

Risk propagation mechanism and prediction model for highway diverging area

Ben Niu*

Logistics Research Center, Shanghai Maritime University, Shanghai, China

* Corresponding author: 905837168@qq.com

ABSTRACT

After an accident occurs, traffic risk propagates upstream along the roadway. Within the spatiotemporal impact area of the accident, traffic flow accident risk experiences fluctuations with uncertainty. Accurately describing the process of traffic accident risk propagation on the road and quantitatively analyzing it will provide a solid theoretical basis for traffic risk management measures. Therefore, this study focuses on the risk propagation mechanism and prediction methods after traffic accidents in highway diversion areas. Firstly, this paper introduces the concept and indicators of risk propagation. Secondly, through simulation experiments, the driving scenarios in the diversion area are simulated with different traffic volumes, and the regularities of risk propagation are studied by observing changes in risk indicators. Finally, using a simulated dataset as input, a GRU (Gated Recurrent Unit)-based accident risk prediction model is constructed. Experimental results indicate that in predicting accident risk, GRU exhibits higher prediction accuracy compared to LSTM (Long Short-Term Memory).

Keywords: risk propagation, diverging area, risk prediction

1. INTRODUCTION

Near the ramp in the diverging area of the highway, the separating behavior of the ramp vehicles might create coordination problems with the mainline vehicles, possibly leading to traffic accidents. However, the impact of traffic accidents is multidimensional. The risks arising from this single factor can spread to other elements of the system, resulting in changes in the state of other elements. Therefore, it is necessary to study the combination of various factors in the analysis of traffic accidents to understand the mechanism of risk propagation.

Scholars from various countries have persistently explored the mechanisms of accident impact propagation, focusing on analyses from both macro and micro perspectives. Macroscopic risks typically investigate the phenomenon of traffic loss caused by disruption propagation and the intensity of accident impacts through traffic wave theory. This includes analyzing congestion scope, delay duration, and other aspects. Microscopic risks predominantly delve into the study of driving risk fields and calculate risk probability based on following distance, expressing risks associated with individual traffic entities. Microscopic risks emphasize analyzing the safety of individual traffic entities, particularly crucial in guiding driver operations. Conversely, macroscopic risks tend to characterize overall driving risks on road sections, bearing significance for road safety engineering and risk management.

Microscopically, Wang et al.[1] proposed a vehicle collision risk algorithm based on a driving safety field model, considering driver behavior, vehicle characteristics, and road conditions. Babu et al.[2] established a risk field comprising collision probability and expected collision energy to estimate driving risks for target vehicles at future positions. Using the critical conditions of highway car-following safety, Lu et al.[3] analyzed rear-end collision mechanism characteristics and derived a quantified indicator for rear-end risk based on rear-end collision motion principles. Zhu et al.[4] analyzed rear-end collision processes from the perspectives of preceding vehicle deceleration and collision induced by it. They used conditional probability to determine the likelihood of rear-end accidents, thereby characterizing follow-up risks. Based on driver vehicle operation and eye movement data, Ma et al.[5] constructed a Bayesian network vehicle operational risk assessment model, quantitatively describing vehicle operational risks under dynamic factor influences.

Macroscopically, Han et al.[6] proposed a probability model for traffic loss at highway merge bottlenecks, analyzing the relationship between failure probability, traffic flow rate, merge gap, merge speed, and driver characteristics. Li et al.[7]

established a traffic flow radiation range estimation model under accident conditions based on TransModeler simulation. Jin et al.[8] categorized accident impact ranges into point, line, and area levels. They quantitatively determined the impact zones at each level using highway maintenance control zone settings, traffic flow theory, and methods comparing path impedance calculations with travel times. To indirectly demonstrate the spontaneous breakdown phenomenon of traffic flow, Hao et al.[9] developed a traffic disturbance evolution model, considering congestion and secondary disturbance. They analyzed the influence of important model parameters on computation results and reasons behind them. Dong et al.[10] established a mathematical model reflecting the number of queued vehicles and total queue delay after traffic accidents, based on traffic flow conservation laws and traffic wave principles. Saeedmanesh et al. [11] established a deterministic queue length model for highways based on arrival and departure rate curves, calculating the total delay due to accidents, queue length, and congestion duration.

Currently, research on traffic accidents mainly focuses on analyzing the factors influencing accidents and assessing and predicting accident risks. There is limited discussion on the propagation of post-accident traffic risks. Moreover, within the scope of limited research on traffic risk propagation, microscopic vehicle operational risks have not been adequately studied in terms of risk propagation characteristics. In the macroscopic risk domain, there is a lack of real-time dynamic quantification research on risk values themselves, with more attention on the impact of unilateral factors on traffic flow after accidents. Comprehensive studies considering traffic flow changes under the influence of risks are scarce. Therefore, investigating and exploring the variations in risk propagation processes on highway sections after accidents holds significance.

2. ACCIDENT RISK PROPAGATION MECHANISM IN HIGHWAY DIVERGING AREA

To explore the mechanism of accident risk propagation in the diverging areas of highways, this study conducts simulation experiments to mimic accident scenarios in highway diverging areas. Through the establishment of risk indicators, and the determination and grading of accident impact areas, we investigate the variations of risk within different levels of impact areas. This approach enables the analysis of the propagation mechanism of traffic accident risk in the diverging areas of highways.

2.1 Concept of Accident Risk Propagation

The diverging area on highways is often regarded as a high-risk area for accidents due to frequent interactions between main road and ramp vehicles. The occurrence of accidents is typically a result of the combined effects and cumulative impacts of numerous risk factors, rather than stemming from a singular cause. Firstly, there exists a competition game for right-of-way among traffic participants. For instance, when ramp vehicles attempt to merge onto the main road, they may contend for space with vehicles already on the main road. Such competitive scenarios escalate the risk of accidents. Apart from right-of-way issues, the information present in the traffic environment places stress on drivers. When main road traffic diverts onto ramps or there are lane changes on the main road, drivers need to process a significant amount of information, including road signs and the behavior of other vehicles. This information load can surpass the driver's processing capacity, leading to errors in judgment or delayed responses.

These risk elements not only exist individually but also exhibit high degrees of interconnectedness. For instance, deficiencies in road design might exacerbate the consequences of drivers' non-standard operations. Cumulatively, these risk factors within the traffic flow could propagate rapidly on the road through phenomena such as congestion, platooning waves, and dissipation waves. This amplification can transform localized risks into larger areas, consequently affecting more traffic participants.

This study assumes that highway traffic follows a car-following pattern, with accident points occurring near the exit of the diverging area, as shown in Figure 1. After an accident occurs, following traffic needs to brake to avoid colliding with the accident-involved vehicles. Simultaneously, the fast lane traffic can be affected by lane changes of ramp traffic. Consequently, vehicles closer to the accident site are more susceptible to its effects, thus increasing their driving risk. On the other hand, vehicles farther away from the accident site experience minimal impact, resulting in nearly unchanged driving risk.

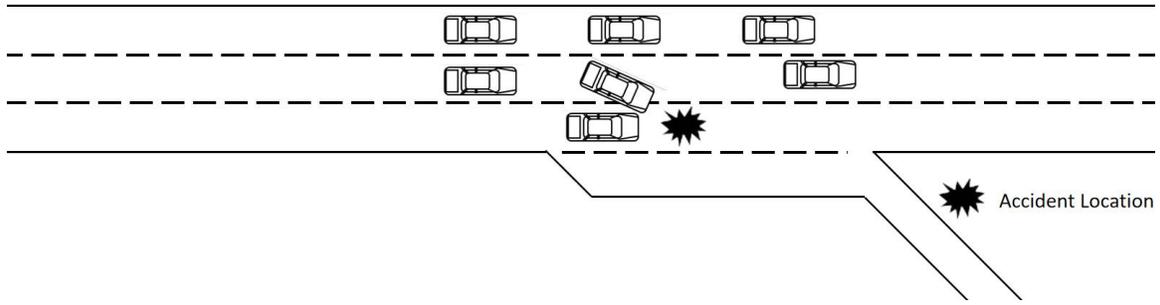


Figure 1. The accident impact illustration of highway diverging area.

However, risk propagation is a dynamic process. In the early stages of an accident, the likelihood of secondary accidents caused by sudden braking is high, increasing the risk of collisions. Over time, traffic flow near the accident site gradually slows down, reducing the risk between vehicles. Concurrently, the risk induced by abrupt braking propagates upstream within the traffic flow, leading to higher risks upstream. As the traffic flow stabilizes, collision risks within the traffic flow dissipate.

2.2 Risk index

The Time to Collision (TTC) is a widely used quantification method within the field of traffic accidents. It serves to assess the relative motion and potential collision risk between two vehicles. Defined as the time that a vehicle would take to collide if it continues to move in its current direction and at its present speed, TTC is instrumental in understanding the immediate threat and response time required to avert a collision. The formula used to calculate TTC is as follows:

$$TTC_i(t) = \frac{S}{V} = \frac{x_{i-1} - x_i - l_{i-1}}{v_i(t) - v_{i-1}(t)} \quad (1)$$

Where $TTC_i(t)$ represents the time to collision between the following vehicle and the leading vehicle; S denotes the headway distance between the leading vehicle and the following vehicle; V represents their relative speed; x_i , x_{i-1} are the respective coordinates of the head positions of the leading vehicle and the following vehicle; $v_i(t)$ represents the instantaneous speeds of the vehicles; and l_{i-1} is the length of the leading vehicle's body.

To define the severity of traffic conflicts, it's essential to determine an appropriate threshold. Many researchers currently use the cumulative frequency curve method to establish this threshold. Specifically, they usually select the 15% and 85% percentiles of the cumulative frequency curve as thresholds for severe and general traffic conflicts, respectively.

According to existing research[12], the threshold for severe traffic conflicts typically falls between 1.5 to 4 seconds, while the threshold for general traffic conflicts ranges from 3 to 8 seconds. Therefore, in this paper, we categorize the severity of traffic conflicts into three distinct levels: Traffic conflicts less than 3 seconds are defined as severe; Those between 3 to 5 seconds are considered general; Those between 5 to 8 seconds are classified as secondary.

However, evaluating the risk within an area should not solely focus on severe traffic conflicts. Instead, a comprehensive assessment needs to consider traffic conflicts of various severity levels. This study provides the following definitions for accident risk:

$$R = \omega_1 \frac{r_1}{r_1 + r_2 + r_3} + \omega_2 \frac{r_2}{r_1 + r_2 + r_3} + \omega_3 \frac{r_3}{r_1 + r_2 + r_3} \quad (2)$$

Where R is the risk indicator within the region for the current time period; r_1 , r_2 , and r_3 respectively represent the frequency of severe, general, and secondary traffic conflicts within the time period; ω_1 , ω_2 , and ω_3 are the weight coefficients.

2.3 Simulation scenario

This paper employs Sumo to create a traffic accident scenario in the diverging area of a highway. Sumo, an open-source software based on micro-level traffic simulation, is used to simulate and evaluate various types of transportation systems, including roads, highways, railways, pedestrian pathways, and bicycles.

Firstly, in the simulation scenario, default traffic flow parameters were employed to model a side-collision accident in a diverging area of a highway, which comprises three main lanes and a 200-meter acceleration lane (as shown in Figure 2).



Figure 2. Simulation scenario.

Next, the accident impact area model is employed to calculate the scope of effects generated by accidents. Wang et al.[13] improved the Gaussian plume model to make it applicable for calculating and dividing the impact area of traffic accidents. The formula is presented as Equation 3.

$$X_i = \sqrt[3]{\frac{\xi Pa}{4\sqrt{\pi}C_i}}, C_i = \frac{C_d}{b_i} \quad (3)$$

Where X represents the scope of accident impact in the special segments of the highway; ξ represents the accident point diffusion traffic ratio (between 0 and 1), determined by the accident type; P represents the potential energy of the accident source point; a is related to the proportion of lanes occupied by the accident; C_d represents the degree of traffic impact at the maximum range of the accident, approaching 0, and can be calculated based on actual data; b is a tuning parameter, dividing the results into three different areas of impact through variations in this parameter(as shown in Figure 3).



Figure 3. Different impact areas in simulation scenario.

After multiple experiments and in combination with literature[13], the values of relevant parameters for the accident impact area model in this simulation scenario are shown in Table 1.

Table 1. Values of parameters for the accident impact area model.

a	P	C_d	ξ	b_i
0.5	1000	5e-7	1	(0.05,0.3,0.1)

Finally, within the traffic flow range of 3600 pcu/h to 4770 pcu/h, and at intervals of 30 pcu/h, multiple simulations are conducted to obtain the results.

2.4 Experiment Results

The simulation results consist of two parts: conflict event data and detector data, both of which are directly exported from SUMO. The detector data records the average speed and flow of vehicles passing through the region during a specific time interval, with a frequency of 10 seconds. The conflict event data logs the coordinates, simulation time, and TTC (Time to Collision) values of events where the TTC is less than 8, and transforms the TTC values for various ranges into risk indicators through the risk index formula.

In this study, the weight coefficients for different levels of risk indicators are set in descending order as 1/3, 2/3, and 1, making the maximum value of the current scenario's risk indicator equal to 1. This maximum value indicates a dangerous state in the region; conversely, a risk index value closer to 0 indicates greater danger.

By merging the data, it is possible to obtain the flow, average passing speed, and risk index in various impact levels every 10 seconds after the accident. This section focuses on analyzing the risk indicators to elucidate the pattern of risk propagation, while the remaining fields will be used for building the accident data set in the next section.

Figure 4 illustrates the variation in regional risk indicators for low traffic volume (3600 pcu/h) and high traffic volume (4770 pcu/h) in the given scenario. From the figure, it is evident that the first-level impact area, being the closest to the accident site, exhibits fluctuating risk values consistently below 1, indicating a persistent state of risk presence. Conversely, the second-level impact area initially remains unaffected due to its greater distance from the accident site. Under conditions of low traffic volume, both the second and third-level impact areas experience an influence throughout the simulation. However, in scenarios with high traffic volume, as the risk indicator of the first-level impact area begins to decline and risks increase, the second-level impact area starts to experience risk after a certain period, denoted as moment t_1 in Figure 4. Similarly, the third-level impact area is influenced by the rising risks in the first and second-level impact areas. Consequently, its regional risk level starts to decrease, but eventually, it experiences an increase in risk, indicated as moment t_2 in Figure 4.

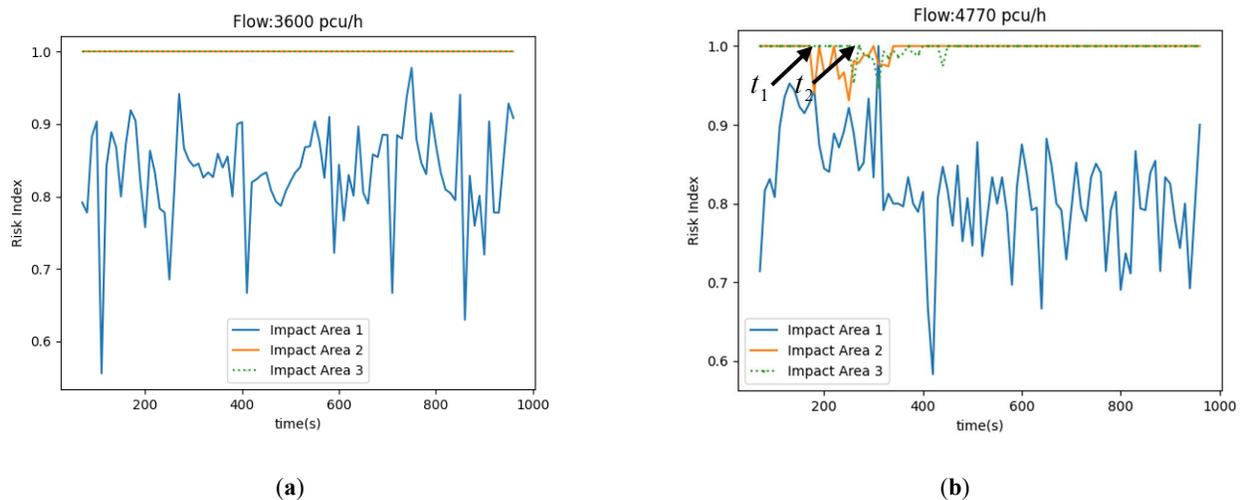


Figure 4. The risk values for first, second, and third-level impact areas. (a)Low flow.(b)High traffic flow.

Due to variations in the input traffic volumes for different experimental scenarios, the points in time at which the second and third-level impact areas begin to exhibit risk fluctuations are also different. Figure 5 depicts the initiation times of risk fluctuations in the second and third-level impact areas under different traffic volumes. When traffic volume is low, the second-level impact area starts to experience influence 800 seconds after the accident. As traffic volume increases, the point in time at which the second-level impact area is affected shifts earlier. In scenarios with higher input traffic volume, the second-level impact area begins to experience influence just 200 seconds after the accident, indicating a reduced time for the propagation of risk from the first-level to the second-level impact area. The third-level impact area follows a similar pattern. With higher traffic volume in the scenario, the point at which the third-level impact area is influenced by the second-level impact area shifts from 700 seconds to 300 seconds.

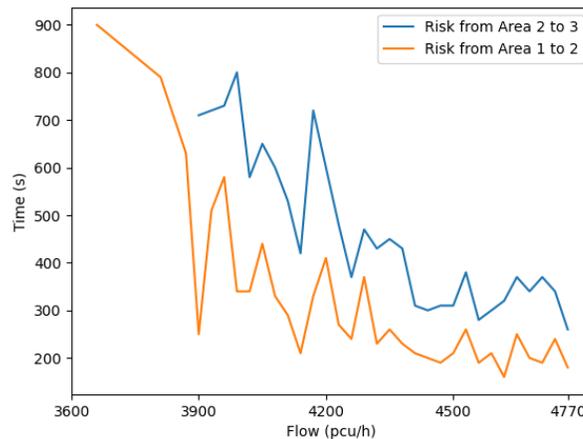


Figure 5. The risk exposure moment of the second and third-level impact areas.

3. TRAFFIC ACCIDENT RISK PREDICTION MODEL

3.1 GRU Model

The Gated Recurrent Unit (GRU) is a specialized variant of the recurrent neural network (RNN) architecture [14]. It was specifically designed to mitigate the vanishing gradient problem faced by traditional RNN when dealing with long sequences. GRU enhance the capabilities of RNN by introducing a "gate" mechanism, allowing them to more effectively capture long-term dependencies in sequence data.

A GRU is composed mainly of two types of gates: the update gate and the reset gate. The update gate is used to determine how much information from the previous hidden state should be retained when computing the new hidden state. This mechanism is akin to the forget gate in LSTM, but the GRU simultaneously deals with the preservation of past hidden states and the incorporation of new information in a single gate, whereas LSTM separate them. On the other hand, the reset gate controls how the past hidden state affects the new candidate hidden state, allowing the model to ignore past information when necessary.

These two gates work together, enabling the GRU to model input sequences more effectively, as they allow the model to decide when to retain past information and when to discard information no longer needed. While similar to LSTM, the simpler structure of GRU often gives them an advantage in terms of parameter count and computational complexity.

3.2 Dataset and Model Parameters

The dataset used in this paper originates from the simulation data in Section 2. Firstly, multiple experiments were conducted with an input flow ranging from 3600pcu/h to 4770pcu/h, increasing by 30pcu/h intervals. Each experiment, under a single flow condition, collected data every 10 seconds over a 15-minute duration, resulting in 90 sets of data per experiment, and a total of 18000 sets of data across all experiments. Since the initial dataset did not include specific accident information, we expanded the dataset by incorporating the following categorical variables: current scene traffic flow, the lane where the accident occurred, the precise location of the accident point within the diverging area.

The dataset was divided into a training set (80%) and a validation set (20%). The Mean Squared Error (MSE) function was employed as the loss function, serving to quantify the disparity between predicted values and actual values during the training process of the GRU model. Within deep learning algorithms, the assessment of the disparity between the learned outcomes and the sample labels stands as a pivotal step. Smaller disparities denote superior learning efficacy. This quantifiable disparity is referred to as the loss value. In this study, the Mean Squared Error (MSE) function was adopted as the loss function, which is particularly well-suited for addressing regression problems. It was utilized to evaluate the predictive performance of the model on both the training and validation datasets. The grid search method was implemented to ascertain the hyperparameter values of the GRU model, with the specific hyperparameter values detailed in Table 2.

Table 2. Values of parameters for GRU model.

Number of Neurons	Optimizer	Activation Function	Learning Rate	Batch Size	Epoch	Dropout
50	ADAM	Sigmoid	0.01	120	50	0.2

3.3 Model Results

In order to validate the predictive performance of the GRU model, this paper selected the LSTM model as a control group for comparison. Table 3 presents a comparison of the predictive results between the LSTM and GRU models. The results indicate that the GRU model outperforms the LSTM model in all evaluation metrics, with a lead of over 10% in each metric.

Table 3. Values of parameters for GRU model.

Impact area	Model	RMSE	MSE	MAE
1	GRU	0.01670	0.00028	0.01319
	LSTM	0.02091	0.00043	0.02498
2	GRU	0.00587	0.00003	0.00364
	LSTM	0.00712	0.00005	0.00458
3	GRU	0.00627	0.00004	0.00350
	LSTM	0.00734	0.00005	0.00486

4. CONCLUSION

This paper introduces the concept of risk propagation in highway diversion areas. To investigate risk propagation at various levels, a series of simulation experiments were conducted under highway diversion scenarios, and the risk propagation patterns were examined through risk indices.

Simulation results indicate that the extent of impact on different levels of affected areas varies under different traffic flows. Regardless of traffic flow, the first-level affected area, which is closest to the accident, remains in a state of risk for 15 minutes after the accident. In contrast, the second and third-level affected areas will be affected after some time has passed since the preceding area's impact. Additionally, due to the greater distance of the third-level affected area from the accident, it remains unaffected by the second-level affected area when traffic flow is low.

Furthermore, this paper presents a highway diversion risk prediction model based on GRU. This model performs well in predicting collision risk and can be utilized by researchers and professionals to predict and monitor risks in highway diversion areas in a potentially more accurate manner, thereby reducing the risk of traffic accidents.

Therefore, after an accident occurs, the affected road segments can be categorized, and dynamic control measures can be devised based on changes in risk within these areas. For instance, when an accident initially happens, the areas in close proximity to the accident site pose the highest risk, necessitating the implementation of more stringent control measures. Conversely, other areas with lower risk levels can have secondary control measures in place. However, as the risk propagates and the risk values begin to rise in areas farther from the accident site, it becomes crucial to adapt the formulation of control measures accordingly.

However, this study has limitations that should be further addressed in future research. Due to the absence of real-world traffic data collected from road segments, the driving patterns of vehicles may deviate from the simulation model, making it challenging to entirely replicate real-life diverging area scenarios. Additionally, while side collisions are the most prevalent collision type on highway diverging areas, other collision types should be considered in future work.

REFERENCES

- [1] Wang, J. Q., Wu, J., Zheng, X. J., Ni, D.H., Li and K.Q., "Driving safety field theory modeling and its application in pre-collision warning system," *Transp. Res. Part C* 72, 306-324(2016).
- [2] Babu, F., Wang, M. and Arem, B.V., "Probabilistic field approach for motorway driving risk assessment", *Transp. Res. Part C* 118, 102716(2020).
- [3] Lu, S.W., Zhang, L.F. and Fang, S.E., "Probabilistic Analysis and Risk Evaluation of Highway Rear-end Collision," *J. Tongji Univ. (Nat. Sci.)* 39(08), 1150-1154(2011).
- [4] Zhu, T., Zhao, Y.H. and Bai, Y., "Modeling real-time car-following risk based on probability computation," *Proc. IEEE ICICTA* 2, 268-272(2008).
- [5] Ma, Y.L., Fang, L.Y., Lv, T.L. and Guo, L., "Quantification classification method for driving risk quantization using Bayesian network," *J. Harbin Inst. Technol.* 52(3), 33-37(2020).
- [6] Han, Y. and Ahn, S. "Stochastic modeling of breakdown at freeway merge bottleneck and traffic control method using connected automated vehicle," *Transp. Res. Part B* 107, 146-166(2018).
- [7] Li, M.H., Zhu, Z.Y. and Ma, L., "Coverage Analysis and Simulation Study for Traffic Flow after Traffic Accidents on Expressways," *Tech. Highw. Transp.* 6, 109-111(2021).
- [8] Jin, S.X., Wang, J.J. and Xu, M.G., "Traffic accident affected zone division and traffic guidance under regional highway network," *J. Chang'an Univ. (Nat. Sci. Ed.)* 37(2), 89-98(2017).
- [9] Hao, Y., Sun, L.J. and Xu, T.D., "Traffic Breakdown Phenomenon and Evolution Model of Traffic Perturbation," *J. Tongji Univ. (Nat. Sci.)* 09, 1178-1194(2009).
- [10] Dong, G.H. and Zuo, Y.L., "Study on Calculation Model of Number of Queuing Vehicles and Total Delay in Traffic Incidents," *Comput. Technol. Autom.* 37(02), 110-116(2018).
- [11] Saeedmanesh, M. and Geroliminis, N., "Dynamic Clustering and Propagation of Congestion in Heterogeneously Congested Urban Traffic Network.," *Transp. Res. Part B* 105, 193-211(2017).
- [12] Zhu, S. and Jiang, R., "Review of Research on Traffic Conflict Techniques.," *J. Highway Transp.* 33, 15-33(2020).
- [13] Wang, J., Wang, S., Long, X., Li, D., Ma, C. and Li, P., "Ellipse-Like Radiation Range Grading Method of Traffic Accident Influence on Mountain Highways.," *Sustainability* 14, 13727(2022).
- [14] Cho, K., Merriënboer, B. V., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y., "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation.," *Proc. EMNLP*, 1724-1734(2014).