Delft University of Technology

Exploring and Enriching the Effect of Different Road Incidents on Traffic Spatiotemporal Characteristics

Movaghar, Mahsa; Behrouzi, Saman; Krishnakumari, Panchamy; Hoogendoorn, Serge; Van Lint, Hans

DOI
10.1109/ITSC57777.2023.10422640

Publication date
2023

Document Version
Final published version

Published in
2023 IEEE 26th International Conference on Intelligent Transportation Systems, ITSC 2023

Citation (APA)
Movaghar, M., Behrouzi, S., Krishnakumari, P., Hoogendoorn, S., & Van Lint, H. (2023). Exploring and Enriching the Effect of Different Road Incidents on Traffic Spatiotemporal Characteristics. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems, ITSC 2023* (pp. 4018-4024). (IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC). IEEE. https://doi.org/10.1109/ITSC57777.2023.10422640

2023 IEEE 26th International Conference on
Intelligent Transportation Systems (ITSC)
24-28 September 2023. Bilbao, Bizkaia, Spain

Exploring and Enriching the Effect of Different Road Incidents on Traffic Spatiotemporal Characteristics

Mahsa Movaghar[1], Saman Behrouzi[1], Panchamy Krishnakumari[1], Serge Hoogendoorn [1], Hans van Lint [1]

Abstract— **Road incidents, including accidents, greatly impact public safety, traffic flow, and overall transportation system functioning. Detecting and predicting incidents is crucial for effective incident management. Accurate algorithms rely on high-quality incident data sets. However, uncertainties exist due to the collection and recording process. To address this, cross-validating incident data with other datasets helps resolve inaccuracies. Additionally, enriching incident data with additional sources enables a more precise analysis of societal costs for planning purposes.**

In this study, we utilize traffic congestion data to examine and quantify the consequences of incidents on the Dutch highway network. First, we map match recorded incidents with related traffic patterns. Then, we label incidents as "congestion" if significant congestion patterns were identified during or after the incidents or as "no-congestion" if no significant congestion pattern occurred. For incidents labeled as congestion, we calculate and associate records with the congestion's duration, location, and Vehicle Loss Hours (VLH). The developed methodology has been implemented on five months of recorded data for the six most significant motorways in the Netherlands. This enriched dataset can be utilized for incident detection algorithms, analysis and management, and policy and decision-making.

I. INTRODUCTION

Road incidents, including accidents, not only have considerable effects on financial and fatality losses, but also their effect on travel time loss and secondary incidents is not negligible [1], [2]. They are considered one of the significant causes of congestion and bottlenecks on the network [3], [4].

Although traffic incidents are highly likely to cause traffic congestion, it is important to note that not all traffic accidents lead to traffic congestion, and congestion itself does not necessarily always stem from incidents [5]. Hence, for labeling the congestion patterns, it is useful to find congestion that is caused by incidents. This labeling process is a time-consuming but beneficial task for developing any data-driven models.

Nowadays, data-driven models are of high interest to researchers for developing models to predict, detect, and/or classify incidents, including accidents [6]. The quality of data-driven methodologies highly depends on the quality of data sources and, moreover, on the quality of sensors/equipment utilized to acquire data [7].

Traffic sensors that monitor cross-sectional traffic flow record data separate from other data sources, such as police reports, surveillance cameras, global positioning systems, and incident data. Developing multi-source data and multi-source data fusion is beneficial for a more accurate and reliable understanding and interpretation of the observed situation by decreasing the inherent uncertainties in individual data sources. This leads researchers and traffic management centers to more reliable and accurate interpretations and decisions.

However, the inherent uncertainties in recorded databases are unavoidable. The effect of these uncertainties is significant in the analysis process and methods, which require estimation and judgment. The main challenge while dealing with uncertainties in databases is handling the data while keeping it useful for data management and further mining applications.

This study aims to enrich the incident dataset with a new label, *congestion* and *no congestion*. Moreover, each record will be matched and enriched with traffic spatiotemporal features of significant congestion. This enriched dataset will provide insights for understanding the behavior of the network in case of any type of incident. This leads to better incident management and more useful action plans and strategies on the horizon.

A. Related Work

Detecting road incidents, including accidents, at an appropriate time can decrease the time lag between the time that incident happens and any further decisions, whether sending emergency services or much-reducing travel time by recommending alternative routes to drivers [8]. This problem is even more challenging due to the rareness of the incidents, especially accidents and lack of data. So any reliable and trustworthy datasets are of high interest for further investigations. Deep learning models have been widely used for real-time incident detection but less for assessing road traffic incidents [5], [9]. Haung et. al (2020) used sensor occupancy data, speed, and volume to extract useful features for Deep Learning models for detecting road crashes [10]. Despite most of the recent studies which use images, Mehrannia et. al used numerical data from automatic traffic recorders (ATR) to detect accidents. They used an LSTM-based model to increase the separability of accident/no-accident class [8].

One of the challenges in developing data-driven models is dealing with incomplete or missing data due to varied recorded features in different datasets. Moreover, integrating different sources may be incomplete due to different data collection techniques and strategies adopted by different sources. These challenges lead to major problems, partic-

[1]Faculty of Civil Engineering and Geosciences, Delft University of Technology (TU Delft), Stevinweg 1, 2628 CN Delft, The Netherlands m.movaghar@tudelft.nl, s.behrouzi@tudelft.nl, p.k.krishnakumari@tudelft.nl, s.p.hoogendoorn@tudelft.nl, j.w.c.vanLint@tudelft.nl

2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC) | 979-8-3503-9946-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ITSC57777.2023.10422640

ularly in nonlinear dynamic problems such as traffic flow prediction. In most related studies, data does not contain information about incidents and events, which lead to unreliable causal effects analysis and predicting long-term traffic predictions [11].

B. Study Outline

Undoubtedly, creating a reliable data set to handle uncertainties, ambiguity, and consistency is the initial and most important step in developing any data-driven, specifically supervised learning research. Although improving data collection techniques is essential for creating an extensive dataset, they are typically labor-intensive, time-consuming, and expensive. Therefore, advanced frameworks for enriching the already existing datasets by integrating different data sources are highly beneficial for creating a reliable and good-quality dataset [12]. The highlights of this study can be summarized as follows:

- A robust framework is proposed to enrich each recorded incident in the dataset with features of traffic congestion.
- To this end, this study tries to connect and fuse the point dataset (incident) with the spatiotemporal dataset (traffic) with map-matching techniques.
- Few studies have focused on the spatiotemporal characteristics of congestion caused by incidents.
- Finally, a better estimation and understanding of the spatiotemporal features of the congestion caused by different incidents will be provided.
- The enriched dataset created in this study can provide significant information for each recorded incident.

To achieve the primal goal of this study, which is enriching the incident data set with spatiotemporal features of traffic, the structure of the next sections is as follows: First, an overview of the steps taken for integrating data sets, and traffic spatiotemporal features added to the dataset is explained in Section II. Then in Section III, the data sets are introduced. The quantifiable effects of incidents on traffic and the corresponding statistics of the added information to the incident dataset will be presented and discussed in IV. The summary of the study and directions for future research are discussed in Section V.

II. Methodology

One of the approaches for reducing uncertainties in datasets is data integration. Through data integration, different data sources will be compared to identify matching entities. Using the matching entities (i.e., mapping), the original data source will be enriched with additional information [13]. To do this, the main challenge is the discovery of correct and meaningful relationships between different data sources. In this study, the incidents point dataset will be qualitatively integrated with the traffic spatiotemporal dataset.

A. Traffic Spatiotemporal Data

Based on the study done by Nguyen et al., clustering methods have been employed, namely point-based and area-based approaches, to extract visual features from each spatiotemporal congestion pattern [14]. Among all the features for each congestion pattern, three main traffic-related and interpretable features have been chosen for enriching the incident dataset [15]:

- Spatial-extent: length of a roadway that is affected by the congestion. The spatial extent of congestion relies on many variables, such as duration, severity of congestion, and/or time of the day. This feature is obtained from a time-space diagram using an image processing technique called contour detection as described in Nguyen et al. [14] and is represented in Figure 1.
- Temporal-extent: duration or time period that the roadway is affected by the congestion. The temporal extent is also influenced by the cause of congestion, traffic flow, and so on. This feature is also obtained using contour detection from a time-space diagram as depicted in Figure 1.
- Total Vehicle Loss Hour (VLH): the total amount of time lost by vehicles because of being stuck in traffic congestion. VLH is computed using spatiotemporal maps of speed u and flow q, smoothed using the Adaptive Smoothing Method (ASM) with equidistant cells of length Δx and duration Δt [16], [17]. Total Vehicle Loss Hour (VLH) can then be estimated as shown in Equation 1.

$$VLH = \sum_{ki} (T_{ki}^{exp} - T_{ki}^{ref}) \tag{1}$$

In which T_k^{exp} is experienced travel time and T_k^{ref} is travel time with free-flow speed.

Free flow speed u_k^{ref}, which is the maximum speed that can be experienced by cars when there is no congestion, is obtained based on the speed limits of motorways. However, this can be more precisely measured by field observations or historical data.

The spatial and temporal extent has been obtained based on the related study conducted by Nguyen et al. [14]. For clustering traffic patterns, they used an area-based method based on domain knowledge. Figure 1 shows the spatiotemporal extent of a recorded congestion pattern with 54 km space extent, 330 minutes temporal extent, and 11840.2 ($veh * hr$) VLH.

B. Incident Point Data

It is important to record and store incident information. However, enriching these datasets with accurate and meaningful information and labels will significantly improve their applications for developing further research or drawing any inferences from the data and further decisions for incident management. Handling large dataset increases the importance of an automated and robust framework in this regard. In general, each point incident dataset record includes the location, type of incident, start and end time. In this work, this record is enriched with congestion information, i.e., spatial and temporal extents. The spatial extent of congestion defines the total length of the road distance being reached by traffic congestion, and the temporal extent defines the duration of the congestion at each accident [14].

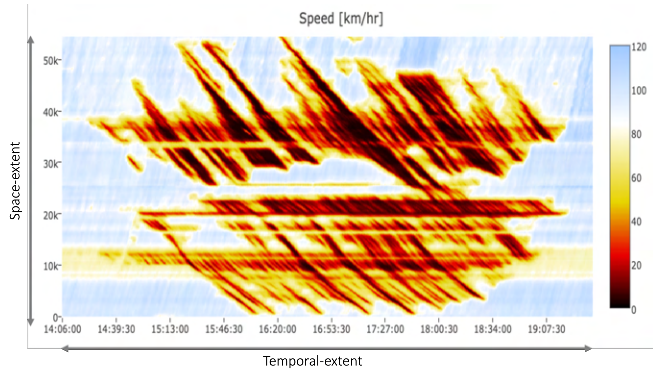


Fig. 1: Spatiotemporal characteristics of a recorded significant congestion on A13, A20, and A15 (October 29, 2019).

C. Map Matching

The main challenge in integrating incident datasets with traffic characteristics is the inherent differences between these two datasets. The incident dataset is a point data set. However, a traffic dataset includes spatiotemporal characteristics since speed and flow change both in time and location. To overcome this challenge, a robust framework is developed in this study. The summary of the framework used in this study is depicted in Figure 2. For each recorded incident, the two most important features have been stored: location and time. This procedure is more challenging for recorded traffic data. Traffic data collected by multiple sensors is recorded per minute and per sensor. To deal with the large recorded traffic dataset, only notable congestion patterns within the dataset have been stored, and the free-flow patterns are disregarded.

To align the incident data with the corresponding speed data from each congestion pattern, we adopted a map-matching approach. Firstly, each incident was associated with the speed data of the motorway segment on which it occurred. For this purpose, we considered the time span during which each incident occurred. By matching the incident time with the recorded speed data timestamps, the appropriate speed dataset for each incident was identified. This step was an exhaustive search that enabled us to precisely link incidents with the corresponding motorway segments and temporal context. To visualize the speed congestion patterns and overlay the incident locations and times, we calculated the distances of the incident locations from the starting point of each motorway segment. This allowed us to effectively represent the relative position of the incident along the motorway. Utilizing this distance information, we plotted the speed congestion patterns and superimposed the incident locations and times onto this speed congestion visualization, providing a comprehensive view of the incidents' spatial distribution and temporal correlation with speed variations.

Based on the observations, two noteworthy scenarios are feasible:

- The recorded incident is before or within a significant recorded congestion pattern. This incident will be labeled *"Congestion"*, and the incident can be considered as one of the probable reasons for this congestion. Moreover, each record will be enriched with three new features: the temporal extent of the related congestion, the spatial extent of the related congestion, and total Vehicle Loss Hour (VLH).
- There is no recorded congestion pattern close (in terms of recorded space and time) to the recorded incident. This incident will be labeled *"No-congestion"*, and no additional feature will be added to the dataset.

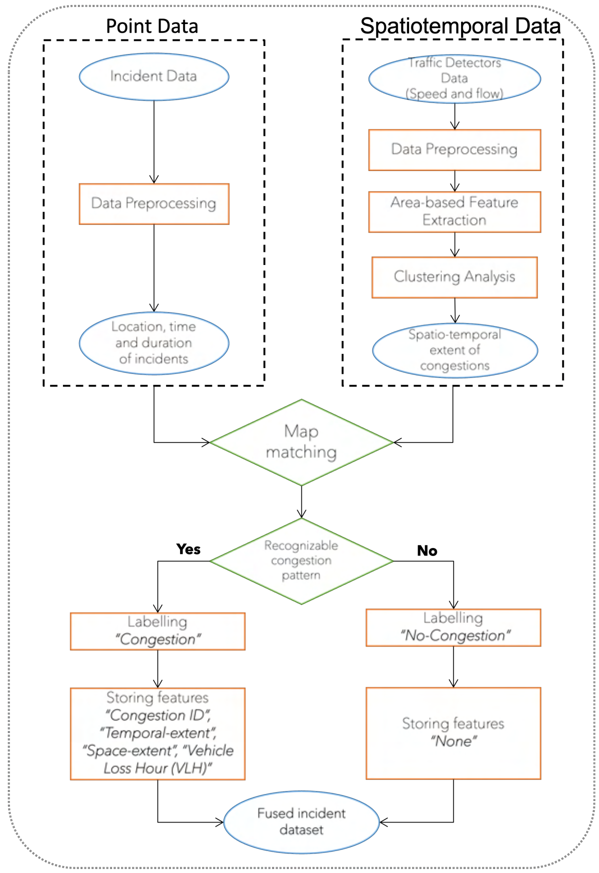


Fig. 2: Developed framework for integrating accident and traffic datasets.

It is worth noting that although a congestion pattern is recorded for the first scenario, the incident may be only one of the causes of identified congestion. This labeled category can be later investigated to estimate the contribution of incidents to congestion.

III. Case Study

In the Netherlands, NDW (Nationaal Dataportaal Wegverkeer) is a national authority responsible for collecting road traffic data, including speed, flow, and travel time [18]. The data has been recorded for over ten years and is available in open source both in real-time and historically. The historical data is being used for traffic analysis and research. Alongside traffic data, there are also extensive records for road incidents for almost ten years of the network available as open data. This dataset comprises data primarily gathered by the road authority through manual methods and integrated with information extracted from police reports, phone calls, and emergency centers. The dataset provides details regarding incidents occurring on national roads within the Netherlands, categorized into three distinct groups:(1) Accidents: including any collision, fire, etc., (2) Vehicle obstruction: including any car or vehicle immobilized in the middle of the road or on the emergency lane, and (3) General obstruction: which is any sudden obstructions that can not be categorized in the other two groups. The incident dataset includes all the unplanned, unexpected, and sudden happenings on the network. Thus, other road disruptions, such as roadwork and bridge openings, are not included in this dataset. For each incident following features have been recorded in the raw dataset: ID, Type, Start time, End time, Longitude, and Latitude.

In this study, a dataset encompassing the entire network in the Netherlands will be explored and analyzed. The data covers a time range of five months, starting from August and concluding in December 2019. Figure 3 shows the frequency of each incident category in the studied dataset.

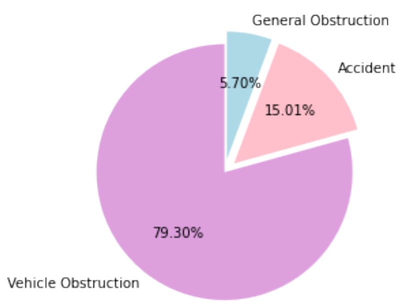


Fig. 3: Frequency of different types of incidents in the Netherlands (August-December 2019).

In the study period, the location of incidents in the Netherlands is depicted in Figure 4.
The density of accidents in the Netherlands in the time range of August to December 2019 is depicted in Figure 5. Despite the dispersed location of incidents in Figure 4, Figure 5 vividly depicts three significant regions in the Netherlands: Rotterdam, Amsterdam, and Utrecht. To have a comprehensive overview of congestion related to incidents, 6 important motorways have been picked for further analysis and comparison: A2, A4, A12, A13, A15, and A20. These motorways hold great significance for transportation planners due to their high density of disruptions and their proximity to major cities and regions. Figure 6 shows the studied motorways in the Netherlands. The next section will present map-matched results of incidents recorded on the studied motorways, focusing on traffic spatiotemporal features. This will be done using the developed methodology, providing visualizations and discussions.

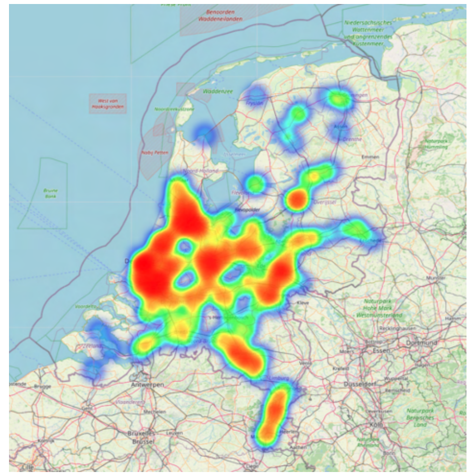


Fig. 4: Density visualization of incidents in the Netherlands (August-December 2019). Warmer hues represent a higher concentration of incidents.

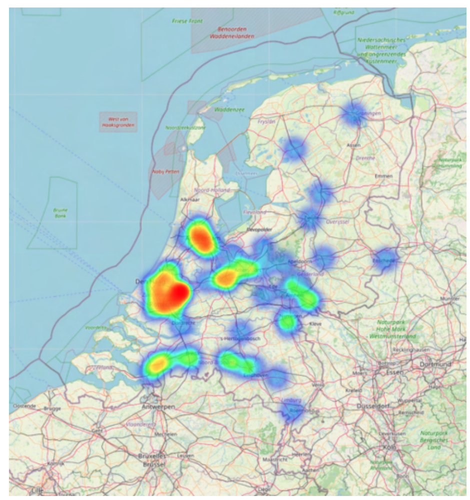


Fig. 5: Density visualization of accidents in the Netherlands (August-December 2019). Warmer hues represent a higher concentration of incidents.

IV. Results and Discussion

As mentioned previously, all recorded incidents will be categorized into two distinct groups: *Congestion* and *No-Congestion*. In this section, results will be shown in three different categories: accidents that are matched with significant congestion patterns, percentage of incidents matched with significant congestion, and spatiotemporal characteristics of the congestion matched with incidents.

Figure 7 shows an accident (horizontal line) matched during an identified significant congestion pattern on the time-space diagram. The horizontal line indicates the location, start time and end time recorded in the incident database. This congestion is stored with 29 km space extent and a time horizon of 377 minutes. Also, 5535.54 ($veh*hr$) total Vehicle Loss Hour (VLH) is estimated based on Equation 1 for this congestion. This accident is recorded during a significant congestion which is probable to be its effect. However, to

Fig. 6: Location of studied motorways in this research.

estimate the more precise contribution of this accident to the congestion, further analysis is required. For instance, the historical and seasonal patterns should be checked on the same day to be able to relate the congestion to peak hours or the geometry of the highway. Moreover, new datasets like weather variables can add insightful information for understanding the contribution of the accident to identified congestion.

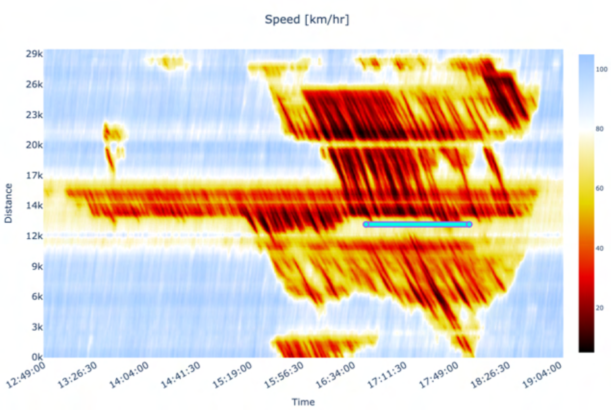


Fig. 7: Example of an accident matched with an identified significant traffic congestion on road number A13 (December 06, 2019).

When estimating VLH, two conditions should be taken into account. First of all, as is shown in Figure 7, only part of the recorded congestion is during the accident period. Therefore, we are probably overestimating the VLH, space, and time extent associated with an incident. On the other hand, since incidents are being matched with only significant congestions, the VLH may be underestimated for the accidents causing insignificant congestion. Another group of accidents, labeled *"No-Congestion"*, highlights the fact that not all accidents will be associated with significant congestion. Analyzing the whole accident data for five months in 2019 for the selected motorways reveals that a significant congestion pattern can be recognized after or during 86 percent of recorded accidents (Figure 8). Therefore, only for 86 percent of accidents, congestion can be considered as one of the effects. And for the other type of incidents (vehicle and general obstructions), this percentage is similar, implying that most incidents have significant congestion associated with them. It is also worth noting that although accidents are showing a notable impact in causing congestion, their exact contribution among other causes, i.e., peak hour, bottlenecks, geometry, and so on, requires further investigation. Conversely, it has been observed that almost 14 percent of

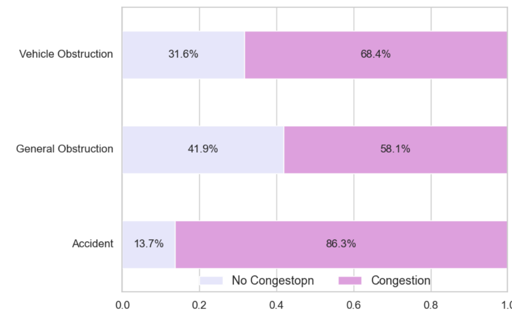


Fig. 8: Incidents matched with congestion (August-December 2019).

recorded accidents do not exhibit any notable congestion patterns. This indicates that these accidents cannot be directly linked to significant instances of traffic congestion. Figure 9 illustrates these categories across different incident types and motorways, shedding light on how incident types and specific characteristics of each motorway contribute to substantial congestion. Among the three types of incidents examined—accident, vehicle obstruction, and general obstruction — accidents demonstrate a higher likelihood of associating with significant congestion compared to the other two categories in all the studied motorways. In addition, Figure

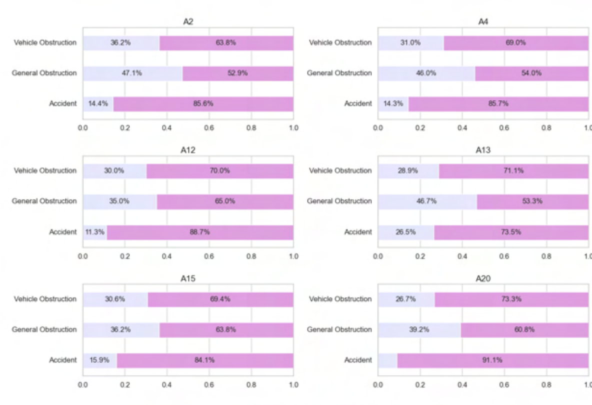


Fig. 9: Percentage of matched incidents with significant congestion per incident category on studied motorways (August-December 2019).

10, demonstrates the box plots for estimated total VLH per incident type. Examining the box plot, the medians, and the spread of estimated VLH among three incidents, insights into their respective impacts on transportation efficiency can be gained. Vehicle obstructions with a median of 1311 and an average of 2698 $(veh*h)$ play the most significant role among the other two incidents. Then, accidents with an average VLH of 2639 $(veh*h)$ can be considered as the most and general obstruction with an average of 2527 $(veh*h)$ as the least costly incidents. This highlights the effective incident management policies in the studied corridors during accidents and vehicle obstructions.

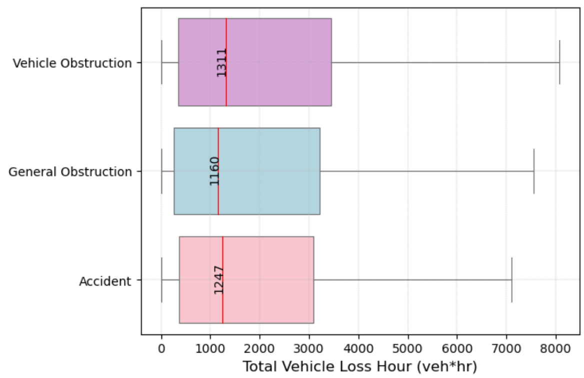
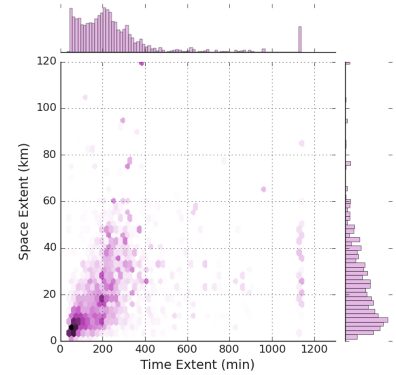
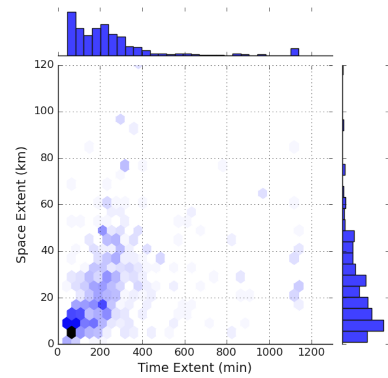


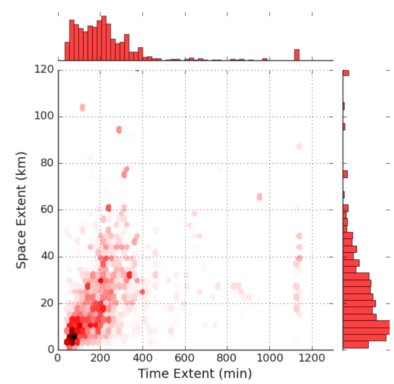
Fig. 10: Estimated total Vehicle Loss Hour (VLH) for different incidents (August-December 2019).



(a) Vehicle Obstruction



(b) General Obstruction



(c) Accident

Fig. 11: Spatiotemporal characteristics of significant congestions matched with incidents (August-December 2019).

Figure 11 demonstrate the joint distributions for spatiotemporal characteristics of congestion matched with incidents. In this Figure, the plot area is divided into hexagonal bins and shows the number of data points falling into each bin with color intensity. Darker colors indicate a higher density of data points. Based on Figure 11, accidents show a high probability of being associated with congestion within a 5-kilometer spatial extent and a 100-minute time extent. Vehicle obstructions are more likely to be matched with congestion of less than 5 kilometers and 100 minutes. Despite almost similar space and time extent characteristics of the congestion associated with all the incidents, their imposing cost to the network, as estimated in VLH, is not negligible. This finding reveals the importance of considering VLH as a significant parameter for any incident management policy and decision-making.

To validate the efficacy of our proposed method, we initially undertook a qualitative analysis to ascertain whether accidents were associated with significant congestion patterns within the motorways under study. It's important to note that this approach does not distinguish between incidents that arose due to congestion and those that subsequently led to congestion. Following the preliminary labeling of incidents associated with congestions, we proceeded with a rigorous validation phase, leveraging both expert domain knowledge and established traffic flow theory.

Besides the approach taken in this study, social media and news can also be interesting sources for enriching incident datasets. For example, on Friday morning, August 02, 2019, a huge traffic jam was recorded on A20 between two cities, Hoek van Holland to Schiedam. A delay of around 1 hour in travel time is also reported for this accident at 8:50 a.m. in the news. When looking at the incident dataset, the location and duration of the accident can be easily validated. However, there might be a couple of accidents reported in the news but not accurately recorded in the dataset. For example, on December 11, 2019, a long traffic jam on A15 after an accident between several vehicles around half past three was reported. However, there are no incident records between 10:00 to 20:00. This highlights the importance of using other data sources for validating the dataset, specifically social media and news, for future studies and investigations.

Although in this study, interesting insights have been extracted, the following limitations should be taken into account:

- Incident dataset is map-matched with significant congestion patterns. This highlights that added label *"No-Congestion"* means that no significant congestion has been recognized close to the incident duration.
- Although matched incidents with congestion patterns add valuable spatiotemporal information for each incident to the dataset, the exact contribution of the incident to the congestion is not clear. To estimate the contribution of the incident to the matched congestion pattern among other factors such as peak hour, bottlenecks, and so far, further investigation is required.
- Estimated VLH may be overestimated since the incidents may cover only part of a matched congestion pattern. And it may be underestimated due to the fact that only significant congestion patterns have been recorded in the traffic spatiotemporal dataset.
- Many other data sources, such as social media and news, can play a significant role in validating and enriching the current dataset.

V. Conclusion

Understanding incidents, including accidents, play a vital role in effective incident management. By gaining insights into accident details, incident management teams can respond promptly and provide necessary assistance, leading to improved public safety. Additionally, understanding incidents helps in implementing efficient traffic control measures, reducing congestion and societal impacts, and minimizing disruptions to maintain smooth traffic flow. The quality of datasets plays an important role in incident detection algorithms and any other data-driven models and inferences. To make accurate datasets, the incident point dataset is map-matched with traffic spatiotemporal characteristics. To do this, significant congestion patterns recorded within or after recorded incidents have been identified. Then, for each matched congestion, space and temporal extents were recorded. Moreover, the total Vehicle Loss Hour (VHL) due to each incident has been estimated.

Results show among different types of incidents; accidents are more likely to be associated with congestion than other types of incidents. Although longer congestion, both in time and space, is expected during general obstruction, accidents and vehicle obstructions contribute to more vehicle loss hours. In summary, this study concludes that various incidents and road disruptions affect traffic spatiotemporal characteristics quite differently. The precise contribution of each incident to the recognized congestion, among other factors, such as peak hour, bottlenecks, and geometry, provides interesting possibilities for future studies. In addition, validating incident datasets with other data sources, such as social media and news, would be a beneficial area for incident management and analysis.

Acknowledgment

The authors would like to acknowledge the Dutch open data warehouse, NDW (Nationaal Dataportaal Wegverkeer), for providing data and support during this project.

References

[1] H. Park and A. Haghani, “Real-time prediction of secondary incident occurrences using vehicle probe data,” *Transportation Research Part C: Emerging Technologies*, vol. 70, 2016.

[2] M.-I. M. Imprialou, F. P. Orfanou, E. I. Vlahogianni, and M. G. Karlaftis, “Methods for defining spatiotemporal influence areas and secondary incident detection in freeways,” *Journal of Transportation Engineering*, vol. 140, Journal of Transportation Engineering.

[3] C. Chen, A. Skabardonis, and P. Varaiya, “Systematic identification of freeway bottlenecks,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1867, pp. 46–52, 1 2004.

[4] H. Li and R. L. Bertini, “Comparison of algorithms for systematic tracking of patterns of traffic congestion on freeways in portland, oregon,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2178, pp. 101–110, 1 2010.

[5] C. Zhao, X. Chang, T. Xie, H. Fujita, and J. Wu, “Unsupervised anomaly detection based method of risk evaluation for road traffic accident,” *Applied Intelligence*, vol. 53, pp. 369–384, 1 2023.

[6] A. Zarindast, S. Poddar, and A. Sharma, “A data-driven method for congestion identification and classification,” *Journal of Transportation Engineering, Part A: Systems*, vol. 148, p. 04022012, 4 2022.

[7] A. Sohail, M. A. Cheema, M. E. Ali, A. N. Toosi, and H. A. Rakha, “Data-driven approaches for road safety: A comprehensive systematic literature review,” *Safety Science*, vol. 158, 2 2023.

[8] P. Mehrannia, S. S. G. Bagi, B. Moshiri, and O. A. Al-Basir, “Deep representation of imbalanced spatio-temporal traffic flow data for traffic accident detection,” *IET Intelligent Transport Systems*, vol. 17, pp. 602–615, 3 2023.

[9] S. Rahmani, A. Baghbani, N. Bouguila, and Z. Patterson, “Graph neural networks for intelligent transportation systems: A survey,” *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[10] T. Huang, S. Wang, and A. Sharma, “Highway crash detection and risk estimation using deep learning,” *Accident Analysis Prevention*, vol. 135, p. 105392, 2 2020.

[11] C. Liu, M. Zhao, A. Sharma, and S. Sarkar, “Traffic dynamics exploration and incident detection using spatiotemporal graphical modeling,” *Journal of Big Data Analytics in Transportation*, vol. 1, pp. 37–55, 6 2019.

[12] N.-E. E. Faouzi and L. A. Klein, “Data fusion for its: Techniques and research needs,” *Transportation Research Procedia*, vol. 15, pp. 495–512, 2016.

[13] N. A.-H. Othman, A. S. Eldin, and D. S. E. Zanfaly, “Handling uncertainty in database: An introduction and brief survey,” *Computer and Information Science*, vol. 8, p. p119, 7 2015.

[14] T. T. Nguyen, P. Krishnakumari, S. C. Calvert, H. L. Vu, and H. van Lint, “Feature extraction and clustering analysis of highway congestion,” *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 238–258, 3 2019.

[15] P. Krishnakumari, T. Nguyen, L. Heydenrijk-Ottens, H. L. Vu, and H. van Lint, “Traffic congestion pattern classification using multiclass active shape models,” *Transportation Research Record*, vol. 2645, pp. 94–103., 2017.

[16] J. W. C. H. V. Lint, “Empirical evaluation of new robust travel time estimation algorithms,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2160, pp. 50–59, 1 2010.

[17] M. Treiber and D. Helbing, “Reconstructing the spatio-temporal traffic dynamics from stationary detector data,” 2002.

[18] Ministerie van Infrastructuur en Waterstaat, “Nationaal dataportaal wegverkeer (ndw),” 2023. [Online]. Available: https://www.ndw.nu/