Index

1. Preface

1.1. Origins

Traffic Safety Indicator project is a small part of a larger project that is been carried out during the last years in the National Taiwan University of Science and Technology Industrial Management Department, called Smart Campus. This wider project involves lots of applications around the campus, and even around the city. This project try to apply technology in different ways to make stakeholders lives easier by studying the human flow, reporting maintenance problems or trying to improve energy consumption, for instance.

NTUST is one of the most important universities in Taiwan and has a very good reputation in terms of technology. In the last years, NTUST Center for IoT Innovation has been working hard in all kind of Internet of Things technologies like: deep learning, image recognition, block chain, sensor technology, etc. In the Figure 1 it is shown a slide of one presentation about this project.

Professor Shuo-Yan Chou, as director of this thesis, proposed this and some other projects to me. After some research about them, 'Traffic Safety Indicator' was the selected one to research into.

In this case, as the project would be applied on external to university traffic safety, this is a topic more close to Smart City concept. Apart from this one, there are several projects applied to Smart City inside Smart Campus, like showing useful information about buses schedules or proposing a shared parking model open to the public.

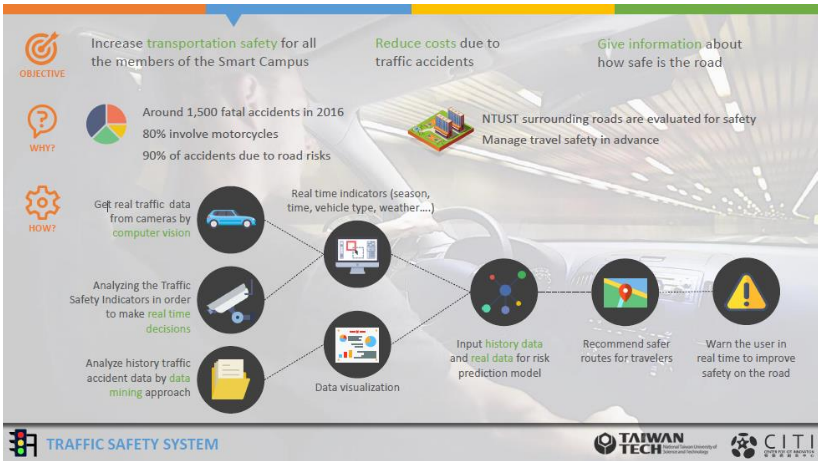

Figure 1 Traffic Safety System in Smart Campus

1.2. Motivation

In addition to my general interest in technology due to my studies and concerns, what have been explained in the previous section is the main element that catches my attention: helping people in their daily life decisions using the most current technology.

The way Internet of Things (IoT) is been applied in many activities of people's routine (like when using phone apps, GPS software or home assistants, between others) is a reality that all professionals involved in technologies world should notice if they have not do it yet.

The interest in this field of studies make me excited and interested in developing this project and expanding my knowledge in it. So, as an engineer, to work in all kind of projects in this scope it is rewarding, even developing the less technical part of them.

1.3. Scope

This project pretends to define a traffic safety index. It means that by installing video cameras in any kind of road, this index could be calculated taking in account some parameters with help of object recognition technology and data extracting.

There are some video cameras installed at NTUST University focused on the street. Thanks to object recognition technology, they are able to generate data about traffic as position, speed and typology of each vehicle on their vision field. Then, data files about traffic will be generated continually. In parallel to this, it is needed to find the way to make this data useful: deciding which data is relevant to define the safety of a road and the way to do it. For this, some parameters are needed to be described for then make up a function model that combining them give safety information as a result, and this is precisely the scope of this thesis.

Parallel technical work to achieve getting useful traffic data from video cameras is taking for granted in this ideal scenario. It means that the way data is obtained and dealing with camera issues does not involve this thesis. It is also assumed that cameras are located in strategic places where the captured date is not influenced by external factors. In addition, it must be borne in mind that the data are purely from traffic, it is not possible to take into account aspects such as the consumption of alcohol and drugs, or other kind of possible distractions for the drivers that cannot be captured in a video file.

Therefore, supposing this idyllic situation in which video file can be extracted and turned into all the necessary data required, this thesis specific goals shown in Figure 2 are to define and quantify parameters and to build the model to calculate Traffic Safety Indicator function.

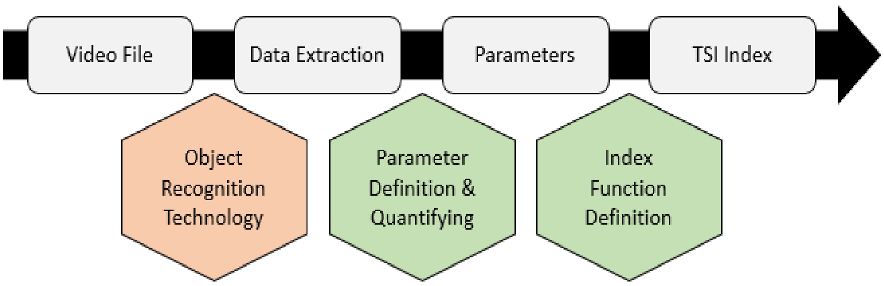

Figure 2 Thesis scope in project outline

But this definition must be in a very clear and exhaustive way. Therefore, is also necessary to quantify this parameters, this is to define the way they are going to be measured. Later, after studying different options, decide which non-very relevant or non-well quantified parameters can be rejected and which most representative of them will remain as candidates.

Once parameters are clearly defined and quantified, next step would be to define the function: deciding the way to combine the inputs (parameters) to take the output (indicator). Probably, there will be some intermedium parameters coming from certain first parameters, that combined will give as a result the TSI index.

Final parameters function will consider should be adapted to the function by normalizing. For each of them, there will be a scale between 0 and 1: the safest situation will be always represented by number 0 and the most dangerous by number 1. Same happens to final indicator, they will show the safety of the road with a 0 to 1 scale.

When this goals has been achieved and this technology has been tested, there will be many application for that index. Then, it would be necessary to collaborate with different companies or organizations interested on the indicator developing or usability. It could be used by the government in terms of accident reduction, or by citizens to get informed of roads safety and decide a way to their destination.

This project is actually on a pilot phase, so we are just building the basis for the index in a simple and generic way. But in a medium term future it would be possible to improve the index by incorporating useful conclusions extracted from crash videos analysis, among others.

1.4. Previous knowledge

This project has two very different parts: the technical one and the analytical one ('Research into the development of a Traffic Safety Indicator'). The first of them, is responsible for extracting data from the video camera files using object recognition technology, while the second one needs to take advantage of this data to convert it into useful information, defining and combining some parameters to analyze traffic safety. [10] [11] [13]

My previous knowledge of the technical part is almost inexistent, and it is not really necessary for developing the second one, because it can be done supposing data extracted properly. Even so, it was so interesting to assist to a class about 'Deep learning' imparted by Richard, one of the members of the Industrial Management department in NTUST. Thanks to that, it was possible to learn some concepts about the work behind the technical part and what are they working with.

The important point is that cameras are the most powerful sensors that technology have developed since nowadays. There are lots of specific sensors that allow us knowing specific information about the environment, but once object recognition technology has been developed, cameras are able to provide us more information at the same time and without human supervision than any sensor ever developed.

Anyway, the important part for this thesis is the second one: how to make data useful for determine traffic safety. For doing it, is very important to define the parameters that are going to be used in a clear way, before quantifying and combining them to get the safety indicator.

Some useful previous knowledge for the progress of the project is related to combinational risk analysis, learned at university previous studies. But there will be also hard work on external research about this topic[14] [15]. Besides, using common sense and mathematical skills will be necessary to define each parameter and the indicator function in a reasonable and strict way.

2. Parameter definition

In this chapter, different possible important parameters are going to be defined and quantified. It is important to do a good and exhaustive definition of each of them to ensure if they are relevant or not, and establish how to quantify them [4].

Some previous comments about necessary assumptions for taking into account in the parameter definition are listed below:

- Nowadays, cameras like the one shown in Figure 3 are getting some useful information, but as told previously, it would be supposed an idyllic scenario in which we could get all the information we want about the traffic situation from video files.
- For a non-influenced analysis it is important to locate the cameras where the traffic is fluent, meaning, not close to intersections or traffic lights. In this way the traffic will be analyzed fairly for each road, based on parameters alien from influence of this facts. As told in the preface, deciding the camera location or the focus of the cameras does not involve this thesis, it is assumed that they simply provide us the information that we need.
- As it will be shown in the following pages, some of the parameters need the presence of at least 1 vehicle during the last n minutes to be calculated, so they cannot be taken into account in the safety analysis. This is the case of the parameters TD, AS, SD, AS and DD. It will be explained later, but LT has a more special condition to be calculated. Furthermore, with a low presence of vehicles, the value of the parameters is not significant.
  For this reason, it has been decided that when the density (DE) of vehicles is under a certain limit value (DE < DE limit), the index will not be calculated and it will be returning 0 as result. This density limit should be defined once the system has been generating data and this data has been conveniently analyzed.

Figure 3 Traffic camera

2.1. Influx of vehicles (IN)

Thanks to image recognition, this is one of the easier parameters to obtain. Video cameras are able to detect and count the number of vehicles passing through its visual camp in a closed period of time. Initially, a road with lots of vehicles per minute will be less safe than other one with barely traffic, but it would be necessary to analyze interaction with other parameters

This parameter will be measured in number of vehicles per minute per lane. The influx of vehicles value considered will be an average of the last n minutes data generation. So, the number of vehicles that has passed in the last n minutes divided by n and by the number of lanes of the road will define the average influx of the road with a decimal value.

$$IN = \frac{NV}{n \times NL} \qquad [1]$$

IN: average influx on camera’s visual field during the last n minutes [vehicles/min. lane]

NV: number of vehicles detected during last n minutes [vehicles]

NL: number of lanes on the road [lanes]

It is relevant to take into account the number of lanes when considering the influx. It is not the same 100 vehicles per minute absolute influx in a 1 lane street in the city center that a 100 vehicles per minute influx in a 5 lanes highway. It will be very important to use the number of lanes static parameter to convert absolute data in relative data, meaning in this case total influx or influx per lane. Same will happen with other parameters we will see in the following pages.

For example, if the camera in the elevated road in front of NTUST University (Jilonglu Elevated Road) northeast direction, has detected 464 vehicles passing through its 2 lanes (NL = 2) in the last n minutes, the average influx parameter will take a value of: IN = 232/n.

2.2. Density of vehicles (DE)

Density of vehicles is a parameter to take into account when talking about traffic safety. A situation with lots of vehicles driving in a road at the same time like in the right side of Figure 4 would be more dangerous than when the road is more empty, like in the left side of it.

Density is an instantaneous parameter, for every moment the number of vehicles on the camera's visual field can be different. Cameras provide the system the number of vehicles detected in the visual field of the camera each second.

In order to get a representative value, an average of vehicles in the image is going to be calculated considering the provided instant number of vehicles for the last 60n seconds. The last 60n seconds instant number of vehicles summation divided by 60n will give as result the last n minutes average number of vehicles. As the instant value, this parameter has as measurement unit number of vehicles.

$$AV = \frac{\sum_{t=-(60n-1)}^{t=0} IVt}{60n} \qquad [2]$$

AV: average number of vehicles in the camera visual field during last n minutes [vehicles]

IVt: instantaneous number of vehicles detected by the camera in the instant t [vehicles]

t = (-(60n-1),0): instant (time in seconds)

Figure 4 Traffic density in a highway

The average number of vehicles divided by the distance in kilometers covered by the camera on the road and by the number of lanes, will give as result the average lineal density of the traffic in vehicles per kilometer of lane:

$$DE = \frac{AV}{DC \times NL} \qquad [3]$$

DE: average lineal density during the last n minutes [vehicles/km. lane]

AV: average number of vehicles in the camera visual field during last n minutes [vehicles]

DC: longitudinal distance (direction of movement) covered by the camera in the road [km]

NL: number of lanes on the road [lanes]

The distance considered in the formula [3] and the number of lanes are both static parameters while the cameras remain in the same position and orientation or unless there are road restructuring works. DC represents the length or longitudinal distance on the road where the camera is able to detect vehicles. As the number of lanes, unless changes in the road or in the camera, is a constant value.

In this way, a one dimension density per lane is obtained, and this allows us to compare different type of roads in a fair way.

For instance, if a camera that is able to detect vehicles in 80 meters (DC = 0,08) in a 3 lane road (NL = 3) is detecting 6,78 vehicles in its visual field as average of the last n minutes (AV = 6,78), the density calculated will be 28,25 vehicles per km of lane (DE = 28,25).

As explained in other chapters of the thesis, this parameter is not only used as the others to describe the safety. If the value of the density is under certain value, the whole system will send a signal that indicates that the TSI index should not be calculated. This limit will be defined in future steps of this project, when there is data available to study with. But it will define situations with very low density values that indicates that the traffic is so insignificant that it makes no sense to evaluate the safety.

2.3. Lane changes (LC)

A high number of lane changes will probably increase the risk of accident, so this parameter is probably one of the most important. As cameras are able to detect vehicles positions, it will be possible to have information about how many times there is a change of lane in a road in a minute. This parameter seems to have no sense when studying the safety of a one lane road, but incorporating and leaving vehicles would be always considered as one change of lane.

Lane changes is also a parameters that has no sense without considering the number of lanes. Certain number of lane change per minute can be very dangerous or safe, depending the dimensions of the analyzed road. Using the number of lanes to turn LC parameter into relative allow us to use it to compare different type of roads in a fair way.

At first, this value is a natural number which represent the number of changes detected by the camera. To take a representative value, it would be convenient to take an average of the last n minutes and divide this number by the number of lanes of the road, so the natural number could probably change to a decimal number. The unit used to measure this parameter will be number of lane changes per minute per lane.

$$LC = \frac{NC}{n \times NL} \qquad [4]$$

LC: average lane changes during last n minutes [changes/min. lane]

NC: number of lane changes detected during last n minutes [changes]

NL: number of lanes on the road [lanes]

For example, if in high speed highway with 4 lanes (NL = 4) the camera has detected 15 changes of lane in its visual field during the last n minutes (NC = 15), the calculated average will be: LC = 3,75/n.

2.4. Typology of vehicle distribution (TD)

If all the vehicles were cars, traffic would be probably safer. Even though, it would be safer a 100% motorbike traffic than a mixture of car and motorbikes.

Thus, in this case it is needed to do a very deep research about relationship between typology distribution and safety, because probably for each situation a different distribution would be the safest, because the optimal combination depends and is affected by many other factors.

Although it is easy to define because cameras can differentiate typology of vehicles passing through their visual field (big size vehicles, sedan, motor bikes and bicycles can be distinguished as different class), this is a very difficult parameter to take into account.

Anyway, for simplifying, for the moment the proposed option to quantify it would be a percentage (or part by one) of how many 2-wheel vehicles are out of the totality:

$$TV = \frac{N2}{NV} \qquad [5]$$

TV: typology of vehicles distribution [2-wheel veh./veh.]

N2: number of 2-wheel vehicles detected during last n minutes [2-wheel vehicles]

NV: number of vehicles detected during last n minutes [vehicles]

For instance, a 0,23 value for this parameter means that 23% of the vehicles are motorbikes or bicycles, in front of 77% of cars or big sized vehicles of 4 or more wheels.

2.5. Average speed (AS)

Speed (Figure 5) is obviously a parameter to take into account for describing safety in traffic. But it is true that sometimes a too slow traffic can be even more dangerous than a faster one, depending on other factors like the density of vehicles or the kind of road. That is the reason why this parameter has no sense by itself and should be combined before being introduced in the TSI function.

For each vehicle, an average speed will be calculated and assigned with time spent and distance covered by camera's visual field.

$$ASi = \frac{DC}{TSi}\ \forall\ i \qquad [6]$$

ASi: average speed of vehicle i [km/h]

DC: longitudinal distance (direction of movement) covered by the camera in the road [km]

TSi: time spent by vehicle i in camera's visual field [h]

i = (0,…,NV): vehicle

NV: number of vehicles detected during last n minutes [vehicles]

Figure 5 Speedometer

A representative value would be obtained by taking an average of all the vehicles driving in the road for the last n minutes measured in km/h.

$$AS = \frac{\sum_{i=0}^{i=NV} ASi}{NV} \qquad [7]$$

AS: average speed during last n minutes [km/h]

ASi: average speed of vehicle i [km/h]

i = (0,...,NV): vehicle

NV: number of vehicles detected during last n minutes [vehicles]

To skip the problem of traffic light or intersection influence on vehicles speed, as told previously, all the cameras must be placed on an adequate location. It means that they should be placed away from them, at a point of the road where vehicles usually drive fluently.

If NV is 1, it means that only one vehicle has passed through the camera in the last n minutes. In this case, the average speed will be equal to the speed of this unique vehicle.

In the case that NV is 0, this parameter cannot be calculated. Same happens with TD, SD, AD, and DD parameters. The case of LT parameter is even more restrictive.

As mentioned in the parameter definition introduction, if some parameters cannot be calculated because of lack of vehicles (DE < DE limit), the TSI index will not be calculated, either due to lack of data to calculate some parameters or because although they can be calculated, traffic is not considered significant enough for safety analysis

2.6. Speed standard deviation (SD)

Speed dispersion of the vehicles is probably more important than speed average. In this case we are talking about a parameter that has sense by itself, without need of combining. Meaning, when all the vehicles have similar speeds, the traffic is more homogenous and the situation use to be safer than when some of them are driving faster than others.

For measuring it, the standard deviation of all vehicles that have passed through the visual field of the camera in the last n minutes will be taken into account. Each of them will have an average speed associated. Considering as many speeds as vehicles, a sample of data is obtained. The standard deviation of this sample will be considered as the speed dispersion of the traffic situation.

As shown in the legend of the formula [8] below, the standard deviation has the same units of measurement than the main variable. In this case, as standard deviation of speed, it is kilometers per hour, because the main value is the average speed of the vehicles.

$$SD = \sqrt{\frac{\sum_{i=0}^{i=NV} (ASi - AS)^2}{NV}} \qquad [8]$$

SD: speed standard deviation during last n minutes [km/h]

ASi: average speed of vehicle i [km/h]

AS: average speed during last n minutes [km/h]

i = (0,…,NV): vehicle

NV: number of vehicles detected during last n minutes [vehicles]

2.7. Average longitudinal distance among vehicles (AD)

The distance with the vehicle in front is always important for traffic safety. That is why in highways is recommended to drive leaving at least 70 meters of distance. If the road is not a high speed one, this limit use to be lower. So, it is intuitive thinking about combining average distance and speed.

This parameter is considering distance in the traffic direction between vehicles in the same lane. To obtain a representative value, the average of the last n minutes will be considered. This value will have meters as measurement unit.

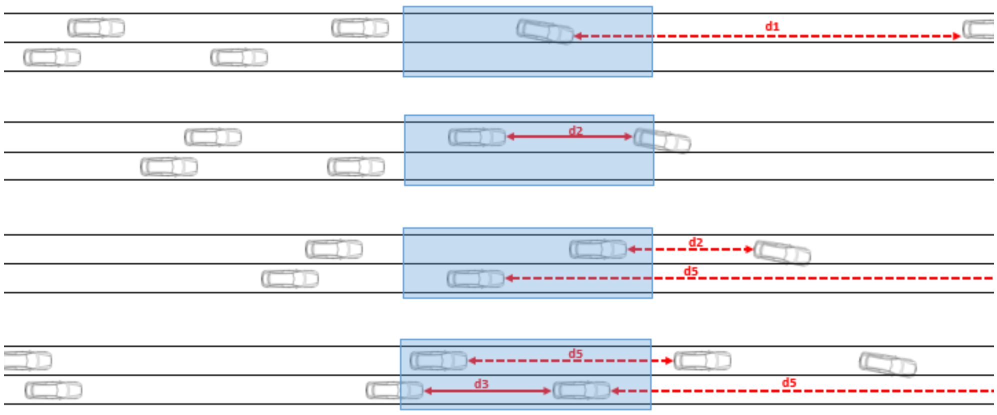

Figure 6 Longitudinal calculating distance scheme

In the Figure 6, the evolution of a traffic situation is shown in 4 different instants. The blue square indicates the camera's visual field. Distances taking into account by the system in each moment are marked in red. Continued lines are used if both vehicles are seen by the camera. It can occur, that two consecutive vehicles in the same lane are not in the visual field at the same time temporally, or even in any moment because they are driving with long distance When the first one leaves the visual field, the system will calculate the distance based on speed and time while the second vehicle is still in the visual field of the camera. These distances calculated by the system are marked in dashed lines in the Figure 6.

Each vehicle entering in the study area has an assigned vehicle (the immediate before in the same lane) with which the distance is going to be considered (like d3 distance in the fourth illustration in Figure 6). So the number of pairs of vehicles considered will be the same as the number of vehicles detected.

So, for each pair of vehicles, an average distance will be calculated based on their instant distance during the time the second one is spending in the study area:

$$ADj = \frac{\sum_{t=0}^{t=TSj} IDjt}{TSj} \; \forall j \qquad [9]$$

ADj: average longitudinal distance between the pair of vehicles j [m]

IDjt: instant longitudinal distance between the pair of vehicles j in the instant t [m]

t = (0,…,TSj): instant (time in seconds)

TSj: time spent by the second vehicle of the pair of vehicles j in camera's visual field [s]

j = (0,…,NV): pair of vehicles

NV: number of vehicles detected during last n minutes [vehicles]

Then, calculating the average distance of the pairs of vehicles that have passed during the last n minutes, a representative average longitudinal distance is obtained:

$$AD = \frac{\sum_{j=0}^{j=NV} ADj}{NV} \qquad [10]$$

AD: average longitudinal distance during last nminutes [m]

ADj: average longitudinal distance for the pair of vehicles j [m]

j = (0,…,NV): pair of vehicles

NV: number of vehicles detected during last n minutes [vehicles]

It should be noted that although the illustrations have used cars to represent vehicles, everything explained above can be extrapolated to all kind of vehicles.

2.8. Longitudinal distance among vehicles standard deviation (DD)

Same as talking about speed homogeneity, we are faced with a parameter that does have meaning by itself. The homogeneity of the traffic in terms of distance is a representative fact about safety. As occurs with speed, standard deviation is going to be used to evaluate the homogeneity of the longitudinal distance among vehicles.

For measuring the standard deviation, every vehicle that have passed through the visual field of the camera will have an associated distance until the vehicle in front in the same lane, as shown in the previous parameter (AD). Associated distances for each pair of vehicles in the last n minutes will be used as a sample to calculate the standard deviation.

$$DD = \sqrt{\frac{\sum_{j=0}^{j=NV} (ADi - AD)^2}{NV}} \qquad [11]$$

DD: longitudinal distance standard deviation during last n minutes [m]

ADj: average longitudinal distance between the pair of vehicles j during last n minutes [m]

AD: average longitudinal distance during last n minutes [m]

j = (0,…,NV): pair of vehicles

NV: number of vehicles detected during last n minutes [vehicles]

As happened in the case of SD parameter, standard deviation has the same measurement units than the main variable. So, as we are in front of a distance deviation, its units are meters.

2.9. Average lateral distance among vehicles (LT)

Lateral distance between vehicles is one the most difficult parameter to consider. When two vehicles are driving in the same lane, considering lateral distance do not make sense. When they are driving in different lanes but with a big longitudinal distance, neither. Same happens when they are driving in not contiguous lanes. So, this parameter has sense if and only if two vehicles are driving in contiguous lanes and not very far longitudinally or when overlapping laterally. [9]

With help of the following Figure 7, it would be easier to understand this conditions more clearly:

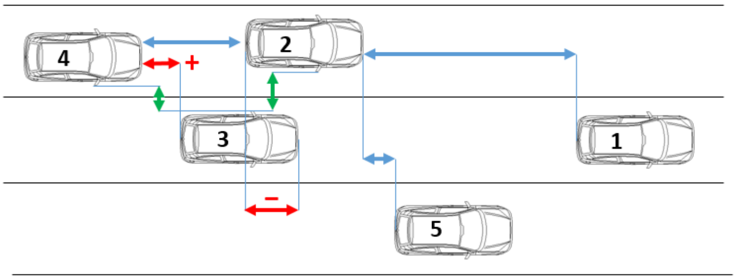

Figure 7 Lateral distances conditions scheme

About the relevance of lateral distance (marked in blue) in the Figure 7:

- Distances between cars 2 and 4 is not relevant because they are driving in the same lane.
- Distances between cars 1 and 2 is not relevant because their longitudinal distance is big enough.
- Distances between cars 2 and 5 is not relevant because they are not driving in contiguous lanes.
- Distances between cars 3 and 4 is relevant because their longitudinal distance is not big enough.
- Distances between cars 2 and 3 is relevant because they are overlapping longitudinally (what means negative distance)

Now, it is necessary to define the limit longitudinal distance to consider the lateral distance, but it depends on the speed of the back vehicle.

A reasonable restriction to delimitate distance between front part of back vehicle and back part of front vehicle could be the following:

$$ld\ (m) < \frac{s\ (km/h)}{10} \qquad [12]$$

ld: longitudinal distance in meters between the pair of vehicles [m]

s : back car speed in kilometers per hour when entering in camera's visual field [km/h]

Then, if the speed is 50 km/h, lateral distance will be only calculated if the longitudinal distance is less than 5 meters; and if the speed is 120 km/h, if it is less than 12 meters.

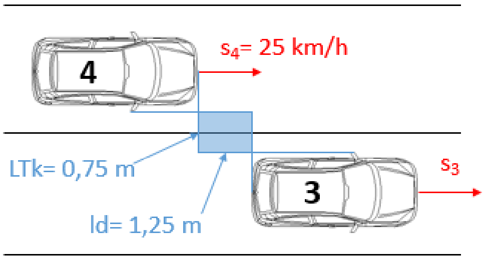

Figure 8 Lateral distance restriction scheme

In the example in the Figure 8, lateral distance (0,75 m) would be considered because longitudinal distance (1,25 m) is lower than the limit distance for 25 km/h (2,50 m).

Thus, lateral distance between 2 vehicles will be considered if and only if their longitudinal distance is lower than the limit one for the corresponding speed. With the distances calculated for each pair of vehicles has passed through the visual field of the camera during last n minutes that meet this restriction, a sample of data is generated.

With that sample a lateral average distance between vehicles is calculated:

$$LT = \frac{\sum_{k=0}^{k=PR} LTk}{PR} \qquad [13]$$

LT: average lateral distance during last n minutes [m]

LTk: average lateral distance for the pair of vehicles k [m]

k = (0,…,PR): pair of vehicles accomplishing the restriction

PR: pairs of vehicles during last n minutes that accomplish the restriction [vehicles]

As told in the introduction of parameter definition, the case of LT is special. For calculating this parameter is necessary that two vehicles that accomplish the restriction has passed through the camera visual field during the last n minutes. This is even a more restrictive condition than the limitation for other parameters (NV = 0). When real data is obtained by the cameras this parameter should be tested exhaustively because of these reason.

2.10. Road wetness (RW)

Weather is obviously relevant in safety. Especially when it is rainy, the traffic become more dangerous. When it is raining or snowing, it is important to drive even in a more cautious way: softer, with less speed and more distance, etc. However, evaluating the weather it is not very useful if it is sunny or cloudy while the floor remains dry. Facts that have influence over the road on traffic safety is precipitation: rain or snow.

For the time being, it is going to be supposed that the cameras are able to detect through the image and provide to the system a parameter about the conditions of wetness of the road. This parameter indicates if the road is dry or wet. If its value is 0 it means that the floor is completely dry and safe to drive, while 1 means that the road is totally wet, the most scenario for this parameter (Figure 9).

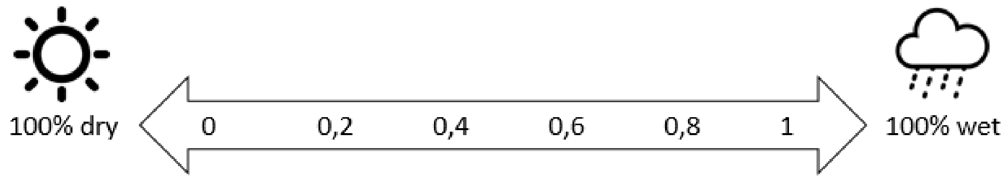

Figure 9 Road wetness scale

$\mathrm{RW} = [0, 1]$: road wetness parameter increasing from dry to wet conditions

For example, when the road is totally wet during a storm the parameter would be located at 1, a sunny or even cloudy but dry day it will be 0, and a cold night without rain but with humidity it would be maybe around RW = 0,3.

2.11. Road visibility (RV)

With visibility happens something similar that with wetness parameter. It can affect to other parameters range of safety. This is, some speed that can seem safe with good visibility in a clear day, could be very dangerous in a situation of a rainy night with a lot of fog. So, it would be good to consider this fact in the safety index.

Thus, considering that cameras can provide us some index of visibility based on the clearness of the captured image, this parameter is already defined. When it has a value of 0 it means that the conditions are completely clear and safe to drive, while 1 means the worst visibility conditions (Figure 10).

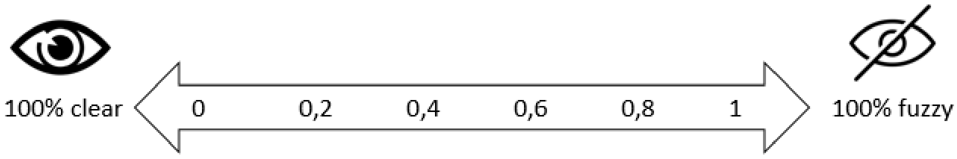

Figure 10 Road visibility scheme

$RV = [0, 1]$: road visibility parameter increasing from clear to fuzzy conditions

This parameter varies depending on the visibility signal coming from each camera. It depends on factors like on the weather (if it is sunny, cloudy, rainy, or especially foggy, like in Figure 11), the street lighting and if it is day or night.

Figure 11 Fuzzy road visibility conditions

3. TSI index definition

3.1. First parameters discarded

Once all possible parameters are already defined, it is necessary to decide which of them are going to be used. First step is to do a first filter and discard some of them.

This thesis is just about the construction a basis for a pilot function. Although is the future this work could be finished and improved incorporating more elements and developing a more complex analysis, for the moment it is more convenient to make it up with some clear and easy to combine parameters than try to use too much of them. [5]

So the first rejected parameters are going to be listed and explained below:

- Influx of vehicles (IN)

The case of parameter is kind of special. It has not been rejected because of lack of information. Actually, the influx of vehicles give us useful information. The problem with this parameter is that it has lineal dependence relationship with other parameters: density and average speed:

$$IN\ \left[\frac{veh.}{min}\right] = DE\ \left[\frac{veh.}{km}\right] \times AS\ \left[\frac{km}{h}\right] \times \frac{1}{3.600}\ \left[\frac{h}{min}\right] \qquad [14]$$

So, using these three parameters at the same time we would be involved in a situation of information redundancy. For this reason, one of them was needed to be rejected, and analyzing possible combinations with other parameters, the chosen one was influx of vehicles.

- Typology of vehicle distribution (TD)

Although this parameter could be very useful in the future, it has been considered as too complex to take in account for the moment. For the moment, TD parameter is not going to be used to calculate the TSI index, but it is left open to be used in possible future improvements of the indicator.

- Road wetness conditions (RW) and Road visibility (RV)

The conditions of the floor in terms of wetness is a very important fact for driving. Same occurs with the visibility conditions. Both of them could be very useful to define the safety of the road although they are not purely traffic parameters. These two parameters describe more the environment conditions for driving than vehicles driving interaction on the road. Anyway, combined with some other traffic parameters can be used licitly in the function.

The reason why these two parameters have been temporarily set aside is about complexity of the function. If they were included, they would be combined with other parameters that will be already used in the function, so there would be parameter repetition. So, for avoiding this situation, W and V parameters will be saved for future improvements, where they will be used with almost total security.

As in the case of influx, density and speed, some kind of lineal relationship was considered between this lineal density and longitudinal distance (AD). But thinking about it, they are not the same, so they must remain as candidates to participate in the construction of the function. For example, when two motorbikes are driving or static in parallel in the same lane, it is considered in density but not in distance parameter. So, this two parameters are similar but not the same, each of them can make sense combined with different ones.

It is important to clarify that, the provisional discarded values could be recovered in the future along with other new parameters to define safety.

3.2. Candidate parameters combination

Now, the number of possible ingredients for making up the TSI index has been reduced to 7:

- Density (DE)
- Lane changes (LC)
- Average speed (AS)
- Speed deviation (SD)
- Average distance (AD)
- Distance deviation (DD)
- Lateral distance (LT)

This parameters are important, and it would be possible to define a safety index simply assigning a weigh to some of them. But some parameters gain importance depending on the value of others. As told before, a certain distance between vehicles can be considered safe enough or not depending on the speed. That is why combining some of the parameters previously explained will generate a better quality index. [6] [7] [8]

So, in this point, for doing the combinatorial analysis to see which parameters can appear alone in the function, which of them must appear combined with others and which are better to set aside, a relationships matrix like Table 1 should be used to study all the possible combinations.

Table 1 Matrix for parameter combination analysis

DE	1						
LC	2	8					
AS	3	9	14				
SD	4	10	15	19			
AD	5	11	16	20	23		
DD	6	12	17	21	24	26	
LT	7	13	18	22	25	27	28
	DE	LC	AS	SD	AD	DD	LT

As shown in the Table 1, even imposing just combinations between 2 parameters, there are 21 possible dual combinations apart from the 7 parameters by itself.

Each of these combinations requires a deep and exhaustive work to decide how relevant is, study different options of combining and normalizing, testing results with data support, etc. Although it would be very interesting, each combination analysis could be object of a whole thesis like this. Besides, the fact that the required data is still not being generated with continuity makes impossible to carry out this task.

For this reason, the next chapters are going to explain some very simplified methods of how to do this using as example a combination of two of the candidate parameters: AS and AD (combination 16, highlighted in Table 1). Furthermore, the generic procedures is going to be explained too.

3.3. Combination quantifying

As told before, each combination considered interesting to define traffic safety should be analyzed. For doing this, it would be necessary to find the way to combine both parameters in a mathematical function to determine the safety of the interaction between them.

Speed and longitudinal distance between vehicles is one of the most important relationships between parameters. They need to be analyzed in terms of risk in case of emergency braking.

For developing a function that take a high value to represent danger and a low one to represent safety, common sense tell us to put speed in the numerator and distance in denominator, because the dangerous situation is when high speed and low distance values are taken.

The question is if one of them has more relevance than the other or they can be combined in a lineal way. For analyzing this fact, it is convenient to analyze lots of situations and make up a risk table.

In this case, a simplified table (Table 2) is shown below with an approximate value of how is wanted to behave the index in each situation:

Table 2 AS and AD interaction analysis

	AS = 25 km/h	AS = 50 km/h	AS = 100 km/h
AD = 60 m	0,0	0,10	0,30
AD = 30 m	0,05	0,25	0,70
AD = 15 m	0,20	0,50	0,90

Initially, speed seems to be more relevant, so it would be appropriate to use an exponential function to combine AS and AD parameters:

$$f(AS, AD) = \frac{AS^{\alpha}}{AD^{\beta}} \qquad [15]$$

This Table 3 represents different tested scenarios from S1 to S6 in columns, each of them has an average speed (AS) measured in km/h and an average distance among vehicles (AD) measured in meters, as shown on the second and third row. These scenarios are arranged from the highest to the lowest risk, so the intermedium safety index should decrease from S1 to S6. The following rows represent several functions parameters with several combinations of α and β.

Table 3 AS and AD combining exponents testing (Excel)

		S1	S2	S3	S4	S5	S6
	AS (km/h)	100	50	60	100	25	40
	AD m)	7	2	20	70	5	30
α	β						
1	1	14,29	25,00	3,00	1,43	5,00	1,33
2	1	1428,57	1250,00	180,00	142,86	125,00	53,33
1	2	2,04	12,50	0,15	0,02	1,00	0,04
2	2	204,08	625,00	9,00	2,04	25,00	1,78
0,5	0,5	3,78	5,00	1,73	1,20	2,24	1,15
0,5	1	1,43	3,54	0,39	0,14	1,00	0,21
1	0,5	37,80	35,36	13,42	11,95	11,18	7,30

As mentioned above, makes sense to give more weight to AS variable by using exponential functions seems to work better. This is, α should be bigger than β. Actually, seeing the results, the function seems to work better because results seem to be more coherent when this is accomplished (highlighted rows in Table 3).

3.4. Combination normalizing

Next step would be to introduce a normalization constant that allows the function to be balanced from 0 to 1 from safest to most dangerous situations:

$$f(AS, AD) = \frac{1}{A} \times \frac{AS^{\alpha}}{AD^{\beta}} \qquad [16]$$

In the Table 4, the two α-β better combinations function are shown with the convenient normalizing value ‘A’: 1.500 and 40, respectively.

Table 4 AS and AD normalized combinations (Excel)

α	β	A	S1	S2	S3	S4	S5	S6
2	1	1500	0,95	0,83	0,12	0,10	0,08	0,04
1	0,5	40	0,94	0,88	0,34	0,30	0,28	0,18

Looking at the range of values offered by the two possibilities, it seems more convenient to take the second one (highlighted in Table 4). The function is getting more continuous and coherent values according to the safety supposed in each scenario. So the proposed function of these two parameter combination could be:

$$|f(AS, AD)| = \frac{1}{40} \times \frac{AS^{1}}{AD^{0,5}} \qquad [17]$$

This function will provide a safety real value between 0 and 1based on average speed and average distance between vehicles.

In order to consolidate the function, it has been calculated for some considered safety braking distance for several speeds (scenarios from S7 to S12)[1]:

Table 5 Braking safety distances testing (Excel)

	S7	S8	S9	S10	S11	S12
AS (km/h)	20	40	60	80	100	120
AD (m)	4	16	36	64	100	144
$f(AS, AD)$	10,00	10,00	10,00	10,00	10,00	10,00
$\|f(AS, AD)\|$	0,25	0,25	0,25	0,25	0,25	0,25

As it shows the Table 5, and as expected, for all these scenarios the function takes the same value. It makes sense because the breaking safety distances are thought specially for each speed, so all of them are supposed to be equal in safety terms. As the distances are to brake safely, the index takes a low value, what also makes sense.

A very common way to normalize this kind of functions would be to use specific prepared functions for it, like for instance the 'logistic function' [12] represented in Figure 12, in its general version, with parameters to regulate:

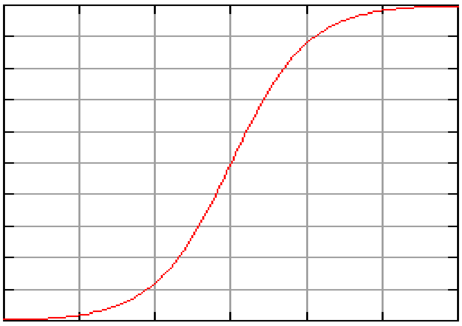

Figure 12 Logistic function graphic

$$f(x) = \frac{1}{1 + e^{-x}} \rightarrow f(x, a, m, n, t) = a \times \frac{1 + m \times e^{\frac{-x}{t}}}{1 + n \times e^{\frac{-x}{t}}} \qquad [18]$$

For normalizing some parameters, like for example, the density of the traffic, this function will fit properly. When a road is almost empty, it is not very important if there are 0, 1, 2 or 3 vehicles per kilometer. Same happens when the road is totally collapsed and its density takes very high values. Where the difference is relevant in safety terms is when the density takes intermedium values. Then, using an adapted version of the logistic function would definitely make sense.

There are many possible methods to normalize values using several mathematical functions to adapt values of parameters to the desired range. Some graphics are shown in Figure 13.

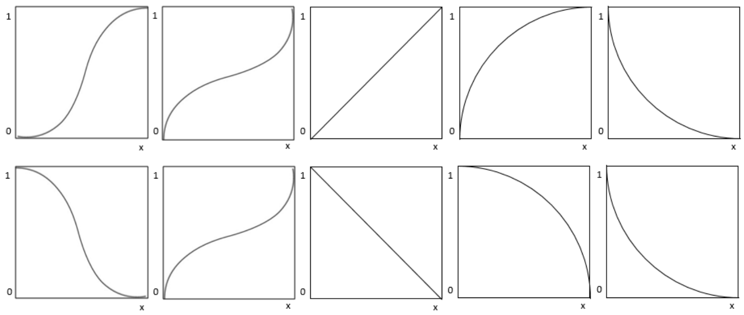

Figure 13 Several normalizing functions graphics

3.5. Study cases

Once the parameters that are going to intervene in the index formula are already normalized, the function has to be built. For doing this, based on how the risk analysis have been historically developed [2] [3], it has been decided that the function has to be a product function of the factors or parameters. Therefore, the weights of the different parameters should be placed as exponents.

Three possible function models have been proposed, studied and tested:

- Weighted safety parameters product:

$$TSI1 = \prod_{i=1}^{n} (Pn^{\alpha n}) \qquad [19]$$

- Weighted risk parameters (1 – Safety) product subtraction:

$$TSI2 = 1 - \prod_{i=1}^{n} (1 - Pn)^{\alpha n} \qquad [20]$$

- Weighted safety parameters subtraction product subtraction:

$$TSI3 = 1 - \prod_{i=1}^{n} (1 - Pn^{\alpha n}) \qquad [21]$$

A resume of some of the tests done in the function model selection is shown in the next pages. There are lots of previous tests done in the first phases of the research that are not going to be shown. These are just some of the last study cases done with proper values for the variables that allow us to see the behavior of the functions.

For doing these final tests, three previous assumptions have been made in this particular case:

- The number of parameters taken in account for the function is 3 (i = 3): P1, P2 and P3.
- Each function has been tested for the same 10 hypothetical scenarios. These scenarios are defined with values for each of the three parameters. The values assigned try to represent and describe several varied characteristic situations and combinations. The value of the 3 normalized parameters (P1 to P3) for each of the 10 scenarios (S1 to S10) are shown in Table 6:

Table 6 Parameter values for each scenario (Excel)

	P1	P2	P3
S1	0,14	0,07	0,18
S2	0,18	0,11	0,44
S3	0,51	0,17	0,43
S4	0,56	0,39	0,43
S5	0,78	0,44	0,49
S6	0,91	0,5	0,81
S7	0,91	0,88	0,83
S8	0,22	0,19	0,77
S9	0,85	0,15	0,82
S10	0,54	0,76	0,28

- The respective weight exponents for this 3 parameters depends on each model. As it will be explained later, each of the analyzed models requires a different range of ‘α’ values. So, in the tests shown in this point, this has been already taken into account, and ‘α’ values fit with the model they have been chosen for. This is, the first tests done for studying the range of ‘α’ are not attached because the value of the index was not relevant.
  The following Table 7 shows the used exponents (α1 to α3) for the tests of each model (TSI1 to TSI3):

Table 7 Alpha values for each model (Excel)

	α1	α2	α3
TSI1	0,21	0,23	0,24
TSI2	0,35	0,4	0,41
TSI3	2,15	2,01	1,9

Because of the mathematical nature of the models, the 'α's' are located in different range of values for each function, but in all the cases the weight given to the parameters is increasing from P1 to P3 (P1 is the least important, P2 the intermediate, and P3 the most important parameter for any of the three functions). Anyway, it has not been wanted to make a big difference in the weights of each one.

But, as can be seen in the Table 7 numbers, the relationship between the value of the exponent and the weight assigned to the parameter has not the same behavior for each model. While in TSI1 and in TSI2, a big value for 'α' means more weight, the opposite occurs in the third one: the more α, the less weight assigned to the respective parameter. This fact is also due to the mathematical disposition of the variables on the different models.

Once the previous assumptions have been properly explained, the 3 functions must be simplified from general (TSIX) to specific study cases (TSIX')

$$TSI1' = P1^{0,21} \times P2^{0,23} \times P3^{0,24} \qquad [22]$$

$$TSI2' = 1 - (1 - P1)^{0,38} \times (1 - P2)^{0,40} \times (1 - P3)^{0,41} \qquad [23]$$

$$TSI3' = 1 - (1 - P1^{2,15}) \times (1 - P2^{2,01}) \times (1 - P3^{1,90}) \qquad [24]$$

Here below are some of the results obtained in the tests, and some conclusions extracted from their analysis:

- Returned results for the first function testing with the specifications previously explained (Table 8):

Table 8 First model (TSI1') returned values (Excel)

TSI1

$$TSI1' = P1^{0,21} \times P2^{0,23} \times P3^{0,24}$$

	P1^α1	P2^α2	P3^α3	TSI1
S1	0,6617	0,5425	0,6626	0,2379
S2	0,6976	0,6019	0,8212	0,3448
S3	0,8681	0,6653	0,8166	0,4717
S4	0,8854	0,8053	0,8166	0,5822
S5	0,9492	0,8279	0,8426	0,6622
S6	0,9804	0,8526	0,9507	0,7947
S7	0,9804	0,9710	0,9563	0,9104
S8	0,7276	0,6825	0,9392	0,4664
S9	0,9664	0,6464	0,9535	0,5957
S10	0,8786	0,9388	0,7367	0,6077

TSI1 test conclusions:

- By simply multiplying the several factors (without exponents) the obtained value for the index are too low, because the parameters are normalized from 0 to 1.
- When the exponents are introduced to weight the different parameters, if they are greater than 1 this problem gets worse.
- By introducing weight exponents smaller than 1, the function seems to return more reasonable results.
- Best results are obtained with exponents with values around 0,15 and 0,3 (Table 8), but they are too high when the parameters take low values (for example S1 and S2).
- In general, especially for extreme parameter values, this function show too moderate values for the index
- This function is rejected because even with the most suitable exponents, it doesn't have the expected behavior for some scenarios

- Returned results for the second function testing with the specifications previously explained (Table 9):

Table 9 Second model (TSI2') returned values (Excel)

TSI2

$$TSI2' = 1 - (1 - P1)^{0{,}38} \times (1 - P2)^{0{,}40} \times (1 - P3)^{0{,}41}$$

	(1-P1)^α1	(1-P2)^α2	(1-P3)^α3	TSI2
S1	0,9486	0,9714	0,9219	0,1506
S2	0,9329	0,9545	0,7884	0,2980
S3	0,7791	0,9282	0,7942	0,4257
S4	0,7503	0,8206	0,7942	0,5111
S5	0,5886	0,7930	0,7588	0,6458
S6	0,4305	0,7579	0,5062	0,8349
S7	0,4305	0,4282	0,4836	0,9108
S8	0,9167	0,9192	0,5474	0,5388
S9	0,5148	0,9371	0,4951	0,7612
S10	0,7620	0,5650	0,8740	0,6237

TSI2 test conclusions:

- In this case, to be more coherent with a risk analysis, the function works with danger/risk (1 – P) instead with safety (P).
- Now, the problem is the opposite: the function is returning too high values when testing without exponents.
- Regulating the function by introducing the exponents, if they are smaller than 1, the results are even bigger.
- When using exponents bigger than 1, the returned values seem to be better.
- Best results are obtained with exponents with values around 0,3 and 0.5 (Table 9), but there are still a problem with low parameter values, although the situation has improved considerably.
- In general, especially for extreme parameter values, this function show too moderate values for the index
- The function is considered acceptable, but it is needed to be compared with another alternative

- Returned results for the third function testing with the specifications previously explained (Table 10):

Table 10 Third model (TSI3') returned values (Excel)

TSI3

$$TSI3' = 1 - (1 - P1^{2,15}) \times (1 - P2^{2,01}) \times (1 - P3^{1,90})$$

	1-P1^α1	1-P2^α2	1-P3^α3	TSI3
S1	0,9854	0,9952	0,9615	0,0570
S2	0,9749	0,9882	0,7898	0,2391
S3	0,7649	0,9716	0,7988	0,4063
S4	0,7125	0,8493	0,7988	0,5166
S5	0,4139	0,8080	0,7421	0,7518
S6	0,1835	0,7517	0,3299	0,9545
S7	0,1835	0,2266	0,2981	0,9876
S8	0,9614	0,9645	0,3914	0,6371
S9	0,2949	0,9779	0,3141	0,9094
S10	0,7341	0,4240	0,9110	0,7165

TSI3 test conclusions:

- By relocating the exponent, we obtain the product of 1 minus parameter to some power, instead of simply a value to a power. This makes more sense and has more complexity but also better characterization.
- Without considering exponents (with α = 1) the results are exactly the same as in TSI2.
- Studying the behavior of the function with different exponents, if they are smaller than 1, the function takes values even higher.
- Testing α > 1 the index takes more reasonable values, especially with exponents around 1,75 and 2,25 (Table 10).
- In this case the objective of reduce the index values for lower parameter values is reached, and the function has also a good behavior for all parameter ranges.
- The objective to extreme results in case of big parameters values is also reached.

Other general test conclusions:

- TS1 returns 0 if one of the parameters is 0. Nevertheless, TSI2 and TSI3 return 1 if there is a parameter that is 1. That occurs because of the mathematical reason that 0 multiplied by any number gives 0 as result.
  This rumination is the extreme case, but in general, TSI1 is a function that (without using exponents to regulate) give always lower value for the index than TSI2 and TSI3 do for same parameter values. This means that TS1 tell us that the situation is safer than TSI2 and TSI3 do. Furthermore, TSI2 y TSI3 takes the same value with all the exponents equal to 1.
- It is also important to highlight that TSI3 use to give more extreme values. This is, when all the parameters are high, the index value given by TSI3 is bigger than those given by the other two models. And with low parameters values is also TSI3 the one that gives the lowest index. Actually, this is one of the multiple reasons why TSI3 is considered the most convenient function.

In the following Table 11 and in Figure 14 graphic, the TSI value obtained with each function is represented. The scenarios have been ordered from more dangerous to more safe (based on the TSI3 criteria) for a better visualization in the graphic:

Table 11 Returned values for each model (Excel)

	TSI1	TSI2	TSI3
S7	0,91035	0,91085	0,98760
S6	0,79469	0,83486	0,95448
S9	0,59565	0,76119	0,90941
S5	0,66219	0,64582	0,75183
S10	0,60773	0,62368	0,71645
S8	0,46642	0,53875	0,63706
S4	0,58223	0,51107	0,51658
S3	0,47166	0,42574	0,40634
S2	0,34479	0,29798	0,23907
S1	0,23786	0,15056	0,05701

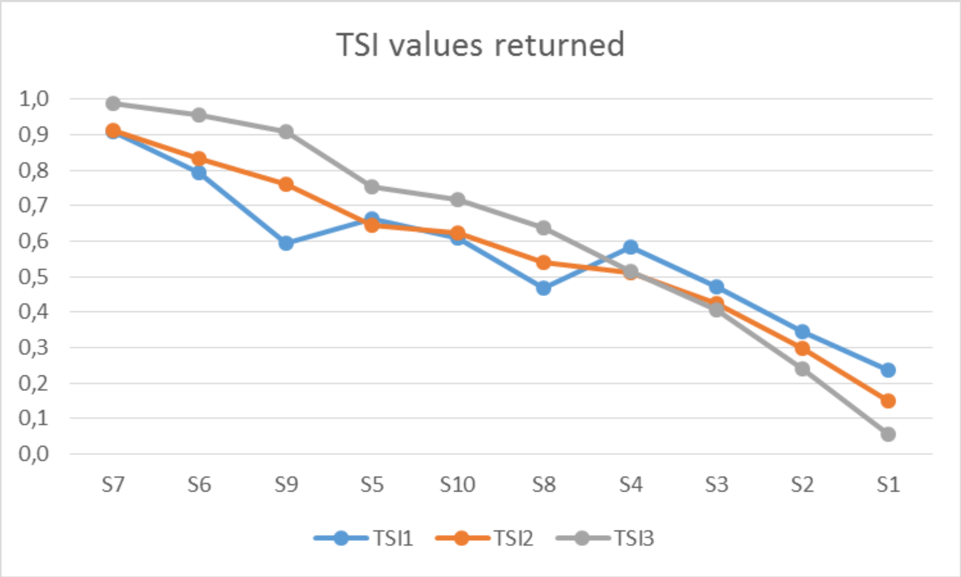

Figure 14 TSI values returned lines graphic (Excel)

As the orange line in Figure 14 indicates, TSI2 sorts the 10 scenarios in the same order than the TSI3 (grey line) does. The difference between these two models resides in that TS2 gives more moderate values than TSI3. Analyzing the scenario S7, that is the most dangerous of the proposed, we can see that the three parameters are bigger than 0,8, what means that the traffic is quite dangerous for three different reasons. When something like this occurs, it means that the general safety situation is very alarming, because in this case, the 3 considered as most important parameters are all showing danger. Same happens with safest situations (like scenario S1), if all the parameters indicates safety for different reasons, it has to be conveniently reflected in the global TSI index.

However, as is shown by the blue line in Figure 14, TSI1 sort the scenarios in a different way. This is caused for the fact that TSI1 is operating directly with the safety parameter values (P) y instead of the risk (1 - P). As explained in the 'TSI2 test conclusions' in Study cases point, in this kind of analysis is convenient to work multiplying risks, rather than safety. This is because when multiplying normalized factors, the mathematical tendency or inclination is to approach to 0 (safety). But when traffic safety is analyzed, one safe parameter cannot camouflage high values of others.

To conclude this chapter, it has to be said that the model that is going to be used hereinafter is TSI3.

3.6. TSI function definition

After analyzing and testing the 3 candidate models and because of the reasons given in the previous point, the taken decision is to use the third model (TSI3). In this way, the function to obtain the traffic safety index is defined as follows:

$$TSI = 1 - \prod_{i=1}^{n} (1 - Pn^{\alpha n}) \qquad [25]$$

Pn: value of the parameter n

αn: weight exponent for parameter n

i = (1,…,n): parameter

n: number of parameters considered by the formula

Note: As explained previously, when DE (density parameter) value is under a certain (still undefined) limit, TSI will not be calculated because the result will be not significant enough to describe traffic safety

As it can be appreciated, the function [25] is in generic, so it is totally open to introduce the convenient 'n' number of parameters 'P' and to give each of them the appropriate weight thanks to the 'α' variables.

The weight of each parameter is assigned through its exponent 'α'. The higher 'α' value, the less weight we will be giving the parameter, and vice versa. The range of values that can take these exponents depends on the correspondent specification of the concrete using model, but they should be always greater than 1.

To specify which parameters introduce in the function, the way to quantify them and the weight they should have is object for future developments and improvements of the index. Then, proposed versions of the function will be necessarily tested with the convenient specifications.

4. Applications

Now that TSI index is already defined in generic and tested with some study cases, it is needed to explain where could be applied in the future.

When already installed cameras are able to provide all the required data and the function has been studied deeper, it will be necessary to locate more cameras in as many roads around the city as possible, in order to obtain the maximum information about the traffic.

For doing this, it would be necessary to collaborate with the responsible public institutions to allow the cameras to be installed, even if the project is financed by private capital. It will be also interesting to collaborate with Taipei Police Department, because they have hundreds of cameras already installed all around the city.

Once the cameras are conveniently located and generating data for the system, it will be possible to study how the index behaves in the whole city road network. When it has been verified that this technology works on a large scale, all the information generated can be used in many possible applications that can be grouped in two:

- Those which are in the scientific or academic development studies closed to public. It can be for studying the traffic safety for public or privet institutions in order to improve
- Those which consists on share obtained information for people to use it in their daily life. The communication channel could be, for instance, a smartphone app.

The proposed application for this technology is in the second of the previously mentioned groups. It consists on sharing the traffic safety information or the roads with people, in order to allow them to make real time decisions about driving routes: ‘TSI for live decisions’.

Thus, following three points are describing how to manage this ‘TSI for live decisions’ specific application.

4.1. Safety ranges of the index

Once the safety of the road is described numerically, it would be convenient to define several ranges to show this information in a more visual way for the users to make live decisions,

The most visual method used in this kind of applications is defining a color code for the different value ranges. In this case, it could be applied something similar to the color code used by multiple map applications to define the level of traffic fluency, indicating in red or orange the level of collapse in the road, as shown in Figure 15.

Figure 15 Example of color code application

The idea would be doing something similar but using colors to indicate the traffic safety based on the TSI value, instead of the fluency like in the Figure 15 case.

In the Table 12 there is a color code and a message displayed proposal for the different TSI values ranges:

Table 12 Safety ranges for TSI values

TSI value range	Message displayed	Color code
0 – 0,3	High safety traffic	Green
0,3 – 0,6	Medium safety traffic	Yellow
0,6 – 0,8	Low safety traffic	Orange
0,8 – 1	Dangerous traffic	Red

Anyway, this is just a proposal example according to the test results for the model chosen with the assumptions done on the Study cases chapter. These safety ranges are open modifications based on the future function given values. This is, they should be defined according to the values taken by the specific used function and the returning results.

4.2. Marketing strategy

If the goal is to achieve a large market share to grow as much as possible for also achieve economic objectives, there are two possibilities to do it:

- Creating an own app and developing a hard marketing plan. The best way to reach success in the arrival to the people is through a mobile app, because of its utility and acceptation by nowadays population. Figure 16 shows a possible logo for this app.

Figure 16 TSI APP logo

- Selling the technology to a big company with an already settled faithful market, either locally, nationally or globally. This is, there is a possibility that a big company already consolidated in the mobile app sector show interest in the product.
  If this occurs, the process of arriving to the people will be faster than by an own app, because of the difference in terms of penetrating power in the market.
  In this kind of cases, it is common that the external interest arrives once the app is already operating.

4.3. Long term goals

In this point, the three main goals for this application thinking in a larger timeline:

- Territorial expansion

If it works and the project start expanding in Taipei, it would be proper to export the technology in many other cities and countries.

- Future TSI predictions

Next step for this technology once there is enough collected data consolidated would be to do predictions based on the historical information and applying also Internet of Things (IoT). Then, people would be able to decide based on future predictions done by the system instead of live information. In other words, before departing, they would have the forecast of a representative value of the road safety in the moment they would be driving over there.

- Balance effect

If the app gets implanted in people daily life, people will get used to decide their routes depending on TSI index and traffic. If this happens, less people would like to drive in roads with a high TSI value, what will change the value of the parameters in there. Then, the TSI of the roads will tend to be balanced. Furthermore, the roads with permanent high TSI could be object of study for the traffic department of the governments to improve the conditions of the traffic.

5. Epilogue

5.1. Conclusions

After reading the whole this thesis, the reader may realize that it is not a project that is closed, but the beginning of something that should be further investigated, tested and improved in the future. Attempts have been made to establish a basis on which to continue working as the data collection part progresses as well.

Therefore, once the system starts obtaining values coming from real situations through the cameras, a wide range of possibilities will be available for studying deep the behavior of the traffic captured in the parameters. Step by step, it will be possible to find lot of interactions between values that will be taken into account for adjusting parameter values. Another advantage from recording data is the fact that there will be an increasing data base with historical information about the traffic in every single road where cameras have been located. This historical data could be useful in the future to make any kind of comparison with a real consolidated reference. It will also be interesting to analyze the status of all the values in crashes situations to evaluate possible common points.

By proposing a TSI index mathematical model in a totally generic way, the option of continue incorporating future new parameters to the traffic safety analysis is open to achieve an increasing accuracy and continuous improvement. There is also the option of adding information coming from other kind of sensors that provide extra information to the system as a complement for the cameras.

But, although the model reached is generic and it has to be defined after several next contributions, some tests have been done. In order to check how the model reacts in front of non-specified parameters with reasonable weights, some assumptions were needed to be imposed. The scenarios used for testing this provisional function has not only given us the option of analyze which range of alphas were the appropriate ones for the function to return values according to the expected but also to study and compare the results from the different proposed models. Thanks to this comparison, the convenient model has been found and defined.

In conclusion, if the whole project continue ahead and some product is taken into the digital market, it is likely that the increasingly accepting mass of users of mobile applications hosted rapidly this technology for daily use. If these happens, the impact into the society should be positive in terms of safety, furthermore be helping people taking decision for their own interests.

5.2. Greetings

This thesis and his author hope that this is just a basis that give support to future contributors in the construction of a useful product that helps to people in their daily live and also to the welfare and citizen security.

I would especially like to thank the help of the Professor Shuo-Yan Chou, who gave me the opportunity to participate in one of the innovative project in which the department is involved. He has also supported me with a high quality advisory, especially in the definition of the parameters and the function model designing, but also in the development of the project in general. He has been able to guide the evolution of this work in a very professional way thanks to his enviable and respectable experience.

Likewise, I would like to thank the people who support me in my daily life in everything I embark on both personal and professional, my family, girlfriend and friends who have also indirectly collaborated as in everything I do: encouraging me

Thanks to all of them.

5.3. Bibliography

[1] Expertoautorecambios.es (2017): 'Calcular la distancia de frenado correctamente – factores que influyen en la distancia de frenado'. (Accessed 15/10/2018) Available at: https://www.expertoautorecambios.es/magazine/calcular-la-distancia-de-frenado-96

[2] Funcionpublica.gov (2012): 'Guía para la construcción de indicadores de gestión (DAFP)'. (Accesed 22/10/2018) Available at: http://www.funcionpublica.gov.co/documents/418537/506911/1595.pdf/6c897f03-9b26-4e10-85a7-789c9e54f5a3

[3] Cepal.org (2009): 'Guía metodológica: Diseño de indicadores compuestos'. (Accessed 22/10/2018) Available at: https://www.cepal.org/ilpes/noticias/paginas/9/35989/Diseno_indicadores_compuestos_ddss.pdf

[4] Dgt.es (2018): 'Portal estadístico'. (Accessed 10/11/2018) Available at: https://sedeapl.dgt.gob.es/WEB_IEST_CONSULTA/

[5] Nai-Wei Lo and Hsiao-Chien Tsai. 'A Reputation Systemfor Traffic Safety Event on Vehicular Ad Hoc Networks' (2009) (Consulted 12/11/2018)

[6] Javier Miñarro and Guillermo Mairal (2018): 'Menu design – optimal incomes combination' (2018) (Consulted 13/11/2018)

[7] Mikel Palacios and Ignacio Peironcely (2018): 'Data analysis – NBA statistics' (2018) (Consulted 13/11/2018)

[8] Ingacio Sanz and Javier Miñarro (2018): 'Communication Aantennas optimization' (2018) (Consulted 13/11/2018)

[9] Circulaseguro.com (2013): 'La 'otra' distancia de seguridad olvidada'. (Accessed 15/11/2018) Available at: http://www.circulaseguro.com/la-otra-distancia-de-seguridad-olvidada/

[10] Tutorialspoint.com (2018): 'Artificial Intelligence fuzzy logic system'. (Accessed 12/12/2018) Available at: https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_fuzzy_logic_systems.htm

[11] Paradigmadigital.com (2017): 'Inteligencia Artificial como servicio: reconocimiento de imágenes'. (Accessed 03/01/2019) Available at: https://www.paradigmadigital.com/techbiz/inteligencia-artificial-servicio-reconocimiento-imagenes/

[12] Quora.com (2017): 'How should I normalize features in machine learning assuming logistic regression?'. (Accessed 03/01/2019) Available at: https://www.quora.com/How-should-I-normalize-features-in-machine-learning-assuming-logistic-regression

[13] Fw2.com (2018): 'CÍCLOPE: detección de vehículos y objetos mediante visión artificial'. (Accessed 04/01/2019) Available at: http://fx2.com.uy/ciclope-deteccion-de-vehiculos-y-objetos

[14] Amsts.tandfonline.com (2011): 'Probabilistic Risk Analysis: Foundations and Methods'. (Accessed 04/01/2019) Available at: https://amstat.tandfonline.com/doi/abs/10.1198/016214502760301264?journalCode=uasa20#.XDNxkVUzbIU

[15] Mondtools.com (2018): 'Risk analysis and risk management'. (Accessed 04/01/2019) Available at: https://www.mindtools.com/pages/article/newTMC_07.htm