



Design of a new measurable approach for the qualification of the behaviour of an autonomous vehicle

Yacine Mezali, Mohamed Khaledi, Loïc Coquelin, Rémi Régner, Jordan Martin

► To cite this version:

Yacine Mezali, Mohamed Khaledi, Loïc Coquelin, Rémi Régner, Jordan Martin. Design of a new measurable approach for the qualification of the behaviour of an autonomous vehicle. European Control Conference, Jul 2022, London, United Kingdom. hal-03603685

HAL Id: hal-03603685

<https://hal.science/hal-03603685>

Submitted on 10 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Design of a new measurable approach for the qualification of the behaviour of an autonomous vehicle

Yacine Mezali¹, Mohamed Idriss Khaledi¹, Loïc Coquelin^{1,2}, Rémi Régnier^{1,2}, Jordan Martin^{1,3}

Abstract—The safety assessment of autonomous driving system is a major challenge in the automotive industry and the role of simulation in development and testing of autonomous vehicles has become predominant to significantly reduce the hundreds of millions of miles required to demonstrate the safety performance of such systems.

In this paper, a novel methodology is presented to assess automated vehicles safety performance based on a multi-factorial analysis of severity indicators in the vicinity of the under test self-driving car. The set of severity indicators includes commonly used time intervals (Inter Vehicular Time, Time to Collision, Time to Steer, ...), distance-based indicators, traffic congestion indicators and a newly developed indicator relying on overlapping geometrical regions. Unsupervised clustering techniques are then used to investigate the correlations, dependence among the whole set of indicators. To address the problem of combining these heterogeneous quantities to derive a global measure of dangerousness for a given scenario, appropriate scaling is performed and various aggregation methods are tested against cut-in, cut-out and cut-through scenarios.

I. INTRODUCTION

At the beginning of the 21st century, the arrival of autonomous vehicles was predicted for 2020 but only several Automated Driving Systems (ADS) aim to ensure driver safety and comfort are on the road today. The most advanced homologated system is the Honda Legend in Japan, which only reaches level 3 autonomy, far from the level 5 hoped for. Beyond the legal and ethical aspects, the lack of tools to ensure the safety of autonomous vehicles is one of the major obstacles. For example since 2014, EuroNCAP has integrated several driver assistance systems during their evaluation process, which are grouped under the title “Safety Assist” but with very small number of use cases (three for AEB), or only to work with data provided by car manufacturers. In view of the diversity of the situations to be tested (considering possible failure, degradation, adverse conditions and attacks, climate, infrastructure, sensors, communication bus etc.), it is increasingly difficult to perform these tests only on real test tracks or open road (real controlled environment). Indeed It is not realistic to test every combinations of sensor input and driving situation from physical testing, only a limited number of use cases that would be encountered on a defined set of test routes is possible because the number of miles an autonomous vehicle would have to accumulate to be tested in all situations is incalculable, but likely be

measured in billions. Consequently, complementary solutions must be found in order to be able to consider a large number of representative situations and data. For this, the use of testing and simulation tools is becoming essential. But for simulation testing, it is necessary to develop procedures, measurement tools, and ground truth. Currently, there are very few indicators and no standardized methods to perform such simulated tests. The topic of the article is to study this issue and to propose evaluation indicators providing a first alternative and effective solution for evaluation and validation of AV by simulation.

New problems are also brought about by simulation. One of them is the need to quantify in a non-binary way the level of success in a test (other than accident/no accident), in a continuous way and allowing to compare between several trajectories which one was the safest.

One of the examples of indicators that is beginning to emerge is the Responsibility-Sensitive Safety (RSS) model was developed and implemented by Intel/Mobileye [SS17] but classically, “Time to X” type measures stay the indicators standard. RSS determine safe following distance to avoid collisions in the worst possible conditions (hard braking of the following vehicle during an acceleration phase of the EGO vehicle). But these different indicators usually only look at one aspect of safety such as the distance to the vehicle in front in the lane and do not analyse the full complexity of a scenario. This is why we propose in this article to write a new metric capable of analysing the dangerousness of a situation in a more global way.

The objective is to obtain a final metric capable of:

- Continuous metric (not binary and not limited to simple thresholds)
- Translates the likelihood of a hazard or severity (accident)
- Translates the criticality of a behaviour associated with a scenario
- Allows the comparison of several trajectories in order to find the best one in a given situation.
- The metric must detect all dangerous sections.

To do this, we will use several microscopic indicators which we will then aggregate into a global macroscopic metric. In this chapter 2, we will present the selected indicators and will come back to the new metric developed in this work. Then we will present the processing required for the aggregation of these indicators and the aggregation procedure itself and finally we will present an example of the use of this methodology for measuring dangerousness.

*This work was supported by Institute of Research and Technology SystemX

² National laboratory for metrology and testing (LNE), France
name.surname@lne.fr
AVSimulation

Name	Type	Variable	Unit	Description	Reference
Inter vehicular time	EGO centric	IVT	(s)	Time required for the EGO vehicle to travel the distance to the target at constant speed	Appendix (A)
Time to Break	EGO centric	TTB	(s)	Time after which a braking maneuver has to be started to prevent the collision. If the TTB is smaller than 0, a collision can not be avoided by braking	[AIMK14]
Time to Steer	EGO centric	TTS	(s)	Time after which an evasive maneuver has to be started to prevent the collision. If the TTS is smaller than 0, a collision can not be avoided by steering	[AIMK14]
Time to collision advanced	EGO centric	TTC_a	(s)	Extension of the original Time To Collision (TTC) [Hay72] in the 2D plane	[JW14]
Distance to Time to collision advanced	EGO centric	$dTTC_a$	(m)	Distance calculated based on TTC_a	Section (E)
Minimum Lateral Distance	EGO centric	MIN_LAT_D	(m)	Minimum distance between the EGO vehicle and the target vehicles driving on its adjacent lanes	Appendix (H)
Region proportion	Region centric	R_PROP	(%)	Occupancy rate of a danger zone	Section (II-B)
Lateral avoidance acceleration	EGO centric	ACC_{lat}	(m.s ⁻²)	Instantaneous lateral acceleration required to avoid a vehicle in its lane	Appendix (F)
Longitudinal avoidance deceleration	EGO centric	DCC_{long}	(m.s ⁻²)	Instantaneous longitudinal deceleration required to avoid a vehicle in its lane	Appendix (G)
Level of Velocity Heterogeneity	Region centric	LVH	(#)	Ratio of targets speed standard deviation and targets mean speed in the vicinity of the EGO vehicle (mean and standard deviation calculations include the EGO vehicle speed)	Appendix (I)
Mask Occupancy Ratio	Region centric	MOR	(%)	Occupancy rate of the road	Appendix (J)

TABLE I
LIST OF MICROSCOPIC INDICATORS USED TO ASSESS THE SEVERITY FOR A GIVEN SCENARIO

II. MICROSCOPIC MEASURES USED TO ASSESS THE SEVERITY OF A SITUATION

A. Selected metrics

Severity indicators are computed at each time step, they mainly fall under two categories.

- **EGO centric** : these indicators assess the criticality impacting the EGO directly, their computation involves the actions and/or the reactions required for the EGO to avoid the collision. The computation of the EGOcentric indicators involves the target in front of the EGO (vehicle in its path).
- **Region centric** : these indicators describe the severity inside the region of interest (ROI) of the EGO, they assess the overall criticality in the ROI, their computation involve all the actors inside the ROI.

The list of severity indicators selected for the design of a global measure of severity is shown in table I. Coming section details a newly developed indicator, called R_PROP , added to the list of more traditional severity indicators.

B. Region proportion metric

In classical dangerousness metrics, there are traditionally two problems:

- They generally consider only one so-called priority target;

- They do not allow for the coupling of lateral and longitudinal dangerousness.

It is to address these two problems that we propose a new approach with the region proportion metric (R_PROP) that can take into account several targets and their relative position with respect to the trajectory of the EGO vehicle.

The construction of this metric is done in several steps:

- 1) At the level of the EGO vehicle will be positioned a region of interest with rectangular or ellipsoidal shape (deformable or not with respect to the curvature of the road or the heading of the EGO vehicle), whose length and width will be dependent on the dynamic parameters of the vehicle (speed, acceleration, steering...). It's the blue region in the figure 1. This area of interest corresponds to a surface where it is considered that the presence of obstacles leads to an increase in the danger of the situation for the EGO vehicle. In our implementation of this metric, we considered a rectangular area (deformable following the shape of the track) of length "D" equal to to the maximum between $dTTC_a$ (Appendix (E)) and a distance corresponding to an IVT of 2.2s.
- 2) For each target crossing the area of interest, we look at the portion of the area rendered unusable for the EGO vehicle and then analyse what remains of the area not impacted by the targets: the hatched area in the figure 1.

- 3) The metric is calculated as 1 minus the ratio of the hatched area and the complete interest region (for a single target, a score equal to 0 meaning no hazard and 1 meaning collision). Let n the number of target, $R(D)$ the hazard interest area and R_i the area made inaccessible by a i target (ground occupation of the target and the area directly in front of the target), R_PROP can be written as :

$$R_PROP = 1 - \frac{R(D) - R(D) \cap \bigcup_{i=1}^n R_i}{R(D)} \quad (1)$$

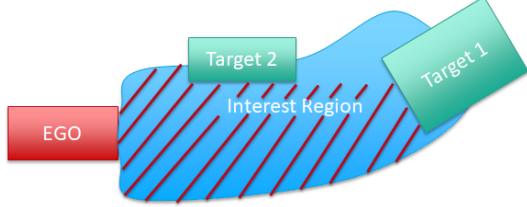


Fig. 1. Example of implementation of the metric, we build the zone of danger here in blue according to the dynamic parameters of the vehicle and we look at the space that remains once we remove the space made inaccessible by the targets around the ego.

Let us illustrate this on a simple implementation case as in the figure 2:

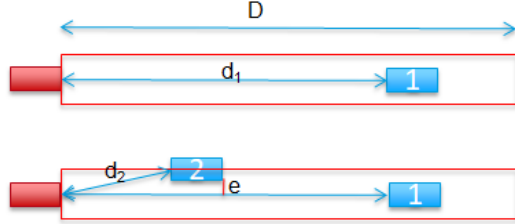


Fig. 2. Example of on a simplified case of generalisation on the lateral

Considering the geometry proposed in figure 2, for the simple configuration of a target at a distance d_1 (top diagram) the formula amounts (because we are in a simple case of a rectangle) here to $S_1 = 1 - d_1/D$. If we add another target off-centre with respect to the trajectory of the EGO vehicle (offset by a parameter e) as on the middle geometry, this lateral target can also be taken into account by our metric. Here we will have in addition to S_1 , a second value to calculate corresponding to this second target $S_2 = (1 - 2 * e/l)(1 - d_2/D)$.

Here the term in $1 - 2 * e/l$ allows us to put a weight on the lateral in relation to the longitudinal. For a smoother approach, this linear term of eccentricity can also be replaced by using a normal distribution. It is interesting to note that the width of the danger zone l can be chosen to be different from the width of the vehicle depending on the choice of the designer and the importance given to taking the lateral into account. The weight makes it possible to give a greater danger to a target at an equivalent distance

from the axis of the EGO trajectory than to a target offset from this axis. Finally, there could also be a second danger zone covering the rear of the vehicle to take account of vehicles behind the EGO vehicle, for which we will add a weight factor.

The advantage of separating each target in this way is to be able to look at the impact of each target on the metric and to know which target is the most dangerous at a given time, but the idea of the metric is still to be able to look at the cumulative impact of all the targets on the interest region.

So in the case of n targets to be taken into account, three analyses can be made:

- 1) An independent target analysis;
- 2) An analysis on the max of S_i ($\max(S_i(t))$) to take into account only the most dangerous target at each time t ;
- 3) An analysis on the area resulting from the set of targets intersecting the entire area to take into account in the set of targets simultaneously (see figure 1).

III. PRE-PROCESSING OF SEVERITY INDICATORS

Severity indicators listed in table (I) are, by definition, heterogeneous: different units, derived from different physical quantities (acceleration, time, distance, ...), etc. Prior to combining such set of indicators, common practice is to simply apply a specific scaling function, ϕ , so that they all lie in a specific range. The shape and support of the scaling function shall be designed carefully at this stage in order to reflect the severity of each indicator individually.

Let $s_k(t)$ denotes the k^{th} indicator value at time t , and $\tilde{s}_k(t) = \phi_k(s_k(t))$ the scaled version of $s_k(t)$ using scaling function $\phi_k : R \rightarrow [0, 1]$ is given by the following equation:

$$\phi_k(s_k(t); \alpha_k, \beta_k) = 1_{\{s_k(t) \in [\alpha_k, \beta_k]\}} \phi(s_k(t)) \quad (2)$$

where the interval $[\alpha_k, \beta_k]$ stands for the studied severity domain for the k^{th} indicator. The shape of the scaling function is purely defined by ϕ in equation (2). For sake of simplicity, ϕ is either the cumulative distribution function (CDF) of the Gamma distribution ($\phi(x) = \frac{\gamma(k, x/\theta)}{\Gamma(k)}$, where $\Gamma(k)$ is the gamma function evaluated at k , γ is the lower incomplete gamma function, k and θ are the shape and scale parameter, respectively) or the CDF of the Gumbel distribution ($\phi(x) = e^{-e^{-(x-\mu)/\beta}}$ where μ and β are the location and scale parameter, respectively). Expert knowledge is clearly required at this stage to set realistic severity domain $[\alpha_k, \beta_k]$ for each indicator and also to model the scaling function (setting (k, θ) for Gamma CDF or (μ, β) for Gumbel CDF). Moreover, it should be noted that in the case of distance-based indicators, the upper bound of the studied severity domain will depend on the EGO speed.

To illustrate the scaling process, let's consider the example of IVT values: an IVT value higher than 2.2 s can be considered as safe which can be modelled as $\text{supp}(\phi_k) = [0, 2.2]$ by setting $\alpha_k = 0$ and $\beta_k = 2.2$. Figure (3) shows the scaled

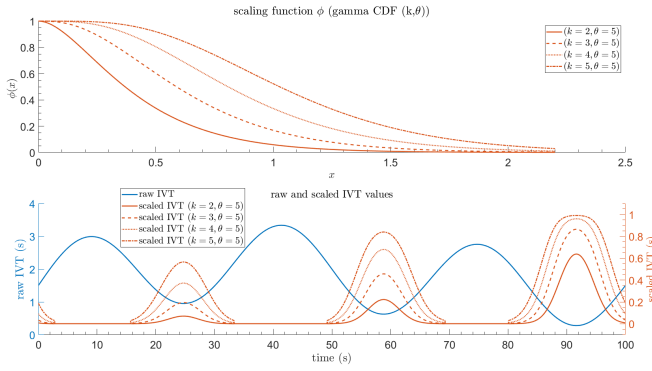


Fig. 3. Illustration of the scaling process for simulated IVT values

version of simulated IVT values when using Gamma CDF as the scaling function for various set of parameters (k, θ) .

Once they are properly scaled, next step is to investigate the relationship among severity indicators prior to their aggregation.

IV. UNSUPERVISED CLUSTERING

A set of microscopic metrics has been defined on which some treatments have been applied to make them comparable, so the last step for the construction of the global metric is to aggregate them. Aggregating many severity indicators to obtaining a single severity score is cumbersome task. In order to address this issue, we need to understand the relationship between the different severity/criticality indicators. In fact, some indicators may express the same criticality, others may bring complementary information. Therefore, understanding the similarity/ dissimilarity of our indicators is crucial. So the first task to do a efficient aggregation of the indicators is to study the correlations, similarities and orthogonalities between each microscopic metric. Two strategies to do this based on time series clustering techniques are described in the rest of this section.

A. Severity indicators clustering using time series imaging

It is well-known that Euclidean distance is not appropriate for time series clustering since it disregards time miss alignment [S.A15]. An alternative to the Euclidean distance is the Dynamic Time wrapping (DTW) [S.A15]. One major downside of DTW is that it is not a distance in the mathematical sense but rather a measure besides DTW is computationally expensive.

Imaging time series, that is transforming time series into images. In fact, time-series can be characterized by a distinct recurrent behavior such as periodicities and irregular cyclicities that can be encoded using The Gramian Angular Fields (GAF) [ZT15]. The GAF matrices captures the temporal correlations $I(t_i), I(t_j)$ of each indicator [ZT15]. The construction of GAF matrices requires normalizing the time series $-1 \leq a < b \leq 1$, computing the polar coordinates of the normalized values by taking the arcos function. Finally it computes the cosine of the sum of the angles for the

Gramian Angular Summation Field (GASF) or the sine of the difference of the angles for the Gramian Angular Difference Field (GADF).

$$\begin{aligned}\phi_i &= \arccos(I_i(t)) \\ GASF_{ij} &= \cos(\phi_i + \phi_j), \forall i, j \in 1..n \\ GADF_{ij} &= \cos(\phi_i - \phi_j), \forall i, j \in 1..n\end{aligned}\tag{3}$$

We use the Kmeans algorithm to do clustering based on the GADF matrix, the number of clusters is validated using the elbow method that indicates two main clusters which are:

- **Cluster 1:** *IVT, TTC_a, dTTC_a, MIN_LAT_D, TTS, DCC_{long}.*
- **Cluster 2:** *MOR, ACC_{lat}, LVH, R-PROP, TTB.*

B. Severity indicators clustering using Kolmogorov Smirnov distance

Another technique that we used for time series clustering is based on KS (Kolmogorov-Smirnov) tests [Hod58] and more precisely (the two-sample KS test). In general a Kolmogorov-Smirnov test is a non-parametric test of the equality of continuous (or discontinuous) one-dimensional probability distributions. This type of test can be used to compare a sample with a reference probability distribution (one-sample K-S test), or to compare two samples (two time series in our case). The two main results provided by that test is the p-value and the D-statistic which represents the maximum distance between the CDFs of the two time series, the closer this indicator is to zero, the more it's probable that our two series are derived from the same distribution. Reasoning based on p-value or D-statistic leads to the same scientific interpretation, but as a choice, we have chosen to work with the D-statistic given its simplicity to explain and present it. Figure 4 shows the KS score matrix with the different D-statistic values for the indicators we have considered.

The next step consists in applying a hierarchical clustering based on the KS-test score matrix that we have calculated. The different steps of this clustering are as follows:

- 1) We start by calculating the dissimilarity between the N indicators (using the KS tests) and we attribute each indicator to a cluster;
- 2) The two closest indicators based on the KS score are grouped together to form a new cluster;
- 3) The dissimilarity between this cluster and the remaining N-2 indicators is recalculated using the KS score and iteratively we repeat the same process from the previous step until only a single cluster is left;
- 4) At each iteration, we are merging/adding the clusters together (that's why this algorithm is also known as "the additive hierarchical clustering"). [Mul11][ZBJ01]

C. Clustering results analysis

Our clustering study uses three simulated drivings of twenty minutes on motorways, each of which contains many

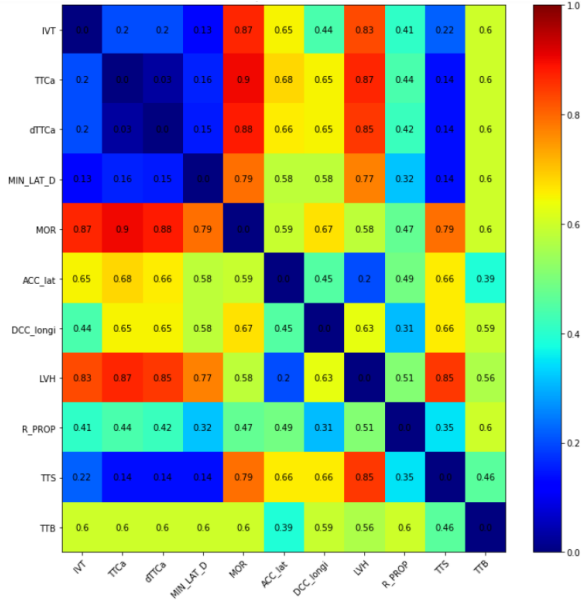


Fig. 4. KS score matrix (D statistic)

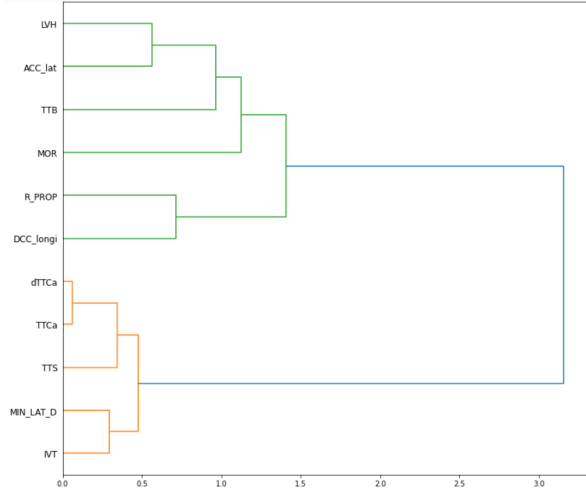


Fig. 5. Results of Time Series Clustering

critical maneuvers including dangerous cut-ins or dangerous cut-outs. Each simulated driving can be regarded as a collection of time series : the dynamics of each actor, the infrastructure, etc. Consequently, each simulated driving is recorded in csv format. The computation of the entire set of indicators is done offline.

We can easily notice that : the results of the two approaches presented above give similar results in terms of clustering, which reassures us about the distribution between the two classes.

Thus, KMeans of GADF and the KS clustering allowed us to separate the indicators into two major classes. The first class contains all the indicators that are related to the notion of Mask Occupancy Ratio, traffic density, target avoidance, and in the second class we find the metrics that are more severe in terms of severity such as collision times, collision

distance. In addition, this clustering provided us with information about the complementarity/orthogonality between the indicators to better aggregate them. The knowledge of the similarity/ dissimilarity of our indicators will allow us to create more complex aggregation models, which we will see in the next section.

V. AGGREGATION TO OBTAIN GLOBAL MACROSCOPIC INDICATORS

This section deals with the problem of combining the pre-processed EGO and region centric severity indicators proposed in section (II)-(III) to a single output time series to reflect the severity at each time step for a studied scenario, denoted $s(t)$. One of the simplest multi-criteria decision analysis is the weighted average model that consists in defining relative weight of importance for each indicator and taking the average of the weighted linear combination as an indicator of the severity ($s(t) = \frac{1}{K} \sum_{k=1}^K w_k s_k(t)$).

Weighted average ensemble technique is related to voting ensemble techniques where the arithmetic mean ($s(t) = \frac{1}{K} \sum_{k=1}^K s_k(t)$) is commonly used in regression problems. Other related data aggregation techniques can be directly derived by replacing the average by other statistical measures such as the median, n^{th} percentile, etc. However, aggregation methods in the context of safety assessment should at least be a robust collision detection estimator. For that reason, naive ensemble models cannot be considered alone. In order to fulfill this requirement, the proposed aggregation technique, denoted empirical aggregation hereafter, is hybrid since it uses high-level priority rules among indicators and also computes weighted combination of indicators.

The proposed empirical method has been derived based on the analysis/interpretation of the results (EGO and region centric indicators) obtained from a large set of scenarios. It also account for the clustering results presented in section (IV). So in this approach, we have four severity classes, namely C_1, C_2, C_3, C_4 ranging from low to high severity:

- C_1 : low severity score varying from 0% to 50% excluded.
- C_2 : medium severity score ranging from 50% to 80% excluded,
- C_3 : high severity score ranging from 80% to 100% excluded (near accident cases),
- C_4 : maximum severity of 100% (collision occurred).

Algorithm 1 details how to calculate the global severity score $s(t)$ at each time step. First, we start by defining $(\delta_l)_{l=1\dots 4}$, $(\Delta_{(l)-(l+1)})_{l=1\dots 3}$ and $(SCl)_{l=1\dots 4}$ representing the severity baseline for each class, the gap between consecutive classes and the severity indicators chosen for each class, respectively (lines 1-11).

The algorithm first checks for an eventual collision: if any of the severity indicators in C_4 has a value of 1 then it returns $s(t) = 1$ (lines 12-14). If no collision occurred, the algorithm computes the severity scores for the remaining classes C_1, C_2, C_3 (lines 15-27) and it returns the global

severity score as the maximum severity over these 3 classes (line 28).

Algorithm 1 Empirical aggregation for a given time step t

```

1:  $\delta_1 = 0$  ▷ baseline severity C1
2:  $\delta_2 = 0.5$  ▷ baseline severity C2
3:  $\delta_3 = 0.8$  ▷ baseline severity C3
4:  $\delta_4 = 1$  ▷ baseline severity C4
5:  $\Delta_{(1)-(2)} = 0.5$  ▷ severity gap between C1 and C2
6:  $\Delta_{(2)-(3)} = 0.3$  ▷ severity gap between C2 and C3
7:  $\Delta_{(3)-(4)} = 0.2$  ▷ severity gap between C3 and C4
8:  $S_{C1} = \{ \text{all indicators at time } t \}$ 
9:  $S_{C2} = \{ IVT, TTCa, MIN\_LAT\_D, R\_PROP, ACC_{lat}, DCC_{long}, TTS, TTB \}$ 
10:  $S_{C3} = \{ IVT, TTCa, MIN\_LAT\_D, R\_PROP \}$ 
11:  $S_{C4} = S_{C3}$  ▷ CLASS C4

12: if any  $(s \in S_{C4}) = 1$  then
13:    $s(t) = 1$  return
14: end if ▷ CLASS C1,  $l = 1, 2, 3$ 

15: for  $l = 1..3$  do
16:    $K_l = card(S_{C_l})$ 
17:    $S_{C_l} = \{ s_k(t), k = 1, \dots, K_l \}$ 
18:    $s_{tmp}^{[k]}(t) = 0, \forall k \leq K_l$ 
19:   for  $k = 1..K_l$  do
20:     if  $\delta_l \leq s_k(t) < \delta_{l+1}$  then
21:        $w_k = 1 - (card(S_{C_l}) - 1) * 0.1$ 
22:        $w_j = 0.1, \forall j \neq k$ 
23:        $s_{tmp}^{[k]}(t) = \delta_l + \Delta_{(l)-(l+1)} \sum_{k=1}^{K_l} w_k s_k(t)$ 
24:     end if
25:   end for
26:    $s^{[l]}(t) = \max_{1 \leq k \leq K_l} (s_{tmp}^{[k]}(t))$ 
27: end for
28:  $s(t) = \max_{1 \leq l \leq 3} (s^{[l]}(t))$ 

```

We have tested our aggregation method against various cut-in, cut-out, cut-through scenarios which were generated using the simulation software SCANer Studio. The simulator enables us to record all the relevant information regarding all the vehicles in the driving simulation. This information includes all the parameters required to compute our severity indicators. At each time stamp, we record the state vector of each vehicle in the driving simulation. Figure (6) illustrates a cut-in scenario, Figure (7) illustrates the same scenario and the associated severity score computed using the empirical aggregation method. In this particular example, the calculated score reaches its maximum value in the middle of the cut-in maneuver.



Fig. 6. Full-Cut-In

VI. CONCLUSION

Vehicle safety assessment in simulation remains a major challenge for the approval of autonomous vehicles. Developing severity indicators and designing high level severity scores help in assessing their safety. The proposed global severity score combines commonly used severity indicators as well as newly developed indicators in order to tackle complex scenarios: multiple targets, longitudinal and/or lateral hazard, ... The presented severity measure is used to assess the under test vehicle safety (EGO vehicle here) yet

the methodology is applicable on the whole set of targets in a given scenario allowing to derive a measure of dangerousness for any given driving scenario.

REFERENCES

- [AIMK14] Carlo Ackermann, Rolf Isermann, Sukki Min, and Changwon Kim. Collision avoidance with automatic braking and swerving. *IFAC Proceedings Volumes*, 47(3):10694–10699, 2014. 19th IFAC World Congress.
- [Hay72] John C. Hayward. Near-miss determination through use of a scale of danger. *Highway Research Record*, 1972.
- [Hod58] J.L. Jr. Hodges. The significance probability of the smirnov two-sample test. *Arkiv fur Matematik*, 3, pages 469–86, 1958.
- [Jan05] J. Janson. *Collision Avoidance Theory with Application to Automotive Collision Mitigation*. PhD thesis, Linköping Universitet, 2005.
- [JW14] Stewart Worrall Eduardo Nebot James Ward, Gabriel Agamenoni. Vehicle collision probability calculation for general traffic scenarios under uncertainty. *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pages 986–992, 2014.
- [MMM01] P. H Bovy M. M Minderhoud. Extended time-to-collision measures for road traffic safety assessment. *Accident; Analysis and Prevention*, pages 89–97, 2001.
- [Mul11] D. Mullner. Modern hierarchical, agglomerative clustering algorithms. 2011.
- [S.A15] T.Wah S.Aghabozorgi, A.Shirkhorshidi. Time series clustering a decade review. *Elsevier : Information Systems*, 52, 2015.
- [S.H18] U.Eberle F.Köster S.Hallerbach, Y.Xia. Simulation-based identification of critical scenarios for cooperative and automated vehicles. 2018.
- [SS17] S. Shalev-Shwartz S. Shammah and A. Shashua. On a formal model of safe and scalable self-driving cars. <https://arxiv.org/abs/1708.06374>, 2017.
- [ZBJ01] T. S. Jaakkola Z. Bar-Joseph, D. K. Gifford. Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics*, 2001.
- [ZT15] Z.Wang and T.Oates. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. 2015.

APPENDIX

A. Inter-Vehicle Time (IVT)

This metric is simply the calculation of the inter-vehicular time between the EGO vehicle and the primary target, i.e. the time required for the EGO vehicle to travel the distance to the target at the current constant speed. Let $v(t)$ be the speed of the EGO vehicle and $d(t)$ the distance to the target then the mathematical formula is :

$$IVT(t) = \frac{d(t)}{v(t)} \quad (4)$$

B. Time to Break (TTB)

Time after which a braking maneuver has to be started to prevent the collision. If the TTB is smaller than 0, a collision can not be avoided by braking [AIMK14].

C. Time to Steer (TTS)

Time after which an evasive maneuver has to be started to prevent the collision. If the TTS is smaller than 0, a collision can not be avoided by steering [AIMK14].

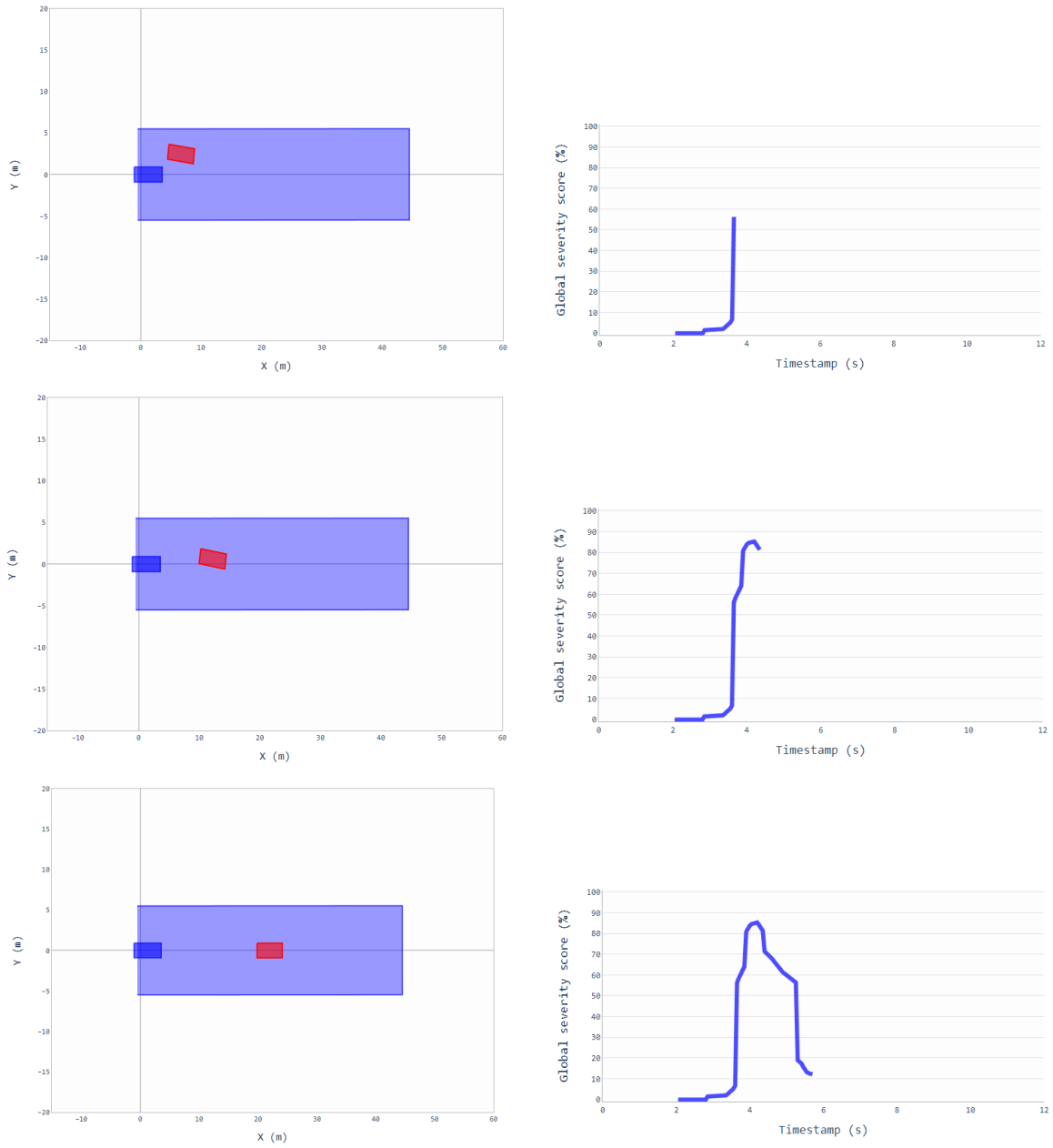


Fig. 7. Cut-in scenario example: calculation of the severity score $s(t)$ at the beginning, middle and end of the maneuver; the blue rectangle represents the ROI for the EGO vehicle.

D. Time to collision advanced (TTC_a)

This metric is the calculation of the time to collision between the EGO vehicle and a target. This extension of the TTC collision time is used in Mobileye systems [MMM01]. Traditionally, the collision time is always calculated on one dimension. One is usually interested in either the longitudinal TTC or the lateral TTC, which is in itself a very important simplification of the problem. [JW14] proposes to calculate the collision time in the more realistic 2D plane between the two closest points of each vehicle as illustrated in the figure 8:

Let $(x(t), y(t))$ the EGO vehicle position vector, $(v_x(t), v_y(t))$ the EGO vehicle speed vector, $(\gamma_x(t), \gamma_y(t))$

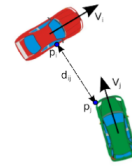


Fig. 8. Calculating the distance between two vehicles

the EGO vehicle acceleration vector, $(x^T(t), y^T(t))$ the target position vector, $(v_x^T(t), v_y^T(t))$ the velocity vector of the target and $(\gamma_x^T(t), \gamma_y^T(t))$ the acceleration vector of

the target, we try to estimate X which corresponds to the collision time, noted t_c , in the following second order equation [JW14] :

$$d_{ij} + \dot{d}_{ij}X + \frac{1}{2}\ddot{d}_{ij}X^2 = 0 \quad (5)$$

E. Distance to Time to Collision advanced ($dTTC_a$)

This metric is the derivative of the previous one in that it is based on the calculation of the pre-collision distance between the EGO vehicle and a target and thus the distance to the TTC_a . The formulation is therefore

$$dTTC_a(t) = TTC_a(t) \times v_x(t)$$

F. Lateral Avoidance metrics

Here we model indicators that highlight the the instantaneous lateral acceleration ACC_{lat} or the instantaneous deceleration required to avoid a vehicle in its lane DCC_{long}

The lateral avoidance acceleration can be written as follow ([Jan05]):

$$\begin{aligned} ACC_{lat}(t) = & \gamma_y^T(t) - \frac{2}{TTC_a(t)^2}(- (y^T(t) - y(t)) \\ & \pm \frac{(W^T \sin(\delta(t)) + W^T \sin(\delta^T(t)))}{2} \\ & - (v_y^T(t) - v_y(t))TTC_a(t)) \end{aligned} \quad (6)$$

With W the width of the EGO vehicle and W^T the width of prior target

G. Longitudinal Avoidance metrics

The instantaneous deceleration required to avoid a vehicle in its lane can be written :

$$DCC_{long}(t) = \min \left(\gamma_x^T(t) - \frac{(v_x^T(t) - v_x(t))^2}{2 \times d(t)}, 0 \right) \quad (7)$$

H. Minimum Lateral Distance

As part of measurement of the lateral dangerousness, we introduce a new metric that allows to calculate the minimum distance between the EGO and the target vehicles that are on its adjacent lanes (figure 9) at each time step t and then take the minimum of these calculated distances, hence the name Minimum Lateral Distance (MIN_LAT_D).



Fig. 9. Minimum lateral distances between the EGO vehicle and its targets

I. Traffic metrics

As part of the measurement of the dangerousness of a situation, we also wanted to put forward indicators on the condition of the traffic which add complementary information to the previous indicators drawing in particular on the work of Hallerbach [S.H18]. The level of velocity heterogeneity indicator (LVH) measures the homogeneity of the speed behaviour of the road actors in the vicinity of the EGO vehicle, as a high degree of heterogeneity increases the chances of having risky situations due to the speed differentials between the different actors.

A simple approach to calculate it is to make a ratio of the standard deviation of speeds to the mean of speeds in the vicinity of the EGO vehicle, i.e. :

$$LVH(t) = \frac{\sigma_v(t)}{\mu_v(t)} \quad (8)$$

with the mean value

$$\mu_v(t) = \frac{1}{N} \sum_{i=1}^N v^{[i]}(t)$$

and the standard deviation

$$\sigma_v(t) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (v^{[i]}(t) - \mu_v(t))^2}$$

J. Mask Occupancy Ratio

Among the metrics that were added, we find “Mask Occupancy Ratio” (MOR), its principle role is to provide information about the number of vehicles that are within a well-defined area around the EGO taking into account its longitudinal speed. MOR is calculated using this mathematical formula:

$$MOR(t) = \frac{1}{nbL} \sum_{i=1}^{nbL} \frac{nbVIL_i(t)}{nbVILmax(t)} \quad (9)$$

with nbL the number of road lanes, $nbVIL_i$ represents the number of vehicles that are within the considered proportion of lane i at timestamp t , $nbVILmax$ is the max number of vehicles that can exist in that proportion of lane length L (e.g. $L = 50m$) at timestamp t and is calculated as:

$$nbVILmax(t) = \frac{L}{Rd(t)} \quad (10)$$

with $Rd(t)$ is the required distance that depends on EGO speed (see. 10).

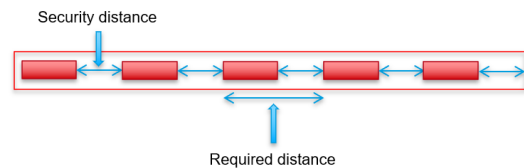


Fig. 10. Security and required distances, respectively Sd and Rd , on a lane section of length L