Inference Method for Abnormal Traffic Events in Highway Monitoring Blind Spots

Guomin Xu[1, a], Sibo Zhang[2, b], Yuxuan Han[3, c], Yi Lin[2, d] and Feng Xiao[4, e]

[1] Sichuan Chengmiancangba Expressway Construction&Development Gruop Co., Ltd., China;

[2] Chang'an University, China;

[3] Sichuan Tibetan Area Expressway Co., Ltd., China;

[4]Sichuan Yakang Expressway Co., Ltd., China;

[a] 342493527@qq.com, [b] 1255087665@qq.com, [c] 1003040275@qq.com, [d] 2939045700@qq.com, [e] 1034921877@qq.com

Abstract. Highway monitoring is critical for traffic safety and efficiency. However, monitoring blind spots—areas between surveillance points—poses significant challenges. This study proposes an inference method to detect abnormal traffic events in highway blind spots by analyzing vehicle counts from adjacent monitored regions. Utilizing Vissim and LumenRT for simulation, we create realistic highway scenarios to collect time series data. Cross-correlation is used to determine the time lag between vehicle counts, and Pearson's correlation coefficient evaluates the correlation between time-shifted data. The results show significant correlation under normal conditions, which decreases noticeably during abnormal events, confirming the method's effectiveness. This approach enhances the ability to monitor and respond to traffic anomalies in blind spots, improving overall highway safety and efficiency.

Keywords: Expressway; Highway Monitoring; Traffic Anomalies; Blind Spots; Time Series Analysis

1. Introduction

The safety and efficiency of highway travel depend significantly on effective monitoring and timely detection of abnormal traffic events. Abnormal traffic events (speeding, reverse traffic, abnormal parking, etc.) can be used as the prior indicators of traffic accidents. The abnormal event detection of expressway traffic is generally identified by manual inspection or the event detector of road monitoring system[2]. However, these systems have limitations, especially in areas where direct monitoring is not feasible, known as blind spots. These blind spots can be stretches of road between monitoring stations where incidents may go undetected until their effects propagate to monitored sections.

The hypothesis guiding this research is that the vehicle count in two contiguous monitoring regions should remain similar over time if there are no entry or exit points between them. Significant deviations from this expected consistency can indicate abnormal traffic events, such as congestion or accidents. This study aims to infer the occurrence of such events in highway monitoring blind spots by analyzing data from adjacent monitored regions.

2. Methodology

2.1 Simulation Scenarios

The simulation scenarios are designed using Vissim and LumenRT to replicate a typical highway environment with specific monitoring setups. The key elements of the simulation scenarios include:

- Highway Layout: A stretch of highway with two monitoring regions, A and B, is simulated. These regions are separated by a segment of road where direct monitoring is not feasible, creating a blind spot.

• Traffic Flow: Various traffic conditions are simulated, including normal flow, peak hours, and incidents such as accidents and congestion.

• Monitoring Tools: The simulation includes various monitoring tools like cameras, radars, loop detectors, and ETC gantries. However, for the purpose of this study, the focus is on data collected from regions A and B without intermediate monitoring.

2.2 Data Collection

This study utilized VISSIM software to conduct traffic flow simulation experiments. The specific experimental method involved constructing a highway model with a 300-meter blind spot section where monitoring is not set. The entrance of this section was designated as Section A, and the exit as Section B. Using VISSIM simulation software, data on traffic flow and speed at Sections A and B were obtained under normal traffic conditions, as well as under traffic anomaly conditions such as congestion and accidents.

Traffic volume refers to the number of vehicles passing through a specific roadway segment within a given unit of time. In practical traffic volume statistics, hourly traffic volume is commonly used for comparison. However, the observation period is often not the standard one hour, necessitating the conversion of software-collected data into hourly traffic volume.

2.3 Identification Indicators

The purpose of time series analysis is to explore and quantify the relationships between time series datasets[7]. It focuses on the degree of correlation between two variables, merely describing the association patterns of these variables over time without implying causality. Currently, the main methods for analyzing the relationships in time series data include correlation coefficient analysis[3], cross-correlation analysis[8], Granger causality tests[9], and dynamic time warping[10]. The primary focus of this section is to establish a causal relationship hypothesis between the time series data from the two monitoring regions (A and B) using correlation analysis. The specific steps involve determining the time lag using cross-correlation and then evaluating the correlation using Pearson's correlation coefficient.

2.3.1 Cross-Correlation Analysis

The correlation number takes into account the lag difference of the data points, and the property of the correlation coefficient is more suitable to show the covariance direction of the two variables over time[4]. Cross-correlation is used to determine the time lag (or delay) between the vehicle counts at regions A and B. This step is crucial as it accounts for the time it takes for vehicles to travel from one monitoring region to the other[5]. The steps involved are:

Compute Cross-Correlation: Calculate the cross-correlation function between the time series data of vehicle counts from regions A and B.

Identify Maximum Cross-Correlation: Identify the time lag at which the cross-correlation function reaches its maximum value.

$$\tau = arg \max_{\tau'} \left[\sum_t (A_t - \bar{A})(B_{t+\tau'} - \bar{B}) \right] \tag{2.1}$$

where τ is the time lag, A_t and B_t are the vehicle counts at time t in regions A and B, respectively, and $\bar{A}$ and $\bar{B}$ are their respective means.

2.3.2 Pearson Correlation Coefficient

After determining the appropriate time lag, the next step is to assess the correlation between the time-shifted vehicle counts using Pearson's correlation coefficient. This measure quantifies the linear relationship between the two time series. The steps are:

Time-Shift Adjustment: Adjust the time series data from region B by the identified time lag. This alignment accounts for the travel time between the regions.

Compute Pearson Correlation: Calculate the Pearson correlation coefficient for the time-shifted data. This coefficient indicates the strength and direction of the linear relationship between the vehicle counts at regions A and B[6].

$$r = \frac{\sum_t (A_t - \overline{A})(B_{t+\tau)} - \overline{B})}{\sqrt{\sum_t (A_t - \overline{A})^2 \sum_t (B_{t+\tau} - \overline{B})^2}} \tag{2.2}$$

where r is the Pearson correlation coefficient, and other symbols are as defined above.

3. Results

3.1 Time Series Analysis

The application of time series forecasting in traffic field mainly focuses on traffic volume forecasting[1]. In this section, we examine the time series data of vehicle counts from monitoring regions A and B, considering the time lag determined through cross-correlation analysis. The objective is to establish a baseline understanding of the traffic flow patterns under normal conditions and to identify any significant deviations that may indicate abnormal events.

3.1.1 Time Lag Analysis: Cross-Correlation Coefficient

In this study, the cross-correlation coefficient calculation method shown in Equation (2-1) was utilized to compute the cross-correlation coefficient between the traffic volumes at Sections A and B. It was observed that the cross-correlation coefficient between the traffic volumes at Sections A and B sharply increased near a lag time of -10 seconds and reached its maximum value. This indicates that the traffic volume at Section A has a significant impact on the traffic volume at Section B 10 seconds later.

3.2.2 Correlation Analysis: Pearson Correlation Coefficient

This study employed the Pearson correlation coefficient calculation method shown in Equation (2-2) to compute the correlation coefficients and corresponding P-values for the traffic volumes at Sections A and B under various lag times.

Table 1. Correlation analysis of volume time series between section A and Section B

Condition	Time series X	Time series Y	Pearson correlation coefficient(r)	P value	Testing result
Normal	Section A A_{t1}	Section B $B_{t1+\tau}$	0.6064	3.8701×10^{-6}	Significant
Abnormal	Section A A_{t2}	Section B $B_{t2+\tau}$	0.1026	0.4784	No Significant

The results indicate that under normal conditions, setting the lag duration to -10 seconds yields a significant correlation coefficient between the traffic volumes of sections A and B, with an r-value of 0.6064. However, in the case of abnormal events, the p-value between the traffic volumes of the two sections reaches 0.4785, which is much higher than the threshold.

3.2 Abnormal Event Identification

In this section, we analyze and verify the flow diagrams of two monitoring sections, A and B, based on the determined lag duration. Through comparative analysis, it is expected that the traffic flow of these two sections should show a significant correlation under normal conditions, whereas during abnormal events, the traffic flow of the two sections would lack such correlation.

First, we adjust the time series data of monitoring section B by applying the previously determined lag duration. Next, we compare the traffic flow diagrams of sections A and B. The specific procedure involves plotting the vehicle counts over time for both sections, with the data for section B time-shifted accordingly.

Under normal traffic conditions, we expect the flow diagrams of sections A and B to exhibit a significant correlation, meaning that the vehicle counts in both sections should display similar trends during peak and off-peak periods. By inspecting these flow diagrams, we can confirm this correlation. Furthermore, we have statistically validated this correlation by calculating the Pearson correlation coefficient, as discussed earlier in this chapter.

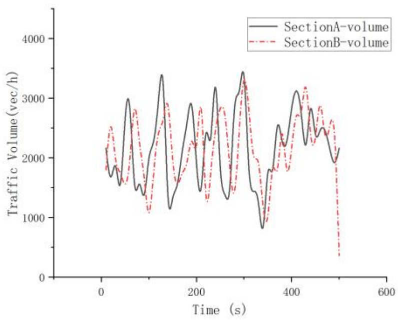

Fig. 1 Normal section time and traffic flow comparison

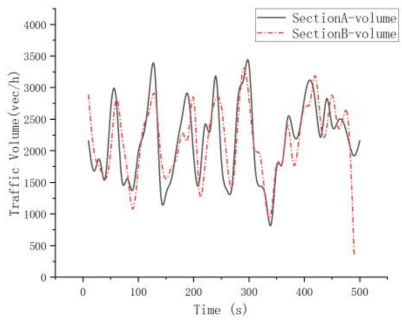

Fig. 2 Normal section time and traffic flow comparison with lag time

During known abnormal events, we expect that the flow diagrams of the two sections will no longer exhibit a significant correlation, meaning that the trends in vehicle counts will no longer be consistent. By inspecting the flow diagrams, we can identify this discrepancy. Additionally, we calculate the Pearson correlation coefficient, as discussed earlier in this chapter, to validate the decrease in correlation under abnormal conditions.

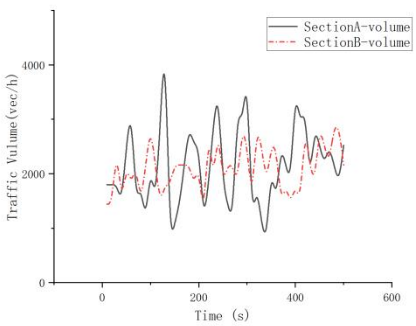

Fig. 3 Abnormal section time and traffic flow comparison

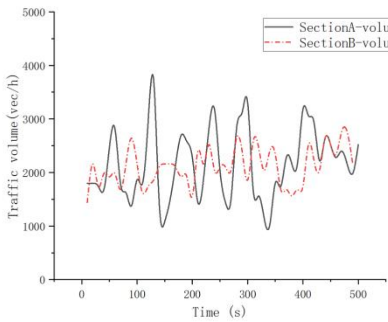

Fig. 4 Abnormal section time and traffic flow comparison with lag time

Through this analytical method, we can validate the following hypothesis: under normal conditions, the traffic flow between the two monitoring sections shows a significant correlation; whereas, during abnormal events, this correlation significantly decreases. This analytical process confirms the effectiveness of inferring abnormal traffic events in highway monitoring blind spots based on time series correlation.

4. Conclusion

In this study, we developed an inference method for detecting abnormal traffic events in highway monitoring blind spots by leveraging the vehicle count data from adjacent monitored regions. We applied cross-correlation analysis to determine the appropriate time lag and used Pearson's correlation coefficient to evaluate the correlation between the time-shifted vehicle counts from the two monitoring regions.The results validated our hypothesis, demonstrating that under normal traffic conditions, the traffic flow between the two regions showed significant correlation, while during abnormal events, this correlation decreased noticeably. This confirms the effectiveness of our proposed method in inferring abnormal traffic events in highway monitoring blind spots.

Acknowledgements

This research was funded by Construction S & T Project of Department of Transportation of Sichuan Province (2022-ZL-04), Joint Technical Research on Tibetan Expressway in Sichuan (Tibetan Expressway Science and Technology 2021-11).

References

[1] Wang Bo, Liu Changhe, Zhang Chi. Crash risk early warning in highway work zone based onroad surveillance camera [J]. Journal of Zhejiang University (Engineering Science),2024,58(06):1221-1232.

[2] Zhao Huaixin, Zhang Yingjie, Deng Ranran. Expressway anomaly event recognition method based on clustering by fast search and find of density peaks [J]. Journal of Chang'an University(Natural Science Edition),2018,38(05):205-212.

[3] Chung Y, Recker W W. A methodological approach for estimating temporal and spatial extent of delays caused by freeway accidents[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(3): 1454-1461.

[4] Erdem O, Ceyhan E, Varli Y. A new correlation coefficient for bivariate time-series data[J]. Physica A: Statistical Mechanics and its Applications, 2014, 414: 274-284.

[5] Delbart C, Valdes D, Barbecot F, et al. Temporal variability of karst aquifer response time established by the sliding-windows cross-correlation method[J]. Journal of Hydrology, 2014, 511: 580-588.

[6] Asuero A G, Sayago A, González A G. The correlation coefficient: An overview[J]. Critical reviews in analytical chemistry, 2006, 36(1): 41-59.

[7] Chandra S R, Al-Deek H. Cross-correlation analysis and multivariate prediction of spatial time series of freeway traffic speeds[J]. Transportation Research Record, 2008, 2061(1): 64-76.

[8] Shen C. Analysis of detrended time-lagged cross-correlation between two nonstationary time series[J]. Physics Letters A, 2015, 379(7): 680-687.

[9] Siggiridou E, Koutlis C, Tsimpiris A, et al. Evaluation of Granger causality measures for constructing networks from multivariate time series[J]. Entropy, 2019, 21(11): 1080.

[10] Izakian H, Pedrycz W, Jamal I. Fuzzy clustering of time series data using dynamic time war** distance[J]. Engineering Applications of Artificial Intelligence, 2015, 39: 235-244.