Traffic Analysis

Highway Incident Detection

Background

Traffic incidents such as accidents, vehicle breakdowns, or sudden obstructions are among the main causes of unexpected congestion, travel time unreliability, and secondary crashes (Coursey, 2024). Rapid and accurate detection of incidents is therefore crucial for reducing their impact by enabling timely response. Incident detection systems should be designed to ensure a high detection rate, a low false alarm rate, and a short detection time to enable timely response.

In recent years, the growing availability of traffic data—from probe vehicles, mobile phones, and other connected sources—has opened new possibilities for data-driven incident detection. These sources provide continuous and high-resolution information on traffic dynamics across the entire network, even on roads without fixed sensors. Combined with machine learning (ML) techniques, they offer a promising way to automatically identify abnormal traffic patterns—such as sudden speed drops, unusually high congestion, or increased variability in travel times—that are indicative of incidents or disruptive events (Yijing, 2023).

Machine learning algorithms are particularly suited to this task because they can capture complex, nonlinear relationships and temporal dependencies that traditional threshold-based methods often fail to represent. One approach to characterize normal traffic behavior is kernel density estimation (KDE), a non-parametric technique used to estimate the probability distribution of traffic variables, such as vehicle speeds or travel times. By modeling the distribution of traffic under normal conditions, KDE provides a flexible baseline against which deviations can be detected (Le, 2022). When observed traffic patterns fall in regions of low estimated probability, they are flagged as potential incidents. KDE is particularly advantageous because it does not assume a specific parametric form for traffic behavior and can naturally handle multi-modal distributions, which are common during rush hours or in heterogeneous traffic networks.

By training ML models and combining them with KDE-based probabilistic thresholds on historical data, it becomes possible to detect incidents effectively: high detection rates can be achieved while keeping false alarm rates low, and unusual traffic patterns can be identified promptly, supporting rapid response and mitigation. Traditional approaches typically rely on data from inductive loops or radar sensors to monitor changes in traffic flow or speed. When deviations exceed predefined thresholds, these methods flag potential incidents. A well-known example of this type of threshold-based detection is the California algorithm, which identifies incidents by comparing real-time traffic conditions to expected patterns and signaling an alert when abnormal drops in speed or flow are detected (Payne, 1978). Although effective in some settings, traditional approaches tend to produce false alarms, miss subtle incidents, and are difficult to scale due to limited sensor coverage. In addition, they are often insensitive to contextual factors such as time of day, weekday, weather, or road type. For example, recurring congestions which follow predictable temporal patterns may be mistakenly identified as incidents. These limitations highlight the need for more adaptive and data-driven approaches that can account for temporal, spatial, and contextual variations in traffic conditions.

Purpose and aim

The motivation behind this project is to develop and evaluate machine learning algorithms for incident detection based on large-scale traffic data. The aim is to improve detection accuracy, reduce response time, and provide traffic management authorities with a more intelligent and scalable tool for proactive decision-making. Furthermore, the project also examines the performance of data-driven incident detection methods in comparison with traditional approaches, such as the California algorithm.

Organization

Project supervisors: Henrik Fredriksson (henrik.fredriksson@liu.se) and David Gundlegård (david.gundlegard@liu.se)

Goals and requirements

The goal of the project is to find suitable machine learnings algorithm for highway incident detection. This process includes several steps:

- A literature study covering machine learning algorithms for incident detection
- Exploratory data analysis of the incident dataset
- Implementation of machine learning algorithms for incident detection
- Implementation of traditional incident detection algorithms, e.g., the California algorithm
- Analysis of methods performance regarding detection rate, false alarm rate, and time to detection

This project will include mathematical modeling and programming related to machine learning and data analytics. The group should therefore be knowledgeable in programming, preferably Python or MATLAB. Real-world incident data sets will be made available for the group.

Project group

The project group should be con`ducted by a group of at least three and at most five students.

Project grading

Except for the common requirements related to course grading specified in the course information, project specific grading is given according to the scale Fail; 3,4, or 5 (or the corresponding ECTS grade).

For grade 3, at least the following steps must be carried out:

- Write a detailed project specification and time plan in correct and readable English.
- Write a final report in English and make an oral presentation of the project results.
- Give an overview on background, perform a literature review and describe related work.
- Perform exploratory analysis of the provided incident dataset.
- Implement and evaluate a threshold-based method on the provided dataset.

For grade 4, the grade 3 steps must be carried out with very good results and the following additional steps must be carried out:

- Implement and evaluate a modified version of the California algorithm and analyze the results

For grade 5, the grade 3 and 4 steps must be carried out with very good results and the following additional steps have to be carried out:

- Implement, train and evaluate an incident detection method based on kernel density estimation.

References

Coursey, A., Ji, J., Quinones Grueiro, M., Barbour, W., Zhang, Y., Derr, T., ... & Work, D. (2024). FT-AED: Benchmark dataset for early freeway traffic anomalous event detection. Advances in Neural Information Processing Systems, 37, 15526-15549.

Yijing, H., Wei, W., He, Y., Qihong, W., & Kaiming, X. (2023). Intelligent algorithms for incident detection and management in smart transportation systems. *Computers and Electrical Engineering*, *110*, 108839.

Le, K. G., Liu, P., & Lin, L. T. (2022). Traffic accident hotspot identification by integrating kernel density estimation and spatial autocorrelation analysis: a case study. *International journal of crashworthiness*, *27*(2), 543-553.

Payne, H. J., & Tignor, S. C. (1978). Freeway incident-detection algorithms based on decision trees with states. *Transportation Research Record*, (682).