Inference Method for Abnormal Traffic Events in Highway Monitoring Blind Spots

Ziqi Tian [1, a], Chenyang Zhao [2, b], Yi Lin [3, c]

[1] Chang'an University; Sichuan Chengle Expressway Co., Ltd., China

[2] Sichuan Chengle Expressway Co., Ltd., China

[3] School of transportation engineering, Chang'an University, Xi'an 71000, China

[a] 764175633@qq.com, [b] 1025175446@qq.com, [c] 2939045700@qq.com

Abstract. Highways have the characteristic of closed-loop traffic flow, presenting unique challenges for traffic surveillance and management. This study focuses on addressing the problem of detecting abnormal events in areas that cannot be directly monitored. These areas typically lack the necessary infrastructure, such as cameras, radar, or other sensing devices, making it difficult to identify traffic abnormal events such as accidents or congestion. To tackle this challenge, this paper proposes a discriminative algorithm for detecting abnormal events in highway surveillance blind spots. Firstly, we simulate traffic scenarios including both abnormal events and normal situations using the VISSIM software, and select speed, density, and occupancy as feature parameters. Subsequently, this study applies the K-means clustering algorithm to judge whether an abnormal event occurs. Experimental results show that the proposed algorithm exhibits high precision (91.4%) in identifying abnormal events, with a false negative rate of only 4.17%. Moreover, the algorithm demonstrates good robustness against disturbances when individual raw parameters become anomalous.

Keywords: Highway, blind spots, abnormal events, indicator selection, clustering.

1. Introduction

Recent studies have focused on improving traffic safety through advanced data analysis methods and technologies. Zhang et al. [1] used Bayesian Networks (BN) to model the causal relationships between traffic events and traffic state parameters, updating event probabilities through bidirectional inference. Wu et al. [2] adopted a deep learning approach for traffic state discrimination, demonstrating higher detection accuracy through simulations in VISSIM software. Aliari et al. [3] employed neural network models and sliding time windows to monitor speed sequences, reducing false alarm rates in imbalanced datasets. C.L. Qiu et al. [4] proposed an improved DBSCAN clustering algorithm to identify accident-prone areas by adjusting parameters epsilon and minPts. Li et al. [5] introduced the Traffic Safety Status Deep Clustering Network (TSDCN) to cluster traffic safety conditions and quantify collision risk levels. Jia et al. [6] combined kernel density estimation (KDE) and spatial clustering techniques to assess land use characteristics, such as Point of Interest (POI) data. Chen et al. [7] conducted real-time traffic state recognition using traffic software simulation and the Fuzzy C-Means (FCM) algorithm. Du [8] improved the FCM algorithm using entropy weighting for feature selection and applied it to train a multi-class Support Vector Machine (SVM). Pi et al. [9] implemented traffic congestion event identification using Hard C-Means (HCM) and FCM algorithms. Li [10] utilized the FCM algorithm to identify different traffic state indicators, demonstrating its high feasibility. Wang et al. [11] proposed a fuzzy logic-based urban traffic accident prediction model and introduced a safety enhancement factor to improve road safety. Alkandari et al. [12] developed an accident detection system based on fuzzy clustering algorithms and demonstrated accident detection scenarios using the FuzzyTech program.

This study addresses the challenge of detecting abnormal events in highway surveillance blind spots. We propose a discriminative algorithm using VISSIM simulation and the K-means algorithm to identify abnormal events. This method aims to fill the research gap in existing traffic surveillance systems and enhance the safety of highway sections in surveillance blind spots.

2. Methodology

2.1. Data Source

Highway abnormal events, such as vehicle breakdowns or traffic congestion due to extreme weather, are simulated using VISSIM software. Scenarios include both abnormal event occurrences and non-occurrences, with a six-lane configuration in each direction. The simulation runs from 10:00 to 12:00, with an abnormal event occurring at 10:30. A total of 100 abnormal events and 20 non-abnormal events are simulated. Data collection points a and b are set up 50 meters upstream and downstream of the affected lane, respectively. Simulation parameters for abnormal event occurrences are detailed in Table 1. The simulation duration is 2 hours with data collected every 5 minutes. An example of the simulation and data collection is shown in Fig. 1.The simulation parameters for abnormal event occurrences are shown in Table 1. An example of the abnormal event occurrence simulation and data collection is illustrated in Fig 1.

Table 1. Parameters for the simulation of abnormal events occurrence

Parameters	Range	Units
Upstream traffic volume	6000, 5900, ..., 4000	veh/h
Number of lanes	3, 2 (one lane closed)	lanes
Proportion of large vehicles	10%, 15%, ..., 30%	%
Length of lane affected by the event	300	m

Figure 1. VISSIM simulation and data acquisition example

2.2. Selection of Discriminatory Indicators

When an abnormal event occurs on a highway, the traffic parameters recorded by the detection equipment exhibit noticeable differences, and the rates of change for these parameters also vary. To quickly detect abnormal events while ensuring that the changes in parameters are substantial, choosing the rate of change as an evaluation parameter is reasonable. As an assessment metric, the rate of change can quantify the speed of parameter changes before and after an abnormal event. Specifically, the rate of change can be used to measure the relationship between the magnitude of parameter changes and time during an abnormal event. Its calculation formula is:

$$CR=\frac{N_2/N_1}{\Delta T} \tag{1}$$

Where: N_2 represents the average value of the parameter after the abnormal event; N_1 represents the average value of the parameter before the abnormal event; ΔT is the time difference of the parameter change, measured in h.

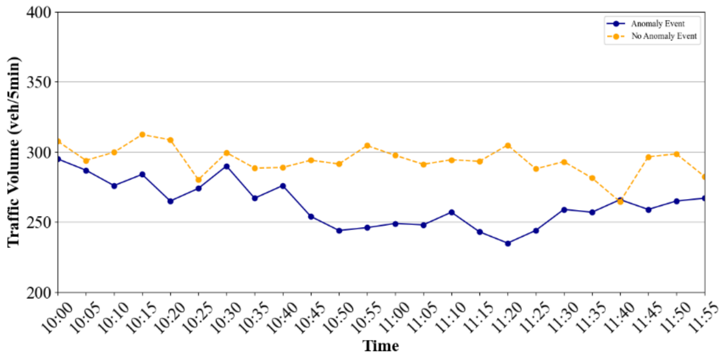

Figure 2. Comparison of changes in traffic volume with or without the occurrence of abnormal events

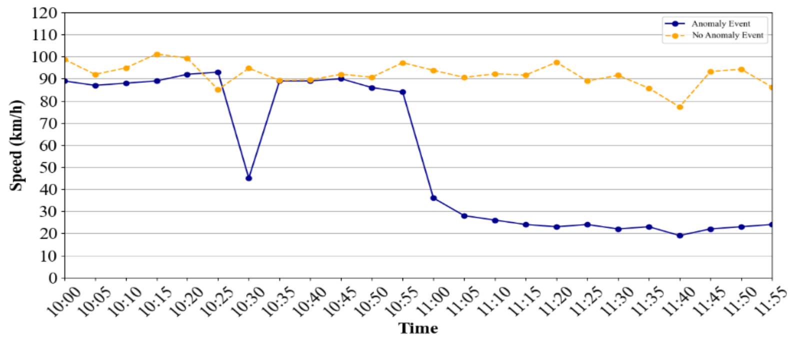

Figure 3. Comparison of changes in speed with or without the occurrence of abnormal events

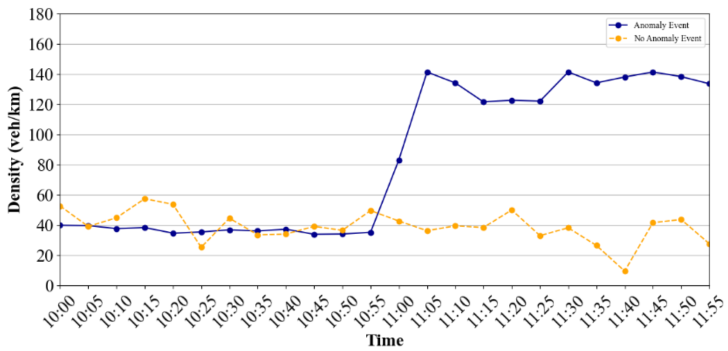

Figure 4. Comparison of changes in density with or without the occurrence of abnormal events

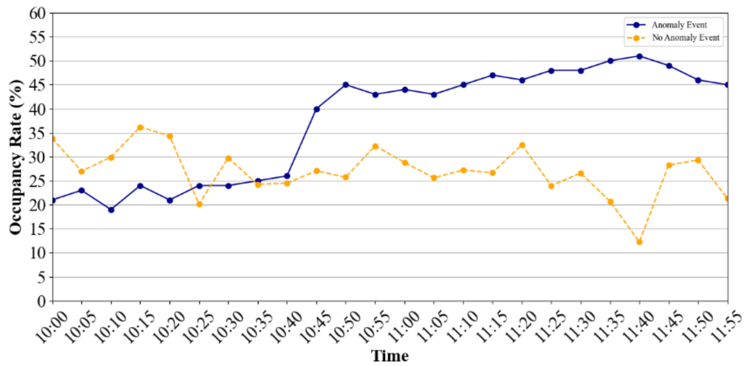

Figure 5. Comparison of changes in occupancy rate with or without the occurrence of abnormal events

Fig. 2, 3, 4, and 5 display the changes in traffic flow, speed, density, and occupancy rate over a 2-hour period for both non-occurrence and occurrence of abnormal events.

Using Equation (1) to calculate the rate of change, we can effectively discern the trend of parameter changes following an abnormal event and respond more promptly to these changes, thereby enabling timely detection of traffic abnormal events. Moreover, a higher rate of change indicates a larger change in the parameter, further enhancing the ability to identify abnormal events.

The rate of change for each parameter is calculated and summarized in Table 2. Among these parameters, speed, density, and occupancy show more significant performance in discriminating abnormal events, indicating that occupancy, speed, and density have stronger sensitivity and representativeness. Therefore, selecting speed, density, and occupancy as feature parameters is more suitable for discriminating whether an abnormal event occurs.

Table 2. CR values corresponding to different transportation parameters

Traffic Parameter Name	Rate of Change
Traffic volume	3.952
Density	13.751
Speed	17.542
Occupancy rate	14.844

2.3. Abnormal Events Occurrence Discrimination based on K-means Clustering Agorithm

2.3.1. Algorithm Selection

Clustering is an unsupervised learning method that groups similar data objects into distinct clusters, maximizing intra-cluster similarity and inter-cluster dissimilarity. Distance measures, such as Euclidean distance, are used to quantify object similarity and serve as the basis for grouping. Common clustering algorithms include k-means, density-based, and hierarchical clustering. For the purpose of discriminating between normal and abnormal events, the k-means algorithm is suitable due to its simplicity and computational efficiency. By setting the number of clusters to 2, this approach effectively differentiates between normal and abnormal traffic states, facilitating the identification of abnormal events.

2.3.2. Discrimination Steps Based on the K-Means Clustering Algorithm for Abnormal Events

The specific process for discriminating the occurrence of abnormal events based on the K-means clustering algorithm is as follows:

First, obtain traffic data for both the non-occurrence and occurrence of abnormal events through VISSIM simulation. Then, through analysis, select density, speed, and occupancy as feature parameters and establish a dataset *D*. Next, preprocess the data and divide the preprocessed dataset containing various sample data into training and testing sets. Set the number of cluster centers to 2 and randomly initialize the cluster centers. Follow the basic steps of the K-means clustering algorithm, iteratively updating the cluster centers until the clustering result converges and the final cluster centers are obtained. Combine the cluster centers for both the non-occurrence and occurrence of abnormal events, and calculate the Euclidean distance l_1 between the input feature data and the cluster center for the non-occurrence of abnormal events and the Euclidean distance l_2 between the input feature data and the cluster center for the occurrence of abnormal events. If $l_1<l_2$, then it is determined that an abnormal event has not occurred; otherwise, it is determined that an abnormal event has occurred, and the discrimination process ends.

In the preprocessing steps, the min-max normalization method was adopted. This method linearly scales the data to a specified range, which enhances the convergence speed and accuracy of the clustering algorithm and ensures that the contribution of different features to the clustering results is balanced.

$$x^*=\frac{x-x_{min}}{x_{max}-x_{min}} \tag{2}$$

To prevent the interference of abnormal data, clustering uses all three feature parameters. When one of the feature parameters encounters issues during the discrimination process, the impact on the overall discrimination is reduced, thus improving the robustness and disturbance resistance of anomaly event detection. The clustering effect of the speed-density-occupancy feature parameters is shown in Fig. 6.

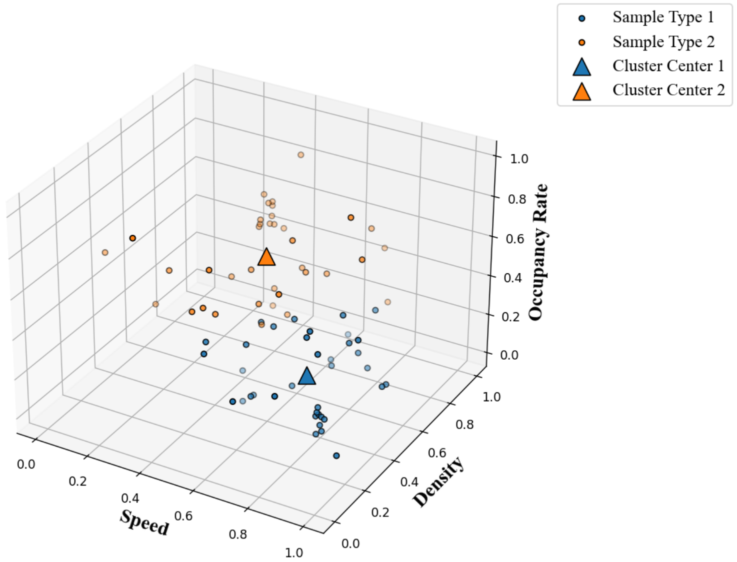

Figure 6. Speed-Density-Occupancy rate three feature parameters clustering effect diagram

The cluster center values obtained were then reverse-normalized using the following reverse-normalization formula:

$$x=x^{*}(x_{max}-x_{min})+x_{min} \tag{3}$$

As shown in Fig. 6, blue sample points represent the class categorized as non-occurrence of an abnormal event, while yellow sample points represent the class categorized as occurrence of an abnormal event. The blue triangle represents the cluster center for the non-occurrence of an abnormal event, and the yellow triangle represents the cluster center for the occurrence of an abnormal event. The specific values after reverse-normalization are shown in Table 3.

Table 3. Cluster Centers for Speed-Density-Occupancy Feature Parameters After Normalization

Traffic State	Speed (km/h)	Density (veh/km)	Occupancy rate (%)
Abnormal event Occurrence	30.6	117.6	43.3
Non-Occurrence of Abnormal event	78.4	46.9	31.8

3. Result

A dataset consisting of 90 instances of abnormal events and 18 instances of non-occurrence of abnormal events was used as the training set, while the remaining 10 instances of abnormal events and 2 instances of non-occurrence of abnormal events were used as the test set. Cross-validation was employed to evaluate the model's performance by repeatedly dividing the training set and test set, effectively reducing any bias in the model due to uneven data distribution. An abnormal event sequence was input, and the speed-density-occupancy three-feature-parameter clustering algorithm was applied for discrimination. The clustering results are shown in the figure below, and the discrimination results for whether an abnormal event occurs are presented in Fig. 7.

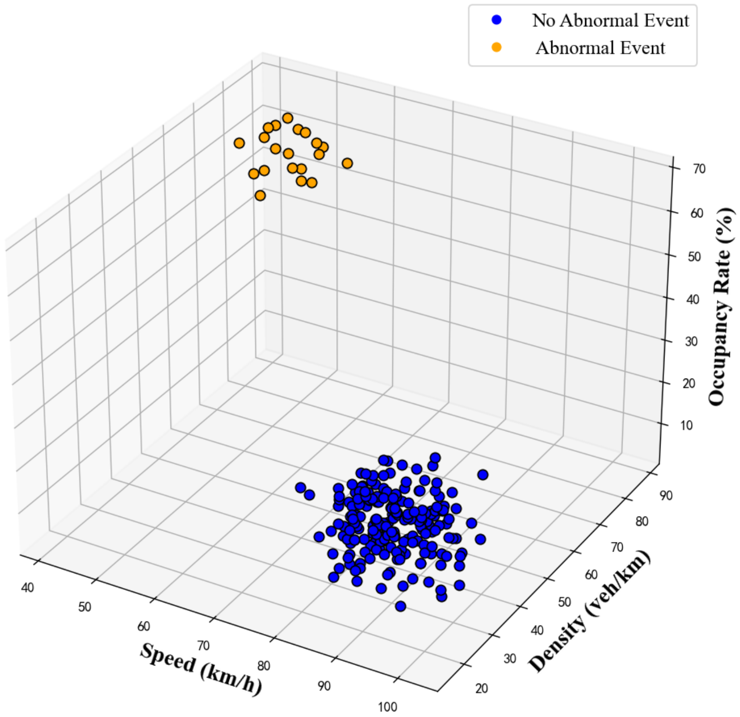

Figure 7. Discriminative results of whether an abnormal event occurs or not based on K-means clustering algorithm

It is evident that the model achieves a precision rate of over 90% in identifying abnormal events, with a false negative rate of 4.17%. This confirms the effectiveness of the proposed clustering algorithm in discriminating highway abnormal events. Furthermore, when one of the original three parameters exhibits abnormal behavior, the impact on the algorithm is minimal, indicating that the algorithm also possesses good anti-interference capabilities.

4. Conclusion

This study has some limitations in abnormal event detection on highways based on the K-means algorithm:

(1) This paper analyzes data simulated by VISSIM, which may present issues such as deviations from real-world scenarios, limitations in the selection of feature parameters, insufficient robustness to disturbances, and problems with data quality and quantity.

(2) The sensitivity of the K-means algorithm to initial cluster centroids can lead to unstable results, and its assumption that clusters are spherical may not align with the actual distribution of data. Additionally, K-means performs poorly with noise and outlier data.

To overcome these limitations, future research could consider incorporating more diverse real traffic data, exploring richer feature parameters, optimizing clustering algorithms to improve handling of noisy data and outliers, and attempting model fusion or ensemble learning methods to enhance the overall performance and stability of the detection system.

References

[1] Zhang K, Taylor M A P. Effective arterial road incident detection: a Bayesian network based algorithm [J]. Transportation Research Part C: Emerging Technologies, 2006, 14(6): 403-417.

[2] Wu Zhiyong, Ding Xiangqian, Ju Chuanxiang. A Discretized Traffic State Discrimination Method Based on Deep Learning [J]. Journal of Transportation Systems Engineering and Information Technology, 2017, 17(05): 129-136.

[3] Aliari S, Sadabadi K F. Automatic detection of major freeway congestion events using wireless traffic sensor data: machine learning approach [J]. Transportation Research Record, 2019, 2673(7): 436-442.

[4] C.L. Qiu, H.Y. Xu, Y.Q. Bao, Modified-DBSCAN Clustering for Identifying Traffic Accident Prone Locations, Intelligent Data Engineering And Automated Learning - Ideal 2016, 2016, pp. 99-105.

[5] H.T. Li, Q.W. Bai, Y.H. Zhao, Z.W. Qu, W. Xin, TSDCN: Traffic safety state deep clustering network for real-time traffic crash-prediction, Iet Intelligent Transport Systems 15(1) (2021) 132-146.

[6] R. Jia, A. Khadka, I. Kim, Traffic crash analysis with point-of-interest spatial clustering, Accident Analysis And Prevention 121 (2018) 223-230.

[7] Chen Huiru. Research and Application of Real-Time Traffic State Discrimination Methods for Basic Sections of Expressways [D]. Chang'an University, 2015.

[8] Du Chong. Study on Real-Time Traffic State Discrimination Methods for Expressways [D]. Beijing Jiaotong University, 2017.

[9] Pi Xiaoliang, Wang Zheng, Han Hao, et al. Application Research on Traffic State Classification Method Based on Information Collected by Inductive Loop Detectors [J]. Journal of Highway and Transportation Research and Development, 2006, (04): 115-119.

[10] Li Xiaolu. Traffic State Discrimination Based on Fuzzy C-Means Clustering for Urban Road Traffic [J]. Journal of Transportation Science and Economics, 2016, 18 (04): 32-36+42. 2016.04.009.

[11] H. Wang, L. Zheng, X.H. Meng, Traffic Accidents Prediction Model Based on Fuzzy Logic, Advances In Information Technology And Education, PT I, 2011, pp. 101-+.

[12] A. Alkandari, F.A. AlAwadhi, Ieee, Dynamic Fuzzy Logic Traffic Light Integrated System with Accident Detection System Using FuzzyTech Program, 2015, pp. 69-75.