Towards Efficient Incident Detection in Real-time Traffic Management

Ferran Torrent-Fontbona[1], Monica Dominguez[1], Javier Fernandez[1], Jordi Casas[1]
[1] Aimsun SLU, Barcelona, Spain

Abstract

Incident detection is a key component in real-time traffic management systems that allows efficient response plan generation and decision making by means of risk alerts at critical affected sections in the network. State-of-the-art incident detection techniques traditionally require: i) good quality data from closely located sensor pairs, ii) a minimum of two reliable measurements from the flow- occupancy-speed triad, and iii) supervised adjustment of thresholds that will trigger anomalous traffic states. Despite such requirements may be reasonably achieved in simulated scenarios, real-time downstream applications rarely work under such ideal conditions and must deal with low reliability data, missing measurements, and scarcity of curated incident labelled datasets, among other challenges. This paper proposes an unsupervised technique based on univariate timeseries anomaly detection for computationally efficient incident detection in real-world scenarios. Such technique is proved to successfully work when only flow measurements are available, and to dynamically adjust thresholds that adapt to changes in the supply. Moreover, results show good performance with low-reliability and missing data.

Keywords: incident detection, situational awareness, data-driven, unsupervised

1 Introduction

Road operators and traffic authorities need to face and adapt to the increasingly complex scenarios of the 21st century. Data shows that immediate action is of paramount importance: traffic accidents in the European Union caused 22.800 deaths and 120.000 seriously injured people in 2019 [ECo19]; traffic congestion costs represent 1 % of GDP each year [ECA19]; and 80 %–95 % greenhouse gas (GHG) emissions need to be reduced by 2050 [EEA21]. In this context, real-time situational awareness systems emerge for traffic management systems to react. Non-recurrent situations, such as incidents, appear as outliers or anomalies in the real-time traffic data feeds. Hence, outlier detection becomes a key task for road traffic incident detection systems[1] that has been considered in the state of the art.

Road traffic incident detection as reviewed by [Pay76] has traditionally been regarded as the start of the California series algorithms. California #7 (probably the best well known in the series) aims to detect incidents comparing occupancy levels between consecutive loop detectors. A modified version of California #7 is proposed in [Sai18], which includes also flow differences between consecutive detectors. Both approaches are limited to highways with sensors measuring occupancy (and flow) with consecutive sensor-pairs close enough to each other to detect differences caused by incidents. California #7 and its derivates rely on static thresholds to detect incidents. Despite these thresholds are robust, a good adjustment of these represents a significant boost in the performance [Nat17]. Therefore, California #7 and its derivates need to be adjusted in a supervised way to reach top performance.

To overcome California #7-related algorithms' limitations, [Zhu19] propose to use convolutional neural networks (CNN), multi-layer perceptron, or random forest to build a supervised model to transform timeseries data to another space where anomaly detection is made easier. The authors test different approaches with synthetic data and found CNN to outperform the others. However, labelled incident data is scarce and prone to errors (e.g. lack of starting or ending time of the incidents, among others). As supervised incident detection models have a low chance of being applied in the real world, unsupervised techniques need to be explored.

[Her18] use flow and occupancy data between consecutive sensors to build gaussian mixture models (GMM) for unsupervised incident detection. Scarce outliers are detected according to the GMM and a threshold. Despite the advantage of the unsupervised approach, its application is limited to highways with sensors measuring flow and occupancy and that are close enough so that the impact of incidents can be detected as differences between sensor-pairs. Similarly, [Men19] propose a similar approach but first converting flow and occupancy data to a 5-level quantization (or congestion rank) of the traffic state. Then, incidents are detected as traffic state differences between sensor-pairs that are non-existent or have a very low probability according to the training dataset.

With the vision to have an unsupervised incident detection system that can work with only flow, [Yan20] used an autoencoder trained over all sensors of the network. The autoencoder reduces the dimensionality. Under the assumption that incidents are structurally different to the normal data and scarce, they would appear as outliers in the space of the autoencoder. However, setting the parameters of an autoencoder is not easy specially when input dimensionality can range from tens of sensors to tens of thousands or even more. Besides, training an autoencoder becomes challenging because it easily finds non-generalizable correlations that overfit the model requiring an amount of data that may not be available.

The above brief literature review shows that efficient incident detection systems are not yet available for downstream application in traffic management systems. To overcome such limitation, this paper presents an incident detection component included in Aimsun Live with

[1] In this paper, we refer to road traffic incidents as sudden changes of the supply.

the following characteristics: i) the system is trained and works in an unsupervised way; ii) it works independently on different timeseries (e.g. traffic flow, or speed, or occupancy); and iii) it works without the need of close sensor-pairs.

2 Methodology

2.1 Data Transformation

In general, one can distinguish between two types of anomalies or outliers: distributional outliers and structural outliers [Her22]. Distributional outliers look like normal data (i.e. inliers), but are in a low-density region in the data (or embedding) space. Since distributional outliers are structurally like inliers, their detection requires a lot of normal data for an accurate estimation of the probability density function and the setting of the probability threshold that defines the frontier of inliers. On the other hand, structural outliers are those that belong to a manifold[2] different from the one formed by inliers. Road traffic incidents belong to the latter category, whose detection is not equally easy for all traffic variables. For example, road traffic occupancy usually offers easier incident detection than traffic flow. Similarly, occupancy difference between sensor-pairs offers an easier detection than single-sensor occupancy. Moreover, different embedding spaces can offer distinct levels of detection power.

The incident detection analyzed in this paper transforms individual timeseries to a space where the distance between structural outliers in the timeseries are magnified. Note that the algorithm applies the same kind of transformation to different input variables, but it works independently on each timeseries.

2.2 Outlier Detection

Structural outliers occur separately from processing normal data. For example, traffic flow is created with the interaction of traffic demand and traffic supply. Traffic demand and supply can change over time and so does the flow. But when there is a sudden change like a lane occlusion due to a breakdown vehicle, this represents a structural change of the process, and it is an infrequent event. Taking advantage of this, we propose a dynamic boundary to detect incidents based on the assumption that incidents are rare. Therefore, we propose to analyze D days (e.g. 30 days of data) for each timeseries, and set the boundary based on where most data are located for all analyzed days.

2.3 Incident Affectation

Since incidents affect traffic in the location of the incident, outlier detection should report outliers for all measured traffic variables recorded by sensors at such specific location. According to this, we propose merging outlier detection from different timeseries in the same

[2] A topological space that is modeled closely on Euclidean space locally but may vary widely in global properties.

segment. Thus, an incident is detected if there is enough outlier detection support. This bottom-up approach enables the system to run independently on what is measured in each segment of the network, but at the same time, to take advantage of the fact that various timeseries are measured in the same segment.

2.4 Incident Duration, Localization, and Severity

The end of an incident is established when traffic returns to normality, i.e. to what is expected in recurrent conditions. According to this, once incident affectation is estimated, the end of the incident is computed as that timestamp after the start of the incident when outlying timeseries return to expected values. Note that the concept of expected values requires an accurate prediction of the timeseries before the incident, i.e. in recurrent conditions or in normal conditions. The presented results are achieved using the data-driven traffic prediction of Aimsun Live.

The system also backpropagates incident trail throughout the network graph to detect the origin of the incident. Severity is estimated considering the difference between what was expected in normal conditions and what is measured with the incident. Incident severity can also be estimated using the distance of the outlier in the embedding space. However, this often leads to results that are difficult to interpret or cause confusion amongst traffic engineers.

3 Results and Discussion

The case of a highly transited highway is presented as an illustration of the proposed methodology. The highway is mostly three or four lanes wide in either direction which carry constant heavy traffic (bound to congestion) during daylight hours, seven days a week. Data from loop detectors as well as labelled incidents from April and May 2022 have been used in this experiment.

Figure 1, Figure 2 and Figure 3 show the measured flow by three different loop detectors, the predicted flow, and the incident detection output as a binary variable to represent whether there is or not an incident. All three examples show how incidents cause a drop of flow (structural outlier) which is detected as an incident. Moreover, Figure 1 shows the ability of the incident detection of working on a timeseries with gaps of data and high variability.

The analyzed period included eight labelled incidents with an impact on the traffic flow. All of them were detected, meaning a recall of 100 %. It is worth mentioning that three of the eight incidents had an impact in the magnitude of the inherent variability of the timeseries, demonstrating the robustness of the system to detect incidents with less severe road affection. Specificity is hard to estimate in this use case, because not all incidents in the period were labelled incidents. Therefore, not all negatives in the ground truth were negatives. However, the positive rate in the period (rate of incident detection) was 0.23 %.

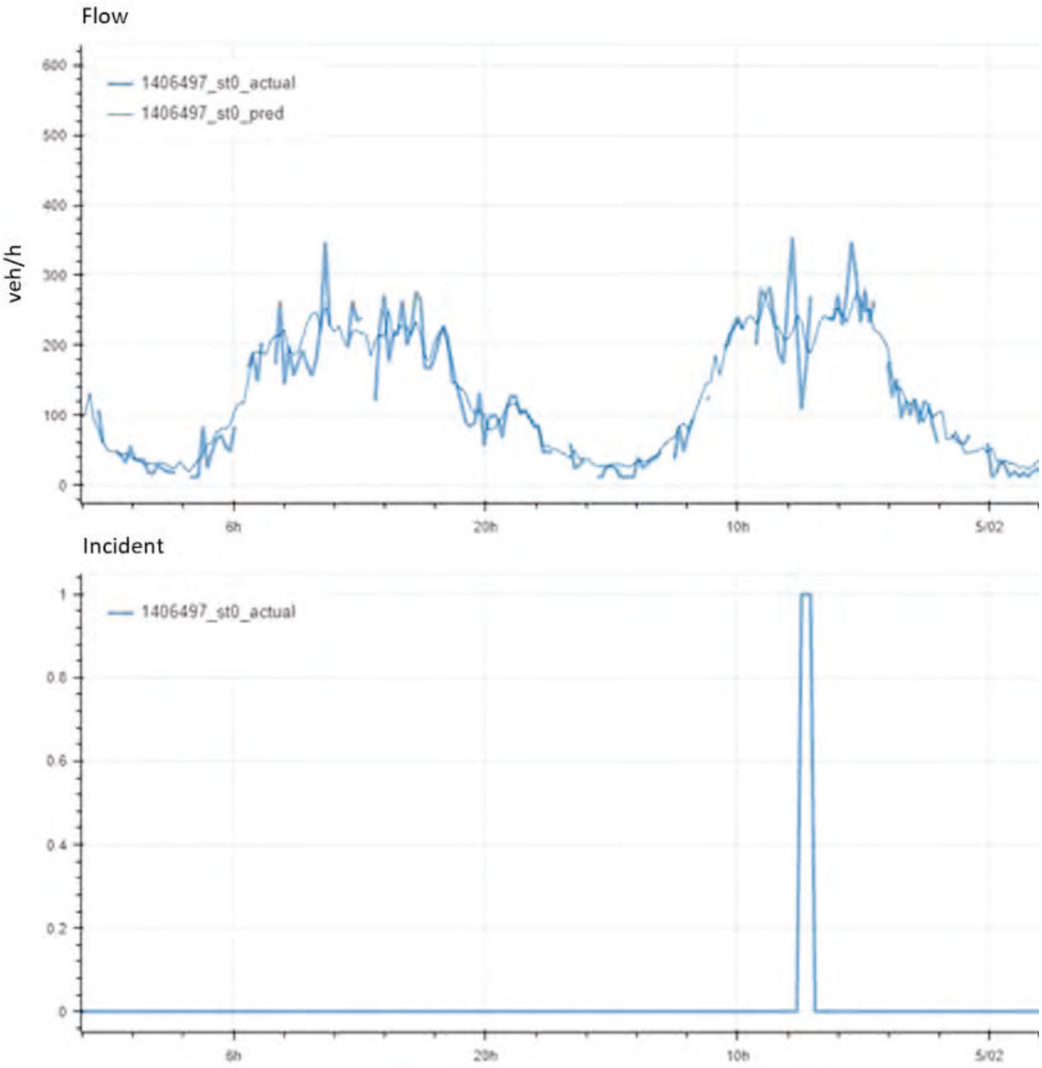

Figure 1: Incident detection of one occurred at 1st of May at 13:46. Top: Measured and predicted flow of the loop detector 1406497_st0. Bottom: incident detection.

This value means a detection of 1 incident every five days in each detector. Such value may rise the suspicion of false positives, however, the lower bound (assuming all positives as false positives) of the specificity is 99.76 %. Results proof the ability of the proposed incident detection methodology of (all) detecting true incidents while keeping the rate of false positives very low.

4 Conclusions

The task of automatically detecting road incidents poses great challenges to both research and industries in the traffic management sector still now-days. Literature in the area often neglects the fact that real-world systems are to deal with unreliable and missing data in a computationally efficient way. Robust unsupervised techniques like the one proposed in this

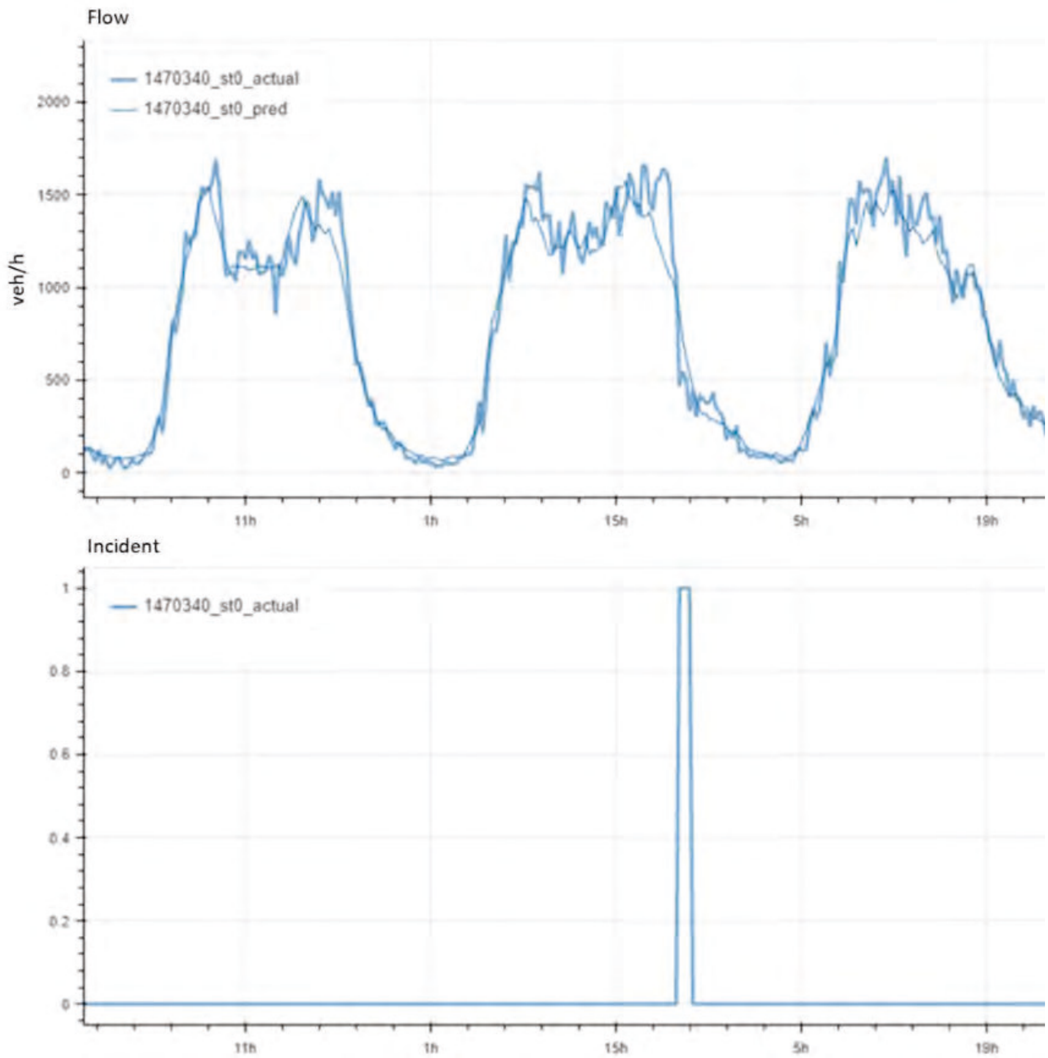

Figure 2: Incident detection of one occurred on 6st of May at 17:28. Top: Measured and predicted flow of the loop detector 1470340_st0. Bottom: incident detection.

paper aim to tackle such problems, and yet, the challenge remains to rely on a common benchmarking system for the scientific community to advance in a steady direction.

This paper proves that incident detection is more robust when methodologies that adapt to the needs of downstream traffic management systems (being flexible, unsupervised, and computationally fast) are integrated. Our future work will explore the combination of several modalities (e.g. data sent by drones) and big data to provide a scalable incident detection module that can be deployed in both traditional and smart cities.

Acknowledgments

The research leading to this publication has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 955317.

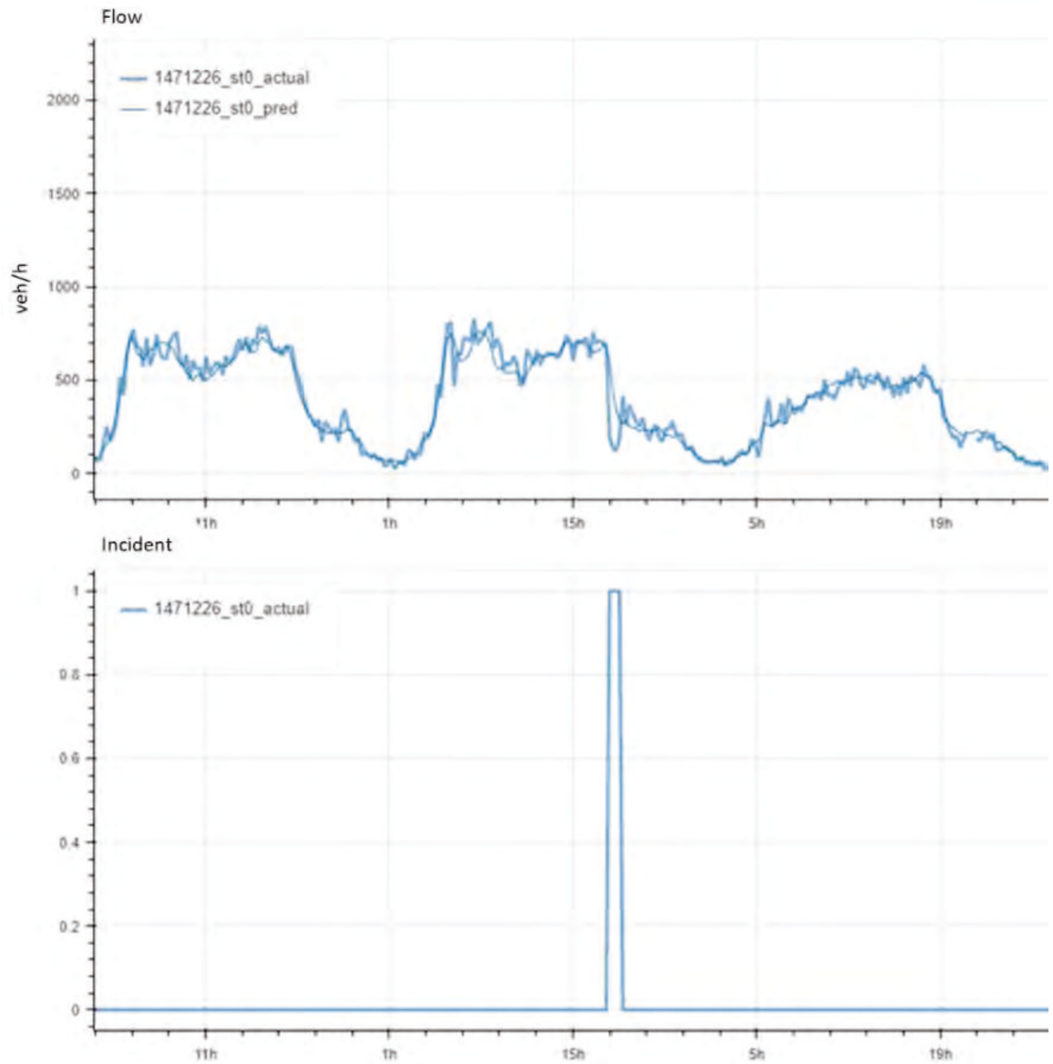

Figure 3: Incident detection of one occurred on 6st of May at 17:49. Top: Measured and predicted flow of the loop detector 1471226_st0. Bottom: incident detection.

References

[ECA19] *Urban Mobility in the EU*. European Court of Auditors, 2019. URL: `https://www.eca.europa.eu/Lists/ECADocuments/AP19_07/AP_URBAN_MOBILITY_EN.pdf`.

[ECo19] *Road fatality statistics in the EU 2019*. European Commission, 2019. URL: `https://www.europarl.europa.eu/news/en/headlines/society/20190410STO36615/road-fatality-statistics-in-the-eu-infographic`.

[EEA21] *Greenhouse gas emissions from transport in Europe*. European Environment Agency, 2021. URL: `https://www.eea.europa.eu/ims/greenhouse-gas-emissions-from-transport`.

[Her18] Y. Hernandez-Potiomkin, M. Saifuzzaman, E. Bert, R. Mena-Yedra, T. Djukic, and J. Casas: "Unsupervised Incident Detection Model in Urban and Freeway Networks". In: *2018 IEEE International Conference on Intelligent Transporta-*

tion Systems (ITSC). Edited by W.-B. ZHANG, A. M. BAYEN, J. J. S. MEDINA, and M. J. BARTH. Maui, Hawai: Institute of Electrical and Electronics Engineers, Nov. 2018, pages 1763–1769. DOI: 10.1109/itsc.2018.8569642.

[Her22] M. HERRMANN, F. PFISTERER, and F. SCHEIPL: *A geometric framework for outlier detection in high-dimensional data*. 2022. DOI: 10.48550/arXiv.2207.00367.

[Men19] R. MENA-YEDRA, J. CASAS, and R. GAVALDÀ: "Probabilistic model for robust traffic state identification in urban networks". In: *2019 IEEE International Conference on Intelligent Transportation Systems Conference (ITSC)*. Auckland, NZ: Institute of Electrical and Electronics Engineers, Oct. 2019, pages 1934–1940. DOI: 10.1109/itsc.2019.8917259.

[Nat17] E. NATHANAIL, P. KOUROS, and P. KOPELIAS: "Traffic volume responsive incident detection". In: *Transportation Research Procedia* 25 (2017), pages 17755–1768. ISSN: 2352-1457. DOI: 10.1016/j.trpro.2017.05.136.

[Pay76] H. J. PAYNE, E. D. HELFENBEIN, and H. C. KNOBEL: *Development and testing of incident detection algorithms : Research methodology and detailed results*. Volume 2. Federal Highway Administration, 1976. URL: https://rosap.ntl.bts.gov/view/dot/744.

[Sai18] M. SAIFUZZAMAN, T. DJUKIC, Y. HERNANDEZ-POTIOMKIN, R. MENA-YEDRA, E. BERT, and J. CASAS: "Understanding incident impact on traffic variables to reduce false incident detection". In: *Proceedings of the ATRF 2018 - Australasian Transport Research Forum*. Darwin, Australia, Oct. 30–Nov. 1, 2018, pages 1–12. URL: https://australasiantransportresearchforum.org.au/wp-content/uploads/2022/03/ATRF2018_paper_52.pdf.

[Yan20] H. YANG, Y. WANG, H. ZHAO, J. ZHU, and D. WANG: "Real-time Traffic Incident Detection Using an Autoencoder Model". In: *2020 IEEE International Conference on Intelligent Transportation Systems (ITSC)*. Rodhes, Greece: Institute of Electrical and Electronics Engineers, Sept. 20–23, 2020, pages 1–6. DOI: 10.1109/itsc45102.2020.9294455.

[Zhu19] L. ZHU, R. KRISHNAN, A. SIVAKUMAR, F. GUO, and J. W. POLAK: "Traffic Monitoring and Anomaly Detection based on Simulation of Luxembourg Road Network". In: *2019 IEEE International Conference on Intelligent Transportation Systems Conference (ITSC)*. Auckland, NZ: Institute of Electrical and Electronics Engineers, Oct. 27–30, 2019, pages 382–387. DOI: 10.1109/itsc.2019.8917015.

Corresponding author: Ferran Torrent-Fontbona, Aimsun SLU, Barcelona, Spain, e-mail: ferran.torrent@aimsun.com