

ORIGINAL PAPER

Open Access



Overview of traffic incident duration analysis and prediction

Ruimin Li^{1*} , Francisco C. Pereira² and Moshe E. Ben-Akiva³

Abstract

Introduction: Non-recurrent congestion caused by traffic incident is difficult to predict but should be dealt with in a timely and effective manner to reduce its influence on road capacity reduction and enormous travel time loss. Influence factor analysis and reasonable prediction of traffic incident duration are important in traffic incident management to predict incident impacts and aid in the implementation of appropriate traffic operation strategies. The objective of this study is to conduct a thorough review and discusses the research evolution, mainly including the different phases of incident duration, data resources, and the various methods that are applied in the traffic incident duration influence factor analysis and duration time prediction.

Methods: In order to achieve the goal of this study, we presented a systematic review of traffic incident duration time estimation and prediction methods developed based on various data resource, methodologies etc.

Results: based on the previous studies, we analyse (i) Data resources and characteristics: different traffic incident time phases, data set size, incident types, duration time distribution, available data resources, significant influence factors and unobserved heterogeneity and randomness, (ii) traffic incident duration analysis methods, mainly including hazard-based duration model and regression and statistical tests, (iii) traffic incident duration prediction methods and evaluation of prediction accuracy.

Conclusions: After a comprehensive review of literature, this study identifies and analyses future challenges and what can be achieved in the future to estimate and predict the traffic incident duration time.

Keywords: Incident duration analysis, Traffic incident duration prediction, Hazard-based duration model, Data mining, Influence factors

1 Introduction

One of the two main types of traffic congestion is non-recurrent congestion, which is mainly due to different events, such as traffic incidents and large-scale sports events. Although non-recurrent congestion is difficult to predict because of its stochastic nature, addressing it in a timely and effective manner is important to reduce its influence on traffic conditions. Incidents normally consist of two intervals: the primary is from the time of occurrence to the time when the incident is cleared, whereas the secondary is from the end of the primary interval to the time when the facility has resumed normal operations. Adler et al. [1] demonstrated that a one-minute duration reduction generates a €57

gain per incident and even considerably higher gains at locations with high levels of recurrent congestion (i.e., approximately €1200 per incident per minute at highly congested locations). A larger number of traffic control centres in cities and highways have deployed the Traffic Incident Management System (TIMS), which is considered as an effective tool to deal with traffic incidents, to alleviate the influence of traffic incidents on traffic conditions [2, 3]. The traffic operators must understand the main factors that influence the traffic incident duration and predict the traffic incident duration accurately to improve the TIMS efficiency. This research field has been examined in terms of two subfields with different techniques: analysis of influence factors of traffic incident duration and prediction of traffic incident duration time with or without the influence factor analysis.

* Correspondence: lrmin@tsinghua.edu.cn

¹Department of Civil Engineering, Tsinghua University, Room 304, Heshanheng Building, Beijing 100084, China

Full list of author information is available at the end of the article

With the development of traffic detection techniques and TIMS over the past decades, researchers can collect data conveniently, conduct a detailed analysis of the influence factors of traffic incident duration time, and predict traffic incident duration time in a highly accurate manner [4]. Traffic incident duration analysis and prediction in TIMS and intelligent transportation systems are currently important topics that have been applied with different results in previous studies. The incident duration time is related to various factors, such as temporal characteristics (e.g., time of day, day of the week, and/or season); incident characteristics (e.g., number of vehicles involved in an incident, truck/taxi/pedestrian involvement, number of deaths and/or injured persons); road characteristics (e.g., incident location and road condition); traffic characteristics (e.g., traffic volume); and weather conditions (e.g., rain, fog, and/or snow).

Various statistical methods have been traditionally applied to analyse and predict the traffic incident duration time. Among these methods are the following: linear/non-parametric regression [5–7], Bayesian classifier [8], hazard-based duration model (HBDM) [9], discrete choice model (DCM) [10], structure equation model (SEM) [11], and probabilistic distribution analyses [12, 13]. A new research field based on data-driven empirical algorithms and supported by unprecedented data availability has recently emerged for traffic incident duration prediction with an increasing amount of published literature. Different data mining (DM)-machine learning (ML) approaches have been employed to estimate and predict the traffic incident duration time; some of these approaches are the following: decision trees (DT) and classification trees model (CTM) [14, 15], artificial neural networks (ANN) [16–18], genetic algorithm (GA) [17], and support/relevance vector machine (SVM/RVM) [19]. Several researchers have recently begun to utilize a hybrid method [20] to predict the traffic incident duration and apply the advantages of the aforementioned methods.

Several reviews have also summarized such studies on traffic incident duration modelling [4, 21, 22], but the rapid development of prediction techniques and available data have presented a new requirement to review the development of traffic incident duration analysis and prediction. This study attempts to review previous studies on several aspects of traffic incident duration analysis and prediction. The main tasks are to compare these previous studies, identify the critical conceptual characteristics of traffic incident analysis and prediction, and discuss the future development tendency of traffic incident duration prediction.

The rest of this paper is organized as follows. First, an analysis of the available literature is conducted to present the current views and describe the development

of the specific research technique from Sections 2, 3 and 4. A critical discussion of the future challenge and direction of traffic incident duration prediction is then presented.

2 Data resources and characteristics

Previous researchers employed different datasets with various characteristics, such as different incident duration time phases, available data types, and dataset sizes, in their studies on traffic incident duration time analysis and prediction.

2.1 Different traffic incident time phases

Generally, traffic incident duration time can be defined as the time difference between the occurrence of an incident and clearance of the incident site. The duration includes four time phases: incident detection/reporting time, incident preparation/dispatching time, travel time, and clearance/treatment time. Most previous studies are limited by data availability, so they focus on the traffic incident duration time that consists of the last three phases. The duration covers the length of time between the reporting of the incident and the clearance of the road. Few studies include incident detection and recovery time [23], as well as define the duration time as the time difference from the time the Freeway Courtesy Patrol (FCP) vehicle arrives on the scene to the time the FCP leaves the scene after clearing the incident [24]. Other studies focus on the clearance time [11, 24–27], response time [28, 29], or different time phases [9, 30]. One study divides the response time into two parts: preparation time of the response team and travel time of the response vehicles [29]. The different divisions or definitions of traffic incident duration time in various studies cause difficulty in comparing their results. The difference in previous studies is also subject to used different data resources. A deeper investigation of traffic incident duration time is possible and necessary with the availability of more detailed data in the future.

2.2 Data size

Traffic incident duration is determined by various factors, including several potential factors that cannot be observed. These factors make the traffic incident duration extremely heterogeneous by nature. Utilizing a larger data set is a possible approach to improve the analysis and prediction accuracy. The adopted datasets in most previous studies includes hundreds or thousands of incident records, some of which are more than 30,000 in number [24, 26, 31, 32]. Only a few studies utilise incident datasets with less than 100 records [16, 17, 33]. Generally, studies with small datasets are more specific, but estimation and prediction of traffic incident duration time benefit more from a dataset with thousands of

records. Larger datasets tend to be better and more comprehensively reflect the characteristics of traffic incident duration.

2.3 Incident types

Most previous studies have obtained their incident/accident data sets from different traffic incident record systems or TIMS; they also have not differentiated the incident types, although the incident data include various incident types such as crashes and other events [13, 30, 34]. For example, 10 incident types are included in the adopted database of two studies [34, 35], namely, broken-down vehicle, broken-down lorry, accident, fire, flooding, fuel spillage, gas leak, police incident, collapsed manhole, and traffic light failure. However, several studies divide the data set into different types to capture the characteristics of the various incident types, such as hazards, stationary vehicles, and crashes [23, 36–38]; disabled and abandoned vehicles [39]; and collision, disabled vehicles, and traffic hazard [40]. Most previous studies also utilize the incident data set from highways or freeways between cities or urbanized regions; few of these studies adopt data from arterial roads and streets in cities. Previous studies [9, 25, 30] revealed that incident location variables significantly influence traffic incident clearance, which imply that locations have different characteristics (such as traffic conditions and geographical attributes) and procedures and training for their local Incident Response Team. Critical analyses of the effects of different incident locations are still limited because of the limited availability of data. The influence of location on traffic incident duration can be further investigated with the support of more detailed data in the future.

2.4 Duration time distribution

The distribution characteristics of the traffic incident duration time are critical for several analyses and prediction models. If the duration time fits a known probabilistic distribution, then modelling the expected value of future incidents will be convenient. Previous studies show that the traffic duration time from different datasets has different distribution characteristics. Several studies reveal that the traffic duration time meets the log-normal distribution [12, 13, 21] or log-logistic distribution [9, 31, 36, 39, 41, 42]. Weibull distribution (or with gamma heterogeneity or random parameters) provides the best likelihood ratio statistics for the used dataset in some other studies [9, 23, 25, 28, 37]. Several other studies report that the generalized F distribution is the best type for the traffic duration time distribution [24, 26]. Several studies have investigated the distribution of different duration phases or incident types and have determined that various distributional assumptions

are appropriate for the different incident duration phase times [9, 30] or incident types [23, 36, 37]. However, Smith, Smith [43] could not demonstrate that the accident clearance time conforms to a convenient probabilistic distribution. Selection of the appropriate distribution is one of the key tasks in the analysis and prediction of traffic incident duration time. Recent research [44] shows that the mixture models may be a potential direction for traffic incident duration time distribution.

2.5 Available data resources

Most of these previous studies only employ the traffic incident dataset, which commonly includes the following information items: time, location, incident type, truck, taxi, or other special vehicle involvement, as well as incident severity (e.g., number of deaths and injured persons) and weather condition. The data records in different traffic incident datasets vary according to the different data collection methods and purposes. For example, several incident datasets include geographical and/or environmental attributes, whereas others do not. Notably, two studies [45, 46] have sequential information available in textual form during the incident process, which can be useful in predicting the duration of traffic incidents.

Owing to limited data availability, only some parts of previous studies employ other types of related datasets, such as the traffic flow data, except for the traffic incident dataset [16, 17, 24, 26, 47]. Ghosh et al. [24] applied traffic flow data from 110 active sensors to study the influence of traffic conditions on the traffic incident duration time. The traffic flow data included speed, volumes by vehicle class, and sensor occupancy information aggregated into 5-min intervals.

We should note that, although this paper specifically focuses on practical dataset, simulated datasets are another source of data for traffic incident duration time estimation and prediction [48]. The relationship between incident clearance time and roadway clearance time for different traffic incident scenarios were explored on the basis of micro-simulation VISSIM modelling [49]. Post-incident traffic recovery time along an urban freeway was estimated via a simulation due to the lack of practical datasets for post-incident recovery time [50]. Simulations should be considered an optional source of basic datasets for traffic incident duration time studies when practical datasets are unavailable.

2.6 Significant influencing factors

Prior studies have generally identified various factors that influence the incident duration time or clearance time, including incident characteristics, environmental conditions, temporal factors, roadway geometry, traffic

flow conditions, operational factors and some other factors, which are shown in detailed in Table 1. Table 1 presents a summary of factors and their significant contributions, as revealed in prior studies, to traffic incident duration analysis and prediction. Factors in Table 1 can be considered as potential factors and predictors for traffic incident duration time analysis and prediction studies, respectively.

Moreover, several studies reveal that the duration of different incident types (i.e., crashes, hazards, or stationary vehicles) respond to various influence factors [37]. The duration of different duration phases (i.e., report time, response time, and/or clearance time) also respond to different influence factors [9, 30]. However, the conclusion from different datasets from different countries or regions in the significant factor analysis is sometimes different. Hojati et al. [37] found no significant effects of the infrastructure and weather on the incident duration, which is different from the findings of many other studies [9, 11, 25, 51]. In some cases, the same factor, such as taxi involvement, has been determined to have an adverse influence on the traffic duration time.

Some factors will influence the duration of traffic incidents, but incident datasets do not always record these factors, for example, the location of emergency and recovery services. Some studies reflected these factors through other factors; for example, the response time can reflect the location of emergency service to an extent. Other studies found that response time influenced the incident duration or clearance time [6, 30, 42]. In many previous studies, however, this kind of information is not included due to the limited availability of the dataset.

2.7 Unobserved heterogeneity and randomness

Limited by the data collection methods, the initial information of an incident obtained by a traffic management centre (TMC) is commonly insufficient. Furthermore, several latent influencing factors for the incident

duration time, such as the real-time traffic flow conditions and the details in characteristic differences of incident locations, cannot often be integrated into the incident dataset. Thus, we must consider several unobserved factors that are not included in the factor vector, which affect the durations and are referred to as unobserved heterogeneity. Two approaches have been adopted in the current traffic incident duration time analysis and prediction to examine the heterogeneity assumption, namely, applying the gamma distribution to incorporate heterogeneity and allowing parameters to vary across observations based on a pre-specified distribution, which is known as the random-parameter duration model [9, 23, 30, 37, 52, 53].

3 Traffic incident duration analysis

The common objective of a traffic incident duration analysis study is to determine the significant influence factors for the duration and/or severity of different types of traffic incidents, which can provide suggestions or recommendations for traffic incident management. The description and key elements of previous studies are listed in Table 2.

When an incident occurs, both the traffic operators and travellers are concerned about how long the incident process will last given that it has already lasted for x minutes, where $x \geq 0$. Thus, the length of time that elapsed from the beginning of incident detection until the end (i.e., duration time or clearance time) is noteworthy in the traffic incident duration analysis. Table 2 shows that many researchers applied various hazard-based models in their previous studies on traffic incident duration analysis. Most of these models are parametric accelerated failure time (AFT) models, which can determine the significant variables that affect the traffic incident duration time. As shown in Table 2, the distribution of accident durations has been found to be different per study and is a basic problem in modelling accident duration analysis. The

Table 1 Factors and their significant contributions to traffic incident duration

Types of Factors	Factors
Incident characteristics	Incident severity, incident type, towing requirements, type of involved vehicles, number of casualties, number of lanes blocked and incident location
Environmental conditions	Rain, snow, dry, or wet
Temporal factors	Time of day, day of week, season, month of year
Roadway geometry	Street, intersection, road layout, horizontal/vertical alignment, bottlenecks, roadway type
Traffic flow conditions	Flow, speed, occupancy, queue length
Operational factors	Lane closures, freeway courtesy service characteristics
Vehicle characteristics	Large trucks, trucks with trailers, taxis, special vehicles, compact trucks, number of vehicles involved
Others	Driver, special events, time that a police officer reaches the site, police response time, report mechanism, accident characteristics reported at accident notification

Table 2 Studies on traffic incident duration time analysis

Method Category	Methodology	Researcher	Data source	Duration time phase	Duration distribution
Hazard-based duration model (HBDM)	AFT hazard-based model	Jones et al. [41]	2156 accidents	Response time + clearance time	Log-logistic
		Nam, Mannering [9]	681 incidents	Detection/reporting, Response time, and Clearance time	Weibull, Weibull, and Log-logistic
		Chung et al. [63]	2940 accidents	Incident duration	Log-logistic
		Alkaabi et al. [25]	583 accidents	Clearance time	Weibull
		Chung, Yoon [21]	1815 accidents	Incident duration	Log-normal
		Ghosh et al. [24]	32,574 incidents	Clearance time	Generalized F
		Kaabi et al. [28]	504 accidents	Response time	Weibull with frailty
		Hojati et al. [37]	4926 incidents	Duration time	Weibull ^a
		Wang et al. [42]	1198 incidents	Incident duration time	Log-logistic
		Chimba et al. [39]	10,187 incidents	Incident duration time	Log-logistic
		Hojati et al. [23]	430 incidents	Incident duration time ^b	Weibull and log-logistic ^c
		Ghosh et al. [26]	32,574 incidents	Incident clearance time	Generalized F
		Chung et al. [53]	3863 accidents	Duration time	Gamma and inverse Gaussian
		Regression and statistical tests	Semi-parametric hazard-based model	Hou et al. [27]	2584 incidents
Shi et al. [64]	7203 incidents			Incident duration	
Log-linear models	Golob et al. [12]		525 accidents	Incident duration	Log-normal
	Statistical tests		Giuliano [13]	512 accidents	Response time + clearance time
Structural equation model	Lee et al. [11]		3147 incidents	Incident clearance time	
OLS regression truncated regression	Zhang, Khattak [31]		37,379 incidents	Event duration ^d	Log-normal or log-logistic distribution
Analysis of variance	Hojati et al. [36]		4926 records	Incident duration time	Log-logistic and log-normal ^e
Mechanism-based approach	Hou et al. [29]		828 incidents	Response time	
Association rule learning algorithm	Lin et al. [65]	999 accidents	Incident clearance time		
Binary probit and switching regression models	Ding et al. [51]	1056 incidents	Response time and clearance time		

^aWeibull AFT models with random parameters for crashes and hazards; a Weibull model has gamma heterogeneity for stationary vehicles

^bThe models include incident detection and recovery time as the components of incident duration

^cWeibull with gamma heterogeneity for crashes; log-logistic with random parameters for hazards and stationary vehicles

^dEvent duration is defined as the "time elapsed from the notification of a primary incident to the departure of the last responder from the event scene after the removal of the primary and associated secondary incidents"

^eLog-logistic distribution for hazards and stationary vehicles during weekdays; log-normal distribution for crashes

differences may have resulted from several factors, including difference in sample size (from several hundred to tens of thousands of accident records), difference in the quality of accident data, difference in countries, and differences in other factors that affect accident duration.

The other previous studies mainly employ various regression methods, for example, ordinary least squares (OLS) regression model [11, 12, 31, 51] and statistical approaches [13, 36] in traffic incident duration analysis. For the time being, various HBDM models have certain advantages in traffic incident duration analysis.

4 Traffic incident duration prediction

Traffic incident duration prediction modelling is considered as a complex problem because of heterogeneity in input data and unobserved elements. In the past two decades, many studies were conducted to investigate proper methodologies to predict traffic incident duration time by using different datasets. Most of the previous studies on traffic incident duration prediction are listed in Table 3.

4.1 Prediction methods

Several approaches have been adopted to model the prediction of the incident duration/clearance time. These

Table 3 Traffic incident duration prediction studies

Method Category	Methodology		Data source	Duration time phase	Accuracy
Regression model	Time sequential method (truncated regression model)	Khattak et al. [5]	109 larger incidents	Duration time	Not test without available dataset
	Regression model	Garib et al. [6]	205 incidents	Incident duration	81% (adjusted R^2)
	Linear regression (LR)	Peeta et al. [7]	835 crashes and 1176 debris	Clearance time	R^2 : 0.234 for crashes; 0.362 for debris
	OLS regression models	Khattak et al. [32]	59,804 incidents	Incident duration	Best MAPE: 37% ^a
	A linear model with a stepwise regression	Yu, Xia [66]	503 records	Incident duration	Acceptable (77.8% predictions have an error within 60 min)
	Cluster-based log-normal distribution model	Weng et al. [67]	2512 accidents	Accident duration	Best MAPE: 34.1%
	Quantile Regression	Khattak et al. [68]	85,000 incidents	Incident duration	RSME: 57.49 min
Fuzzy system	Fuzzy system model	Kim, Choi [69]	2457 incidents	Incident service time	Average error: 0.3 min
	Fuzzy logic (FL) model	Wang et al. [70]	457 records	Incident duration	Average performance
	Fuzzy duration model	Dimitriou, Vlahogianni [71]	1449 accidents	Accident duration	Best MAPE: 36%.
Classification Tree Method (CTM)	Decision tree	Ozbay, Kachroo [22]	650 incidents	Clearance time	60% less than 10 min
	Non-parametric regression and CTM	Smith, Smith [43]	6828 accidents	Clearance time	Not good (correct rate 58%)
	CTM	Knibbe et al. [72]	1853 incidents	Incident duration time	Theoretical reliability: 65%
	Hybrid tree-based quantile regression	He et al. [40]	1245 incidents	Incident duration	MAPE: 49.1%.
	M5P tree algorithm	Zhan et al. [15]	2585 incidents	Lane clearance time	MAPE: 42.7%.
Artificial neural networks	CTM	Chang, Chang [73]	4697 cases	Incident duration	Accuracy of classification: 75.1%.
	FL and ANNs	Wang et al. [74]	695 vehicle breakdowns	Incident duration	RMSE: about 20%
	ANNs	Wei, Lee [33]	39 accidents	Accident duration	MAPE: 20%–30%
	ANN-based models	Wei, Lee [16]	24 incidents	Incident duration	MAPE mostly under 40%.
	A sequential forecast based on two ANN-based models	Lee, Wei [17]	39 accidents	Accident duration	The MAPE value at each time point is mostly under 29%.
	Multiple LR; DT; ANN; SVM/RVM; K nearest neighbour (KNN)	Valenti et al. [19]	237 incidents	Incident duration	MAPE of the five models: 34%–44%.
	Four adaptive ANN-based models	Lopes et al. [56]	10,762 incidents	Clearance time	Model 4: 72% incidents: <10 min error; 92%: <20 min error
	Topic modelling and ANN-based models	Pereira et al. [45]	10,139 accidents	Incident duration	A median error of 9.9 min in the best model
ANN models	Vlahogianni, Karlaftis [18]	1449 accidents	Accident duration	Accuracy defined in the paper is about 10%	
Bayesian networks	Bayesian ANNs	Park et al. [57]	13,987 incidents	Incident duration	MAPE: 0.18–0.29.
	Bayesian networks	Ozbay, Noyan [75]	700 incidents	Incident clearance times	Accuracy of approximately 80%
	Probabilistic model based on a naïve Bayesian classifier (NBC)	Boyles et al. [8]	2970 incidents	Incident duration	Classification is correct half of the time.
	Bayesian decision model	Ji et al. [76]	1853 incidents	Incident duration	Theoretical reliability of 74%
	Tree-augmented NBC and a continuous model based on latent Gaussian NBC	Li, Cheng [77]	2973 incidents	Incident duration	The frequency of the correct classification is below 0.5.
	Bayesian network	Shen, Huang [78]	2629 incidents	Incident duration	

Table 3 Traffic incident duration prediction studies (Continued)

Method Category	Methodology		Data source	Duration time phase	Accuracy
					overall classification accuracy is 72.6%
hazard-based duration model	Time sequential procedure with HBDM	Qi, Teng [55]	1660 incidents	Remaining incident duration	Accuracy increases with more information
	Log-logistic AFT model	Chung [58]	4869 accidents	Accident duration	MAPE: 47%.
	Log-logistic AFT model	Hu et al. [35]	5362 incidents	Incident duration	MAPE: 43.7%.
	Weibull AFT model	Kang, Fang [79]	1327 incidents	Incident duration	MAPE: 43%.
	KNN and Log-logistic AFT model	Araghi et al. [34]	5362 incidents	Incident duration	MAPE: KNN: 41.1%; AFT: 43.7%
	HBDM	Ji et al. [38]	24,604 incidents	Clearance and arrival time	39.68% of incident: <10 min error
	Competing risk mixture HBDM	Li et al. [52]	12,093 incidents	Incident duration	MAPE: 45% for >15 mins
SVM	G-component mixture model	Zou et al. [44]	2584 incidents	Clearance time	MAPE: 39%
	Ordered probit model and SVM	Zong et al. [80]	3914 cases	Accident duration	MAPE: 22%
Combined/hybrid	SVM	Wu et al. [81]	1853 incidents	Incident duration	Total accuracy: 70%
	Ordered probit model and a rule-based supplemental module	Lin et al. [10]	22,495 incidents	Incident duration	Duration less than 60 min is 82.25% (within 10-min error)
	CTM and Rule-Based Tree Model (RBTM), DCM	Kim et al. [14]	4 years' worth of data	Incident duration	The overall confidence is more than 80%.
	A hybrid model that consists of a RBTM, MultiNomial Logit model (MNL), and NBC	Kim, Chang [20]	6765 records	Incident duration	Performed satisfactorily for incidents that last from 120 to 240 min
	Combined M5P tree and HBDM	Lin et al. [54]	602 accident records	Accident duration	MAPE: 36.2% for I-64 and 31.87% for I-190.

^aThe best mean absolute percentage error (MAPE) is 37% for the incidents that lasted for approximately 15 min

approaches can be divided into several groups based on the different classification standards.

4.1.1 Single and combined models

The majority of previous studies generally adopt one basic technique to develop the traffic incident duration prediction model. However, one method cannot suit all of the incident duration time ranges, so several researchers combined two or more methods to predict the traffic incident duration. Lin et al. [10] predicted incidents with less than 60-min duration by utilizing the ordered probit model and employed a rule-based supplemental module to predict incidents with longer than 1-h duration, which is similar to the method used by Kim et al. [14]. Kim, Chang [20] developed a hybrid model that consists of RBTM, MNL, and NBC. Lin et al. [54] constructed an M5P-HBDM (hazard-based duration model) model in which HBDMs are adopted as the leaves of the M5P tree to improve the ability of the original M5P tree algorithm to predict the traffic duration time. Vlahogianni, Karlaftis [18] applied a fuzzy entropy feature selection methodology to determine the redundant factors and Artificial Neural Network (ANN) models to predict the incident duration time.

4.1.2 Sequential and one-time models

Many previous studies assume that all information is available when predicting the traffic incident duration because these studies were conducted by utilizing a historical dataset. These models are called one-time models. In fact, obtaining all information when the traffic incident was reported to the centre is almost impossible. Thus, the traffic incident duration time prediction model must accommodate new information as it arrives in its own time sequence. Several studies have considered this challenging problem. A time sequential methodology was developed by Khattak et al. [5] to predict the incident duration as the TMC receives the incident information based on a dataset of 109 large-scale incidents. Khattak et al. [32] developed dynamic incident duration models to predict the incident duration more accurately because additional information can be obtained as an incident progresses. Wei, Lee [16] developed a time sequential traffic incident duration prediction procedure utilizing ANN-based models and data fusion techniques. Lee, Wei [17] then employed ANNs and genetic algorithms to construct two models to provide a sequential prediction of accident duration from the accident notification to clearance. Qi, Teng

[55] developed a time sequential procedure that included different hazard-based duration regression models with different variables for each stage according to the specific information available. Lopes et al. [56] developed four adaptive ANN-based models to be activated with the incoming data to improve the predictive performance. Pereira et al. [45] also developed sequential models to obtain more reliable predictions by using a radial basis function network.

4.2 Evaluation of prediction accuracy

The prediction accuracy is generally evaluated by comparing the detected traffic duration time and predicted traffic duration time. The MAPE is the most frequently applied measurement to investigate the accuracy of the predictions. Root mean squared error (RMSE) and mean percentage error (MPE) are also used in some cases. The lower the RMSE and MAPE values are, the more accurate the prediction model becomes. The MPE shows prediction bias. Notably, the MAPE has several drawbacks. For example, the MAPE increases when the observed value is lower, and even has no upper limit to the percentage error. The mean absolute error and mean squared prediction error can also be employed [57].

Another frequently utilized measure of effectiveness in traffic incident duration prediction is related to a certain tolerance of the prediction error [15, 20, 43, 58]. Similarly, Qi, Teng [55] stated that an incident duration is correctly predicted if the percentage of the relative error tolerance of an incident is less than a given value. Park et al. [57] defined the proportion of the underestimated prediction to reveal what percentage of incident has been underestimated.

5 Challenges and future work

The challenges of traffic incident duration analysis and prediction are summarized in Table 4 and explained as follows.

5.1 How to combine multiple data resources

Several previous studies [6, 15, 41] have revealed that except for the observed factors, several latent factors can affect the traffic incident duration. Thus, obtaining more detailed and various types of data is necessary for a more accurate analysis and prediction of traffic incident duration time.

First, although the incident databases in many countries are relatively extensive, they still have the limitation of no-data field that provides the exact occurrence time of the incident. In particular, we can only obtain the time stamp when the operator first recorded an incident into the database. The incident detection/reporting time is an important phase in traffic incident duration and can affect the duration time of the following phases. Obtaining the incident exact occurrence time based on an intelligent vehicle system, such as the eCall system [59, 60] in Europe and the OnStar system of General Motors, is possible in the future.

Second, several studies [16, 17, 40] prove that the traffic flow condition can affect the traffic incident duration time; thus, how to integrate the increasing data on traffic flow condition is also a critical topic in future studies on traffic incident duration analysis and prediction. Traffic condition information was previously sourced from the section detector, and the parameters mainly included traffic flow volume, average spot speed, and occupancy. Owing to the recent development of floating cars and smartphones, several traffic information service companies can now provide the travel time information, which can be considered as an information resource.

Third, new data resources, such as crowdsourcing technology (e.g., Waze, Twitter and Weibo), can also provide information on traffic incident conditions. Gu et al. [61] studied a method based on natural language processing to extract incident information from tweets on highways and arterial roads. Kurkcu et al. [62] determined that Web-based social media data can be applied for more

Table 4 challenges of traffic incident duration analysis and prediction

Challenges	Potential methods	Previous research
Combining multiple data resources	Intelligent vehicle system (for example, eCall)	Sdongos et al. [59]; Oorni, Goulart [60]
	Traffic condition detection information	Wei, Lee [16]; Lee, Wei [17]; He et al. [40]
	Crowdsourcing technology	Gu et al. [61]; Kurkcu et al. [62]
Time sequential prediction model	Based on response term's report	Khattak et al. [5]; Pereira et al. [45]; Li et al. [46]
	Based information from social media	Gu et al. [61]
Outlier prediction	Different models for different duration ranges	Lin et al. [10]; Valenti et al. [19]
	A time sequential prediction model	Qi, Teng [55]; Pereira et al. [45]; Li et al. [46]
Improvement of prediction methods	Machine Learning	Zhan et al. [15]; Lin et al. [54]; Park et al. [57]; Ma et al. [82] et al.
	Updated HBDM	Li et al. [46] et al.
Combining recovery times	Combine new data resource	Hojati et al. [23]
Influence of unobserved factors	Randomness model	Nam, Mannering [9]; Hojati et al. [23]; Li et al. [52]

effective real-time incident responses and obtain time-critical incident-related information. Utilizing such information involves several challenges, such as how to obtain more useful records and adopting such information accurately because they can be vague and limited by the text size. Therefore, how to combine such emerging information sources with traffic incident duration analysis and prediction is also a challenging topic in future studies. Text analysis tools, such as topic modelling and sentiment analysis, show good potential for discovering useful information for analysis and prediction.

Overall, the first important step for future studies in traffic incident duration analysis and prediction is to combine extensive information from connected vehicles, traffic information providers, and social media to increase the amount of datasets available for study. Information from various sources should also be acquired from incidents and constantly updated to correct prediction results. Prediction accuracy may be improved through the integration of more data.

5.2 Time sequential prediction model

The traditional methods that analyse and predict the traffic incident duration time employ the historic dataset of traffic incidents with or without other dataset types, such as the traffic condition dataset. These methods assume that when a model is employed to analyse or predict the traffic incident duration time, all the possible information has already been obtained. However, when an incident is reported to the traffic control centre, information on the incident (e.g., location, time, weather, and traffic conditions) is provided by the reporting persons with considerable limitations. After the traffic response team arrives at the incident location, further information is sent to the traffic control centre [45], which can help understand the traffic incident more accurately.

Two possible data types can provide sequential useful information on an incident. One type is the report from the incident response team, as previously mentioned. After the team arrives at the incident location, the incident record is updated in several aspects, including affected lanes, traffic condition, and size of rescue force. The other type is from crowdsourcing platforms. Travelers who pass through the incident site can post information about the incident on Twitter or other platforms, thereby providing useful information [61]. Thus, determining appropriate methods to mine useful information from these different data resources, such as text analysis technique and machine learning techniques, can be a challenging subject of future studies.

A time sequential prediction model needs to be developed based on various basic models, such as HBDM, various ANN models, and some other models, to

accommodate new information chronologically. Time sequential prediction models can predict the elapsed time of an incident more accurately in support of the appropriate traffic management and traveller information services by using continually updated information.

5.3 Outlier prediction

Traffic incident duration prediction currently faces difficulties in predicting outliers accurately. Most previous studies show that the probability distribution of incident duration has a long tail, which prevents several duration prediction (i.e., statistical) models from predicting extreme values properly. For example, the HBDM models are disadvantaged by their inability to predict extreme values. The reason is that the statistical models tend to capture the central tendency in the data rather than the outliers to a certain extent. For example, several studies [30, 32] show unreasonable predictions that are longer or shorter than the average range with the same prediction model. Valenti et al. [19] compared five different models for traffic incident duration time prediction and found that only the ANN-based model can predict an incident longer than 90 min. Lin et al. [10] employed different models for different duration ranges; an embedded discrete model is utilized on incidents with a duration of less than 60 min, whereas a rule-based supplemental module is adopted for incidents that can last for more than 1 h. In reality, the longer the traffic incident duration time, the higher its influence on the traffic system. Thus, predicting a longer outlier traffic incident duration as accurately as possible is important. Pereira et al. [45] reported that a time sequential model with continuously updated information can be an alternative method to predict the longer traffic incident duration, particularly through the incremental analysis of incoming textual messages. Qi, Teng [55] determined that the accuracy of the incident duration prediction increased as more information is incorporated into the models. Thus, a time sequential model can be a feasible prediction method for longer outliers.

5.4 Improvement of prediction methods

The appropriate method is key to the accurate prediction of the traffic incident duration time. The two main types of utilized methods in the past are statistical and data-driven methods. The former are mainly regression and hazard-based models, whereas the latter are mainly neural networks and decision tree models. However, the accuracy measurements (e.g., MAPE) show that the prediction of most methods is only reasonable and few are very good. A few methods are suitable partly because of the randomness of the traffic incident duration. Several studies investigate the combination of two or more methods, as previously mentioned, to overcome the

limitations of a single model. The results indicate a slight but insignificant improvement. Machine learning has recently developed rapidly and can provide a potential direction to explore prediction methods for traffic incident duration. Machine learning can conduct data-driven predictions from sample inputs by constructing an algorithm that can learn from the data. Several machine learning methods, such as DT learning, SVM, Bayesian networks, and genetic algorithms, have been applied in predicting traffic incident duration time [15, 17, 54, 57]. It needs to be noted that each of these approaches has its own advantages and disadvantages. For example, DT learning may consider many possible outcomes but the final decisions based primarily on expectations, which could lead to unrealistic results. SVM/SVR is powerful for solving problems of classification, regression, but is more time consuming if dealing with very large datasets. Bayesian networks can accommodate incomplete information but computing posterior distribution may be extremely difficult. In traffic incident duration prediction, genetic algorithms help to reduce the input features but the time taken for convergence maybe longer.

The prediction methods need to focus on the following aspects in future practical applications:

- 1) The critical function of the traffic incident duration time prediction model is to support real-time traffic management and traveller information service, so the prediction model has to be run online and must be less time-consuming.
- 2) The prediction model must adopt incomplete information because when an incident is reported, only part of the information on the incident can be obtained for incident duration prediction and even until the incident is cleared. Obtaining all the information that influences the traffic incident duration time is impossible. For example, if no traffic detector is present near the incident location, then obtaining the volume of traffic that passes through the incident location is almost impossible. Thus, the traffic incident duration prediction model to be developed should have the ability to consider incidents with incomplete information.

In traffic incident duration estimation and prediction, both the traffic operators and travellers are concerned with the length of time between detection and clearance of an incident; that is, how long the entire process will last given that it has already lasted for several minutes. The hazard-based duration model can provide effective techniques to estimate and predict traffic incident duration time as shown by previous studies. HBDM remains a significant, potential method for future work, but it

needs to consider heterogeneity, variation in time, and randomness in modelling. Furthermore, with the combination of different data resources and larger datasets, more advanced machine-learning and other potential methods can be explored in the future to predict traffic incident duration (e.g., deep learning approach and self-learning method). Several text-mining tools should be employed in data processing to deal with more useful, textual data resources from social media or from reports of incident responders [45].

5.5 Combining recovery times

Two previous studies [23, 50] show that longer traffic incident duration can result in longer recovery times, leading to severe congestion. Travelers must generally know how long the recovery time will be so that they can select the suitable route to their destination. Detecting the recovery time was previously difficult because of the limitations in the fixed traffic detectors; few studies consider the recovery time [23]. The development of several emerging traffic-condition detection techniques currently provides an opportunity to detect or infer the recovery time duration. For example, INRIX or Baidu in China can provide real-time traffic conditions mostly based on floating car data of taxis, trucks, coaches, and other vehicle types. Such information can be used to infer the recovery time duration of an incident, and sometimes the simulation dynamic traffic assignment tool is also needed. One of the difficulties with this inference is how to identify the congestion cause, that is, whether the congestion is due to the incident independently or caused by other factors (e.g., recurrent congestion). Investigating the significant factors that influence the recovery time are possible with the recovery time data, which can be helpful in adopting appropriate traffic management strategies to reduce the incident influence. Thus, determining a proper method to infer or detect the recovery time and corresponding method to analyse and predict it can be a future topic. An appropriate traffic theory model or method based on simulations may provide effective means to infer the recovery time of traffic flow conditions.

5.6 Influence of unobserved factors

Many previous studies show that except for several recorded factors, several unobserved factors affect the traffic incident duration. The prediction model must deal with unobserved factors. Several researchers [9, 23, 52] have recently investigated methods dealing with unobserved heterogeneity, such as the duration model with random parameter. The reason for heterogeneity cannot be easily understood. For example, different response patterns will result in different traffic incident duration times even for incidents with similar factors. Several

countries, including China, have deployed a quick clearance policy for minor accidents, such as those without injuries or vehicles that are still functional. In fact, drivers who become involved in incidents can negotiate among themselves before the incident response team arrives at the scene. The drivers can also fill in the necessary insurance forms and take photos as evidence to reduce the incident duration. However, other drivers will stay at the incident scene and wait for the incident response team even for minor incidents, thereby resulting in a longer traffic incident duration time. This difference is related to several characteristics of different drivers, such as psychological traits, experiences, and knowledge, which are difficult to consider in the modelling. Thus, control for randomness, heterogeneity, and the time-varying variables in the traffic incident duration estimation and prediction provide avenues for future work.

6 Conclusion

To effectively support different traffic incident management strategies and applications, an appropriate method that can determine the significant factors for the traffic incident duration and prediction techniques to match various circumstances and data resources in a timely manner to predict traffic incident duration must be applied. This study reviews the literature on traffic incident duration analysis and prediction. It also analyses the different data resources and characteristics, including traffic incident time phase, data set size, incident types, duration time distribution, available data resources, significant influence factors, unobserved heterogeneity, and randomness. We then investigated the various techniques employed in traffic incident duration analysis and prediction. Finally, we analysed several challenges in future research and application, such as how to combine extensive data resources, the time sequential prediction model, outlier prediction, improvement of prediction methods, combining recovery times, and influence of unobserved factors.

Traffic detection techniques, social media platforms, and machine learning techniques have all been promoted rapidly in the past few years, thereby providing new opportunities for traffic incident duration time analysis and prediction in many ways. Different traffic incidents are still the main reason for traffic congestion in urban road networks and highways between cities. Thus, exploring new methods to analyse and predict traffic incident duration more accurately is necessary in the future to support the adoption of appropriate traffic operation strategies for traffic management under various traffic incident conditions. Future studies may combine recovery time with traffic incident duration time and various data sources, focus on the outlier value prediction and experiment with novel predictive methodologies, or investigate the effects of unobserved factors to improve prediction accuracy.

Acknowledgments

This study was supported by the National Natural Science Foundation of China under Grant No. 71361130015 and Beijing Natural Science Foundation under Grant No.8162024.

Authors' contributions

All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Civil Engineering, Tsinghua University, Room 304, Heshanhang Building, Beijing 100084, China. ²Department of Management Engineering, Technical University of Denmark, DTU Bygningstorvet 116B, 2800 Kongens-Lyngby, Denmark. ³Department of Civil and Environmental Engineering, MIT, Room 1-181, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.

Received: 12 November 2017 Accepted: 22 May 2018

Published online: 31 May 2018

References

- Adler MW, Ommeren JV, Rietveld P (2013) Road congestion and incident duration. *Econ Transp* 2(4):109–118. <https://doi.org/10.1016/j.ecotra.2013.12.003>.
- Schrank D, Lomax T (2009) 2009 urban mobility report. Texas Transportation Institute, College Station
- Owens N, Armstrong A, Sullivan P, Mitchell C, Newton D, Brewster R, Rego T (2010) Traffic Incident Management Handbook. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C
- Wang W, Chen H, Bell MC (2005) A review of traffic incident duration analysis. *J Transp Syst Eng Inf Technol* 5(3):127–140.
- Khattak AJ, Schofer JL, Wang M-H (1995) A simple time sequential procedure for predicting freeway incident duration. *IVHS J* 2(2):113–138.
- Garib A, Radwan AE, Al Deek H (1997) Estimating Magnitude and duration of Incident delays. *J Transp Eng* 123(6):459–468. [https://doi.org/10.1061/\(ASCE\)0733-947X\(1997\)123:6\(459\)](https://doi.org/10.1061/(ASCE)0733-947X(1997)123:6(459)).
- Peeta S, Ramos JL, Gedela S (2000) Providing Real-Time Traffic Advisory and Route Guidance to Manage Borman Incidents On-Line Using the Hoosier Helper Program. Joint Transportation Research Program, 1284 Civil Engineering Building, Purdue University, West Lafayette, Indiana 47907-1284.
- Boyles S, Fajardo D, Waller ST (2007) A Naive Bayesian Classifier for Incident Duration Prediction. Paper presented at the TRB 86th Annual Meeting Compendium of Papers CD-ROM, Washington DC, United States.
- Nam D, Mannering F (2000) An exploratory hazard-based analysis of highway incident duration. *Transp Res A* 34(1):85–102. [https://doi.org/10.1016/S0965-8564\(98\)00065-2](https://doi.org/10.1016/S0965-8564(98)00065-2).
- Lin P-W, Zou N, Chang G-L (2004) Integration of a Discrete Choice Model and a Rule-Based System for Estimation of Incident Duration: a Case Study in Maryland. In: CD-ROM of Proceedings of the 83rd TRB Annual Meeting, Washington, D.C.
- Lee J-Y, Chung J-H, Son B (2010) Incident clearance time analysis for Korean freeways using structural equation model. *J East Asia Soc Transp Stud* 8: 1850–1863.
- Golob TF, Recker WW, Leonard JD (1987) An analysis of the severity and incident duration of truck-involved freeway accidents. *Accid Anal Prev* 19(5): 375–395. [https://doi.org/10.1016/0001-4575\(87\)90023-6](https://doi.org/10.1016/0001-4575(87)90023-6).
- Giuliano G (1989) Incident characteristics, frequency, and duration on a high volume urban freeway. *Transp Res A* 23(5):387–396. [https://doi.org/10.1016/0191-2607\(89\)90086-1](https://doi.org/10.1016/0191-2607(89)90086-1).
- Kim W, Chang G-L, Rochon SM (2008) Analysis of Freeway Incident Duration for ATIS Applications. In: 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting, New York NY.
- Zhan C, Gan A, Hadi M (2011) Prediction of lane clearance time of freeway incidents using the MSP tree algorithm. *IEEE Trans Intell Transp Syst* 12(4): 1549–1557.

16. Wei C-H, Lee Y (2007) Sequential forecast of incident duration using artificial neural network models. *Accid Anal Prev* 39:944–954.
17. Lee Y, Wei C-H (2010) A computerized feature selection method using genetic algorithms to forecast freeway accident duration times. *Copmut Aided Civil Infrastruct Eng* 25:132–148.
18. Vlahogianni EI, Karlaftis MG (2013) Fuzzy-entropy neural network freeway incident duration modeling with single and competing uncertainties. *Copmut Aided Civil Infrastruct Eng* 28(6):420–433. <https://doi.org/10.1111/mice.12010>.
19. Valenti G, Lelli M, Cucina D (2010) A comparative study of models for the incident duration prediction. *Eur Transp Res Rev* 2(2):103–111.
20. Kim W, Chang G-L (2012) Development of a hybrid prediction model for freeway incident duration: a case study in Maryland. *Int J Intell Transp Syst Res* 10(1):22–33. <https://doi.org/10.1007/s13177-011-0039-8>.
21. Chung Y, Yoon B-J (2012) Analytical method to estimate accident duration using archived speed profile and its statistical analysis. *KSCE J Civ Eng* 16(6):1064–1070.
22. Ozbay K, Kachroo P (1999) Incident management in intelligent transportation systems. Artech House Publishers, Norwood.
23. Hojati AT, Ferreira L, Washington S, Charles P, Shobeirinejad A (2014) Modelling total duration of traffic incidents including incident detection and recovery time. *Accid Anal Prev* 71:296–305. <https://doi.org/10.1016/j.aap.2014.06.006>.
24. Ghosh I, Savolainen PT, Gates TJ (2012) Examination of factors affecting freeway incident clearance times: a comparison of the generalized F model and several alternative nested models. *J Adv Transport*. <https://doi.org/10.1002/atr.1189>.
25. Alkaabi AMS, Dissanayake D, Bird R (2011) Analyzing clearance time of urban traffic accidents in Abu Dhabi, United Arab Emirates, with hazard-based duration modeling method. *Transp Res Rec* 2229:46–54. <https://doi.org/10.3141/2229-06>.
26. Ghosh I, Savolainen PT, Gates TJ (2014) Examination of factors affecting freeway incident clearance times: a comparison of the generalized F model and several alternative nested models. *J Adv Transport* 48(6):471–485. <https://doi.org/10.1002/atr.1189>.
27. Hou L, Lao Y, Wang Y, Zhang Z, Zhang Y, Li Z (2014) Time-varying effects of influential factors on incident clearance time using a non-proportional hazard-based model. *Transp Res A Policy Pract* 63:12–24. <https://doi.org/10.1016/j.tra.2014.02.014>.
28. Kaabi AA, Dissanayake D, Bird R (2012) Response time of highway traffic accidents in Abu Dhabi investigation with hazard-based duration models. *Transp Res Rec* 2278:95–103. <https://doi.org/10.3141/2278-11>.
29. Hou L, Lao Y, Wang Y, Zhang Z, Zhang Y, Li Z (2013) Modeling freeway incident response time: a mechanism-based approach. *Transp Res C* 28:87–100. <https://doi.org/10.1016/j.trc.2012.12.005>.
30. Li R (2015) Traffic incident duration analysis and prediction models based on the survival analysis approach. *IET Intell Transp Syst* 9(4):351–358. <https://doi.org/10.1049/iet-its.2014.0036>.
31. Zhang H, Khattak AJ (2010) Analysis of cascading incident event durations on urban freeways. *Transp Res Rec* 2178:30–39. <https://doi.org/10.3141/2178-04>.
32. Khattak A, Wang X, Zhang H (2012) Incident management integration tool: dynamically predicting incident durations, secondary incident occurrence and incident delays. *IET Intell Transp Syst* 6(2):204–214.
33. Wei C-H, Lee Y (2005) Applying data fusion techniques to traveler information services in highway network. *J East Asia Soc Transp Stud* 6:2457–2472.
34. Araghi BN, Hu S, Krishnan R, Bell M, Ochieng W (2014) A comparative study of k-NN and hazard-based models for incident duration prediction. In: 2014 17th IEEE international conference on intelligent transportation systems, ITS-C 2014, pp 1608–1613. <https://doi.org/10.1109/ITS-C.2014.6957923>.
35. Hu J, Krishnan R, Bell MGH (2011) Incident duration prediction for in-vehicle navigation system. Paper presented at the Transportation Research Board annual meeting, Washington DC.
36. Hojati AT, Ferreira L, Charles P, bin Kabit MR (2012) Analysing freeway traffic incident duration using an Australian data set. *Road Transp Res* 21(2):19–31.
37. Hojati AT, Ferreira L, Washington S, Charles P (2013) Hazard based models for freeway traffic incident duration. *Accid Anal Prev* 52:171–181. <https://doi.org/10.1016/j.aap.2012.12.037>.
38. Ji Y, Jiang R, Qu M, Chung E (2014) Traffic incident clearance time and arrival time prediction based on hazard models. *Math Probl Eng* 2014. <https://doi.org/10.1155/2014/508039>.
39. Chimba D, Kutela B, Ogletree G, Horne F, Tugwell M (2014) Impact of abandoned and disabled vehicles on freeway incident duration. *J Transp Eng* 140(3). [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000635](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000635).
40. He Q, Kamarianakis Y, Jintanukul K, Wynter L (2011) Incident duration prediction with hybrid tree-based quantile regression. IBM research report.
41. Jones B, Janssen L, Mannering F (1991) Analysis of the frequency and duration of freeway accidents in Seattle. *Accid Anal Prev* 23(4):239–255. [https://doi.org/10.1016/0001-4575\(91\)90003-N](https://doi.org/10.1016/0001-4575(91)90003-N).
42. Wang J, Cong H, Qiao S (2013) Estimating freeway incident duration using accelerated failure time modeling. *Saf Sci* 54:43–50. <https://doi.org/10.1016/j.ssci.2012.11.009>.
43. Smith K, Smith BL (2001) Forecasting the Clearance Time of Freeway Accidents. Center for Transportation Studies, University of Virginia, Charlottesville.
44. Zou Y, Henrickson K, Lord D, Wang Y, Xu K (2016) Application of finite mixture models for analysing freeway incident clearance time. *Transportmetrica A Transp Sci* 12(2):99–115. <https://doi.org/10.1080/23249935.2015.1102173>.
45. Pereira F, Rodrigues F, Ben-Akiva M (2013) Text analysis in incident duration prediction. *IEEE Intell Transp Syst Trans Mag* 37:177–192. <https://doi.org/10.1016/j.trc.2013.10.002>.
46. Li R, Pereira FC, Ben-Akiva ME (2015) Competing risk mixture model and text analysis for sequential incident duration prediction. *Transp Res C* 54:74–85. <https://doi.org/10.1016/j.trc.2015.03.009>.
47. Sullivan EC (1997) New model for predicting freeway incident and incident delays. *J Transp Eng* 123(4):267–275.
48. Knoop VL, Hoogendoorn SP, van Zuylen H (2010) Stochastic Incident Duration: Impact on Delay. In: Transportation Research Board 89th Annual Meeting, Washington DC, United States.
49. Zhou H, Tian X (2012) Modeling analysis of incident and roadway clearance time. *Procedia Soc Behav Sci* 43:349–355.
50. Jeahani M, James P, Saka AA, Ardeshiri A (2015) Traffic recovery time estimation under different flow regimes in traffic simulation. *J Traffic Transp Eng Engl Ed* 2(5):291–300.
51. Ding C, Ma X, Wang Y, Wang Y (2015) Exploring the influential factors in incident clearance time: disentangling causation from self-selection bias. *Accid Anal Prev* 85:58–65. <https://doi.org/10.1016/j.aap.2015.08.024>.
52. Li R, Pereira FC, Ben-Akiva ME (2015) Competing risks mixture model for traffic incident duration prediction. *Accid Anal Prev* 75:192–201. <https://doi.org/10.1016/j.aap.2014.11.023>.
53. Chung YS, Chiou YC, Lin CH (2015) Simultaneous equation modeling of freeway accident duration and lanes blocked. *Anal Methods Accid Res* 7:16–28. <https://doi.org/10.1016/j.jamar.2015.04.003>.
54. Lin L, Wang Q, Sadek AW (2016) A combined MSP tree and hazard-based duration model for predicting urban freeway traffic accident durations. *Accid Anal Prev* 91:114–126. <https://doi.org/10.1016/j.aap.2016.03.001>.
55. Qi YG, Teng HH (2008) An information-based time sequential approach to online incident duration prediction. *J Intell Transp Syst Technol Plann Oper* 12(1):1–12.
56. Lopes J, Bento J, Pereira FC, Ben-Akiva M (2013) Dynamic forecast of incident clearance time using adaptive artificial neural network models. Paper presented at the Transportation Research Board 92nd annual meeting Washington DC, 2013-1-13 to 2013-1-17.
57. Park H, Haghani A, Zhang X (2016) Interpretation of Bayesian neural networks for predicting the duration of detected incidents. *J Intell Transp Syst Technol Plann Oper* 20(4):385–400.
58. Chung Y (2010) Development of an accident duration prediction model on the Korean freeway systems. *Accid Anal Prev* 42:282–289.
59. Sdongos E, Bolovinou A, Tsogas M, Amditis A, Guerra B, Manso M (2017) Next generation automated emergency calls - Specifying next generation ecall & sensor-enabled emergency services. In: 2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC), 8-11 Jan, pp 1–6. <https://doi.org/10.1109/CCNC.2017.8015368>.
60. Oorni M, Goulart A (2017) In-vehicle emergency call services: eCall and beyond. *ICOMM* 55(1):159–165. <https://doi.org/10.1109/MCOM.2017.1600289CM>.
61. Gu Y, Qian Z, Chen F (2016) From Twitter to detector: real-time traffic incident detection using social media data. *Transp Res C* 67:321–342. <https://doi.org/10.1016/j.trc.2016.02.011>.
62. Kurkcu A, Morgul EF, Ozbay K (2015) Extended implementation method for virtual sensors: web-based real-time transportation data collection and analysis for incident management. *Transp Res Rec* (2528):27–37. <https://doi.org/10.3141/2528-04>.

63. Chung Y, Walubita LF, Choi K (2010) Modeling accident duration and its mitigation strategies on South Korean freeway systems. *Transp Res Rec* 2178:49–57. <https://doi.org/10.3141/2178-06>.
64. Shi Y, Zhang L, Liu P (2015) Survival analysis of urban traffic incident duration: a case study at shanghai expressways. *J Comput (Taiwan)* 26(1):29–39.
65. Lin L, Wang Q, Sadek A (2014) Data mining and complex network algorithms for traffic accident analysis. *Transp Res Rec* 2460. <https://doi.org/10.3141/2460-14>.
66. Yu B, Xia Z (2012) A methodology for freeway incident duration prediction using computerized historical database. In: *CICTP 2012: Multimodal Transportation Systems - Convenient, Safe, Cost-Effective, Efficient - Proceedings of the 12th COTA International Conference of Transportation Professionals*, pp 3463–3474. <https://doi.org/10.1061/9780784412442.351>.
67. Weng J, Qiao W, Qu X, Yan X (2015) Cluster-based lognormal distribution model for accident duration. *Transportmetrica A Transp Sci* 11(4):345–363. <https://doi.org/10.1080/23249935.2014.994687>.
68. Khattak AJ, Liu J, Wali B, Li X, Ng M (2016) Modeling traffic incident duration using quantile regression. *Transp Res Rec* 2554:139–148.
69. Kim HJ, Choi H-K (2001) A comparative analysis of incident service time on urban freeways. *J Int Assoc Traffic Saf Sci* 25(1):62–72.
70. Wang W, Chen H, Bell M (2002) A Study of the Characteristics of Traffic Incident Duration on Motorways. Paper presented at the Traffic And Transportation Studies, Guilin, China,.
71. Dimitriou L, Vlahogianni EI (2015) Fuzzy modeling of freeway accident duration with rainfall and traffic flow interactions. *Anal Methods Accid Res* 5–6:59–71. <https://doi.org/10.1016/j.jamar.2015.04.001>.
72. Knibbe WJJ, Alkim TP, Otten JFW, Aidoo MY (2006) Automated estimation of incident duration on Dutch highways. In: *Proceedings of the 2006 IEEE intelligent transportation systems conference, Toronto, Canada*, pp 870–874.
73. Chang H, Chang T (2013) Prediction of freeway incident duration based on classification tree analysis. In: *Proceedings of the Eastern Asia Society for Transportation Studies*.
74. Wang W, Chen H, Bell MC (2005) Vehicle breakdown duration modelling. *J Transp Stat* 8(1):75–84.
75. Ozbay K, Noyan N (2006) Estimation of incident clearance times using Bayesian networks approach. *Accid Anal Prev* 38:542–555. <https://doi.org/10.1016/j.aap.2005.11.012>.
76. Ji YB, Zhang X, Sun L (2008) Traffic incident duration prediction based on the Bayesian decision tree method. In: *Proceedings of transportation and development innovative best practices 2008, Beijing*, pp 338–343.
77. Li D, Cheng L (2011) Bayesian Network Classifiers for Incident Duration Prediction. Paper presented at the Transportation Research Board 90th Annual Meeting, Washington DC,.
78. Shen L, Huang M (2011) Data mining method for incident duration prediction. *Appl Inform Commun Commun Comput Inf Sci* 224(1):484–492.
79. Kang G, S-E F (2011) Applying survival analysis approach to traffic incident duration prediction. In: *First International Conference on Transportation Information and Safety (ICTIS), Wuhan, China*, pp 1523–1531.
80. Zong F, Zhang H, Xu H, Zhu X, Wang L (2013) Predicting severity and duration of road traffic accident. *Math Probl Eng* 2013. <https://doi.org/10.1155/2013/547904>.
81. Wu W, Chen S, Zheng C (2011) Traffic incident duration prediction based on support vector regression. In: *Proceedings of the ICCTP 2011*, pp 2412–2421.
82. Ma X, Ding C, Sen L, Wang Y, Wang Y (2017) Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method. *IEEE Trans Intell Transp Syst* 18(9):2303–2310. <https://doi.org/10.1109/TITS.2016.2635719>.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com