SAFETY METRICS ASSESSMENT USING LOGGED VEHICLE TRAJECTORY DATA

Xintao Yan
Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor

Shuo Feng
David J. LeBlanc
Carol Flannagan
Henry X. Liu
University of Michigan Transportation Research Institute

Alrik L. Svenson
National Highway Traffic Safety Administration
United States

Paper Number 23-0224

ABSTRACT

Assuring safe operation remains as one of the main challenges for developing and deploying automated driving systems (ADS). Real-time safety risk metrics may play important roles in calculating a risk measure of driving situations. Although a number of safety metrics have been proposed previously, it is difficult to compare different safety metrics and assess their performance because different behavioral assumptions underly for each. In this paper, a method to assess the behavior of safety risk metrics by determining the subject vehicle (SV) situational safety using logged vehicle trajectory data is proposed. Specifically, it is examined whether the SV is in a collision unavoidable situation at each moment, given the near-future trajectories of all surrounding principal other vehicles (POVs) recorded in the dataset after this moment. The main benefit of using logged vehicle trajectory data is the elimination of behavior prediction errors caused by model assumptions and approximations. This establishes a ground truth for crash outcomes independent of the risk metrics. Using the proposed methodology, the performance of different real-time safety metrics can be evaluated using simulated and/or real-world vehicle trajectories. The proposed methodology also has the potential to be applied in scenarios with vulnerable road users (VRU) interactions. In the case study, three real-time safety metrics are considered: time-to-collision (TTC), the PEGASUS Criticality Metric (PCM), and the Model Predictive Instantaneous Safety Metric (MPrISM). The results can help practitioners to better understand the characteristics and applicability of different safety metrics for different situations. The evaluation results can also help researchers improve and refine existing safety metrics.

INTRODUCTION

The overall idea of real-time safety metrics is to evaluate the situational safety risk based on the predicted future trajectories of both the subject vehicle (SV) and surrounding principal other vehicles (POVs). Utilizing vehicle state information (e.g., position, speed, heading, etc.) at or up to the evaluation moment, different behavior assumptions can be made to infer future trajectories, and these will lead to distinct safety metrics outcomes. An example is shown in Figure 1, where the blue vehicle represents the SV, and the green vehicle is the POV. The arrows represent different potential actions that SV and POV might take in the future. If the POV does not sense the SV in its blind spot and changes lanes in front of the SV, and the SV does not slow down or make an evasive lane change, then the resulting situation potentially can be dangerous. Alternatively, if the POV remains in its current lane, then the current moment in time is safe. This example demonstrates that different behavior assumptions can lead to different safety metrics results.

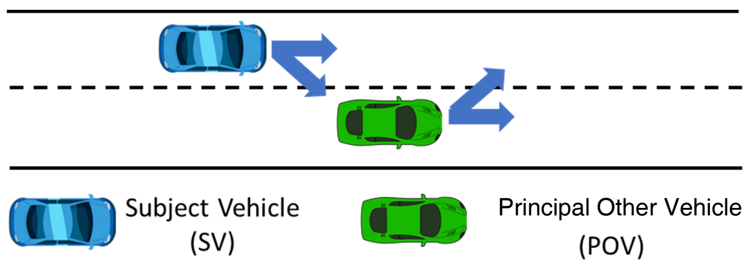


Figure 1. An illustration that different behavior assumptions can lead to different safety metrics results

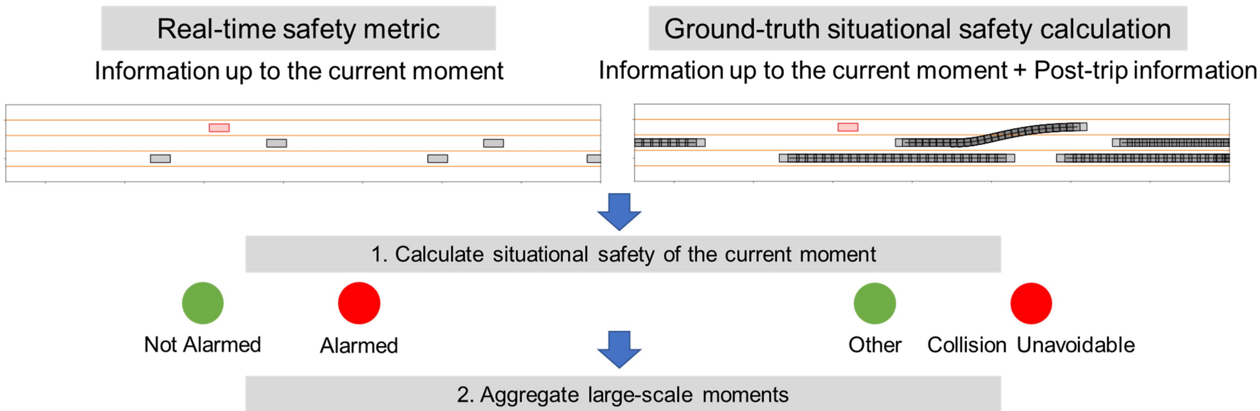
Over the past few decades, a number of safety metrics were proposed in the literature. One of the most commonly used safety metrics is Time-to-Collision (TTC) [1]. It assumes both the SV and its leading vehicle will maintain their current velocities. Based on TTC, Modified Time-to-Collision (MTTC) [2], and Worst Time-to-Collision (WTTC) [3], etc., were also proposed, attempting to capture different behavior assumptions including the acceleration and deceleration processes of both SV and POV. Junietz et al. proposed a Criticality Metric [4], which adopts similar behavior assumptions as in TTC regarding POVs behavior but assumes the SV will take evasive action to avoid the collision. Based on these assumptions, a model predictive control (MPC) method was applied to examine whether the current situation is at risk for the SV. Since it has been developed as part of the PEGASUS project, it is abbreviated here as PCM (PEGASUS Criticality Metric). The Model Predictive Instantaneous Safety Metric (MPrISM) was proposed in [5]. It also applies the MPC-based approach to generate a high-dimensional model predictive TTC (MPrTTC) that can consider both longitudinal and lateral possible risks. It assumes there exists one challenging POV, which will take the worst-case behavior trying to challenge the SV, and the SV will take the best response trying to avoid the collision. To account for the stochastic nature of human drivers, several probabilistic-based metrics have been proposed [6-8]. The collision probability is usually estimated to evaluate situational safety. In order to calculate the collision probability, multiple predicted future trajectories are generated to estimate the collision expectation using different methods, e.g., Markov chain [6], Monte Carlo simulation [7], and Bayesian network [8], etc. The potential stochastic behaviors of the SV and POVs are considered in these methods so they can better capture interactions between vehicles. However, the computational burden of probabilistic methods is heavier which might limit their real-time application.

The key challenge of evaluating a safety metric's performance is the lack of ground-truth measurement of risk at each moment in time because the future behavior of POVs is uncertain. Most existing evaluation methods rely on expert knowledge or heuristics, which suffers from subjective biases. The objective of this research is to propose an optimization-based method to utilize the post-trip information from vehicle logged trajectory data to evaluate and compare the performance of different safety metrics. In this way, a fair comparison of different safety metrics can be made, as a safety metric should alarm in advance to the collision-unavoidable moment. When trajectory data from a large number of trips are available, different metrics' statistical performance can be systematically evaluated and compared.

METHOD

In this study, a systematic evaluation framework has been applied to assess the performance of real-time safety metrics. The method can leverage post-trip information to calculate the situational risk objectively, as illustrated in Figure 2. For a real-time safety metric, information at or up to the current moment can be used to evaluate the SV's (shown by the red rectangle in the figure) situational safety. But when the trip is played back to determine the situational safety of the SV at any moment during the trip, trajectories of POVs are given, and no behavior assumptions are needed. By incorporating this post-trip POV trajectory information, the prediction errors of POV behaviors caused by assumptions and approximations of real-time safety metrics can be eliminated. Note that the post-trip information is obtained from logged trajectories, which can be either simulated trajectories or real-world recorded trajectories. Therefore, the safety metric output and the ground-truth can be compared to evaluate its effectiveness. Specifically, collision-unavoidable moments can be identified when evasive maneuver(s) to avoid the crash do not exist. They are objective without relying on expert knowledge or heuristically defined rules. The goal of a safety metric is to alarm no later than the collision-unavoidable moment. To identify the collision-unavoidable moment, an optimization problem can be formulated to examine whether there exists an evasive trajectory for the

SV, given the near-future trajectories of all POVs after this moment. It is assumed the POV behaviors will not be influenced by the SV actions within a short look-ahead horizon, so they will follow the observed trajectories. Therefore, the current moment is collision unavoidable if the SV has no possible evasive trajectory. The detailed formulation of the proposed algorithm can be found in [9].



Confusion Matrix		Real-time Safety Metric	
		Alarmed (Positive)	Not alarmed (Negative)
Ground-truth situational safety	Should alarm (Positive)	True Positive (TP)	False Negative (FN)
	Should not alarm (Negative)	False Positive (FP)	True Negative (TN)

Figure 2. Illustration figure of the proposed evaluation framework

Note that each observed trip is a realization of the real-world stochastic driving environment. Therefore, in the second step, a large number of trips need to be aggregated to construct the confusion matrix. The trajectory dataset is expected to reproduce the real-world driving environment and cover both normal and safety-critical driving situations in order to investigate the safety metric performance. Note that not only collision-unavoidable moments but also some moments before the earliest collision-unavoidable time need to be included in the "should alarm" category to examine the safety metric predictive capability. Based on the confusion matrix obtained from large-scale trajectories, the safety metric performance can be systematically evaluated. The evaluation results can not only be used to compare different metrics, but also to help build a feedback improvement process to iteratively tune, refine, and improve the safety metric performance.

CASE STUDIES

Introduction of three evaluated real-time safety metrics

Three real-time safety metrics, including TTC, PCM, and MPrISM, were selected to demonstrate the performance of the evaluation framework using simulated trajectories. TTC is one of the commonly used safety metrics to evaluate vehicle situational safety. PCM and MPrISM are state-of-the-art safety metrics proposed to evaluate ADS safety. Detailed information on the three metrics is introduced in the following paragraphs.

The first selected safety metric is TTC, which is defined by

$$TTC = \frac{x_{i-1} - x_i - l}{v_i - v_{i-1}}, \#(1)$$

Where x_{i-1} and x_i denote the longitudinal position of the leading and following vehicles, respectively. v_{i-1} and v_i denote their corresponding longitudinal speed. l denotes the length of the leading vehicle. In this study, the TTC is considered alarmed if it is smaller than 1 second.

The second selected safety metric is PCM. It formulates the situational safety evaluation problem into a SV trajectory planning task. An optimization problem is developed to find the best SV trajectory as defined by an objective function with four weighted terms. The decision variables of the optimization problem are the SV action sequences. The authors assume POVs will follow their current speeds and headings during the prediction horizon and the objective is to minimize the SV "criticality" within the prediction horizon. The optimization formulation is as follows

$$\min_{\boldsymbol{u}=[u(0)\ \dots\ u(N-1)]} \sum_{k=1}^{N-1} \left(w_x R_x(k) + w_y R_y^2(k) + w_{ax} \frac{a_x^2(k)}{(\mu_{max} g)^2} + w_{ay} \frac{a_y^2(k)}{(\mu_{max} g)^2} \right), \#(2)$$

$$s.t. \quad s(k+1) = \boldsymbol{A}(k)s(k) + \boldsymbol{B}(k)u(k), \forall t = 0, \dots, N-1, \#(3)$$

$$c_r(k) \le y(k) \le c_l(k), \#(4)$$

$$x(k) \le c_f(k), \#(5)$$

$$\boldsymbol{G}u(k) \le \boldsymbol{h}. \#(6)$$

The objective function, i.e., the criticality, is composed of four parts, which include the longitudinal margin $R_x(k)$, lateral margin $R_y(k)$, longitudinal acceleration $a_x(k)$, and lateral acceleration $a_y(k)$. These terms are weighted using four pre-defined parameters w_x, w_y, w_{ax}, and w_{ay}. The decision variable is the SV action at the k-th moment $u(k)$ which includes the longitudinal and lateral accelerations $a_x(k)$ and $a_y(k)$. N denotes the number of look-ahead steps. μ_{max} denotes the road friction coefficient. The constraints of the optimization problem include the SV vehicle dynamics (Equation 3), where $s(k)$ denotes the SV state at the k-th timestep. Eqs. 4-5 denote the SV safety constraints with surrounding POVs during the prediction horizon. c_l, c_r, and c_f denote the left, right, and front boundaries, respectively, which are calculated based on the predicted trajectories of surrounding POVs. The last set of constraints (Equation 6) characterizes the admissible action space (i.e., the Kamm's circle as illustrated in Figure 3) of the SV. By solving this optimization problem at each time step, the safety metric can be obtained using the maximum estimated acceleration during the prediction horizon. If the maximum acceleration exceeds certain pre-defined thresholds (i.e., the maximum acceleration of the SV), then this moment is considered dangerous and the PCM is alarmed. The model parameters have been set according to the original paper, where $w_x = 1, w_y = 1, w_{ax} = 0.1, w_{ay} = 1$, time resolution 0.1s, look-ahead steps N =20 [4]. The PCM is considered alarmed if the maximum expected acceleration is greater than 8 m/s^2.

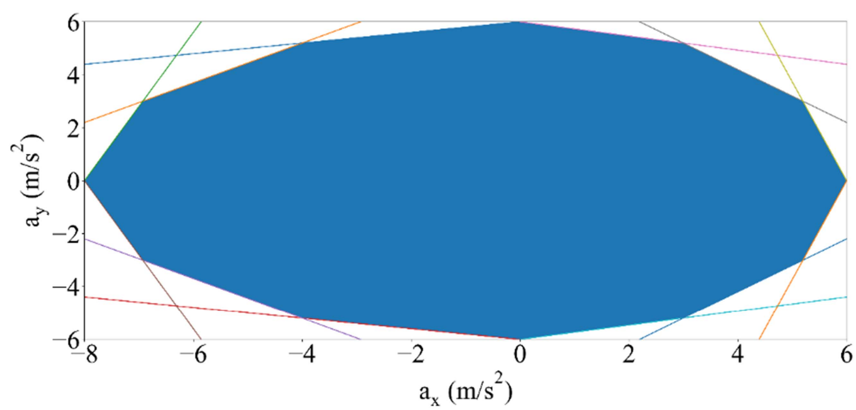


Figure 3. Vehicle admission action space (i.e., Kamm's circle)

The third selected metric is MPrISM. To evaluate the situational safety, it considers the pairwise interaction between the SV and each surrounding POV. It assumes the POV will take the worst-case behavior trying to collide with the SV. At the same time, it assumes the SV will take the best response trying to avoid the collision. Therefore, for each moment, to find the riskiest POV action and the best SV response, a min-max calculation is made. This is done every moment with the SV state (including vehicle position, speed, and heading) s_0, and the POV state, s_i. For each look-ahead step $N = 1, \dots, T$, the Euclidean distance between the SV and the POV at the end of the prediction horizon N can be obtained by solving the following minimax problem:

$$\min_{\boldsymbol{u_i}} \max_{\boldsymbol{u_0}} \quad \sqrt{\left(x_i(N) - x_0(N)\right)^2 + \left(y_i(N) - y_0(N)\right)^2}, \#(7)$$

$$s.t. \quad s_0(k+1) = \boldsymbol{A_0}(k)s_0(k) + \boldsymbol{B_0}(k)u_0(k), \forall t = 0, \dots, N-1, \#(8)$$

$$s_i(k+1) = \boldsymbol{A_i}(k)s_i(k) + \boldsymbol{B_i}(k)u_i(k), \#(9)$$

$$\boldsymbol{G_0}u_0(k) \leq \boldsymbol{h_0}, \#(10)$$

$$\boldsymbol{G_i}u_i(k) \leq \boldsymbol{h}_i, \#(11)$$

where x_0, y_0 and x_i, y_i represent the longitudinal and lateral positions of the SV and POV. $\boldsymbol{u_0}$ and $\boldsymbol{u_i}$ denote action sequences of the SV and the POV, respectively. For the constraints, Equation 8 and Equation 10 denote the system dynamics and Kamm's circle constraints of the SV, respectively. Equation 9 and Equation 11 denote the corresponding constraints for the POV. The predicted time-to-collision equals to $N\Delta$ (Δ is the time resolution) if the distance between SV and POV is smaller than a pre-defined collision threshold C at the end prediction horizon $N\Delta$. After iterating over all POVs, the least time-to-collision value across all POVs can be obtained and denoted as the model predictive time-to-collision (MPrTTC), which is used as the safety metric output. The model parameters have been set according to the original paper, where time resolution $\Delta = 0.1s$, look-ahead steps $N = 10$, and collision threshold $C = 4m$ [5]. The MPrISM is considered alarmed if the MPrTTC is smaller than 1 second.

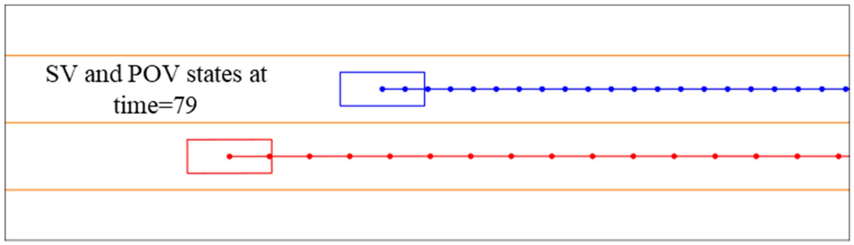
Simulation results

In this study, the performance of the proposed method is demonstrated using three simulated scenarios: SV overtaking POV, POV cutting in front of the SV, and POV and SV both moving into the same lane. The scenarios are generated using simulation models. In these scenarios, the real-time safety metrics may fail, for example, generating false alarms, or misidentifying potentially dangerous situations. The results show that the proposed method can successfully identify potential weaknesses caused by metric assumptions, approximations, and parameters.

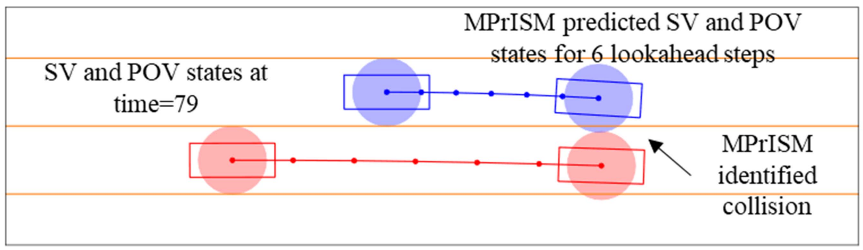
Scenario 1: SV overtakes POV

The first scenario is a normal overtaking case, where the SV is overtaking the POV on the adjacent lane and there is no crash happening in this scenario. The logged SV trajectory (shown by red curve and box) and POV trajectory (shown by blue curve and box) starting from the timestep 79 are shown in Figure 4a. The real-time safety metric results and the ground truth calculated with logged POV trajectory are shown in Figure 5. The red color moments in the ground truth denote the collision unavoidable moments. The red color in the safety metric results indicates that it alarmed at that moment. The MPrISM produces false alarms in this scenario.

As shown by the ground truth, it can be seen that moment 79 is not dangerous, and all metrics except for the MPrISM do not alarm. The MPrISM predicted trajectories of SV and POV are shown in Figure 4b. The shaded area denotes the assumed vehicle geometry by MPrISM. The MPrISM predicts that the POV will steer into the SV's lane since it makes worst-case behavior assumption regarding POV behavior. Therefore, a collision is predicted to occur in 0.6 seconds even if the SV takes the best response to avoid the collision. According to the collision definition in MPrISM, a collision is considered to happen when the Euclidean distance between two vehicles (considered as mass points with a certain radius) is smaller than the collision threshold. From the results as shown in Figure 4b, it is apparent that the predicted SV and POV circles overlap, so a collision is predicted by MPrISM. However, the two vehicles are not colliding with each other if considering their real geometry as shown by the two rectangles. Therefore, a false positive case is produced by the MPrISM due to both the worst-case behavior assumption and the single-circle approximation. Note that there is a trade-off for the MPrISM collision threshold. Using a larger threshold will cause more false alarm cases as shown in this scenario, whereas using a smaller threshold will cause false-negative cases (i.e., miss dangerous situations) as shown in Scenario 3.



(a)



(b)

Figure 4. (a) the logged SV (Red) and POV (Blue) trajectories starting from timestep 79, (b) the MPrISM predicted SV (Red) and POV (Blue) trajectories starting from timestep 79 of Scenario 1. The shaded area denotes the assumed vehicle geometry by the safety metric.

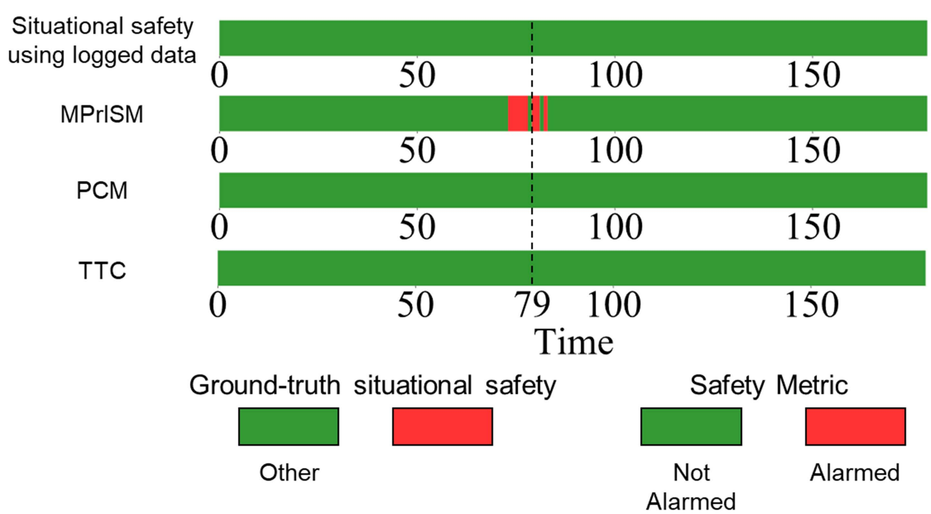
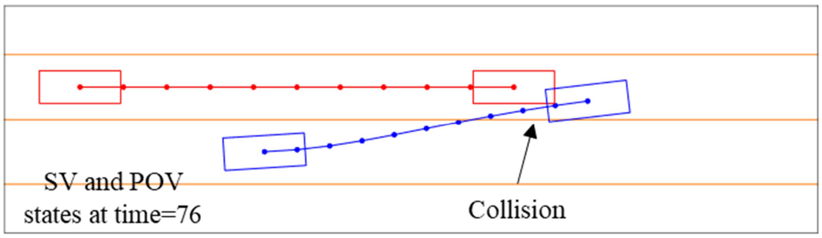


Figure 5. The ground-truth situational safety calculated using post-trip logged trajectory data and real-time safety metrics results of Scenario 1.

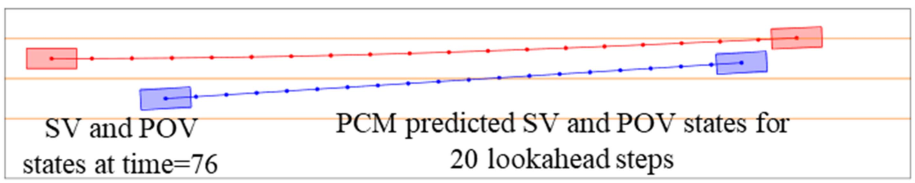
Scenario 2: POV cut-in SV

In the second scenario, the POV cuts in front of the SV, and a crash happens during the process. The logged trajectories starting from timestep 76 are shown in Figure 6a. The post-trip calculation and safety metrics results are shown in Figure 7. The PCM and TTC produce false-negative cases in this scenario.

From post-trip calculation results, starting from moment 76, the state of the SV is collision unavoidable, given the actual POV actions after that moment. The MPrISM successfully alarms starting from the moment 76. It does not generate false positives as in the previous case, since the worst-case behavior assumption regarding the POV can accurately characterize the aggressive cut in behavior. However, both PCM and TTC did not activate at the moment and generated false negatives. The PCM predicted SV and POV trajectories are shown in Figure 6b. The PCM assumes POV will maintain its current speed and heading within the prediction horizon. As a result, the SV can avoid the crash by turning to the adjacent lane. However, in the real situation, the POV is cutting in more aggressively compared with PCM behavior assumptions. Therefore, the PCM fails to activate at this collision unavoidable moment. For the TTC, it fails to consider the cutting in POV since they are not in the same lane at the current moment. Thus, the SV does not consider the POV as its leading vehicle and therefore produces false-negative cases.



(a)



(b)

Figure 6. (a) the logged SV (Red) and POV (Blue) trajectories starting from timestep 76, (b) the PCM predicted SV (Red) and POV (Blue) trajectories starting from timestep 76 of Scenario 2. The shaded area denotes the assumed vehicle geometry by the safety metric.

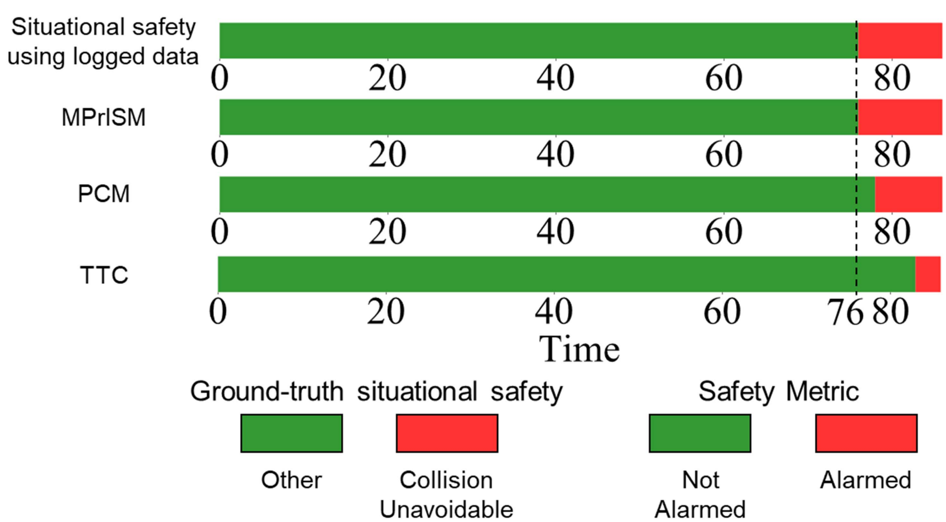
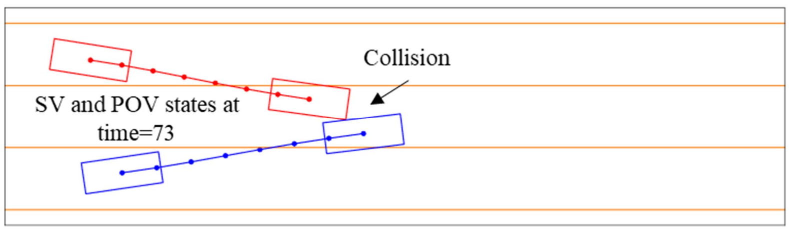


Figure 7. The ground-truth situational safety calculated using post-trip logged trajectory data and real-time safety metrics results of Scenario 2.

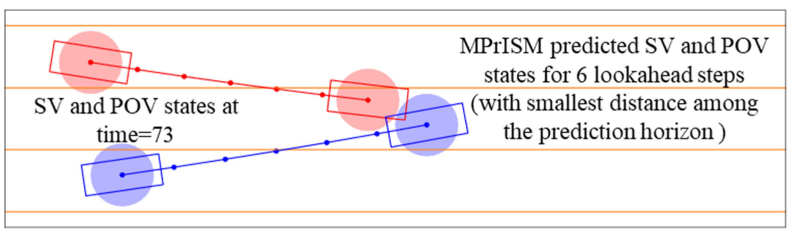
Scenario 3: POV and SV both move into the same lane

In the third scenario, the POV and SV are lane-changing to the same target lane and a crash happens during this process. All three metrics produce false-negative cases in this scenario. The logged trajectories starting from timestep 73 are shown in Figure 8a. The post-trip calculation and safety metrics results are shown in Figure 9.

From post-trip calculation results, starting from moment 73, the SV is collision unavoidable. However, all three metrics fail to alarm before the moment and cause false negatives. To analyze the MPrISM result, the MPrISM predicted trajectories of SV and POV are shown in Figure 8b. Within the 10 looking-ahead steps of MPrISM, the minimum predicted distance between the SV and POV occurs at the 6 look-ahead steps. As shown in Figure 8b, the two vehicles have already collided with each other as shown by the rectangles. However, due to the single circle approximation adopted by the MPrISM, it fails to identify the collision. One method to improve accuracy is to use multiple circles to represent the SV and POV. The benefit of using three circles rather than one circle is that it can more accurately capture the vehicle geometry and determine whether a collision happens or not. For the PCM, the reason why it fails to detect the crash in this scenario is similar to that discussed in Scenario 2. For the TTC, the reason why it fails to detect the crash in this scenario is that the leading vehicle velocity is higher than the following vehicle velocity.



(a)



(b)

Figure 8. (a) the logged SV (Red) and POV (Blue) trajectories starting from timestep 73, (b) the MPrISM predicted SV (Red) and POV (Blue) trajectories starting from timestep 73 of Scenario 3. The shaded area denotes the assumed vehicle geometry by the safety metric.

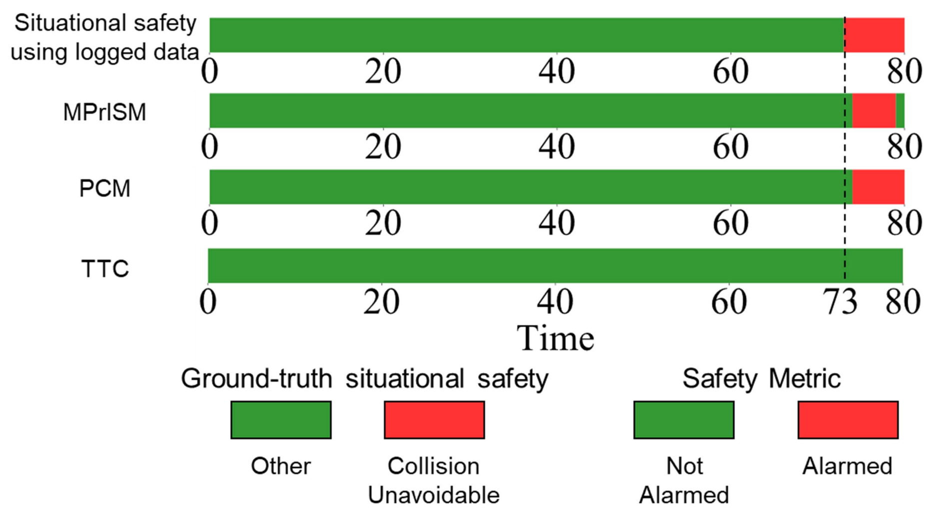


Figure 9. The ground-truth situational safety calculated using post-trip logged trajectory data and real-time safety metrics results of Scenario 3.

CONCLUSIONS

In this study, an assessment framework of real-time safety metrics performance is developed, utilizing post-trip information from vehicle logged trajectory data. Specifically, the logged POVs trajectories are utilized to determine whether the SV is in a collision unavoidable situation, which is an objective reflection of the situational safety, and this can serve as the ground truth of safety metrics. The goal of a safety metric is to alarm no later than the collision unavoidable moment to indicate the situational safety of the SV, but also without over-alarming (i.e., without too many false-positives). By analyzing the safety metric outputs and ground truth, one can identify the safety metrics' failed scenarios and investigate their potential weaknesses caused by model assumptions, approximations, parameters, etc. These results can help developers improve the performance of existing safety metrics. It can also help researchers and practitioners better understand the characteristics of different safety metrics, which may result in a choice of more appropriate metrics for different application purposes. This study has shown the benefit of a large number of logged trajectories for conducting a systematic analysis of safety metric performance. These can be used to compare the safety metric results with the calculated ground truth in order to obtain their statistical performance evaluation. It is recommended that logged trajectories cover diverse scenarios and locations to examine safety metrics performance under different circumstances. In this study, simulated trajectories were used to showcase the evaluation results. In the future, with the development of sensing technologies and data acquisition systems, real-world trajectories can be recorded and used to enrich and complement the simulated datasets.

STUDY LIMITATIONS

- In this study, it is assumed that perception systems are perfect in identifying POVs and their relevant attributes. Metric performance implications of uncertainties and inaccuracies that may be associated with a perception system's performance are not explored in this study.
- The metrics are assessed using three relevant but simple scenarios involving a single POV in each case. There are other complexities (e.g., computational) associated with these metrics when calculating a crash risk measure in real-world scenarios where multiple safety relevant objects may operate within the future time horizon of a metric's consideration. These are not explored in this paper.
- The crash risk metrics studied in this paper make assumptions on the bounds of how POVs may behave. In this study, those parameters are set as recommended by their authors with citation. However, metric performance is very sensitive to these settings. The use of the cited parameters does not imply agreement on these as reasonable or acceptable settings.

REFERENCES

[1] Hayward, J. C. (1972). Near miss determination through use of a scale of danger.

[2] Ozbay, K., Yang, H., Bartin, B., & Mudigonda, S. (2008). Derivation and validation of new simulation-based surrogate safety measure. Transportation research record, 2083(1), 105-113.

[3] Wachenfeld, W., Junietz, P., Wenzel, R., & Winner, H. (2016, June). The worst-time-to-collision metric for situation identification. In 2016 IEEE Intelligent Vehicles Symposium (IV) (pp. 729-734). IEEE.

[4] Junietz, P., Bonakdar, F., Klamann, B., & Winner, H. (2018, November). Criticality metric for the safety validation of automated driving using model predictive trajectory optimization. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 60-65). IEEE.

[5] Weng, B., Rao, S. J., Deosthale, E., Schnelle, S., & Barickman, F. (2020, June). Model predictive instantaneous safety metric for evaluation of automated driving systems. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 1899-1906). IEEE.

[6] Althoff, M., Stursberg, O., & Buss, M. (2009). Model-based probabilistic collision detection in autonomous driving. IEEE Transactions on Intelligent Transportation Systems, 10(2), 299-310.

[7] Althoff, M., & Mergel, A. (2011). Comparison of Markov chain abstraction and Monte Carlo simulation for the safety assessment of autonomous cars. IEEE Transactions on Intelligent Transportation Systems, 12(4), 1237-1247.

[8] Schreier, M., Willert, V., & Adamy, J. (2016). An integrated approach to maneuver-based trajectory prediction and criticality assessment in arbitrary road environments. IEEE Transactions on Intelligent Transportation Systems, 17(10), 2751-2766.

[9] Yan, X., Feng, S., LeBlanc, J. D., Flannagan, C., & Liu, X. H. (2022) Evaluation of automated driving system safety metrics with logged vehicle trajectory data. Under Review.