METAVERSE

CE784A Literature Review: Semester 2021-22 (II)

Traffic Incident Duration Analysis and Prediction

Aditya Gupta(190060), Chirag Sharma(190248), Prakhar Pradhan(190618),

Turnitin accuracy: 16%

1. Introduction

Prediction of traffic incident duration and analysis has become very important in Traffic Management. As the number of vehicles is increasing on-road, the need for the accurate forecast to take necessary steps to reduce traffic congestion in case of any incident is essential, and reducing incident duration can be seen as one of the main targets of traffic incident management. Many cities are using the Traffic Incident Management System (TIMS) for managing traffic incidents but improving its efficiency requires a clear basis of the factors affecting the traffic incident. One of the significant factors used widely in many models is incident duration, which is the time a traffic incident takes from reporting to clearance. Thus, an incident duration can be described as consisting of three parts, namely, reporting time, response time, and clearance time as shown in Fig1. Predicting the incident duration on the basis of various conditions, such as regional and local traffic conditions, time of day, etc., is critical in improving the efficiency of the predictions and the present systems.
Various models such as linear regression, bayesian classifier, OLS, Fuzzy system model, etc., have been utilized to predict the duration. Several studies have also used data-driven empirical algorithms and various machine learning approaches such as decision trees, artificial neural networks, classification trees, K-nearest neighbors, and Support Vector Machine models to predict and analyze the incident duration. Numerous studies have deployed combined/hybrid models to predict and analyze the term.
This literature review focuses on various studies previously done in the prediction and analysis of incident duration and attempts to review several aspects.

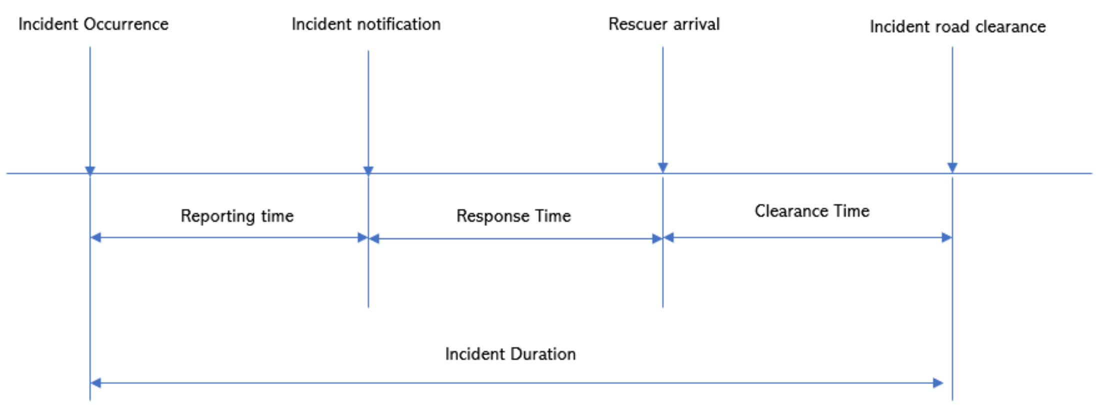

Fig 1. Components of incident duration

2. Type of factors and their role in traffic incident duration

2.1. Incident characteristics:

Characteristics such as severity, type, towing requirements, type of vehicle, number of casualties, number of lanes blocked in the incident, and the location of the incident

2.2. Environmental conditions:

Conditions such as rain, snow, dry, or wet : (heavy snow during winter) plays important role in predicting incident duration

2.3. Highway geometry and type:

Road layout, type of street, intersection, bottlenecks, roadway type are factors that influence incident duration. For eg., the business of a highway would directly influence the incident duration.

2.4 Traffic flow conditions :

- Vehicle speed: can be defined as the rate of motion at which drivers operate their vehicles.
- Traffic flow (volume) can be defined as the quantity which represents the number of vehicles passes over each point on the given highway in a certain period of time. The units can be represented in terms of vehicles per time period like 5-min flow or hourly flow.
- The Average Daily Traffic (ADT) can be defined as the average of the daily volumes (flow) for each day of the selected period of interest.
- Annual Average Daily Traffic (AADT) is simply the total vehicle volume (flow) per day, averaged over a year-long period.
- VMT(vehicle miles traveled) can be defined as the total miles driven by all of the vehicles over a roadway segment during a specified time period.
- VHT(vehicle hours traveled) can be defined as the total amount of time spent by all of the vehicles over a roadway segment during a specified time period
- Q can be represented as the sum of VMT divided by the sum of VHT. For a highway segment over time, it can represent a measure of the efficiency of the transportation system.

2.5. Operational factors:

These factors include lane closures, freeway courtesy service characteristics

2.6 Vehicle characteristics:

Characteristics of a vehicle such as large or small, with or without trailers, taxis, special vehicles, number of vehicles involved

2.7 Temporal factors:

- Time of day: if an incident happens during peak hours (for eg., 8-10 am and 5-8 pm) duration may increase
- Day of the week: weekends generally are more congested
- Season or month of the year: heavy load during Christmas season

3. Incident Duration Prediction

The objective of the study is to determine the factors which affect the incident duration. When an incident occurs at time X minutes. Suppose the authority gets a notification at Y minutes. The time at which the rescuer arrives is Z minutes and the time at which incident road clearance takes place, let's say it is W minutes. The time interval between the occurrence of the incident(X) and notifying the incident (Y), is called reporting time. The time interval between reporting of incident(Y) and the arrival of the rescuer(Z) is called response time. The time interval between the arrival of the rescuer (Z) and the clearance of the road (W) is called clearance time.

Nam, Mannering[1] statistically determined influence factors for reporting time, response time, and clearance time using Hazard Based duration model and obtained the Weibull distribution for reporting and response time and log-logistic distribution for clearance time. However, data in-availability in this study limits the ability to draw temporal conclusions. There is a need to study parametric, non-parametric parameters in case of both imperceptible heterogeneity and hazard function and increase the data length, which will result in non-biased coefficients for factors affecting incident duration.

Chung[2] used an AFT log-logistic model for identifying the factors which can affect the incident duration. However, the report was based on only one year's data and did not account for freeway designs. Ghosh[3] studied and developed a hazard-based model to determine the factors that influence the time needed to clear traffic incidents on the Michigan network and found Generalised F distribution. The study found that clearance times are longer during winter times and shorter during weekends.

Models using statistical methods such as regression were used in various studies to analyze and forecast the incident duration. Khattak [4] used the truncated regression model, constructed the time-sequential incident clearance time prediction models, and used four time-sequential models with a 5-minute interval between two consecutive models and concluded that as time increases, the model accuracy improves. Garib [5] presented a regression model for predicting the incident duration and showed that 81% of variations could be expected by the number of lanes affected, vehicles participated, time, authority response time, and environmental conditions.

Kim[6] implemented a fuzzy incident response model and predicted the incident duration with an average error of 0.3 minutes. However, the model did not consider the significant factors like traffic volume, time and day of the week, and weather. A study by Knibbe [7] reported the feasibility of predicting the incident duration time using the classification tree method on data collected from freeway networks of the Netherlands and found an accuracy of 65%. The CTM approach proved to be flexible and improved by increasing the data size. On the other hand, He[8] used quantile regression because of its ease of handling categorical variables and proved it is crucial to predict spatial-temporal incidents, which considers how congestion traverses over spatial geometry. At the same time, Zhan[9] studied predicted lane clearance time using M5P instead of incident duration. The advantage of using M5P is that it deals with categorical variables with missing values. The author recorded a MAPE of 42.7% and found that the number of blocked lanes, time of day, numbers, and types of vehicles were significant variables that affected lane clearance time.

Various studies have also utilized different machine learning approaches such as ANN, Bayesian networks, SVM, etc. Wang[10] analyzed the characteristics of vehicle breakdown and the relationship between duration, vehicle type, time, and location. This study included two models: fuzzy logic system (FL) and artificial deep neural networks (ANN) used to predict vehicle malfunction time. ANN performed better than FL. However, both fuzzy logic and ANN suffered in forecasting outliers. Valenti[11] found MLR (multiple linear regression) to be the best-suited model for short-term incidents. SVM is best suited for the medium and long incident term, and an Artificial neural network is the only model suited for incidents duration greater than 90 minutes. However, all other models seem to have low accuracy for the more significant incident duration. Pereria [12]showed a framework that used topic modeling (LDA) text analysis to predict the incident duration and got an average error of 9.9 minutes.

Park [13] introduced the Bayesian Network for predicting the duration of the incident which used a hybrid monte Carlo algorithm for updating network parameters and got a MAPE of 29%. It utilized TREPAN, a pedagogical rule extraction algorithm, extracted comprehensible representation using M and N expressions and used it to determine the relationship between variables and a series of decisions that were previously used. Also, in order to determine what percentage of incidents have been underestimated, it has defined the weightage of the miscalculated prediction in evaluating prediction accuracy. Boyes [14] on the other hand developed the probabilistic model that was based on a naive Bayesian Classifier to predict the incident duration. The advantage of using a naive classifier is that it accommodates incomplete information and is robust to outliers.

Hazard-based duration models (HBDM) were used in previous research where they have employed time-sequential procedures along with HBDM[15] or have used Log-logistic AFT model [16,17] or Weibull AFT model[19] or KNN and Log-logistic AFT model[20] in predicting remaining incident duration or incident duration. Qi, Teng [15] developed a time-sequential procedure, dividing the process of the incident management into stages on the basis of the available information along with the different hazard-based duration regression models with different variables for each stage and determined that the accuracy of the incident duration prediction is proportional to the information available for the models and the truncated median-based prediction is more accurate than that the truncated mean based prediction. Traffic duration time is a critical characteristic for the analysis and prediction models. Studies[16,19,20] divulged that the traffic term time fits the log-logistic distribution [16,17,20].

Li et al.[22] used the competing risk mixture model which considers the uncertainty in clearance methods that occurred whose probability is specified by using a multinomial logistic model. This model has better accuracy than the traditional accelerated failure time model with a MAPE of 45% for a duration greater than 15 min. Also, the obtained MAPE is greater than what has been obtained while using the log-logistic model (MAPE: 43.7%[17]) or Weibull AFT model (MAPE: 43%[19]) or using KNN (MAPE: 41.1%[20]).

Zong et al.[23] showed that using goodness-of-fit of Ordered Probit model is higher than that of SVC model in severity modeling where the three severity indicators are taken were the number of injuries and casualties and property damage. Thereby taking the accident severity as an important determinant in the accident duration.
There were also various studies that focused on combined or hybrid models such as the probit model (Ordered) and a rule-based supplemental module[24], CTM and Rule-Based Tree Model (RBTM), and CTM and Rule-Based Tree Model (RBTM). Lin et al.[26] showed that the M5P-HBDM model identified many meaningful and significant variables than either M5P or HBDMs while also having the lowest MAPE error of 36.2% for I-64 and 31.87% for I-190

4. Conclusion and Future Work

4.1 Combining new data sources:

We can obtain the exact location of the incident based on intelligent location [27]. Sdongos[27] used IoT and sensor-based technologies to intelligently determine the exact location of the incident along with many other parameters. Kurchu[29] showed that data from social media can be used for real-time incident duration prediction. Gu[28] studied a method NLP task to get information related to incidents from tweets.

4.2 Improvement in Machine learning techniques:

Lin[30] presented the novel approach by constructing the M5P-HBDM model on the Virginia dataset, in which nodes of the M5P tree algorithm are HBDMs in place of linear regression which offers the advantage of minimizing data heterogeneity. This study can further be improved by involving random parameter HDBM.

Ma[31] used gradient boosting decision tree (GBDT) on the Washington dataset to forecast clearance time as it inherits both statistical and AI approaches to tackle complex relations.

4.3 Outlier prediction:

Valenti [11] found only ANN to be best suited to predict incident duration greater than 90 min out of the implemented five models.

5. References

[1] 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.

[2] 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.

[3] 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

[4] Khattak AJ, Schofer JL, Wang M-H (1995) A simple time-sequential procedure for predicting freeway incident duration. IVHS J 2(2):113–138

[5] 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).

[6] 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.

[7] Knibbe WJJ, Alkim TP, Otten JFW, Aidoo MY (2006) Automated estimation of incident duration on Dutch highways. In: Proceedings of the2006 IEEE intelligent transportation systems conference, Toronto, Canada, pp 870 –874.

[8] He Q, Kamarianakis Y, Jintanakul K, Wynter L (2011) Incident duration prediction with hybrid tree-based quantile regression. IBM research report

[9] Zhan C, Gan A, Hadi M (2011) Prediction of lane clearance time of freeway incidents using the M5P tree algorithm. IEEE Trans Intell Transp Syst 12(4): 1549–1557.

[10] Wang W, Chen H, Bell MC (2005) Vehicle breakdown duration modeling. J Transp Stat 8(1):75 –84.

[11] 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.

[12] 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.

[13] 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.

[14] 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,.

[15] 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

[16] Chung Y (2010) Development of an accident duration prediction model on the Korean freeway systems. Accid Anal Prev 42:282–289.

[17] 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

[18] 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

[19] 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

[20] 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, ITSC 2014, pp 1608–1613. https://doi.org/10.1109/ITSC.2014.6957923.

[21] 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

[22] 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

[23] 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

[24] 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..

[25] 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.

[26] Lin L, Wang Q, Sadek AW (2016) A combined M5P 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.

[27] 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.

[28] 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.

[29] 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.

[30] Lin L, Wang Q, Sadek AW (2016) A combined M5P 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.

[31] 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 .

6. Work Distribution

6.1 Prakhar Pradhan (190618):

- Reviewed various research papers in incident duration prediction on regression models, fuzzy systems, classification tree method, artificial neural networks, Bayesian networks.
- Prepared conclusion and future work
- Formatted Incident Duration prediction section
- Prepared Figure 1

6.2 Aditya Gupta (190060):

- Reviewed various research papers in incident duration analysis on the hazard-based duration regression model, support vector machine, and combined/ hybrid models.
- Prepared and formatted Introduction
- Formatted the Literature View and prepared the final draft

6.3 Chirag Sharma (190248):

- Prepared type of factors and their contributions to traffic incident duration.