Australasian Transport Research Forum 2023 Proceedings
29 November – 1 December, Perth, Australia
Publication website: http://www.atrf.info

Area-wide traffic incident duration prediction: A multi-task learning approach

Thanh Tran[1], Dan He[1], Jiwon Kim[1], and Mark Hickman[1]
[1]School of Civil Engineering, The University of Queensland, St. Lucia, QLD 4072
Email for correspondence (presenting author): thanh.tran@uq.edu.au

1. Introduction

Urban traffic congestion remains a persistent and significant challenge, attributed to factors such as population growth, workforce concentration, and inadequate public transport. Non-recurrent incidents, such as crashes and weather conditions, account for nearly 60% of traffic congestion (Schrank and Lomax, 2005). Accurately predicting their duration can optimize route planning, minimize operational costs, and reduce traffic congestion. However, predicting the duration of traffic incidents is challenging due to their stochastic nature (Schrank and Lomax, 2005). Moreover, existing incident duration prediction models (Kuang et al., 2019; Grigorev et al., 2022) have limited applicability, as they are tailored to specific road networks with the same road type and require different models for different purposes (e.g., classification of incident duration or estimation), resulting in suboptimal prediction performance. Furthermore, incorporating network structure in learning traffic data plays a vital role in improving machine learning models for transportation (Cui et al., 2020). Unfortunately, existing works have not adequately captured this essential information. Consequently, developing an advanced model that can effectively predict incident duration across different road types and networks, irrespective of their complexity, while accurately predicting traffic incidents is desirable for real-time incident management and duration prediction.

In the literature, various statistical methods have been applied to traffic incident duration prediction models (Tang et al., 2020; Grigorev et al., 2022). Recently, machine learning models have gained popularity due to their ability to capture non-linear, spatial, and temporal correlations in traffic data more accurately. In particular, the majority of studies in the literature have focused on applying state-of-the-art machine learning models primarily for classifying incident severity (Nguyen et al., 2017) or duration (Li et al., 2018). Additionally, to predict incident duration in minutes (regression task), the most similar work was published in Kuang et al.'s (2019) study, which relied only on one dataset, one method for classification (Bayesian network), one method for regression (K-nearest neighbors), and authors selected a static threshold (30 min) to alleviate the class-imbalance problem. Grigorev et al.'s (2022) study demonstrated that the majority of prior works had studied the prediction of incident duration on specific types of roads and very few applied the prediction strategies on normal arterial roads due to high modeling complexity and location mismatching; most traffic incident duration analysis studies focus only on one type of road network (arterial roads, highways, etc.). Hence, Grigorev et al.'s (2022) study contributed a method to predict traffic duration using multiple machine learning (ML) models on multiple datasets with unique features such as arterial roads data or motorways data, along with outlier removal and joint optimization.

However, existing studies have not proposed a single model that can address different road types all at once (e.g., motorways, arterial roads) or successfully tackle both classification and regression tasks within a unified framework. This limitation results in the need for building different machine learning models for various traffic networks with diverse road types and separate models for classifying traffic incident duration as short or long-term (classification) and predicting traffic incident duration in minutes (regression problem). These fragmented approaches not only increase the complexity of traffic incident management but also hinder the scalability and adaptability of existing models to new or evolving traffic scenarios. Therefore, there is a growing need for a single, robust model that can effectively manage real-world traffic networks with various road types and handle both the classification and prediction of traffic incident duration.

Motivated by the lack of a comprehensive model that can accurately predict traffic incident duration regardless of road network or complexity, this paper proposes a single Multi-task learning Graph-based Network (MT-GN) model applicable to different road types and networks. Our proposed model leverages heterogeneous data collected from various road types, including arterial roads and motorways, making it suitable for any traffic network without the need to build separate models for each road type. The MT-GN model is designed to capture both the spatial and topological information inherent in traffic networks, allowing it to adapt to different road infrastructures and conditions more effectively. By incorporating multi-view learning techniques, our model can better understand the complex relationships among traffic features, incident characteristics, and network topology, resulting in more accurate predictions of traffic incident duration. To evaluate the effectiveness of our model, we conduct experiments on historical incident records. Specifically, we train and test our model using traffic data from Brisbane, Australia (2017), to predict incident duration at both fine-grained (predicted in minutes) and coarse-grained (classified as short-term and long-term) levels. To achieve this, we propose a shared loss or objective function that optimizes the model based on historical incident records, allowing the model to jointly learn both classification and regression tasks. This unified approach not only simplifies the model training process but also promotes better generalization and transferability across various road networks and types. Our results demonstrate the effectiveness of the model architecture and the potential for jointly predicting traffic incident duration at both fine-grained and coarse-grained levels using a single model. These findings highlight the value of our proposed MT-GN model in addressing the challenges faced by existing traffic incident duration prediction models, paving the way for more efficient and reliable real-time incident management and duration prediction. The full paper version will provide a detailed description of the model, experimental setup, and in-depth analysis of the results, showcasing the superiority of our approach compared to existing methods in the literature.

2. Problem definition

According to Grigorev et al.'s (2022) study, traffic incident duration refers to the time interval between the detection of an incident and the clearance of the incident, which can be broken down into several phases, including incident detection time, incident response time, incident clearance time, and incident recovery time. While the total incident duration includes all of these phases, our study focuses specifically on the time between incident detection and clearance, as reported in traffic logs provided by local authorities.

Classification task of traffic incident duration. A binary classification modelling with the purpose of identifying short versus long-term incident duration, split by the incident clearance threshold T_c. Thus, our task is to predict y_i^c, where Yc takes one of the binary values:

$$\begin{cases} y_i^c = 0, & \text{if } y_i \le T_c, \quad short-term\ incidents \\ y_i^c = 1, & if\ y_i > T_c, \quad long-term\ inidents \end{cases}$$

where the threshold T_c is varied and can be chosen.

Regression task of traffic incident duration. The regression task in traffic incident duration aims for a more precise prediction of incident duration, with a fine-grained approach that considers incidents with reported duration mainly ranging from 0 to 30 minutes. The need for this level of precision is due to the fact that incidents falling within this range require different handling procedures, with a 5-minute accident requiring a different response than a more severe accident lasting 30 minutes (Grigorev et al., 2022). Hence, the regression task is to predict the traffic incident duration y_i^r based on sub-network representations of the incidents, as described in Tran et al.'s (2021) and Tran et al.'s (2023) research.

3. Model overview

In this section, we introduce the overall framework of our model. Figure 1 shows the architecture of .

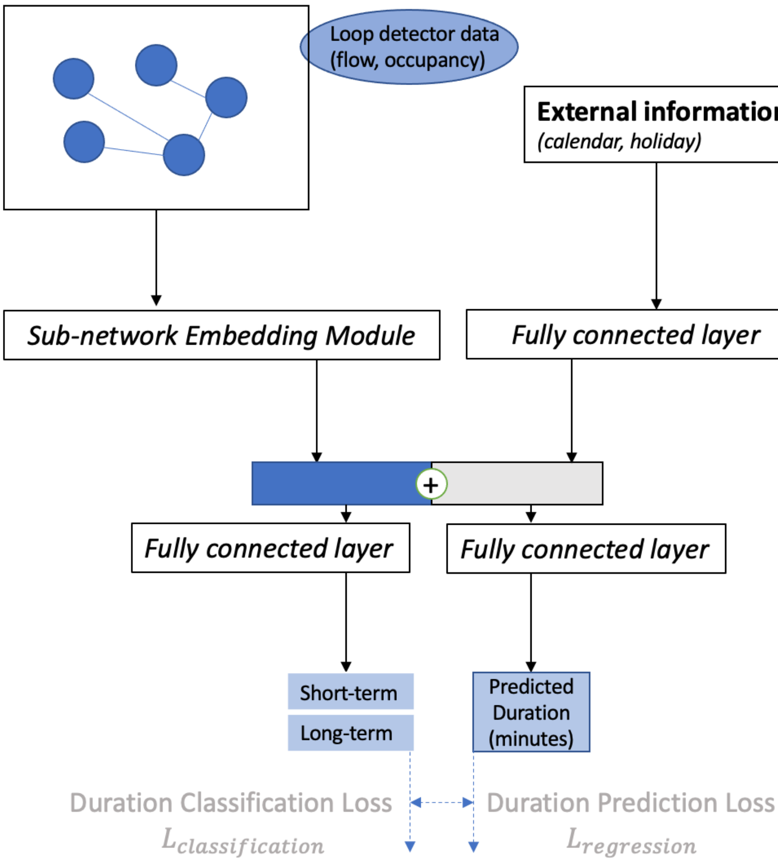

Figure 1: The proposed MT-GN. Dash arrows indicate operations used in the training phase only.

Figure 1 shows the overall structure of the MT-GN. The input data for the model is prepared by integrating multiple data sources using the approach presented in Tran, et al.'s (2021) study. For each incident, a sub-network with a graph structure representing a geographical region is constructed, where nodes represent road links and edges indicate connectivity among links. The sub-network is then attached to a group of features, including traffic conditions 30 minutes before the incident and external information such as day of the week, month, time of day, and holiday information.

The Subgraph Embedding Module from the Hybrid Graph-based Neural Network (HGNN) introduced in our previous work (Tran, et al., 2021) is utilized to learn the sub-network embeddings. The learned embeddings are concatenated with the external information (which is captured by a fully connected layer (FC)) and fed into two different FC layers to generate the final prediction results. The output of the prediction results consists of two parts: the first is a binary classification result, indicating whether the incident is short-term or long-term duration within a given region (500m-radius sub-areas); the second is a regression result that predicts how long the duration of the incident will be in minutes ahead of time. Dash arrows in Figure 1 indicate operations used in the training phase only.

4. Objective function for multi-task learning

Our goal is to predict the fine-grained (predicted how long will incident last in minutes) and coarse-grained (classified as short-term and long-term incident) traffic incident duration simultaneously. Thus, the two prediction tasks can be conducted jointly under a multi-task learning framework which is commonly used in Wang et al.'s (2021) study. The objective function ($\mathcal{L}_m$) of the designed multitask learning model contains two parts (Figure 1): the prediction loss for the coarse-grained data ($\mathcal{L}_{classification}$) and the prediction loss for the fine-grained data ($\mathcal{L}_{regression}$).

$$\mathcal{L}_m = \lambda_1 \mathcal{L}_{classification} + \lambda_2 \mathcal{L}_{regression}, \qquad (1)$$

where λ_1, λ_2 are the parameters of the loss function which can balance the importance of the two losses. $\mathcal{L}_{classification}$ (Binary Cross-Entropy loss) and $\mathcal{L}_{regression}$ can be formulated as follows.

$$\mathcal{L}_{classification} = -(ylog(p) + (1-y)\log(1-p)), \qquad (2)$$

where y is the ground-truth label for the word (either 0 or 1), p is the predicted probability that the word belongs to the positive class (i.e., $y = 1$), log is the natural logarithm function. This loss function (2) measures the dissimilarity between the predicted probability and the true label for a binary classification task. The aim is to minimize this loss function during training so that the predicted probability becomes closer to the ground-truth label.

$$\mathcal{L}_{regression} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \qquad (3)$$

where y_i is the ground-truth value for the i-th example, $\hat{y}_i$ is the predicted value for the i-th example, n is the number of examples in the dataset. This loss function (3) measures the average squared difference between the predicted and true incident duration across all examples in a regression task. The aim is to minimize this loss function during training so that the predicted values become closer to the ground-truth values.

5. Preliminary experiments and discussion

To verify the effectiveness and efficiency of our model, we conducted experiments on a real-world dataset. Specifically, we utilized incident data from Brisbane, Australia in 2017.

5.1 Experimental settings

In our experiment, our model is evaluated by five metrics: AUC (Area Under the Curve), Acc (Accuracy), Precision, Recall and F1 (F1-score), which are formulated as follows (TP:True positive; TN:True negative; FP:False positive; FN:False negative).

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}, \qquad \mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}},$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \qquad \text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Recall} + \text{Precision}}$$

We adopted the same traffic data sources, methods for constructing features, and parameter settings as in our previous work, i.e., HGNN (Tran, et al., 2021). Following the settings in our previous work, we generated 1600 samples, which were split into training, validation, and testing sets with a ratio of 70%, 10%, and 20%. At this stage, we used a threshold (T_c) of 31.5 minutes (short-term incident duration <= 31.5 minutes, while long-term incident duration > 31.5) calculated from the median of historical traffic incident duration. This helps to address the imbalance of classes as well. Our classification task aims to predict whether a detected incident will be short-term or long-term using up to 30 minutes of historical traffic data. Additionally, our regression task seeks to predict the duration of the detected incident in minutes. For more details on data preprocessing and feature extraction, please refer to HGNN (Tran, et al., 2021). We used the Adam optimizer with a learning rate of 0.001, a batch size of 32, and a total of 100 epochs with early stopping to accelerate the training process. A random search strategy was applied to find the best hyperparameters for the models. Min-max normalization was used to normalize traffic incident duration records to a range of [0-1].

5.2 Prediction performance

Table 1: Classification performance with $T_c = 31.5\ minutes$

AUC	Acc	Precision	Recall	F1-score
0.9155	0.8249	0.8376	0.8062	0.8216

Table 2: Regression performance

MAPE	R-squared	MAE	MSE	
37.64%	0.7803	0.025622	0.0015	**MAPE**: Mean Absolute Percentage Error **R-squared**: the coefficient of determination **MAE**: Mean Absolute Error; **MSE**: Mean Squared Error

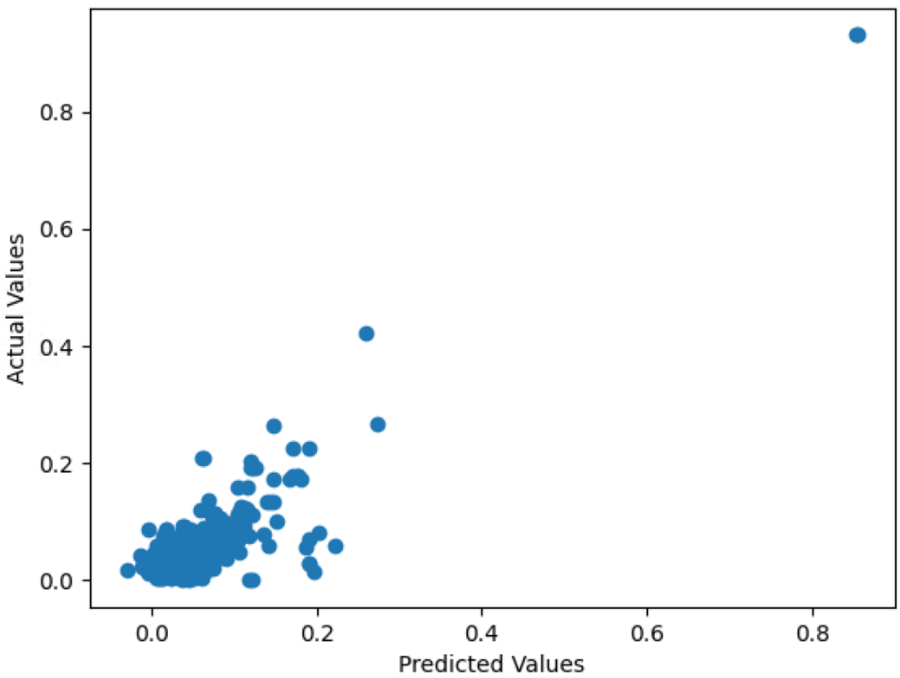

Figure 2: A scatter plot of the predicted values (x-axis) against the actual values (y-axis) from testing set (scaled between 0 and 1). Each point on the plot represents an observation in the dataset. If the predicted values perfectly match the actual values, then all the points would lie on the diagonal line y=x.

5.3 Discussion

We report the preliminary results and highlight the promising performance of MT-GN model. Our model achieved good classification performance for incident duration, as well as acceptable regression performance with low MAE and MSE values, and high variability explained. These results suggest that MT-GN can perform both tasks effectively on a road network with different road types, as demonstrated in the experiments conducted on the Brisbane dataset. In addition, Figure 2 shows a comparison between the actual traffic incident duration and the predicted traffic incident duration using the MT-GN model. The x-axis represents the predicted incident duration, while the y-axis represents the actual incident duration. Figure 2 shows that the MT-GN model has relatively accurate predictions, with most of the predicted durations falling close to the actual durations. However, there are some instances where the predicted duration is significantly different from the actual duration, indicating room for improvement in the model's predictive performance. While MAE and MSE are low, MAPE is relatively high. This discrepancy can be attributed to the absence of outlier removal in the data preprocessing, as was done in Grigorev et al.'s (2022) study. Based on these promising results, future work could focus on extending the model's capability to incorporate multi-view information, such as road maps, to capture more important information. Another could be to leverage the multi-task approach to learn various tasks such as incident severity in addition to duration enabling the model to make more accurate predictions about incident impacts. Additionally, experiments on different road networks or across multiple years of data could further validate the model's performance, as recommended by Grigorev et al.'s (2022) study.

6. Acknowledgements

The authors gratefully acknowledge the support of Australian Research Council (ARC), Queensland Department of Transport and Main Roads, and Transmax Pty Ltd under the ARC Linkage Project on Real-time Analytics for Traffic Management (LP180100018).

References

Cui, Z., Ke, R., Pu, Z., Ma, X., & Wang, Y. (2020). Learning traffic as a graph: A gated graph wavelet recurrent neural network for network-scale traffic prediction. Transportation Research Part C: Emerging Technologies.

Kuang, L., Yan, H., Zhu, Y., Tu, S., & Fan, X. (2019). Predicting duration of traffic accidents based on cost-sensitive Bayesian network and weighted K-nearest neighbor.

Li, R., Pereira, F. C., & Ben-Akiva, M. E. (2018). Overview of traffic incident duration analysis and prediction. European Transport Research Review.

Lomax, T. J., & Schrank, D. L. (2005). The 2005 urban mobility report.

Nguyen, H., Cai, C., & Chen, F. (2017). Automatic classification of traffic incident's severity using machine learning approaches. IET Intelligent Transport Systems.

Tang, J., Zheng, L., Han, C., Yin, W., Zhang, Y., Zou, Y., & Huang, H. (2020). Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review. Analytic Methods in Accident Research, 27, 100123.

Tran, T., He, D., Kim, J., & Hickman, M. (2021). Data-driven traffic incident prediction with hybrid graph-based neural network. Australasian Transport Research Forum.

Tran, T., He, D., Kim, J., & Hickman, M. (2023). MSGNN: A Multi-structured Graph Neural Network Model for Real-time Incident Prediction in Large Traffic Networks. Transport Research Board (TRB).

Wang, S., Zhang, J., Li, J., Miao, H., & Cao, J. (2021).Traffic accident risk prediction via multi-view multi-task spatio-temporal networks. IEEE Transactions on Knowledge & Data Engineer