MD-17-SHA-UM-4-19

MARYLAND DEPARTMENT OF TRANSPORTATION
STATE HIGHWAY ADMINISTRATION

MARYLAND DEPARTMENT OF TRANPORTATION
STATE HIGHWAY ADMINISTRATION

EXECUTIVE SUMMARY

A KNOWLEDGE-BASED SYSTEM FOR ESTIMATING INCIDENT CLEARANCE DURATION FOR MARYLAND I-95

A CASE STUDY FOR THE PROJECT OF MD-17- SHA/UM/4-19

"DEVELOPMENT OF A TRAFFIC MANAGEMENT DECISION SUPPORT TOOL FOR FREEWAY INCIDENT TRAFFIC MANAGEMENT (FITM) PLAN DEPLOYMENT"

MINSU WON, HYEONMI KIM, AND GANG-LEN CHANG

UNIVERSITY OF MARYLAND
COLLEGE PARK

DECEMBER 2017

Abstract

For the incident response operations to be appreciated by the general public, it is essential that responsible highway agencies be capable of providing the estimated clearance duration of a detected incident at the level sufficiently reliable for motorists to make proper decisions such as selecting a detour route. Depending on the estimated clearance duration, the incident response center can then implement proper strategies to interact with motorists, ranging from providing incident information only to executing mandatory detouring operations. This study presents a knowledge-based system, based on the detailed incident reports collected by the Maryland CHART (Coordinated Highway Action Response Team) program between the years 2012 through 2016, for such needs. The proposed system features its use of interval-based estimates, derived from the knowledge of the historical data with different confidence levels for each estimated incident clearance duration, and its rule-based structure for convenient updates with new data and available expertise from field operators. Since some key variables associated with incident duration often only become available as the clearance operations progress, the developed system with its sequential nature allows the users to dynamically revise the estimated duration when additional data have been reported. Our preliminary evaluation results have shown the promise of the developed system which with its invaluable historical information can circumvent many data quality and availability issues which have long plagued the applicability of some state-of-the-art models on this subject.

Keywords: incident clearance duration, knowledge-based model, association-rule mining, hybrid modeling method

RESEARCH BACKGROUND

To contend with the non-recurrent congestion, caused mainly by various types of incidents, most responsible highway agencies over the past decade have implemented reasonably efficient response plans, ranging from providing the information of incident location to the motorists to deploying full-scale detouring operations. Such practices, however, are not sufficient in the ITS (Intelligent Transportation Systems) era because most en route drivers would like to have more information to make a proper route choice decision. Control center operators, aside from devoting full efforts to clear incidents, also need to decide the most cost-effective management strategy, such as offering travel time information of alternate routes for either advisory or mandatory detour operations. The difficulty in advancing the state of such practices to have those desirable functions mainly lies in the lack of real-time estimated clearance duration (or entire incident duration) for detected incidents, which consequently prevents the responsible agency from estimating the resulting queue lengths, impacts on the ramp flows, surrounding local arterials, and travel times.

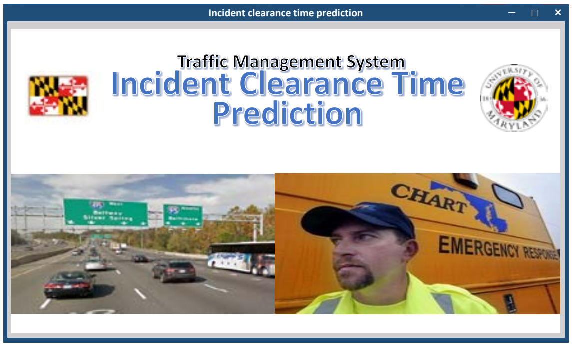
Over the past several decades, transportation researchers have devoted considerable efforts on estimating incident duration with various techniques, including continuous statistical models (*1-16*), discrete/classification methods (*17-22*), neural-network approaches (*23-30*), and hybrid modeling techniques which integrate different methods (*31-37*). Also, some of them developed time-sequential models to account for the fact that incident information at the operations center is acquired over the life of the incident (*4*, *7*, *24*, and *35*). The accuracy of those models is highly dependent on available variables, collected by the field response agencies, and the data quality as well as their statistical properties of available data samples. Hence, depending on the underlying assumptions of such models and their consistency with the distributions of collected incident data, such models mostly suffer from lacking the transferability and the difficulty for use and update by the highway agencies. Thus, despite their contributions to advancing state of the arts on this subject, providing estimated incident duration during non-recurrent congestion or guiding responsive traffic operations remains to be done in most highway networks.

Furthermore, in review of the incident response practices in most state highway agencies *(38)*, one may recognize that a variety of factors and their complex interrelations have limited the field applications of various existing models or tools for estimating incident duration. Examples of those challenges include: 1) a large number of variables (e.g., incident type, lane blockages, available resources) may have significant impacts on the clearance duration of a detected incident; 2) many of those critical factors, such as incident type (i.e., collision with fatality, collision with property damage, or collision with personal injury), responders (i.e., the need of towing, medical, and/or fire units), and environmental conditions, are qualitative in nature and often neither available nor collectable during the incident response and clearance process; 3) available models or tools may not be sufficiently robust to contend with the inevitable missing data or data errors due to human-factor related issues; 4) the quantitative output from some elegant advanced models may not directly correspond to the available management actions during incident response operations; and 5) most commonly and critically, many responsible highway agencies do not have an effective system to collect various critical data associated with incident response operations in real time at the desirable level of accuracy.

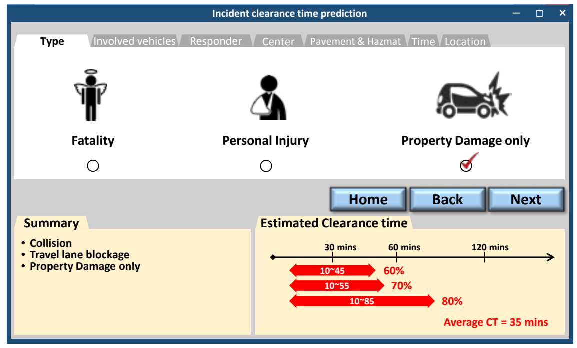
In response to the above concerns and in view of the uniquely rich incident response data available in the Maryland CHART (Coordinated Highway Action Response Team) program over the past 20 years (*39*), this research has developed a knowledge-based system for robustly estimating the incident clearance duration, allowing control center staff to select the most effective ATIS (Advanced Traveler Information System) and response strategy when tackling the non-recurrent congestion. The selection of the knowledge-based models was due to the following concerns: 1) most variables associated with incident duration are mainly collected by the responsible agencies via approximation, not precise measurement; 2) due to the emergency nature of incident response and clearance many vital data are qualitative and/or not collected; 3) the experience and expertise of some senior responders could be very valuable in estimating the required incident duration, especially when facing large missing data; 4) the distribution of key variables in the incident database is often inconsistent with the underlying assumptions of many available statistical models; and 5) the employed method for such needs ought to be transparent and convenient for field operators to understand, evaluate, and update with most recent available information.

The proposed model, designed for field application, has the following distinct features: 1) using the information and knowledge embedded in a large body of prior incident records to derive effective estimation rules which are convenient for update, and also flexible for incorporating with the experience/judgement of field operators or respondents; 2) constructed with a knowledge-based structure that allows efficient enhancement with additional incident cases and rules; and 3) an interval-based output to correspond to available traffic operations and information distribution strategies during the incident response process. A graphical interface of the proposed system is shown in Figure 1. Development procedures for all estimation rules are detailed in the ensuing sections.

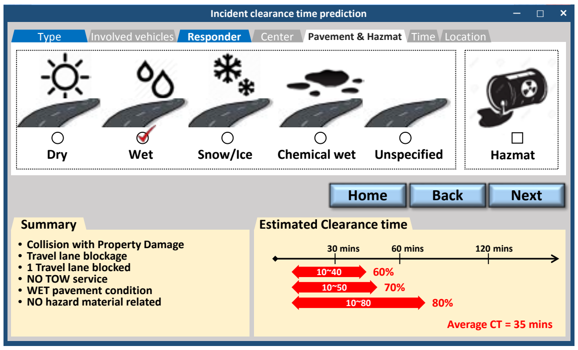
This study defined the incident clearance time (CT) as the time elapsed from arrival of the first response unit at the scene to the complete clearance of the incident. Note that the selection of clearance duration, instead of the entire incident duration, is since most incident response times under the Maryland-CHART highway patrolling strategy are quite stable, varying only in a very indifferent range.



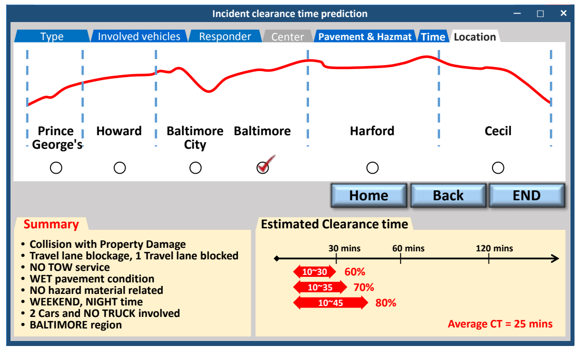
(a) Main cover (intro)



(b) Incident type input and updated CT



(c) Pavement and Hazmat related input and updated CT



(d) Location input and updated CT

FIGURE 1 Examples of the software interface

STUDY SCOPE: I-95 in Maryland

To develop an incident clearance time estimation model, the research first applied the incident dataset of years 2012-2016 from CHART II Database for derivation of estimation rules. The data of the first 4 years (2012~2015) were used for model calibration, and the final year (2016) was used for model evaluation. The data from the first six months (January-June) in 2017 were used for the model update. As a prototype system, its spatial scope covers only MD I-95 from exit 27 to exit 109, and the data include incidents with collision only. Similar models for incidents resulting in only shoulder-lane blockage have also been developed in the final system, but not reported hereafter.

Table 1 shows those variables recorded during each incident event, which are incident type, location, time, lane blockage information, involved vehicle information, environmental conditions, response unit information, etc.

TABLE 1 List of Key Factors Associated with the Response/clearance of a Detected Incident

Category	Variables	Classification
Incident type	Incident type	Collision with Fatality (CF), Collision with Personal Injury (CPI), Collision with Property Damage (CPD)
Time	Hour indicator	AM-peak (7AM~10AM), Day time (10AM~4PM), PM-peak (4PM~7PM), Night time (7PM~7AM)
	Weekend indicator	Weekend, Weekday
	Holiday indicator	Holiday, Non-Holiday
	Season indicator	Spring, Summer, Fall, Winter
Location	County indicator	Prince George, Howard, Baltimore, Baltimore City, Harford, Cecil
	Direction indicator	Northbound, Southbound
	Exit number indicator	Exit 27, …, Exit 109
Environmental condition	Pavement condition indicator	Dry, Wet, Snow/Ice, Chemical wet, Unspecified
	Hazard material related	Yes, No
Operation center	Center indicator	AOC, TOC3, TOC4, TOC5, SOC, Others
Lane blockage	# of blocked lanes*	1, 2, 3, 4…
	# of blocked shoulders	0, 1, 2, 3…
	# of blocked travel lanes**	0, 1, 2, 3…
	# of blocked traffic lanes	0, 1, 2, 3…
	# of blocked auxiliary lanes***	0, 1, 2, 3…
	Travel lane blocked in tunnel	Yes, No
	Travel lane blocked in toll	Yes, No
Involved vehicle	# of total involved vehicles****	1, 2, 3, 4…
	# of involved passenger cars	0, 1, 2, 3…
	# of involved trucks	0, 1, 2, 3…
	# of involved motorcycles	0, 1, 2, 3…
Responder	# of total response units	1, 2, 3, 4…
	# of arrived CHART	0, 1, 2, 3…
	# of arrived Police	0, 1, 2, 3…
	# of arrived Fireboard	0, 1, 2, 3…
	# of arrived Medical service	0, 1, 2, 3…
	# of arrived TOW service	0, 1, 2, 3…
	FIRST responder	CHART, Police, Fireboard, Medical, Tow

* Lanes = Shoulders + Travel lanes
** Travel lanes = Traffic lanes + Auxiliary lanes
*** Auxiliary lane includes on-ramp, off-ramp, acceleration lane, deceleration lane, and collector/distributor lane
**** Vehicle includes passenger car, truck, bus, cyclist, pedestrian, and motorcycle

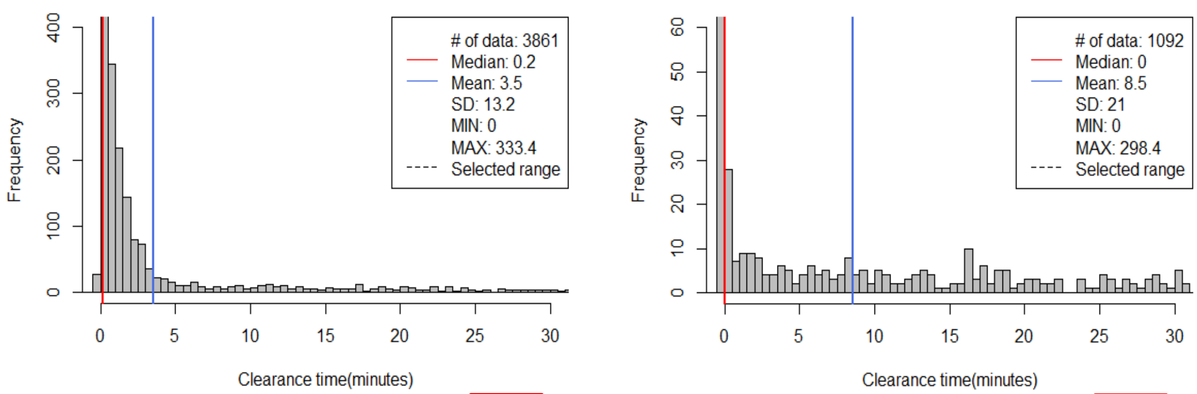
DATA PREPROCESSING

To remove potential data errors by system operators, which often occurred during the real-time incident response and management process, the research team has taken the following steps after consulting with field operators.

Step 1) If the difference between the event-cleared time and the all-blocked-lane-reopened time, including shoulders, is larger than 5 minutes (see Figure 2 (a)), the all-blocked-lane-reopened time is used as the event-cleared time in computing the incident clearance time.

Step 2) If the difference between the all-blocked-lane-reopened time and the all-blocked-travel-lane-reopened time is larger than 32 minutes, the sum of the all-travel-lane-reopened time plus 32 mins is used as the event-cleared time in computing the incident clearance time. The logic behind this step can be seen from Figure 2 (b) where the 90% of the open time of all blocked lanes, including shoulders, for all recorded incidents are within 32 minutes difference from the open time of blocked travel lanes.

Step 3) Since some recorded incident events show an unreasonably short clearance time due to various reasons, the 10% quantile is used as the lower boundary to trim the available data.



Quantile	10%	30%	50%	70%	80%	90%	95%
Time(mins)	0.05	0.08	0.17	0.83	1.63	6.77	20.53

- Incident type: Collision
- Data period: January of Year 2012 ~ December of Year 2015

(a)

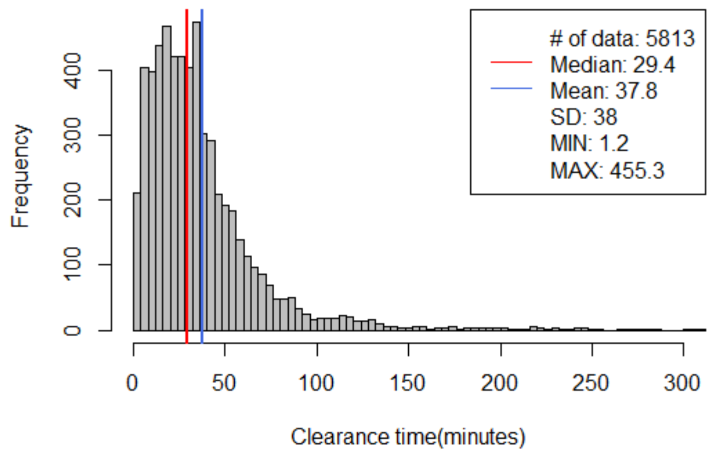
Quantile	10%	30%	50%	70%	80%	90%	95%
Time(mins)	0.00	0.00	0.00	0.67	12.13	32.44	48.07

- Incident type: Collision
- Data period: January of Year 2012 ~ December of Year 2015

(b)

FIGURE 2 (a) Distribution of time differences between the all-lane-reopen times and the event-cleared times, (b) distribution of time differences between the travel-lane-reopen times and the all-lane-reopen times.

The final dataset after the data preprocessing is shown in Figure 3.



CT (mins)	Sample size		
	Original	After step 1-2	After step 3
< 30	3,475 (54%)	3,609 (56%)	2,959 (51%)
30~60	1,943 (30%)	1,990 (31%)	1,990 (34%)
60~120	831 (13%)	682 (11%)	682 (12%)
≥ 120	214 (3%)	182 (3%)	182 (3%)
Total	6,463	6,463	5,813
Mean	*36.8 mins*	*34.1 mins*	*37.8 mins*

Quantile	10%	30%	50%	70%	80%	90%
Time(mins)	7.7	18.4	29.4	41.9	52.5	71.2

- Incident type: Collision
- Data period: Year 2012 ~ Year 2015

FIGURE 3 Distribution of clearance times after the data preprocessing.

MODEL DEVELOPMENT

In view of the large number of qualitative factors, the team first categorized the incident data based on those major factors, and then searched for classification rules to divide the incident data into several categories based on incident clearance time (CT), using the association rule mining method and hybrid modeling method which fit well into the knowledge-based modeling. Association rule mining method is commonly used to analyze data for frequent if/then patterns and to identify the most important relationships *(40-41)*. It has the advantages for users to interpret the resulting relationships and to implement the results in practice. Hybrid modeling method is mainly selected to integrate different methods based on incident duration natures *(38)*. It enables the developed system to provide the estimate in a sequential manner, and to deal with an incident with missing data. Moreover, from each category of incident clearance time, several ranges of estimated clearance duration were selected according to approximately 60%, 70%, and 80% of probability for field operators to make a proper assessment. Note that the selection of confidence thresholds can be specified by the users based on the available incident response resources and the quality of available data recorded during the incident clearance process.

Incident Categorization

Note that key factors associated with the resulting incident clearance time have complex relationships with one another. To sort out such relationships, all available incident data were categorized with the following steps: 1) dividing all incident data into two categories: shoulder-only-blockage and travel-lane-blockage; 2) providing the mean incident clearance time and its estimated ranges with approximately 60%, 70%, and 80% of probability; and 3) further dividing the travel-lane-blockage cases into several groups, based on the incident types and number of blocked travel lanes. Note that no further classification for shoulder-only-blockage cases was taken, because those incidents all lie in a very stable and small range. Additionally, the incident data for Collision with Fatality (CF), resulted in the travel-lane-blockage, was not further classified with the number of blocked lanes due to the small sample size.

Such an initial estimate for each detected incident (see Figure 4) shall be useful for the operators to assess the potential impacts when the information for the full picture of the detected incident is not available in the clearance process yet.

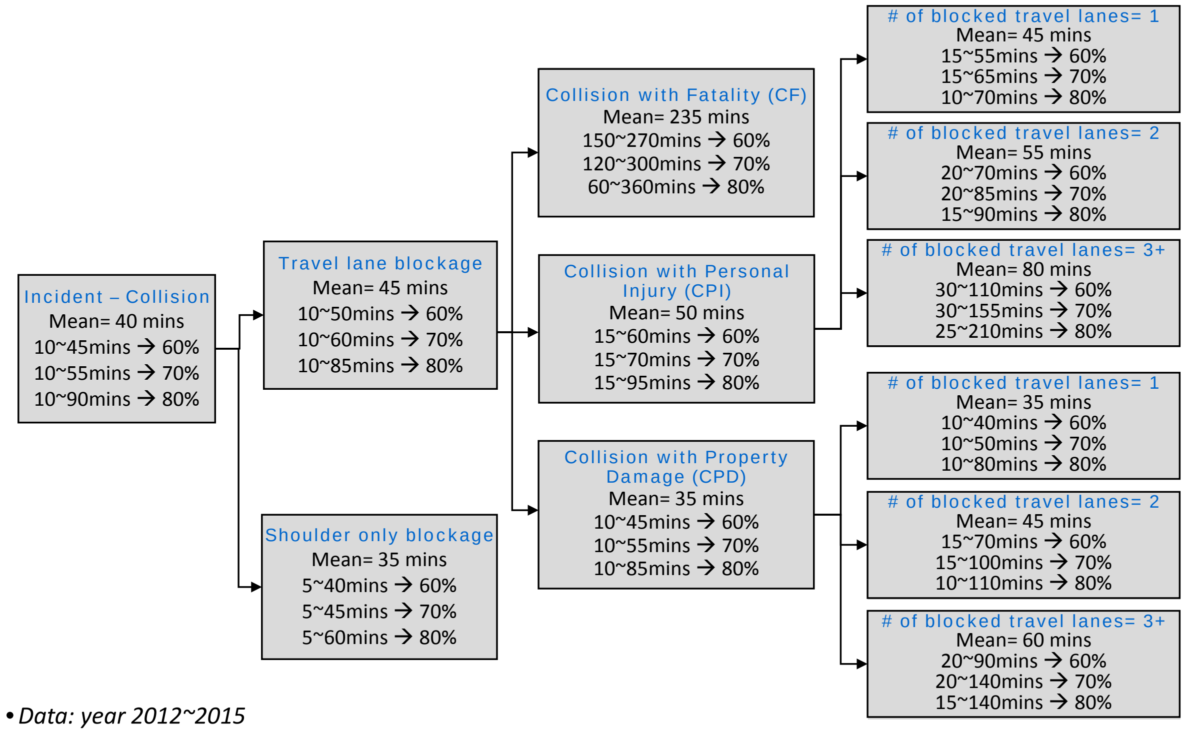


FIGURE 4 Initial incident categorization and estimated clearance duration.

Development of Classification Rules

Based on the incident categorization in Figure 4, the research can provide the initial clearance time information for each incident case. In the subsequent steps, both the association mining method and hybrid modeling method are applied to the last seven categories in the travel-lane-blockage cases to increase the estimation accuracy. Note that the clearance duration for most shoulder-only blockage cases is around the interval of 30 minutes, so the team only considered the travel-lane-blockage cases in exercising these modeling steps for the further classification.

Collision with Personal Injury (CPI) and Collision with Property Damage (CPD) in the Travel Lane Blockage Cases

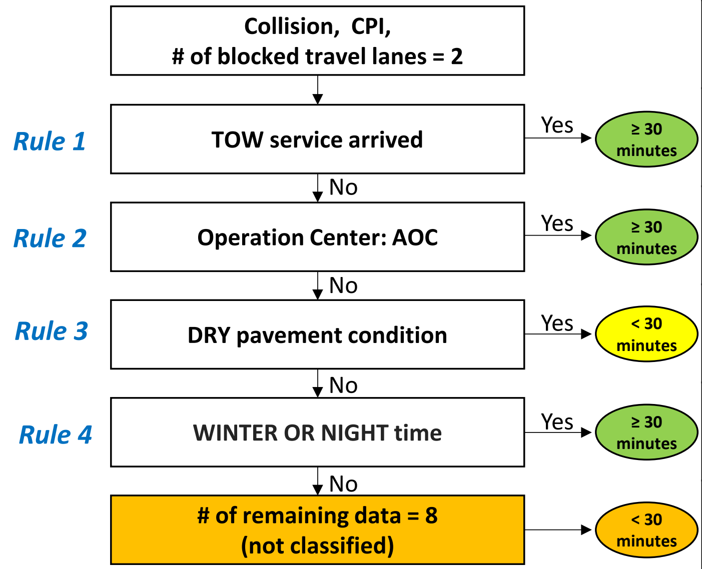
There are six categories of CPI and CPD in the travel-lane-blockage cases, which are divided by the number of blocked travel lanes. To estimate the range of incident clearance time for each category for operation's needs, the team first classified the data into "< 30 minutes" and "≥ 30 minutes" using the association rule mining method. Then, from the "≥ 30 minutes" subset, the team searched other classification rules to classify the data into "< 60 minutes" and "≥ 60 minutes". Similarly, the team found classification rules to further categorize the expected incident clearance times for "< 120 minutes" and "≥ 120 minutes". Finally, based on the distributions of these categories, the team produced three intervals of expected incident clearance time under the 60%, 70%, and 80% confidence levels.

The procedures to model the CPI and CPD in the travel-lane-blockage events to generate classification rules for "< 30 minutes" and "≥ 30 minutes" as follows:

- Step 1) From the dataset, search for classification rules to classify the data into "< 30 minutes" and "≥ 30 minutes" using the association rule mining method.
- Step 2) Select a critical rule with approximately more than 75% confidence level and the highest support level.
- Step 3) Filter out the data associated with the selected rule from the dataset and proceed with further classification for the remaining data.
- Step 4) Stop the classification procedures if no further rule for classification can be derived from the remaining data.
- Step 5) Otherwise, go to Step-1 and repeat the same procedures.

Figure 5 shows an example of CPI with two blocked travel lanes. Using such hybrid rules for classification, the team found four classification rules for "< 30 minutes" and "≥ 30 minutes" and set up the estimated interval for incident clearance time for the remaining data.

• *Training set: year 2012~2015, Test set: year 2016*



Set	Accuracy*	Cumulative Accuracy	Cumulative Coverage
Training	83%(83/100)	83%	46%
Test	80%(24/30)	80%	53%
Training	74%(57/77)	79%	82%
Test	80%(12/15)	80%	79%
Training	79%(22/28)	79%	94%
Test	60%(6/10)	76%	96%
Training	100%(4/4)	79%	96%
Test	-%(0/0)	76%	96%
Training	75%(6/8)	79%	100%
Test	100%(2/2)	77%	100%

* Numbers in parenthesis represent "the number of data that are correctly estimated by the rule in the remaining dataset / the number of data that is categorized by the rule in the remaining dataset"

FIGURE 5 An example of the classification rules for "< 30 mins" and "≥ 30 mins" in the CPI with two blocked travel lanes.

In the next step, the team applied the same classification procedures to the dataset classified in the subset of "≥ 30 minutes" in the previous step to further divide the data into "< 60 minutes" and "≥ 60 minutes". Same procedures were also employed for "< 120 minutes" and "≥ 120 minutes" categories. Figure 6 shows the results of this hybrid classification procedure for six intervals of estimated incident clearance time, such as "< 30 minutes", "≥ 30 minutes", "< 60 minutes", "≥ 60 minutes", "< 120 minutes" and "≥ 120 minutes", as an example of application for CPI with two-travel-lane blockage.

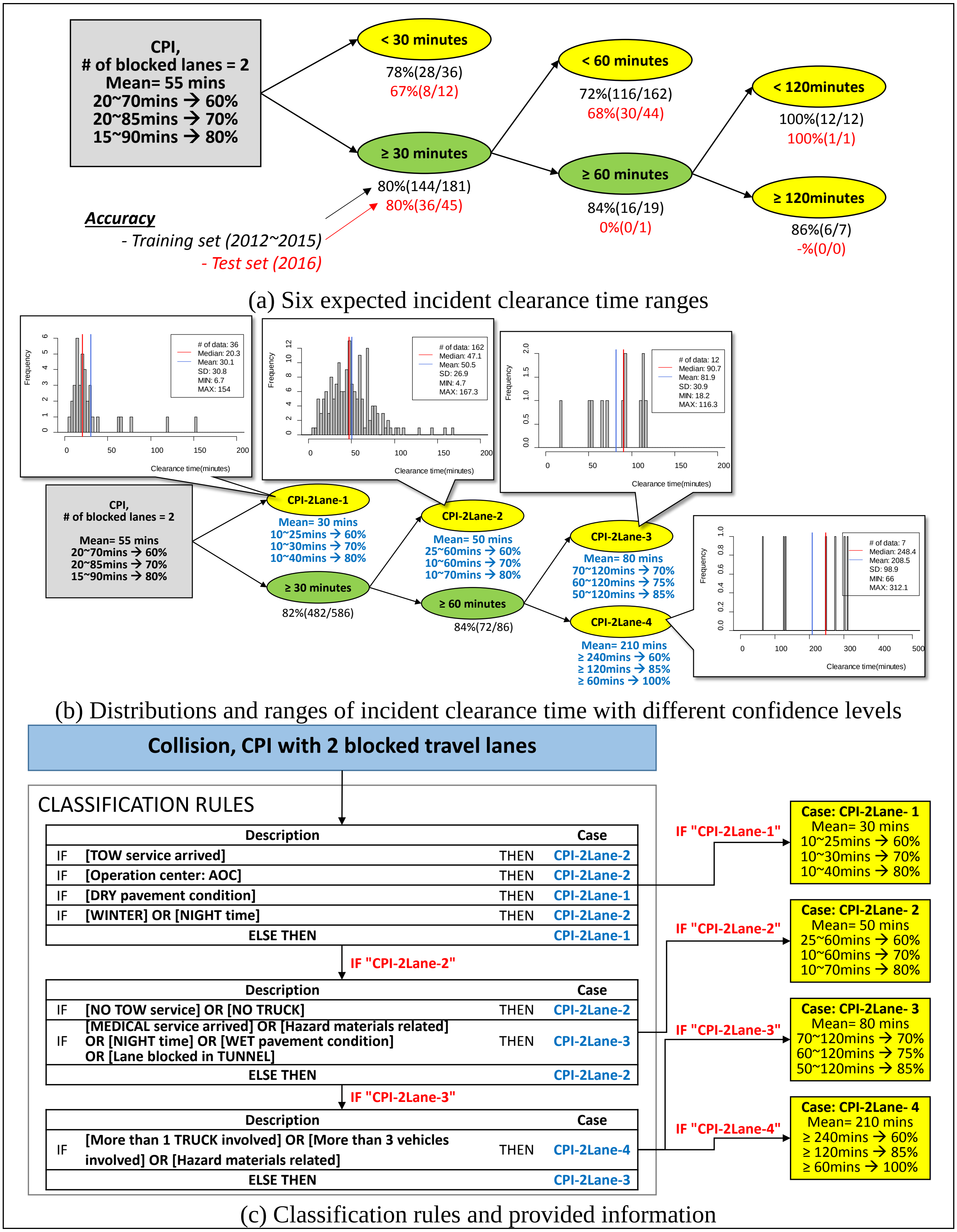


FIGURE 6 An example of application for CPI with two-travel-lane blockage.

Figure 6 (c) shows the rule application process for the CPI with two blocked travel lanes. If an incident with CPI and two blocked travel lanes was reported, this event will go to the rules in the first table in Figure 6 (c). Based on the rules in the first table, this event will be estimated to belong to “CPI-2Lane-1” or “CPI-2Lane-2”. If this event is confirmed from the field report to belong to “CPI-2Lane-1”, the model will then provide the expected incident clearance time information associated with “CPI-2Lane-1”. If not, this event will go to the rules in the second table for further classification. This is the process to provide the expected incident clearance time information sequentially using the proposed system. Note that this hybrid classification model is to be executed on a sequential process, so it can provide an initial estimate in real time, even if full information for a detected incident from the field is not available.

Collision with Fatality (CF) in the Travel Lane Blockage Cases

There are only 25 and 5 events in the training set (2012~2015) and test set (2016), respectively, for this type of incidents. Thus, the modeling procedure for the CPI and CPD in the travel lane blockage cases is not suitable for those CF cases. In addition, most CF cases have incident clearance times longer than 120 minutes, so different boundaries were set for the clearance time ranges for such incidents. Rule-generation procedures used for analyzing CF incidents are summarized below:

- Step 1) Calculate the median value of the incident clearance time of the dataset and divide the dataset into two groups using the median value.
- Step 2) Search the classification rules to classify the data into those two groups using the association rule mining method.
- Step 3) Select a critical rule with approximately more than the 70% confidence level and the highest support level.
- Step 4) Divide the data into two subsets according to the selected rules and set up the estimated incident clearance time ranges, based on the distribution of incidents in each subset.
- Step 5) Stop the procedures if the estimated incident clearance time range is sufficiently robust for use by control center operators.
- Step 6) Otherwise, go to the first step and repeat the procedures for incident data in each subset.

This classification mining method can sequentially classify the dataset into two groups, based on the median value. Thus, it can continuously reduce the distribution and range of the incident clearance times from each subset. Figure 7 shows the classification results and the estimated intervals for incident clearance time. Most estimated ranges for clearance time for collision-fatality incidents have an interval of 60 minutes with approximately 70% of confidence.

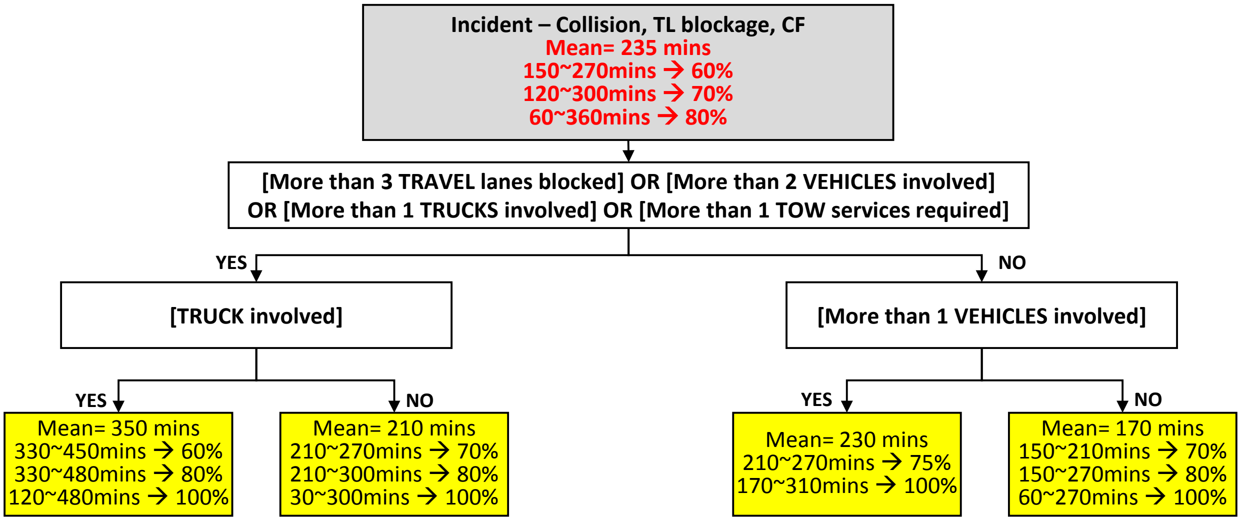


FIGURE 7 Classification rules and incident clearance time ranges in the CF with travel lane blockage.

MODEL EVALUATION AND UPDATES

Model Evaluation

In this study, the team used the data over four years (2012~2015) as the training set for model calibration. Based on the developed model results, the team then evaluated the model using the data in year 2016 as the test set. Table 2 shows the model calibration and validation results. The resulting accuracies for the estimated travel lane blockage incidents in the training set and test set are approximately 80% and 75%, respectively.

TABLE 2 Model Validation Results

	Travel lane (TL) blockage cases							
	CF	CPI			CPD			Total
		1 TL block.	2	3+	1	2	3+	
Training set (2012~2015)	**100%** (21/21)	**79.6%** (399/501)	**82%** (178/217)	**85.9%** (79/92)	**77.2%** (700/907)	**77.8%** (140/180)	**80%** (32/40)	**79.1%** (1549/1958)
Test set (2016)	**83.3%** (5/6)	**77.5%** (131/169)	**71.9%** (41/57)	**62.5%** (15/24)	**74.5%** (222/298)	**70.3%** (45/64)	**75%** (6/8)	**74.3%** (465/626)

* Numbers in parenthesis represent “the number of data whose clearance time is correctly estimated by the model / the total number of data in the category”

Model Updates

Since the new incident data are recorded daily to the database, to enhance the model's performance, the team has updated the developed model with the new dataset from the first six months (January-June) in 2017.

The model update process is focused on those misestimated cases with more than 5 minutes of difference between the actual and the estimated incident clearance times. First, the team searched for a rule to discriminate those misestimated cases. Then, the team applied any newly derived rule as a supplemental rule or to revise the initial model. The difference between adding the supplemental rule and revising the initial model is that a newly derived rule will serve as a supplemental rule if it covers a small portion of cases or the estimated incident clearance time by this selected rule is notably different from the incident clearance time, estimated with the initial model rules. Otherwise, the initial model will be updated with newly derived rules. In addition, to prevent the model from over fit in the update process, the team assessed the impacts of the newly update rules on the early dataset. Table 3 shows the estimation results for the six sequential model updates.

TABLE 3 Model Update Results

Model	**Training set**	**Test set**	**Additional datasets for model updates**					
	2012~2015	**2016**	**Jan. 2017**	**Feb. 2017**	**Mar. 2017**	**Apr. 2017**	**May. 2017**	**Jun. 2017**
Before updates	**79.1%** (1549/1958)	**74.3%** (465/626)	**75.8%** (25/33)	**77.5%** (31/40)	**72.7%** (24/33)	**67.7%** (21/31)	**78.3%** (36/46)	**79.3%** (23/29)
After update with Jan. 2017	**79.2%** (1550/1958)	**74.4%** (466/626)	**87.9%** (29/33)	**85.0%** (34/40)	**72.7%** (24/33)	**67.7%** (21/31)	**78.3%** (36/46)	**79.3%** (23/29)
After update with Feb. 2017.	**79.8%** (1562/1958)	**75.7%** (474/626)	**87.9%** (29/33)	**95.0%** (38/40)	**75.8%** (25/33)	**67.7%** (21/31)	**78.3%** (36/46)	**79.3%** (23/29)
After update with Mar. 2017	**79.8%** (1563/1958)	**76.0%** (476/626)	**87.9%** (29/33)	**95.0%** (38/40)	**100%** (33/33)	**67.7%** (21/31)	**78.3%** (36/46)	**79.3%** (23/29)
After update with Apr. 2017	**80.0%** (1566/1958)	**77.0%** (482/626)	**87.9%** (29/33)	**95%** (38/40)	**100%** (33/33)	**87.1%** (27/31)	**78.3%** (36/46)	**79.3%** (23/29)
After update with May. 2017	**80.1%** (1569/1958)	**77.2%** (483/626)	**87.9%** (29/33)	**95%** (38/40)	**100%** (33/33)	**87.1%** (27/31)	**93.5%** (43/46)	**79.3%** (23/29)
After update with Jun. 2017	**80.2%** (1570/1958)	**77.2%** (483/626)	**87.9%** (29/33)	**95%** (38/40)	**100%** (33/33)	**87.1%** (27/31)	**93.5%** (43/46)	**93.1%** (27/29)

* Numbers in parenthesis represent "the number of data whose clearance time is correctly estimated by the model / the total number of data"

In brief, the developed knowledge-based model can achieve more than 85% accuracy after been updated with new data from six additional months. For example, the model accuracy for February 2017 (77.5%) is increased to 85.0% by the model updated with January 2017 data. The model accuracy for March 2017 (72.7%) is increased to 75.8% by the model after updated with the data of February 2017.

SYSTEM APPLICATIONS

It should be noted that the proposed system has two unique features to account for the uncertainty associated with the nature of incident response and clearance: 1) using intervals with different confidences for estimating the clearance duration and 2) using three thresholds of 30 minutes, 60 minutes, and 120 minutes to proceed with the classification of available incident data. Both key features were actually developed in response to the needs of field operators, and to be in consistent with their selection of potential response plans. For example, the responsible staff in the traffic control/response center are likely to provide only the incident nature and location to the general public if the clearance duration is estimated to be less than 30 minutes. However, more information about the detected incident and resulting traffic conditions, such as travel time information on available alternate routes, would be expected if the estimated clearance duration extends to between 30 and 60 minutes. Also, responsible traffic agencies may have to exercise advisory detouring plans and share all related information to the general public if the clearance duration of the detected incident is expected to prolong to between 60 to 120 minutes. In general, mandatory detour operations may be implemented for severe incidents (e.g. involving fatalities or hazardous material) if it is likely to take more than two hours to recover the traffic conditions.

CONCLUSION

This study has presented a knowledge-based system for estimating the interval of a detected incident's clearance time, using the incident data from Maryland-CHART between Years 2012 and 2015 as the calibration set and Year 2016 as the evaluation set. The proposed system features its use of interval-based estimates, derived from the knowledge of the historical incident response data with different confidence levels for incident clearance time, and the rule-based structure for convenient updates with new data and for effective incorporation of expertise from field operators to revise the initial estimate. Its sequential nature allows the users to dynamically revise the estimate when additional data have been reported, because some key variables, affecting the duration of a detected incident, often only become available as the clearance operations progress.

~~Our~~ The preliminary evaluation results have shown the promise of the proposed system which with its invaluable historical knowledge can circumvent many data quality and availability issues plaguing the applicability of some state-of-the-art models on this subject. The proposed system can also be used to estimate reliable travel time during incidents, given the approximated or observed queue length. Furthermore, the proposed interval-based estimates with three thresholds (i.e., 30, 60, and 120 minutes) offer the information consistent with the potential ATIS/ATMS strategies (e.g., providing incident information only, advisory or mandatory detouring operations) for the responsible highway agencies to interact with en route motorists and the general public. Further research along this line includes: 1) extending the knowledge-based rules and embedded models to neighboring highway systems to evaluate its transferability and 2) developing a reliable and convenient update mechanism for experienced field operators to enhance and enrich the knowledgebase.

ACKNOWLEDGEMENT

This study was supported by Maryland-CHART program. We would like to thank the MDOT State Highway Administration and their staff, Egua Igbinosun, Mohammed Raqib, Jason Dicembre, and Sharon Hawkins, for their support.

REFERENCES

1. Golob, T. F., W. W. Recker, and J. D. Leonard. An Analysis of the Severity and Incident Duration of Truck-involved Freeway Accidents. *Accident Analysis & Prevention,* Vol. 19, No. 5, 1987, pp. 375-395.
2. Giuliano, G. Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway. *Transportation Research Part A: General,* Vol. 23, No. 5, 1989, pp. 387-396.
3. Khattak, A. J., H. M. Al-Deek, and R. W. Hall. Concept of an Advanced Traveler Information System Testbed for the Bay Area: Research Issues. *Journal of Intelligent Transportation Systems,* Vol. 2, No. 1, 1994, pp. 45-71.
4. Khattak,A. J., J. L. Schofer, and M.-h. Wang. A Simple Time Sequential Procedure for Predicting Freeway Incident Duration. *Journal of Intelligent Transportation Systems,* Vol. 2, No. 2, 1994, pp. 113-138.
5. Garib, A., A. E. Radwan, and H. Al-Deek. Estimating Magnitude and Duration of Incident Delays. *Journal of Transportation Engineering,* Vol. 123, No. 6, 1997, pp. 459-466.
6. El-Basyouny, K. and T. Sayed. Comparison of Two Negative Binomial Regression Techniques in Developing Accident Prediction Models. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1950, 2006, pp. 9-16.
7. Qi, Y. and H. Teng. An Information-based Time Sequential Approach to Online Incident Duration Prediction. *Journal of Intelligent Transportation Systems,* Vol. 12, No. 1, 2008, pp. 1-12.
8. Chung,Y. Development of an Accident Duration Prediction Model on the Korean Freeway Systems. *Accident Analysis & Prevention,* Vol. 42, No. 1, 2010, pp. 282-289.
9. Khattak,A., X. Wang, and H. Zhang. Incident Management Integration Tool: Dynamically Predicting Incident Durations, Secondary Incident Occurrence and Incident Delays. *IET Intelligent Transport Systems,* Vol. 6, No. 2, 2012, pp. 204-214.
10. Wang,X., S. Chen, and W. Zheng. Traffic Incident Duration Prediction Based on Partial Least Squares Regression. *Procedia-Social and Behavioral Sciences,* Vol. 96, 2013, pp. 425-432.
11. Junhua,W., C. Haozhe, and Q. Shi. Estimating Freeway Incident Duration Using Accelerated Failure Time Modeling. *Safety science,* Vol. 54, 2013, pp. 43-50.
12. Hojati,A. T., L. Ferreira, S. Washington, and P. Charles. Hazard Based Models for Freeway Traffic Incident Duration. *Accident Analysis & Prevention,* Vol. 52, 2013, pp. 171-181.
13. Li, R. and P. Shang. Incident Duration Modeling Using Flexible Parametric Hazard-based Models. *Computational intelligence and neuroscience,* Vol. 2014, 2014, p. 33.
14. Zou,Y., K. Henrickson, D. Lord, Y. Wang, and K. Xu. Application of Finite Mixture Models for Analysing Freeway Incident Clearance Time. *Transportmetrica A: Transport Science,* Vol. 12, No. 2, 2016, pp. 99-115.
15. Khattak, A. J., J. Liu, B. Wali, X. Li, and M. Ng. Modeling Traffic Incident Duration Using Quantile Regression. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2554, 2016, pp. 139-148.
16. Li,X., A. J. Khattak, and B. Wali. Large-Scale Traffic Incident Duration Analysis: The Role of Multi-agency Response and On-Scene Times, In *Transportation Research Board*

96th Annual Meeting, CD-ROM. Transportation Research Board of the National Academies, Washington, D.C., 2017.

17. Ozbay, K. and P. Kachroo. *Incident Management in Intelligent Transportation Systems*. Artech House Inc., Massachusetts, 1999.
18. Ozbay, K. and N. Noyan. Estimation of Incident Clearance Times Using Bayesian Networks Approach. *Accident Analysis & Prevention*, Vol. 38, No. 3, 2006, pp. 542-555.
19. Boyles,S., D. Fajardo, and S. T. Waller. A Naive Bayesian Classifier for Incident Duration Prediction. , In *Transportation Research Board 86th Annual Meeting*, CD-ROM. Transportation Research Board of the National Academies, Washington, D.C., 2007.
20. Xiaoqiang, Z., L. Ruimin, and Y. Xinxin. Incident Duration Model on Urban Freeways Based on Classification and Regression Tree. in *Intelligent Computation Technology and Automation, 2009. ICICTA'09. Second International Conference on*, 2009.
21. Chang, H.-L. and T.-P. Chang. Prediction of Freeway Incident Duration based on Classification Tree Analysis. *Journal of the Eastern Asia Society for Transportation Studies,* Vol. 10, 2013, pp. 1964-1977.
22. Ma, X., C. Ding, S. Luan, Y. Wang, and Y. Wang. Prioritizing Influential Factors for Freeway Incident Clearance Time Prediction Using the Gradient Boosting Decision Trees Method. *IEEE Transactions on Intelligent Transportation Systems*, 2017.
23. Wang,W., H. Chen, and M. C. Bell. Vehicle Vreakdown Duration Modelling. *Journal of Transportation and Statistics*, Vol. 8, No. 1, 2005, pp. 75-84.
24. Wei, C.-H. and Y. Lee. Sequential Forecast of Incident Duration Using Artificial Neural Network models. *Accident Analysis & Prevention*, Vol. 39, No. 5, 2007, pp. 944-954.
25. Valenti, G., M. Lelli and D. Cucina. A Comparative Study of Models for the Incident Duration Prediction. *European Transport Research Review,* Vol. 2, No. 2, 2010, pp. 103-111.
26. Guan,L., W. Liu, X. Yin, and L. Zhang. Traffic Incident Duration Prediction Based on Artificial Neural Network. in *Intelligent Computation Technology and Automation (ICICTA), 2010 International Conference on*, 2010.
27. Wu, W.-w., S.-y. Chen, and C.-j. Zheng. Traffic Incident Duration Prediction Based on Support Vector Regression. in *ICCTP 2011: Towards Sustainable Transportation Systems,* 2011, pp. 2412-2421.
28. Vlahogianni, E. I. and M. G. Karlaftis. Fuzzy Entropy Neural Network Freeway Incident Duration Modeling with Single and Competing Uncertainties. *Computer Aided Civil and Infrastructure Engineering*, Vol. 28, No. 6, 2013, pp. 420-433.
29. Park,H., A. Haghani and X. Zhang. Interpretation of Bayesian Neural Networks for Predicting the Duration of Detected Incidents. *Journal of Intelligent Transportation Systems,* Vol. 20, No. 4, 2016, pp. 385-400.
30. Park,H. J., S. Gao and A. Haghani. Sequential Interpretation and Prediction of Secondary Incident Probability in Real Time. In *Transportation Research Board 96th Annual Meeting*, CD-ROM. Transportation Research Board of the National Academies, Washington, D.C., 2017.
31. Lin, P.-W., N. Zou, and G.-L. Chang. 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, DC,* 2004.

32. Kim, W. and Chang, G.L., 2012. Development of a hybrid prediction model for freeway incident duration: a case study in Maryland. International journal of intelligent transportation systems research, 10(1), pp.22-33.
33. Ji,Y. Prediction of Freeway Incident Duration Based on the Multi-model Fusion Algorithm. In 2011 International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE), 2011.
34. He, Q., Y. Kamarianakis, K. Jintanakul, and L. Wynter. Incident Duration Prediction with Hybrid Tree-based Quantile Regression. In *Advances in Dynamic Network Modeling in Complex Transportation Systems*, Springer, 2013, pp. 287-305.
35. Li,R., F. C. Pereira, and M. E. Ben-Akiva. Competing Risk Mixture Model and Text Analysis for Sequential Incident Duration Prediction. *Transportation Research Part C: Emerging Technologies*, Vol. 54, 2015, pp. 74-85.
36. Zhu,W., Z. Zeng, Y. Wang, and Z. Pu. Predicting Incident Duration based on Spatiotemporal Heterogeneous Pattern Recognition, In *Transportation Research Board 96th Annual Meeting*, CD-ROM. Transportation Research Board of the National Academies, Washington, D.C., 2017.
37. Pettet,G., S. Nannapaneni, B. Stadnick, A. Dubey and G. Biswas. Incident Analysis and Prediction Using Clustering And Bayesian Network. In *2017 IEEE International Conference on Smart City Innovations*, 2017.
38. Kim, W. Development of a Traffic Incident Management System for Contending with Non-recurrent Highway Congestion. In *Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR)*, 2014.
39. Chang, G.-L. and S. Rochon. *Performance evaluation and benefit analysis for CHART*. Coordinated Highways Action Response Team, Maryland Department of Transportation, 2016.
40. Agrawal, R., T. Imieliński, and A. Swami. Mining Association Rules between Sets of Items in Large Databases. In *Acm sigmod record*, 1993.
41. Hahsler, M., B. Grün, and K. Hornik. A Computational Environment for Mining Association Rules and Frequent Item Sets. 2005.