2024 5th International Conference on Electronic Communication and Artificial Intelligence (ICECAI) | 979-8-3503-8694-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICECAI62591.2024.10674835

A Method for Extracting Cut-in Scenario Based on Real-world Road Data Collected by ASEva

*Yanhong Fu
CARTAC Intelligent and Connected Technology Co., Ltd.
Tianjin, China
*Corresponding author's e-mail: fuyanhong@catarc.ac.cn

Bo Wei
CARTAC Intelligent and Connected Technology Co., Ltd.
Tianjin, China
e-mail: weibo@catarc.ac.cn

Xin Hu
CARTAC Intelligent and Connected Technology Co., Ltd. Tianjin, China
huxin@catarc.ac.cn

Pengchao Zhao
CARTAC Intelligent and Connected Technology Co., Ltd. Tianjin, China
zhaopengchao@catarc.ac.cn

***Abstract*—Based on the real-world driving data collected by the ASEva scene data acquisition system of ZEER Company, this paper proposes a method to extract dangerous cut-in scenarios from the driving data. This method analyzes the original data of ASEva to obtain effective sample data corresponding to a certain moment, such as the ego vehicle data, lane line data, and the target object data. Based on these data and algorithms, the front key vehicle is extracted, while excluding the misjudgments caused by excessive curvature due to the big curves on the road or the curvature caused by the turning of the ego vehicle. Finally, it is determined whether the latitudal absolute velocity of the target vehicle meets the threshold of the dangerous cut-in scenario. If the parameters of the target vehcile meet the cut-in scenario, the segment data between the start and end time is extracted. Finally, the accuracy of the extraction of dangerous cut-in scenarios is verified through the scenario generation method.**

Keywords-scenario extraction; front key vehicle; overlap rate between the target object and the ego vehicle

I. INTRODUCTION

Common methods to construct dangerous cut-in scenarios for autonomous driving simulations involve the use of some mainstream visualization simulation software, such as VTD (Virtual Test Drive)[3]. The Road Designer module of VTD is used to draw OpenDrive[8] static road network maps, while the Scenario Editor[6] module is used to model dynamic scenarios. However, scenarios constructed through visualization simulation software lack real road data, which is not conducive to automakers verifying various aspects of real road network environments. To address the above issues, this paper proposes a method to extract dangerous cut-in scenarios based on real road data. After extracting the cut-in scenario fragment data through this method, it can be used in our company's scenario generation tool to generate OpenDRIVE and OpenSCENARIO standard files for the generation of simulation test cases[4].

II. DESIGN OF SCENARIO EXTRACTION METHODS

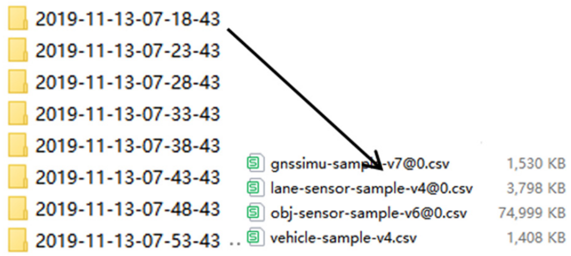
A. Overview of extraction algorithm

ASEva is a testing system independently developed by ZEER company, which is lightweight and of high freedom, and can provide customized services for customers[7]. This system is committed to providing overall solutions for the comprehensive performance evaluation of active safety on open roads, as well as the establishment of a real-world scene database. The data sources in this paper are the data files collected by the ASEva system on real roads during normal driving. The key information of the data used in this article is shown in Table 1.

Tablel 1. ASEva key fields

Data Name	Key fields
vehicle-sample	velocity, yaw rate, driving curvature, longitudinal acceleration, latitudinal acceleration, width, length, heigth, front overhang, steering wheel
lane-sensor-sample	0th-order&first-order&second-order&third-order coefficient of the y-curve equation for the nth lane line
gnssimu-sample	positional mode, longitude, latitude, altitude, pitch angle, roll angle, sideslip angle
obj-sensor-sample	ID of the nth target object, category of the nth target object, absolute velocity of the nth target object in the x-axis and y-axis directions

Besides the key feilds in Table 1, these data share the following fields: session, sync state, CPU tick, ego posix, guest posix, server posix, Gnss posix, and time. The original data file is a 5-minute interval, and the naming form of the data folder is timestamp year month day hour minute second, as shown in Figure 1 (a). The specific content form within the folder is shown in Figure 1 (b). The data segments used in this paper are 613, with a time span of 2019-11-13-07-18-18-43 to 2019-11-17-18-14.



(a) (b)

Figure 1. Raw data form

As shown in Figure 2, the flowchart shows the key steps of the scenario extraction method described in this paper.

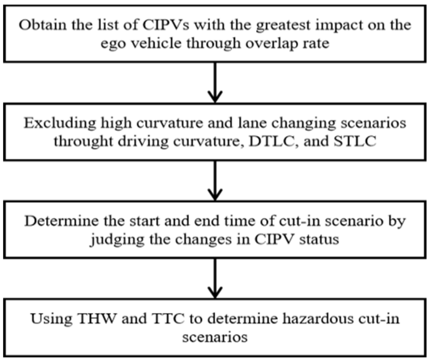


Figure 2. The flowchart of cut-in extraction algorithm

We define a special target object, which is the forward target object closest to the vehicle in this lane. We call this target object CIPV. When a target object changes from a non CIPV to a CIPV, it indicates that the target object has entered the current lane from an adjacent lane and is closest to the ego vehicle. This process is defined as cut-in.

B. Calculate the front vehicle-CIPV

1) Calculate the front vehicle CIPV

As shown in Figure 3, we introduce a concept: the overlap rate between the ego vehicle and the target vehicle, which is the percentage of the overlapping part of a target object to the width of the ego vehicle. When the overlap rate meets a threshold, this target object is considered to be in the lane of the ego vehicle. The target object with the closest distance to the ego vehicle among all the target objects that meet the overlap rate threshold is the CIPV.

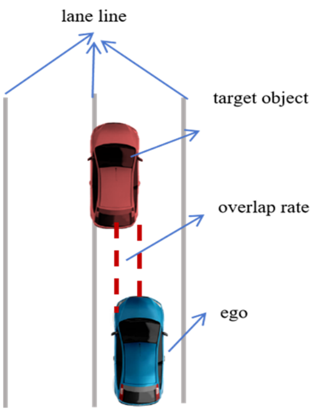


Figure 3. Overlap rate between the ego vehicle and the target object

The following is the detailed algorithm process for calculating the overlap rate:

- Obtain the width parameters of the ego vehicle and the target object. The width of the ego vehicle is directly extracted from the vehicle width parameter in the vehicle-sample. If the position mode of the target object is the nearest point, different width values are defined for different types of target objects in this method, such as the width of a small car is 1.9m and the width of a pedestrian is 0.5m. If the position mode of the target object data is the center of the box, then calculate the eight points around the target object based on the position and width of the target object, and obtain the width of the target object by subtracting the minimum value from the maximum value.

- Calculate the position of the ego vehicle and the target object in the y-axis direction[10].

$$\begin{aligned} &dyEgo = 0.5 * curvEgo * objInfo.posx^{2} + \\ &curvEgo * (fo + wb) * objInfo.posx \\ &dyObj = objInfo.posy \end{aligned} \tag{3}$$

As shown in formula (3), dyEgo represents the data of the ego vehicle in the y-axis direction at a certain time, dyObj represents the data of the target object in the y-axis direction, curvEgo represents the driving curvature of the ego vehicle, objInfo.posx and objInfo.posy are the position coordinates of the target object, fo is the front overhang of the ego vehicle, and wb is the wheelbase of the ego vehicle.

Calculate the overlap rate between the ego vehicle and the target vehicle.

$$\begin{aligned} &overlap = \min(dyEgo + 0.5 * widthEgo, \\ &dyObj + 0.5 * widthObj) - \max(dyEgo - \\ &0.5 * widthEgo, dyObj - 0.5 * widthObj) \end{aligned} \tag{4}$$

$$\begin{aligned} &overLapRate = overLap / (widthEgo \text{ if } \\ &overLap < 0 \text{ else } \min(widthEgo, widthObj)) \end{aligned} \tag{5}$$

In order to optimize the above algorithm, we need to exclude some scenarios, such as lane changing scenarios and large curvature scenes, and the specific methods are as follows.

① Excluding lane changing scenarios

As shown in Figure 4, when the vehicle changes lanes, there is an overlap between the vehicle and the target vehicle in the lateral position, and the lane changing scene is easily mistaken as cut-in scenario. Lane-changing scenarios can be categorized into two situations: there are obvious turn signal data in the sample; or it is judged based on some parameter values and actual driving conditions.

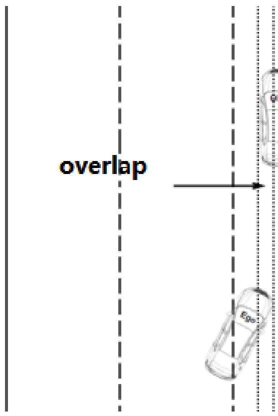


Figure 4. The ego vehicle is changing lanes

The steering field data in the vehicle-sample data can assist us in obtaining the current turn signal status of the vehicle.

Specifically, the values are as follows: 0, unknown; 1, no lights on; 2, left turn signal; 3, right turn signal; 4, hazard lights.

Before making judgments based on parameter thresholds, we need to calculate two parameters: the distance between the wheel and the lane line (DTLC for short), and the speed of the wheel approaching the lane line (STLC for short). The specific calculation logic is as follows:

- DTLC

The specific calculation formula is shown in formula (6), where vehiFrontOH is the front suspension of the ego vehicle, and a&b&c&d are the coefficients of the first-order cubic equation of the lane line in lane-sensor-sample, as shown in Table 1. vehiWidth is the width of the ego vehicle, and this parameter can also be obtained from vehicle-sample.

$$\begin{aligned} & x = -1 * \text{vehiFrontOH} \\ & y = a * x^3 + b * x^2 + c * x^3 + d \\ & \cos\theta = \text{abs}(\cos(\arctan(c))) \\ & \text{DTLC} = \cos\theta * (y - \text{vehiWidth} * 0.5) \end{aligned} \quad (6)$$

Note: If the value of d is negative, it indicates that the target object is in the front right of the ego vehicle, and then the y value should be multiplied by -1.

- STLC

As shown in formula (7), vehiVX represents the ego vehicle's velocity. Similarly, if the value of d is negative, the STLC value needs to be multiplied by -1.

$$\begin{aligned} & \sin\theta = \sin(\arctan(c)) \\ & \text{STLC} = \sin\theta * \text{vehiVX}/3.6 \end{aligned} \quad (7)$$

If DTLC or STLC meets any of the following conditions, it is considered that the ego vehicle is changing lanes or has a tendency to change lanes.

Condition 1:DTLC<-0.1&&abs(STLC)>0.

Condition 2:DTLC<0.1&&abs(STLC)>0.3

Condition 3:abs(STLC)>0.5

② Excluding scenarios with large curvature

As shown in Figure 5, in a large curvature scene, both the ego vehicle and the target vehicle are driving in their respective lanes. Due to the large curvature of the road and the unrecognized lane markings, there is an overlap between the ego vehicle and the target vehicle in the lateral position. Therefore, the high curvature scene is prone to be mistaken as cut-in scenario, so it is necessary to filter the high curvature scene. Extract the driving curvature data from the vehicle-sample. If the data is greater than 0.01, it is considered that the curvature of the data is too large. Here, the driving curvature obtained by default refers to the degree of curvature of the road. According to experience, when the curvature is 0.01, the radius of curvature is 100m. When the driving curvature of vehicle-sample is greater than 0.01, the radius of curvature of the road will be less than 100m. Therefore, we consider the scene with a curvature radius of less than 100m as a scene with excessive curvature, and it is necessary to exclude this section of data from the sample.

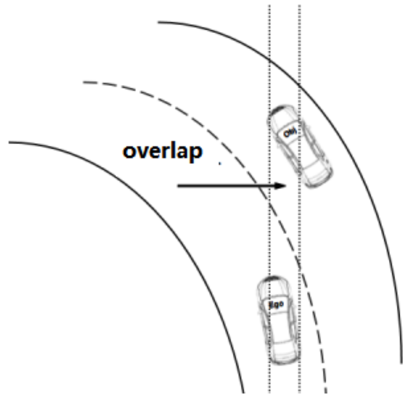


Figure 5. Road of large curvature

2) Determine the cut-in scenario

We can determine the target object, the start time and end time of cut-in through the state changes of CIPV. The situation without lane markings is shown in Figure 6.

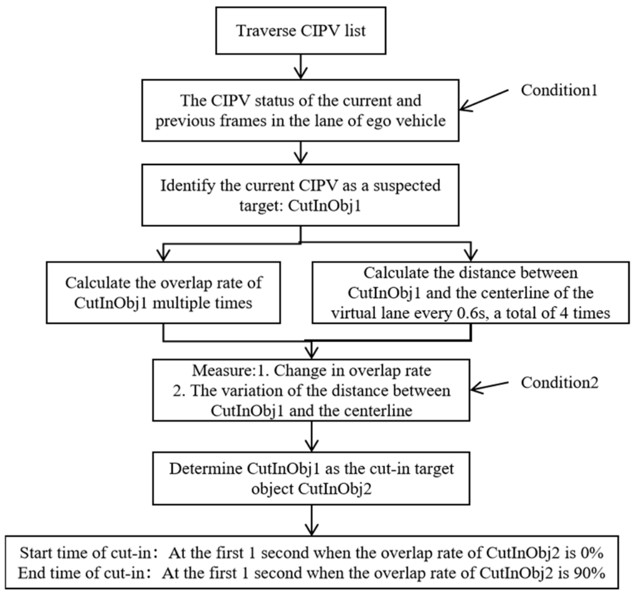


Figure 6. No lane markings

Condition1 in Figure 6:

- CIPV_ID(n-1)=NAN and CIPV_ID(n)≠NAN;
- CIPV_ID(n)≠CIPV_ID(n-1)

Condition2 in Figure 6:

- The overlap rate value increases over time;
- The distance from the centerline of the lane gradually decreases, with a maximum distance greater than 0.8m and a variation amplitude exceeding 0.35m

The situdtion with lane markings is similar to Figure 6, and the differences are:

- Condition2:The overlap rate increases over time and varies according to -100, 50, 100
- End time of cut-in: At the first 1 second when the overlap rate of CutInObj2 is 100%

3) Using THW and TTC to determine dangerous scenarios

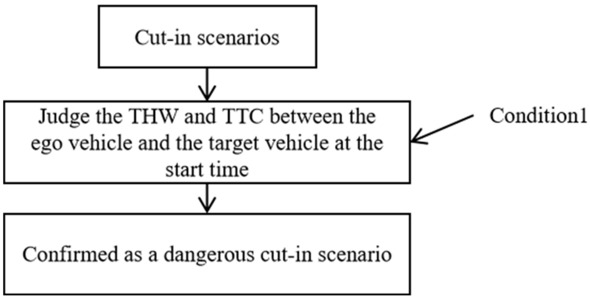


Figure 7. Determine the dangerous cut-in scenario

Condition1 in Figure 7:

- Ego vehicle's speed is greater than 20km/h and THW(Time headway) is less than 1s
- Ego vehicle's speed is greater than 20km/h and the TTC(Time to collision) is less than 2s

III. EXPERIMENTAL RESULTS

The technical stack for implementing the scenario extraction method mentioned in this paper is Python[9] and its ecosystem. Programming implementation is carried out based on the above-mentioned scenario extraction method and design concept. As shown in Figure 8, the information saved in the database includes scene_name, scene_type, start_time, and keep_time. There are 613 data segments in this experiment, and a total of 152 cut-in scenarios were extracted. By comparing the road sampling video data, it is confirmed that the cut-in extraction rate for this time span is 80%[1].

id	scene_name	scene_type	start_time	keep_time
144	2019-12-17-07-13-44	LeftCutIn_tube	296.3	2.0
145	2019-12-17-07-13-44	LeftCutIn_tube	307.55	2.75
146	2019-12-17-07-13-44	LeftCutIn_tube	339.2	1.55
147	2019-12-17-07-13-44	RightCutIn_tube	349.75	1.7
148	2019-12-17-07-13-44	LeftCutIn_tube	360.25	2.95
149	2019-12-17-07-13-44	LeftCutIn_tube	364.0	2.45
150	2019-12-17-07-08-44	LeftCutIn_tube	28.4	2.1
151	2019-12-17-07-08-44	RightCutIn_lane	103.0	3.05
152	2019-12-17-07-08-44	RightCutIn_tube	103.0	3.7
NULL	NULL	NULL	NULL	NULL

Figure 8. The basic infos of the extracted cut-in scenarios

Figure 9 shows the information saved in the csv file for the extracted cut-in scenario 2019-11-13-07-18-43. It is obvious that the file size has changed.

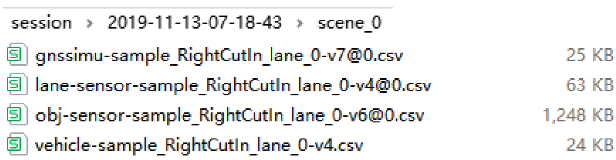


Figure 9. The csv files of the extracted cut-in scenarios

Generate XODR and XOSC files using the company's self-developed generation software[5], and the effect using the open-source software esmini is shown in Figure 10.

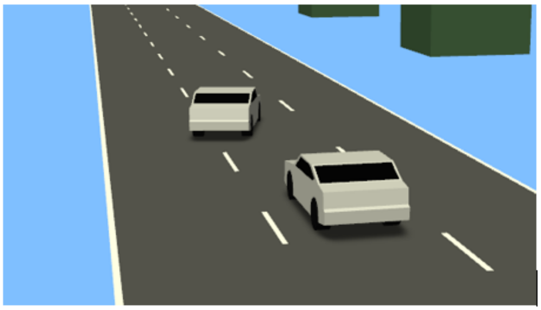


Figure 10. A cut-in scenario extracted by csv files

IV. CONCLUSIONS

The advantage of the scenario extracted by this method compared to the scenario generated by VTD is that the data is real, but the disadvantage is that the restoration degree of the scenario may be lower, and the accuracy of the extraction still needs to be improved. There are still many deficiencies in this paper's algorithm. For instance, the setting of some judgment parameters's threshold is currently based on empirical assumptions, lack of rigor. IIn the future, we will use some deep learning methods to make these parameter values adaptive[2], in order to improve the accuracy of the algorithm's extraction.

REFERENCES

[1] Kodani Y ,Sekine K ,Tanaka Y , et al. A study on the reliability and validity of the Japanese version of the Scenario Test for people with chronic stroke-induced aphasia: A cross-sectional study. [J]. International journal of language & communication disorders, 2024

[2] Lijie L ,Zhaoguang Z ,Chunfeng Q . Research on Induction Measurement of Operators Approaching Charged Bodies in Complex Operation Scenarios [J]. Journal of Physics: Conference Series, 2023, 2450 (1)

[3] Foretellix; ASAM's OpenSCENARIO 2.0 Concepts Demonstrated Using Foretellix M-SDL [J]. Journal of Transportation, 2020

[4] Patrick D ,R. S B ,Hartmut K . Shared Memory Parallelism in Modern C++ and HPX [J]. SN Computer Science, 2024, 5 (5):.

[5] Sunday O B ,Nduka O S . MySQL Database Server: Deploying Software Application to Enhance Visibility and Accountability [J]. Current Journal of Applied Science and Technology, 2023, 42 (4): 16-23.

[6] Companies Team to Bring Virtual Drive Test Technology to Enhance Mobile Experience on Italy's Roads and Railway Lines [J]. Telecomworldwire, 2024,

[7] Li N ,Chen L ,Huang Y . Research on Specific Scenario Generation Methods for Autonomous Driving Simulation Tests [J]. World Electric Vehicle Journal, 2023, 15 (1):

[8] B. S ,H. T K . VALIDATION OF PARAMETRIC OPENDRIVE ROAD SPACE MODELS [J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2022, X-4/W2-2022 257-264.

[9] Ryzhkov V F ,Ryzhkova E Y ,Elinson N M . Python tools for structural tasks in chemistry. [J]. Molecular diversity, 2024,

[10] Kumar R ,Bera C ,Kumar A . Optimization of BLDC-based electric vehicles: Vehicle dynamics modelling through dual-motor approach and designing a novel augmented TLBO algorithm for PID control [J]. Engineering Research Express, 2024, 6 (2)