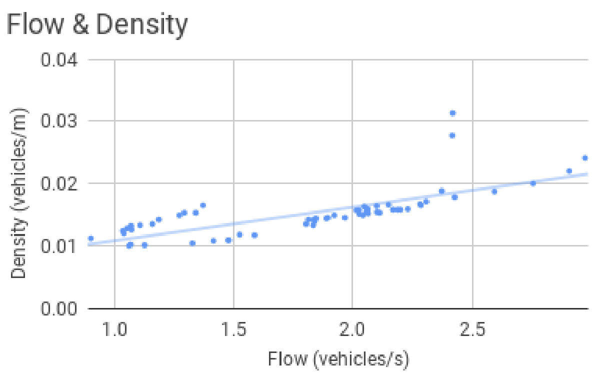

Figure 4. The traffic flow plotted against the traffic density. Each point represents a recording in the data set.

6.4 CAM generations and average speed

When plotting the average speed against the CAM generations per second, one would expect a positive linear relationship, as when vehicles move at a higher speed they will reach the 4 meters of displacement sooner than when travelling at a slower speed. However, no such relationship can be observed in Figure 5.

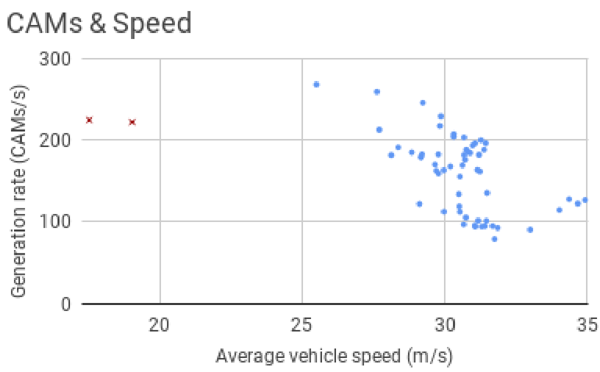

Figure 5. The CAM generations per second plotted against the average speed. Each point represents a recording in the data set. The two points that are represented using red crosses are a subgroup of the data which is discussed in Section 6.2.

When plotting the average speed of the recordings against the traffic density, a negative linear relationship can be observed (see Figure 6). This means that with a higher average speed, the traffic density is lower. The positive linear relationship found in 6.2 explains the strange distribution in Figure 5; a higher speed should increase the CAM generations per second, but this is compensated by the low density, which decreases the CAM generations per second.

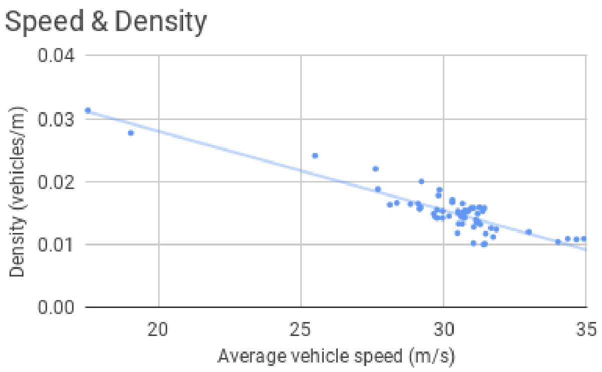

Figure 6. The traffic density plotted against the average speed. Each point represents a recording in the data set.

7. CONCLUSION AND FUTURE WORK

This research aimed to write a script that would allow the analysis of the relationship between macroscopic traffic parameters and the overall Cooperative Awareness Message generations. More specifically, we were interested in what trends could be observed between them. The graphs produced from the output data of this script suggest that there exists a positive linear relationship between traffic density and CAM generations, as well as between the traffic flow and CAM generations. However, there also exists a positive linear relationship between the traffic density and flow, meaning that the density-CAM and flow-CAM relationships amplify one another. Thus, it may be interesting to run this script on data where either the traffic flow or density is kept constant. Interestingly, there is no clear relationship between the average traffic speed and the CAM generations, as a higher speed also correlates to a lower traffic density and flow. The low density and flow values thus counteract the high speed as the former results in fewer CAM generations and the latter in more. Again, an interesting continuation of this research may lie in keeping some of these variables at a constant value. Another possible direction for future work lies in analysing what CAM generation condition has a major influence on CAM generations in different traffic scenarios. This could be approached by implementing some minor additions to my tool.

8. REFERENCES

[1] C. Campolo, A. Molinaro, and R. Scopigno. From today's VANETs to tomorrow's planning and the bets for the day after, 7 2015.

[2] S. Hoogendoorn and V. Knoop. *Traffic flow theory and modelling*. Edward Elgar Publishing Limited, 2012.

[3] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein. The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, volume 2018-Novem, pages 2118–2125. Institute of Electrical and Electronics Engineers Inc., 12 2018.

[4] H. S. M. Lim and A. Taeihagh. Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities. *Sustainability (Switzerland)*, 11(20), 10 2019.

[5] N. Lyamin, A. Vinel, M. Jonsson, and B. Bellalta. Cooperative awareness in VANETs: On ETSI EN 302 637-2 performance. *IEEE Transactions on Vehicular Technology*, 67(1):17–28, 1 2018.

[6] F. van Jaarsveld. Frankavj/camEstimation.

[7] M. Weber. Where to? A History of Autonomous Vehicles - CHM.