ABSTRACT
This paper will explain how the extra time between the orange and red light change at a traffic light could change a driver’s driving behavior. Thus avoiding dangerous traffic situations. The drivers’ behavior will be categorized into three classes. This classification has been done by discerning driver patterns, and with a training/learning algorithm a percentage of accuracy is found from the experiment’s results. These accuracies where then analyzed and found that with an accuracy of $\approx 50\%$ a driver’s driving behavior can be determined. The data has been recorded with a smartphone and a data logger application.

Keywords
Smartphone Sensors, Yellow Time, Traffic Situations, Traffic Lights, Driver Patterns, Driver Behavior, Driver Categorization.

1. INTRODUCTION
Since March 2016, the Netherlands has an additional second of orange time at the traffic lights. This has been implemented because, since 2011, each year there have been between 17 and 25 traffic light accidents by drivers ignoring a red light. The idea was that a better approximation of the ideal orange time could give the drivers, who are slow with deciding whether to brake or to speed up, a better chance to traverse the traffic lights without creating an unsafe situation.

The research done by L.H.A. Prinsen, M. Dicke-Ogenia and L. Krol [5] has shown, in terms of their behavior at traffic lights, drivers can be roughly divided in three groups: The stoppers, speeders and debaters. The stoppers usually let go of the gas when anticipating that the red light will go on soon and otherwise speed up. For the stoppers, this will be leisure deceleration so they do not need to brake suddenly. And for the debaters it depends on the timing they arrive.

1. Deceleration:
Drivers want to decelerate at a leisure speed, not suddenly. So if the drivers can help it, they either want to speed up if they can’t brake on time, or brake relatively early. In case of the speeder class, this would mean that they will either brake only when they feel the red light will go on soon and otherwise speed up. For the stoppers, this will be leisure deceleration so they do not need to brake suddenly. And for the debaters it depends on the timing they arrive.

2. Approaching speed:
The speed at which the drivers approach traffic lights partially depends on the driver’s behavior, but also on the road they are situated on. In a urban area (50 km/h), the chance that a driver finds himself waiting for a traffic light is different than on an rural road (80 km/h). At some urban and rural roads, pressure sensors in the ground detect cars and stimulate the throughput of traffic, the so called “green wave”[3], and thus encourages speeding up/keeping the same speed.

3. Perception-response time:
The last parameter that is used to calculate the orange time is the perception-response time (PRT). This is the time between the traffic light turning orange and the driver reacting to this change. This reaction depends on what road the driver is on. There can be a lot of distractions in an urban area, whereas on a rural road there could not be much to see. Also, the drivers who have a low PRT usually have the tendency to slow down gradually, where the drivers with a long PRT compensate this slow reaction with faster deceleration.

Although it may vary by municipalities, these guidelines have been applied all over the Netherlands. But whether the extra second helped decreasing the unsafe traffic light situations, is yet to be determined.

In this research these parameters will be examined and then put to the test, to see in what forms they present themselves. The first parameter, deceleration, can be obtained from acceleration data, while the second, approaching speed, can be obtained from acceleration-combined with GPS-data. The PRT, though, is very personal and will not be a part of this experiment.
1.1 Problem Statement

The extra second has now been implemented for 3 months and according to the research from Prinsen et al.(2016) the problem of drivers ignoring a red light should be halved. The reason behind this is that the driver who, as before has been described as the debater class,ponders about speeding up or slowing down, now has the little extra time to think. The question is, the ones who were just on time or a little late before the extra orange time, will they now make a different decision? The ones who were the edge case here and drive fully through the red light, will they now act differently? And how about the other drivers? Those debating drivers, what do they think, now that they have an extra second?

All these changes can be discerned using the three parameters mentioned in the introduction, but how can this be measured? The latest trend in the new cars is that they are connected with the driver's smartphone, or at least, almost 80 percent of the drivers carry their smartphone with them while driving[4]. So, is there a possibility for the smartphone to learn the drivers braking habits and thus categorizing the driver in the earlier named classes. Also is it possible to assist the drivers who usually can barely brake on time, because of e.g. lower response time or distraction, to lower their speed earlier?

1.2 Research Questions

Based on the problem statement, one main research question was made and then divided into three subquestions:

1. Can data provided by the smartphone sensors help the driver make the right decisions at certain traffic light situations?
   (a) Can this data be used to categorize a driving profile?
   (b) Can this data assist to inform the driver to brake on time?
   (c) Does the longer orange time make people change their habits in the three phases at the traffic lights?

2. RELATED WORK

Chen et al.(2015)[1] have researched abnormal driving behaviors, including weaving, fast U-turns and sudden braking. They have used their own smartphone application to identify these situations and how to recognize these. The data of their research has been put into graphs and show, among other things, the acceleration data while braking suddenly. This data can be used to already classify a braking situation at traffic lights. Also the research states to be careful with this data because some patterns are not so distinct as e.g. weaving.

In the research of Han et al.(2014)[2], they have used an own application as well, but this research was about how to measure very accurate speed data from combining smartphone sensors. Seeing as the GPS is not very precise, as GPS data can vary within a few meters and could have delay, this is not enough to calculate speed. This has to be combined with e.g. the accelerometer and a gyroscope sensor of the smartphone.

Prinsen et al.(2016)[5] have completed their research on the orange time before the extra second was implemented. But they have only questioned experts and analyzed videos of participants in their investigation. With this data they created a conclusion to combine with their own literature investigation. This can be of use to manually classify the participants of this research, if only to make it an approximation.

3. RESEARCH METHODS

At the start of the research, a few possible methods were considered to accumulate the data needed to make a detailed assessment about the drivers’ behaviors. As said earlier, a smartphone was used and from this device some sensors were considered which’s data were to be used. The sensors that came to mind were: the accelerometer, gyro sensor, magnetometer and GPS. The data logger application provided by the University of Twente[6] logs the data from all of these sensors, so all of these sensors could be used in the comparison. But after comparing the data of one driver, the accelerometer had the most detailed data needed for comparison. The GPS was also used, the combination of the timestamps per coordinate determined the location of the traffic lights in the accelerometer data.

4. METHODOLOGY

4.1 Predictions

The first step of the research was to describe the possible situations at a traffic light. These situations will be created to categorize drivers in the three classes mentioned earlier. The traffic light situations that were used are listed below and also the expected results of the accelerometer data are included:

1. Red light at arrival and halting:
   There can be three versions: The speeders will probably wait until the last second to brake, because there is a possibility they can immediately continue if the light turns green on arrival, so a quick drop in speed before halting(before point three). The second version is for the stoppers, they will let go of the gas when nearing the traffic light, so a leisure decrease of the speed. The debaters will wait until the last moment to decide, a fluctuating speed/acceleration is expected.

2. Green light at arrival and continue.
   At this point there can also be three versions: The speeders will keep/increase their speed to assure continuing, so leveled/rising acceleration/speed. The stoppers will keep their speed or maybe already let go of the gas if they think the light could change any moment, so leveled or decreasing speed. The debaters will, again, wait until the last moment to decide, a fluctuating speed/acceleration is expected.

3. Red at arrival, but immediately turns green.
   The speeders will quickly continue, so a hard peak in acceleration, where the stoppers will slowly accelerate to return to their former speed. The speed of the debaters will fluctuate as said in point two and the acceleration/speed can be in either of the debater/stopper class.

4. Orange light when passing.
   Similar to the situation with the green light at arrival, the speeders will speed up, whereas the stoppers have already slowed down and thus can stop. For the debater class it will depend on the time the light has been orange until arrival, here as well, a fluctuating acceleration/speed will be expected.

5. Orange/red light and halting.
   Here the speeder class will either pass a red light, or make a really hard brake. So, either a sudden drop in the speed/acceleration or a raise in speed will be

2
expected here. For the debater class it will be the last moment decision here, so it will be in either of the speeder/stopper class.

The last two situations are particularly interesting and can tell the most about the driver’s driving behavior around traffic lights.

4.2 Gathering Data
After the predictions, the experiments were held in two following weekends with an LG G5 (LG-H815) smartphone with Android 6.0. The smartphone with the DataLogger app from the University of Twente[6] was put in the car of the participant, facing the driving direction, e.g. in Figure 1 and logged all the data in the 18-20 minute lasting experiment. In total 12 participants have partici-

4.3 Data Processing
The data processing has been done with a self-written Java program in the following matter: The data from the all participants was separated into five parts which represent the five situations mentioned in the Methodology. This was done by hand, as seen in Appendix C.3, based on the observations made while driving with the participant in combination with the GPS data. The processing calculates a total of four parts:

1. The Maximum of the data.
2. The Minimum of the data.
3. A threshold (half of the Minimum).
4. And the amount of times the data is under the threshold.

These parts are stored in an array so they can later be compared to the test data. The test data was then compared to the train data, also by four parts:

1. Comparison of the Maximum.
2. Comparison of the Minimum.
3. Comparison of the amount of times the data went under the Threshold.
4. Comparison of the total time the data was under the Threshold.

All comparison were made as a percentage of all the train data, so if, for example, the maximum of the speeder / debater / stopper class at situation 1 would be respectively: 1.5; 1.1 and 0.5 and the test data maximum was 1.4, the resulting comparison for the speeder/debater/stopper would then respectively be 1.4/1.5 ≈ 0.933, 1.1/1.4 ≈ 0.79 and 0.5/1.4 ≈ 0.36. The comparison ensures that the train value closest to the test value has the highest percentage. This comparison is done for all data of the participants. All the collected data would be too much to include in this paper, so parts of the data are included in Appendix C. Here the averages per participant are divided per situation and shown in a bar plot. Also for all the participants, for all five situations, the comparison has been done by hand. As seen in Appendix C.3 the first three figures 3a, 3b and 3c show the graph around situation 3 for the three classes, figure 3d shows the test data. Here the comparison is fairly easy, as the test data looks most like the speeder data, because of the spike towards the minimum and the amount of times the data goes under the threshold. By making this comparison for all of the situations for all participants an accuracy can later be calculated, this by comparing the program’s results with this manual classification.

5. RESULTS
All the results are calculated as mentioned in the previous section and shown in Appendix C. After this, an analysis of this data has be made, this has been done for every situation. From there, every class has been compared together with the gender of the participant. In the comparison the corresponding participant will be shown as (Part.1-9):

5.1 Situation 1
With situation 1, there was a clear difference in distinctiveness of the classes. If the debater’s percentage was the highest, it would be above the rest with around 0.10 more. Whereas, if the stopper / speeder was the highest, the speeder / stopper respectively would follow right behind it. This resulting in a situation where both of the classes could be the case. Although this is pretty peculiar, because this would mean that the speeder and stopper class have the same tendencies when approaching a red light. When looking at the data from a male/female perspective something does come up. All but one male participants have a debater class, the next value being either the stopper or the speeder class. The other male clearly has the speeder class (Part.1). On the other hand, with the females this is not the case: The debater class here is always followed up by a stopper class and in two cases (Part.2&8) the stopper class has the highest percentage. So this would mean that females have the tendency to have a more careful driving style.

5.2 Situation 2
In situation 2 the debater class is, yet again, easy to determine, because if the debater class has the highest percentage, the other two classes would be left far behind. On the other hand if it is an other class, the debater class would have a very low percentage. This would yet again mean that the speeder and stopper class look alike in this situation. Which would be strange as one can imagine that
the stopper class would slow down at a green light whereas the speeder would keep the speed / increase it. The logical explanation would be that in both classes the participants would keep driving at the same speed, whereas the debater would in-/decrease the speed a few times. In the male / female comparison, the conclusion is rather the same as the class comparison. The stopper / debater classes are with both genders near each other and if the highest percentage is from the debater class, it would clearly show so. With the females there is, this time, only one clear stopper (Part.2), but also some speeders (Part.4). The males also have one clear speeder (Part.7), but also a clear stopper (Part.6). With situation 2 the classes are equally spread.

5.3 Situation 3
With situation 3 the speeder class is easily recognized e.g. (Part.1&2) or really close with a debater class (Part.9). If the highest percentage is a debater class, this is also clearly visible (Part.3&4) and also always followed by the stopper class and after that the speeder class. The only class in this situation that is not seen as much is the stopper class (Part.7) which only barely has a higher percentage. When looking at the males and females, the only thing with both the males and females is that if the speeder class is highest, which in both cases are present, the other percentages fall really far behind, for example (Part.1&2). For situation 3 it would mean that the speeder and debater class are the most seen patterns and they are clearly recognizable.

5.4 Situation 4
In situation 4, the speeder and stopper class both are easily recognized by the comparator. If either of these classes has the highest percentage, the other classes are almost always far behind. The speeder (Part.3&6) is always followed by the debater class, although there is no distinct follower with the stopper class. As for the gender comparisons, the only thing that comes to show is that if the highest percentage is from the speeder class it will be followed by the debater with both genders. For situation 4 this would mean that the stopper and speeder class are clearly visible. The debater class, however, does appear most (Part.5&7&8&9), but then one of the other classes is not far behind. So the debater will be hard to recognize at an orange light situation, which is yet again strange, because a fluctuation would be very apparent in comparison to a decreasing / increasing line for a stopper / speeder respectively.

5.5 Situation 5
For the last situation, two classes appear, again, that clearly peak above the rest, namely the debater (Part.7&8&9) and the stopper class (Part.1&2&4). If the highest percentage is the speeder, it is only barely higher (Part.3&5) than the stopper class, but if it is the stopper class, it always clearly jumps above the rest. When looking at the genders, the charts from the females always show a distinct highest percentage (Part.2&4&7&8&9). Whereas for the males this is not always the case. At a longer orange light situation turning into a red light, this would mean that the female’s driving type can be recognized more easily than for a male.

6. DISCUSSION
6.1 Can the data be used to categorize a driving profile?
After looking at all the distinct situations, we can say that for some part it is possible to determine a driving profile with the smartphone data. The only question that remains is, how accurately can this be done?

6.2 Can the data assist to inform the driver to brake on time?
The data cannot just be used as a brake assistant, because for the speeders who break very late, the acceleration will have a certain peak as seen in Figure 3a, which can be used to assess the situations in which the comes late. This, however, will not be enough to inform the driver, as the program does know when the driver is approaching a traffic light. An extension on the application could be made to learn the behaviors of a driver. This should then be combined with the GPS sensor to know when the person is approaching a traffic light and thus the can the application inform the driver, based on prior data, when to slow down.

6.3 Has the orange time influence on the driver behavior?
At almost the end of the trip the participants were asked if they had knowledge of the extended orange time. Only four of the participants had no knowledge of this and two of them did show changes in their driving behavior. As seen in the participant table in Appendix A and in Appendix C.4 participant 2 and 4. The female of age 56 and the male of age 53 were a stopper/debater class respectively according to the manually and program categorization, but in the graphs of 4 differences are seen. In figures 4a and 4c the situation while approaching a traffic light turning orange without knowledge of the extra second was a stopper / debater class respectively. But after casually telling the drivers about the extra orange time, their attitude changed. Both of them have seen another traffic light who turned yellow (and also turned red), and now they both showed, as seen in figures 4b and 4d an attitude matching the speeder class.

7. CONCLUSION
The conclusion is that the extra second of orange time definitely has SOME influence on the driver patterns, but it cannot be said yet if the debater classes are halved. This, not only because lack of data, but also because the categorization of the drivers can be improved. If an accuracy of almost 100% can be reached, an accurate approximation of the amount of drivers in each class can be made and thus also compared if the debater class has halved.

7.1 Future Work
As said earlier, future work could be an extension on the application to create a self-learning program to assist drivers to brake earlier if needed. Another thing could be more detailed comparing based on not only accelerometer data, but also including magnetometer / gyroscope.

8. REFERENCES

APPENDIX

A. PARTICIPANTS

<table>
<thead>
<tr>
<th>Number</th>
<th>Gender</th>
<th>Age</th>
<th>Expected Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>39</td>
<td>Speeder</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>63</td>
<td>Stopper</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>56</td>
<td>Stopper</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>57</td>
<td>Speeder</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>59</td>
<td>Debater</td>
</tr>
<tr>
<td>6</td>
<td>Female</td>
<td>56</td>
<td>Stopper</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>21</td>
<td>Debater</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>53</td>
<td>Debater</td>
</tr>
<tr>
<td>9</td>
<td>Female</td>
<td>49</td>
<td>Stopper</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>20</td>
<td>Stopper</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>55</td>
<td>Stopper</td>
</tr>
<tr>
<td>12</td>
<td>Male</td>
<td>58</td>
<td>Speeder</td>
</tr>
</tbody>
</table>

Table 1: Participant data.

B. FIGURES

C. RESULTS

Below, in Appendix C.4, are all the result graphs created from the measured data. The data is represented in percentages per situation divided per class.

For easy reading: These were the different situations:

1. Red light and halting.
2. Green light while continuing.
3. Red light at arrival, but immediate green light.
4. Orange light while passing.
5. Orange/red light just before halting.

C.1 Manual Classification

Table 2 shows the manual classification of each situation for all participants. With Speed, Deb, Stop being the speeder, debater and stopper class respectively.

<table>
<thead>
<tr>
<th>Nr</th>
<th>Sex</th>
<th>Age</th>
<th>Sit1</th>
<th>Sit2</th>
<th>Sit3</th>
<th>Sit4</th>
<th>Sit5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>39</td>
<td>Speed</td>
<td>Deb</td>
<td>Speed</td>
<td>Speed</td>
<td>Deb</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>56</td>
<td>Speed</td>
<td>Stop</td>
<td>Speed</td>
<td>Stop</td>
<td>Deb</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>53</td>
<td>Deb</td>
<td>Speed</td>
<td>Speed</td>
<td>Stop</td>
<td>Deb</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>20</td>
<td>Speed</td>
<td>Speed</td>
<td>Deb</td>
<td>Speed</td>
<td>Speed</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>55</td>
<td>Stop</td>
<td>Deb</td>
<td>Stop</td>
<td>Deb</td>
<td>Stop</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>58</td>
<td>Stop</td>
<td>Speed</td>
<td>Speed</td>
<td>Speed</td>
<td>Speed</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>59</td>
<td>Speed</td>
<td>Speed</td>
<td>Stop</td>
<td>Deb</td>
<td>Stop</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>56</td>
<td>Deb</td>
<td>Speed</td>
<td>Stop</td>
<td>Deb</td>
<td>Deb</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>49</td>
<td>Stop</td>
<td>Deb</td>
<td>Stop</td>
<td>Deb</td>
<td>Stop</td>
</tr>
</tbody>
</table>

Table 2: Manual Classification Table.

C.2 Program Accuracy

In table 3 there are three columns per situation, these mean the following:
a: Amount of correctly detected this class by the program.
b: Total amount of this class in this situation.
c: Correctly detected percentage.

C.3 Data Example Graphs
### Table 3: Accuracy of the program.

<table>
<thead>
<tr>
<th>Class</th>
<th>Sit 1</th>
<th>Sit2</th>
<th>Sit3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>Speeder</td>
<td>1</td>
<td>4</td>
<td>25%</td>
</tr>
<tr>
<td>Debater</td>
<td>1</td>
<td>2</td>
<td>50%</td>
</tr>
<tr>
<td>Stopper</td>
<td>0</td>
<td>3</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sit4</td>
<td>a</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>Speeder</td>
<td>1</td>
<td>4</td>
<td>25%</td>
</tr>
<tr>
<td>Debater</td>
<td>2</td>
<td>2</td>
<td>100%</td>
</tr>
<tr>
<td>Stopper</td>
<td>1</td>
<td>3</td>
<td>33%</td>
</tr>
</tbody>
</table>

**C.4 Comparison Graphs**

The comparison graphs are shown below:

![Comparison Graphs](image)

Figure 3: Data examples of the five situations.
Figure 4: Comparison graphs before and after knowledge of the extra orange time.