TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
1–12

DOI: 10.1177/03611981251320398
journals.sagepub.com/home/trr
Sage

Incident Duration Sequential Predictions Using Reinforcement Learning for Advanced Traveler Information System Applications

Ray (Zirui) Huang[1], Xiaofeng Li[2], Yi-Chang Chiu[3], and Yao-Jan Wu[4]

Abstract

Accurate incident duration prediction is crucial for the effectiveness of advanced traveler information systems. Sequential prediction is the process by which an incident duration is predicted at its inception, and potential follow-up predictions or revisions to previous predictions are made thereafter. However, one scenario that requires sequential predictions but has rarely been discussed is when the incident remains unclear despite the estimated incident duration time that has elapsed. To address this issue, we proposed a methodology for training a reinforcement learning (RL) agent to produce sequential predictions, and apply it to Houston TranStar incident data. Results indicate that compared to a one-time prediction model, the proposed RL-based method produces more accurate follow-up predictions, and compared to other models in sequential prediction, such as the artificial neural network and random forest, the RL agent achieves lower mean absolute error by 1.5 min. In addition, we have found that in 63.6% of incidents, the agent only needs to make a single prediction, indicating that the agent primarily serves as a one-time prediction method. However, in cases where a one-time prediction incurs a significant error, the agent can generate additional predictions to address the limitations of the one-time prediction approach. The proposed methodology enables traffic operators to provide travelers with updated incident duration information, empowering both operators and users to make informed decisions.

Keywords

incident duration, reinforcement learning, sequential prediction, traffic safety, traffic incident management

An efficient road system is essential for the economic development and expansion of a society, but congestion caused by incidents is a significant impediment; thus, developing an intelligent traffic incident management (TIM) system to lessen the negative effects of incidents on travelers is imperative. The TIM system's main goal is to coordinate various agencies, including transportation agencies, law enforcement, fire and rescue, and towing and recovery, to clear incidents quickly. This is crucial since, on average, a 1-min traffic lane closure results in a 4-min delay for each incoming vehicle and a 2.8% increase in the likelihood of a secondary incident (*1*). The TIM system also collects incident-related information that traffic operators can communicate to travelers before they depart or when they are close to an incident location. In 1995, a telephone information service was launched, serving northern Kentucky and Cincinnati, providing both en route and pre-trip information. A survey conducted after the service's launch revealed that more than 99% of those surveyed thought the service helped them avoid traffic, save time, and

[1]Cambridge Systematics, Inc., New York, NY
[2]School of Travel Industry Management, Shidler College of Business, University of Hawaii at Manoa, Honolulu, HI
[3]Metropia, Inc., Houston, TX
[4]Department of Civil and Architectural Engineering and Mechanics, The University of Arizona, Tucson, AZ

Corresponding Author:
Xiaofeng Li, xiaofeng.li@hawaii.edu

arrive at their destinations on time (*2*). With the rapid development of communication technology, the channels through which traffic operators and travelers communicate with each other have significantly expanded, from traditional text messages, radio, and dynamic message signs (DMSs) to mobile phones and on-board navigation systems. In recent decades, developing advanced traveler information systems (ATISs) that disseminate information to travelers has garnered much attention.

In an incident situation, the incident location, estimated duration, and anticipated delay are the notifications that travelers expect to receive the most (*3*). While incident location retrieval is relatively straightforward for the TIM system, estimating incident duration and delay is more complex and requires prediction. Incident delay models have shown that the duration of an incident can account for up to 85% of the variation in incident delay, which is influenced by the number of affected lanes, the number of vehicles involved, and traffic demand before the incident (*4*); Khattak et al. (*5*) have attempted to predict incident delays using the estimated incident duration and deterministic queuing model; Javid and Javid (*6*) have suggested a travel time variability paradigm based on incident clearance time. Therefore, accurately forecasting incident duration is essential in helping travelers make informed decisions about the best routes and times to depart and reduce the negative impacts of incidents on traffic flow.

The research on predicting incident duration can be classified into two categories: one-time prediction and sequential prediction. One-time prediction involves predicting the incident duration only at the start of an incident without any further revisions to the initial prediction. In contrast, sequential prediction refers to predicting the incident duration at the beginning of an incident and making subsequent predictions or revisions as the incident progresses. Historically, the majority of incident duration prediction techniques have been of the one-time prediction type, mainly because of the limitations of the TIM system, which has lacked the capacity to monitor incidents and track their current status. Even if the TIM system had the capability to monitor incidents, it did not possess the communication channels necessary to promptly convey this information to the public or engage with travelers regularly. However, recent advancements in Traffic Management Centers (TMCs) have led to the refinement of incident response standard operating procedures (SOPs) and the assignment of traffic operators to document the live status of incidents. In addition, some transportation agencies have launched websites and mobile apps to disseminate their most recent updates and interact with travelers. These enhancements have generated impetus for the formulation of sequential prediction techniques, which are instrumental in providing travelers with precise information for making informed decisions.

Various incident duration prediction models are used for one-time prediction. Hazard-based models, originally developed to assess patient survival in medical studies, have been applied to incident duration prediction, with clearance as the primary event of interest (*7–9*). Regression models, which describe the relationship between independent variables (incident characteristics) and a response (incident duration or percentiles of incident duration), have been used for incident duration prediction (*10–12*). With the popularity of artificial intelligence (AI) techniques, various AI models, such as the artificial neural network (ANN) (*13*) long short-term model (LSTM) (*14*), and support vector machine (*15*), have been developed to predict incident duration. However, because of differences in the incident records used and their varying data quality, the performance of these methods, with respect to mean absolute percentage error (MAPE), varies between 20% and 50% (*16*). In the sequential prediction group, Wei and Lee (*13*) created two adaptive ANN models, one of which is used to predict the duration time at the time of incident notification and the other to predict information being updated after incident notification; Qi and Teng (*17*) divided an incident management process into multiple stages based on the availability of specific information and created models for each stage; Li et al. (*18*) developed a mixture model that takes into account the possibility of the independent variables and discovered that the suggested model outperforms the non-mixture model; Pereira et al. (*19*) predicted the incident duration by applied topic modeling techniques on real-time incident reports as the contents are updated.

However, these previous studies have not investigated the application of sequential prediction when an incident has not been cleared despite the estimated incident duration time that has elapsed. For example, if an incident is predicted to be cleared in 20 min, but 20 min later, the incident remains uncleared, then the initial prediction is obviously underestimated, and a further prediction is necessary. Nowadays, many TMCs dedicate resources and staff to collect and update incident information in real-time, allowing operators to monitor incident clearance progress. This real-time information can provide continuous inputs for using sequential prediction in such cases. In this study, we proposed a sequential prediction framework that allows the TIM system to generate new predictions as soon as the results from the previous prediction are proven underestimated. Specifically, this study made the following contributions to the body of knowledge in this research domain.

- Sequential prediction is developed and applied in incident duration when an incident has not been cleared despite the elapsed time specified in the incident duration estimation, which has rarely been explored in the existing literature or practice.

- We have proposed a methodology to train a reinforcement learning (RL) agent to make sequential predictions by trial and error.
- As a case study, we have implemented our framework using the Houston TranStar incident data and provided incident duration estimations sequentially. Our framework enables the TIM system to quickly adapt to incorrect prior predictions and leverage real-time incident updates.

The remainder of this paper is structured as follows. The *Data and Study Location* section introduces data that were utilized in this study, which are collected from the Houston TranStar incident management system. The *Methodology* section describes the framework development process and the training of the RL agent for providing sequential predictions. The *Implementation and Results* section presents the experimental setup and highlights the key findings and performances of the case study. Finally, the *Conclusions* section concludes the paper and discusses potential directions for future research.

Data and Study Location

The incident dataset is collected from Houston TranStar, the transportation and emergency management center for Greater Houston. In the Houston area, the center's operators continuously monitor roadway incidents and log the live status of each incident. These incident records provide rich information that can be used to identify patterns and trends in incident clearance times. An example of the type of information that can be collected in incident records is shown in Table 1, and all information is categorized into four groups as follows.

- **Incident characteristics**
  This category describes the type of incident, the number of vehicles involved, the number of blocked lanes, and other related information.
- **Spatial characteristics**
  These characteristics show where the incident occurred, which can be used to map each incident to the specific road where it happened.
- **Temporal characteristics**
  These characteristics track the time when the incident is detected and cleared, and the duration of the incident can be calculated by taking the difference between these two timestamps.
- **Other characteristics**
  This category includes information gathered from external sources, such as weather, precipitation, and other relevant factors that could affect incident response.

We gathered incident records from the year 2022 and removed the data with duration of less than 1 min or more than 6 h, resulting in 9378 incidents for this case study. Figure 1, *a* and *b*, shows the durations of these incidents on a linear and log scale, respectively. The log-normal distribution of incident durations can be inferred from the curve in Figure 1*b*, which is consistent with observations in other literature (*13*, *16*). Figure 2 displays a heatmap of incident locations, revealing a higher concentration of incidents on I-10 East, I-69 East, and I-45 North, among other locations, across the greater Houston area highway and freeway network in 2022.

Methodology

In the event that the prior predictions become underestimated, meaning that the incident has not been cleared despite the predicted duration having elapsed—for instance, the duration prediction was 30 min, but the incident had not been cleared after 30 min—it is imperative to generate new forecasts and apprise travelers of the prevailing situation. In our study, we proposed a methodology that leverages RL to generate sequential predictions for incident duration. RL has demonstrated remarkable performance in acquiring action sequences (*20*), making it a promising approach for incident duration sequential prediction. In the subsequent two sections, we provide a detailed explanation of the fundamental concepts of RL and the deep Q-network (DQN) algorithm. We also present the specific settings we used in this study to implement our proposed methodology.

Reinforcement Learning

Figure 3 illustrates the typical framing of an RL scenario. The process can be summarized as follows: the agent takes an action, and the environment responds by transitioning to a subsequent state representation at the next time step, and providing a corresponding reward. The primary goal of RL is to develop a policy that maximizes the cumulative rewards over time. In this study, we have adopted an innovative approach to address the issue of sequential incident duration prediction. Specifically, we have formulated the problem as a game in which the RL model is trained to produce a new incident duration in case the previous one underestimates the actual time required for incident resolution. An underestimation occurs when the predicted duration for an incident is shorter than the time required to resolve it, as indicated by the real-time TIM system. For instance, if the prior predicted duration is 30 min, but after 30 min, the

Table 1. Attributes of Incident Records from Houston TranStar

Attribute	Definition	Data type	Example
Incident characteristics			
ID	A unique identifier	Integer	914232
VEHICLES_INVOLVED	# Vehicles involved	Integer	2
HAZMAT_SPILL	If there is a hazardous materials spill	Boolean	1
HEAVY_TRUCK	If there is a heavy truck involved	Boolean	0
HIGH_WATER	If there is a flood	Boolean	0
INC_ON_ROADWAY	If there is ice on the road	Boolean	0
LOST_LOAD	If a vehicle's load is lost	Boolean	0
ROAD_DEBRIS	If there is road debris	Boolean	0
STALL	If there is a stalled vehicle	Boolean	0
VEHICLE_FIRE	If the vehicle is on fire	Boolean	0
BUS	If there is a bus involved	Boolean	0
CRASH	If two or more vehicles collide	Boolean	1
CONSTRUCTION	If the incident occurs in a construction zone	Boolean	0
OTHER	Not belonging to any type above	Boolean	0
MAINLANES_BLOCKED	# Main lanes blocked	Integer	2
FRONTAGE_LANES_BLOCKED	# Frontage lanes blocked	Integer	0
RAMP_LANES_BLOCKED	# Ramp lanes blocked	Integer	0
HOV_LANES_BLOCKED	# HOV lanes blocked	Integer	0
SHOULDER_LANES_BLOCKED	# Shoulder lanes blocked	Integer	1
Spatial characteristics			
LATITUDE	Incident location latitude	Float	29.77810
LONGITUDE	Incident location longitude	Float	−95.28400
ROADWAY_NAME	Name of the road where the incident occurred	String	"IH-10 East"
ROADWAY_TYPE	Type of road where the incident occurred	String	"Freeway"
DIRECTION	The direction of the road where the incident occurred	String	"Eastbound"
NUMBER_LANES	# Lanes of the road where the incident occurred	Integer	5
CROSS_STREET_NAME	Name of the cross street near the incident location where the incident occurred	String	"MC CARTY ST/US-90 ALTERNATE"
Temporal characteristics			
DETECTION_TIME	Time at which the incident was detected	Timestamp	2022-01-0101:21:52
CLEARED_TIME	Time at which the incident was cleared	Timestamp	2022-01-0101:49:09
Other characteristics			
TEMPERATURE	Temperature in °F when the incident occurred	Integer	76
CONDITION	Weather condition when the incident occurred	String	Light rain
PRECIPITATION	Precipitation near the incident location when the incident occurred	Float	0.0

Note: HOV = High-Occupancy Vehicle.

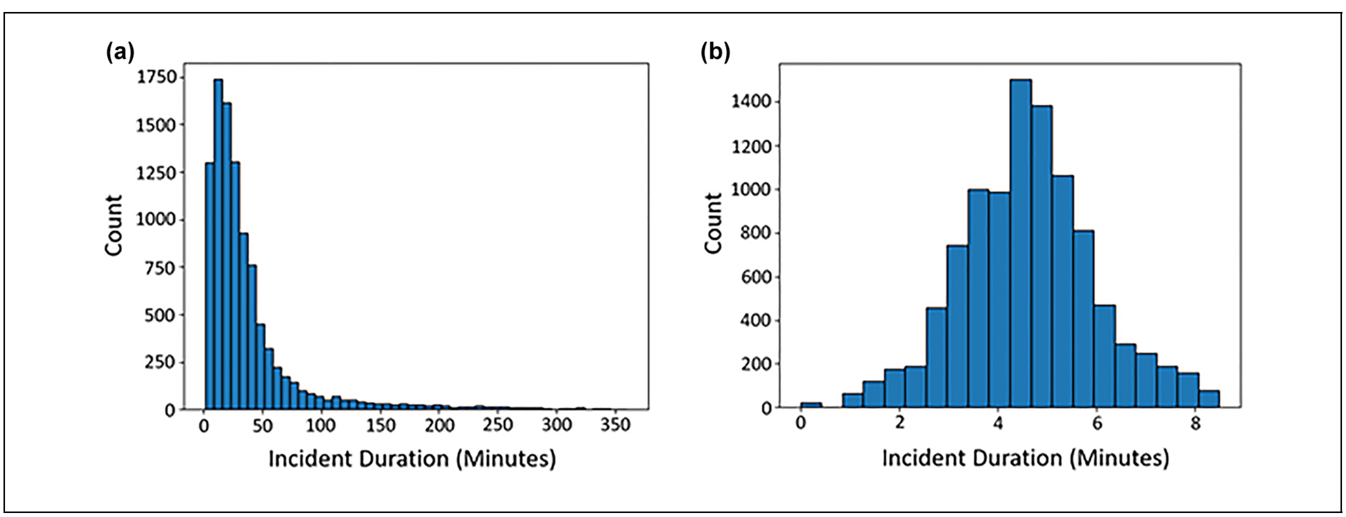

Figure 1. Histograms of incident durations in (*a*) a linear scaled *x*-axis and (*b*) a log scaled *x*-axis.

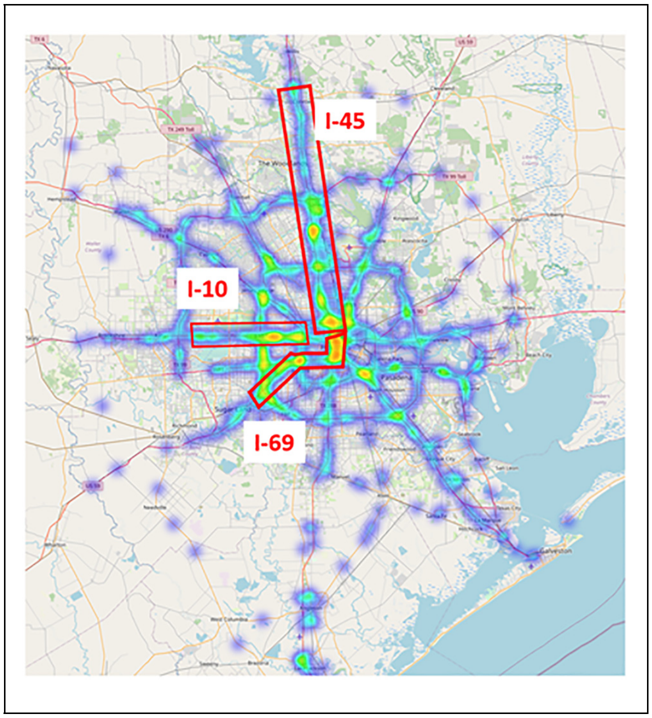

Figure 2. Heatmap of incident locations.

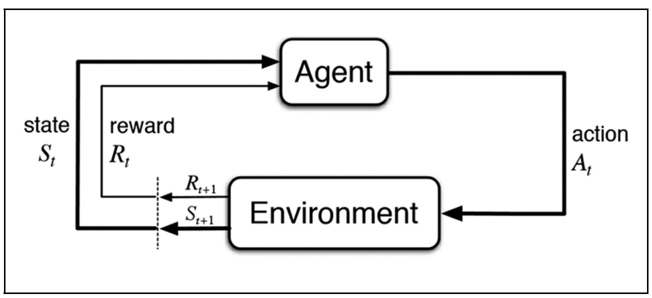

Figure 3. The typical framing of a reinforcement learning scenario.

incident is still not cleared yet according to the real-time TIM system information, then the previous prediction underestimates. In such cases, we revise the previous prediction by increasing its estimated duration to better reflect the actual time required for resolution. The fundamental components of this RL problem can be interpreted as follows.

Agent. The RL model is the agent that is trained to predict the duration of the incident and take actions based on the information acquired from the environment.

Environment. The TIM system acts as the environment in which the RL model operates, providing the agent with real-time incident status updates, such as its location, severity, type, and clearance status, among other pertinent details. Moreover, the system offers feedback to the agent through rewards or penalties, depending on the accuracy of the agent's predictions and actions.

State. State space S consists of multiple attributes that affect the agent's action selection. These attributes include the following: (1) incident characteristics $(x_1, x_2, \cdots, x_n)$ that have significant impacts on incident duration; and (2) the elapsed time since the incident occurred p. The value of p is initially set to 0 and subsequently updated by the agent as the incident resolution process progresses. Specifically, a state can be represented as $s = (x_1, x_2, \cdots, x_n, p)$. In this study, the Cox proportional-hazards (CPH) model was utilized to analyze the Houston TranStar dataset and identify factors that significantly affect incident duration. The model identified 12 characteristics that have a statistically significant impact on the incident duration, with a p-value of less than 0.005, resulting in a state vector s of length 13. These significant characteristics are summarized in Table 2.

Reward. The reward function serves as an incentive mechanism that utilizes rewards and punishments to communicate with the agent and encourage it to take actions that will help achieve the goal. In this study, the goal is to maximize the total rewards while making a sequence of incident duration predictions. Each incident in the Houston TranStar dataset has a recorded duration d, which we consider as the ground truth. At each time step t, the state of the agent is represented by $s^t = (x_1^t, x_2^t, \cdots, x_n^t, p^t)$. After taking the action a^t, the state changes to $s^{t+1} = (x_1^{t+1}, x_2^{t+1}, \cdots, x_n^{t+1}, p^{t+1})$. To incentivize the agent to make accurate predictions, we designed a reward function that rewards the agent when its prediction is within 10 min of the ground truth duration d and penalizes it otherwise. Specifically, if the predicted duration p^{t+1} is within 10 min of the ground truth duration d, the agent receives a reward of 100. If p^{t+1} deviates from the ground truth d by more than 10 min, the agent receives a penalty of $-|d - p^{t+1}|$, as indicated by Equation 1:

$$r^{t+1} = \begin{cases} 100 \text{ if } |d - p^{t+1}| \leq 10 \\ -|d - p^{t+1}| \text{ otherwise} \end{cases} \quad (1)$$

Action. The RL agent revises the incident duration through the agent's actions, denoted by a. In this RL problem, all actions involve adding more minutes to the previous prediction since the agent is only making another prediction when the previous one

Table 2. Cox Proportional-Hazards Model Regression Results

Covariate	Coefficient	Hazard ratio	*P*-value (significance)
HAZMAT_SPILL	−0.84	0.43	<0.005 (***)
HEAVY_TRUCK	−0.48	0.62	<0.005 (***)
HIGH_WATER	−0.77	0.46	<0.005 (***)
ROAD_DEBRIS	0.34	1.40	<0.005 (***)
VEHICLE_FIRE	−0.53	0.59	<0.005 (***)
CRASH	0.39	1.47	<0.005 (***)
VEHICLES_INVOLVED	−0.05	0.95	<0.005 (***)
MAINLANES_BLOCKED	−0.02	0.98	<0.005 (***)
FRONTAGE_LANES_INVOLVED	−0.17	0.84	<0.005 (***)
RAMP_LANES_BLOCKED	−0.12	0.88	<0.005 (***)
HOV_LANES_BLOCKED	−0.78	0.46	<0.005 (***)
SHOULDER_LANES_BLOCKED	−0.17	0.84	<0.005 (***)

Note: HOV = High-Occupancy Vehicle.
***indicates that the variable is highly statistically significant.

underestimates. The difference between actions is the number of minutes added to the previous prediction. By designing the actions in this way, there is no upper bound on the estimated duration, allowing the model to handle extreme cases where the incident takes longer to clear than any previous incident, ensuring that the model provides predictions in all scenarios.

RL agents can have either a discrete or a continuous action space. The former involves choosing a distinct action to carry out from a finite action space, which provides computational efficiency and implementation simplicity advantages. By using a discrete action space, the agent can provide a range of possible durations rather than attempting to predict an exact duration, since incident duration is influenced by several unpredictable factors, making precise predictions difficult to make. Using a discrete action space allows for more flexibility in the agent's decision-making process.

In the preliminary analysis of incident duration, more than 95% of incidents have a duration under 160 min. Therefore, we created a discrete action space with 32 actions, ranging from adding 5 min to the previous prediction to adding 160 min to the previous prediction, as listed below:

- a_0: increase the current prediction by 5 min;
- a_1: increase the current prediction by 10 min;
- …
- a_{30}: increase the current prediction by 155 min;
- a_{31}: increase the current prediction by 160 min.

Policy. The strategy of deciding the next action is based on the current state, essentially a mapping from the agent's states to actions.

Time Steps. It's important to clarify that time steps in this study correspond to the moments when the agent is required to make a new prediction. Unlike in typical gaming scenarios where the agent must act at every time step, in our problem, the agent only performs actions when the prior prediction is shown to be an underestimate. We define the discrete time steps as follows.

- $t_0 = 0$: at the beginning of the incident, the first prediction must be made.
- t_i: the moments when the prior prediction's time has elapsed, but the incident has not been cleared yet, and the prior prediction is proven to be an underestimate. For instance, an example of our constructed RL scenario, as illustrated in Figure 4, involves making a sequence of three predictions.

1. Initially, the RL agent receives the incident state S_0 including incident characteristics and the elapsed time since occurrence (i.e., 0) and takes action A_0 to predict the incident duration as t_0.
2. At time t_0, if the incident is still ongoing, the agent clearly underestimates and receives a penalty for the inaccurate prediction. The agent then updates the current state to S_1, which includes incident characteristics at that point in time and the elapsed time since occurrence (i.e., t_0), and takes action A_1 to revise the duration to t_1.
3. At time t_1, the incident is still not cleared, the agent receives another penalty and revises the duration prediction to t_2.
4. Finally, at time t_2, when the incident is cleared, the agent receives a reward for making an accurate prediction, and the loop ends.

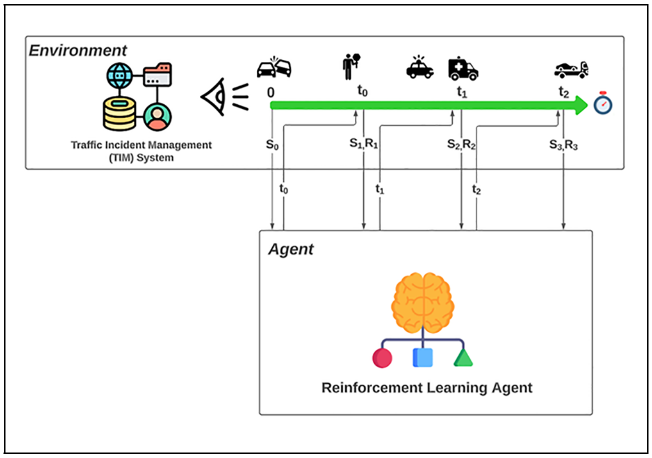

Figure 4. The framing of a reinforcement learning scenario for incident duration sequential prediction.

Q-Learning and Deep Q-Networks

The action value function $Q(s, a)$ is a crucial concept in RL as it represents the expected reward received when taking action a from the state s. The action value function is used to determine the optimal action to be taken in different states, and is thus a fundamental component of the RL framework.

In vanilla Q-learning, Q values are stored in a table with states as rows and actions as columns. The agent retrieves the associated Q values for a given state–action pair by looking up the table. However, the table-based approach is inefficient for large state and action spaces. For instance, a table with 1000 states and 1000 actions would require a million cells. To overcome this challenge, we approximate Q values using machine learning models such as neural networks. This is the foundation of DQNs, which can handle discrete action spaces. However, they face challenges when dealing with a continuous action space. The deep deterministic policy gradient (DDPG) addresses this issue by employing an actor–critic architecture. In this study, we opted to use DQNs and discretized the action space. This decision was motivated by several factors. Firstly, discrete action spaces are often simpler and easier to implement, especially when the number of actions is small. This can lead to faster training and better convergence to optimal policies. Secondly, discrete action spaces are computationally more efficient than continuous action spaces, allowing for better optimization. Overall, using a discrete action space has several advantages with respect to computational efficiency, interpretability, and reward function design, making it a more practical and effective approach for many RL methods.

Figure 5 depicts the action–reward feedback loop for our problem, which was solved using DQNs to train the RL agent. DQNs utilize neural networks to approximate Q values, where the state is the input and the Q values for all possible actions are the outputs. The neural network in our DQN model is configured as follows based on the dimensions of the state and action space:

- the input dimension is 13, obtained by concatenating 12-dimensional incident characteristics and one-dimensional time elapsed since an incident occurs;
- the output dimension is 32, representing each action in the action space, from adding 5 min on prior prediction to adding 160 min on prior prediction;
- the neural network includes a single hidden fully connected layer with 16 neurons and a rectified linear unit (ReLU) activation function between the input and output layers.

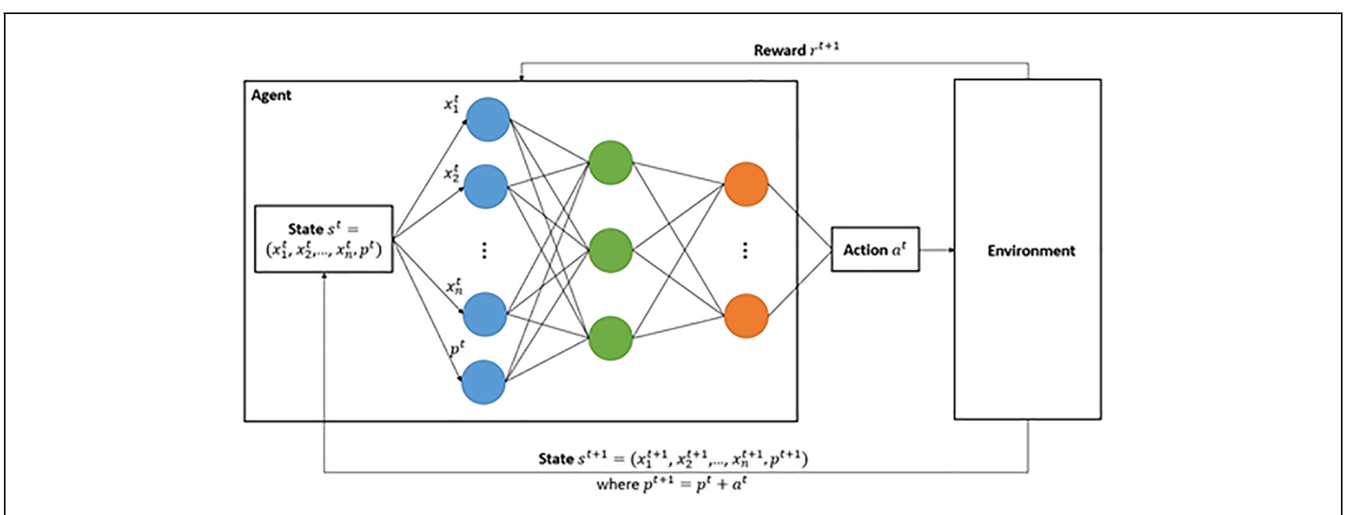

Figure 5. Loop of reinforcement learning interactions in incident duration sequential predictions.

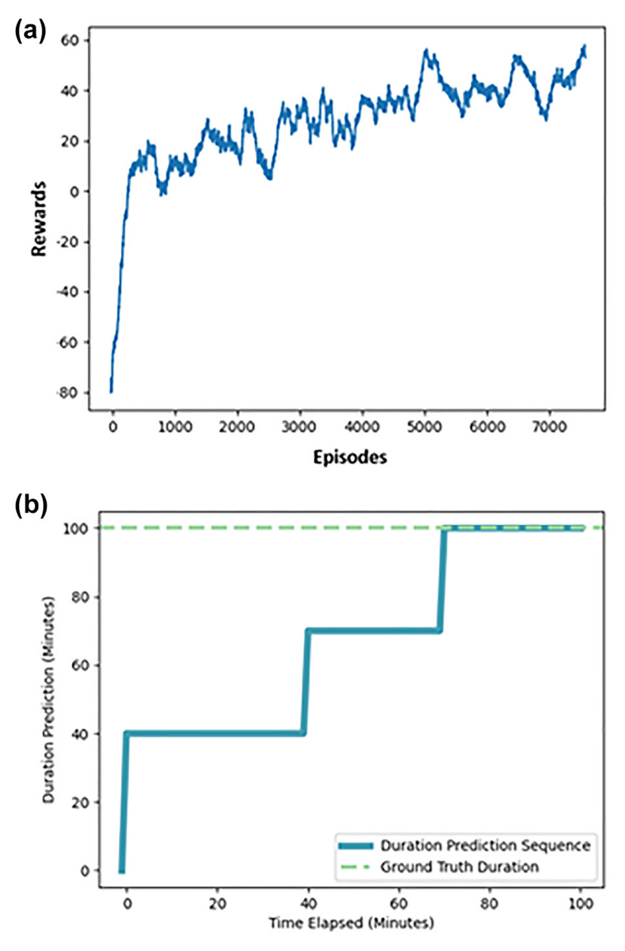

Figure 6. (*a*) Average accumulative rewards in training.
(*b*) Incident duration predictions sequence for incident 202202631.
Note: Color online only.

The neural network in DQNs is kept relatively simple because of the small number of incident records. In this case study, we only have 9378 incident records, with 80% of them being used as training data, resulting in 7500 training samples. Given the total number of parameters in the DQN neural network is 496, the number of training samples is sufficient for training this scale of neural network. However, if more incident records become available, the dimensions and number of hidden layers in the neural network of DQNs can be expanded to better capture the relationships between states and actions.

Implementation and Results

The CPH-based model identified 12 characteristics that have a statistically significant impact on the incident duration, with a p-value of less than 0.005. These characteristics are summarized in Table 2, resulting in a state

Table 3. Cox Proportional-Hazards Model Regression Results

Attribute	Value
HAZMAT_SPILL	0
HEAVY_TRUCK	0
HIGH_WATER	0
ROAD_DEBRIS	0
VEHICLE_FIRE	1
CRASH	0
VEHICLES_INVOLVED	1
MAINLANES_BLOCKED	3
FRONTAGE_LANES_INVOLVED	0
RAMP_LANES_BLOCKED	0
HOV_LANES_BLOCKED	0
SHOULDER_LANES_BLOCKED	1

Note: HOV = High-Occupancy Vehicle.

vector s of length 13 (the detailed data descriptions, along with their respective types, are provided in Table 1).

Figure 6*a* displays the average cumulative reward of the RL agent for each episode during training, demonstrating a significant improvement in the cumulative rewards as the training progressed. Initially, the rewards were low, with a value of -80, indicating that the agent's actions were not generating accurate duration estimation. However, as the agent gained more experience and adapted its behavior accordingly, the cumulative rewards gradually increased, reaching a value of 60 at the end of the training period. This upward trend in the rewards reflects the agent's ability to make better duration predictions over time, leading to more favorable outcomes and higher rewards, which demonstrates the effectiveness of the RL model in learning from its environment and achieving its goal of maximizing rewards.

To demonstrate the progress of the agent producing a sequence of duration predictions, we can use incident record *202202631* as an example, with its attributes listed in Table 3. The ground truth duration of this incident was 100 min, as shown by the green dashed line in Figure 6*b*. The initial duration prediction was 40 min, which was disseminated through the ATIS to travelers for making informed travel decisions before 40 min had passed. After 40 min, the agent checked the TIM system and found that the incident was not cleared yet, triggering a revision to the prior prediction. The agent updated the prediction to 70 min and synced up with transportation agencies and travelers. Similarly, after 70 min, the incident was still not cleared, and the agent updated the prediction again to 100 min. Finally, at 100 min timestamps, the incident was cleared, and the agent finished its work.

The RL agent was implemented using the PyTorch framework on an Apple M1 chip, which features an eight-core central processing unit (CPU) and a seven-core graphics processing unit (GPU). We trained the RL

Table 4. Results Comparison between Reinforcement Learning (RL) and the Cox Proportional-Hazards (CPH) Model

	RL							CPH (*21*)	ANN	RF
	MAE_1	MAE_2	MAE_3	MAE_4	MAE_5	MAE_6	MAE	MAE_1	MAE	MAE
Group 1	10.3	NA	NA	NA	NA	NA	10.3	10.0	10.4	10.4
Group 2	16.1	11.9	NA	NA	NA	NA	14.0	17.4	14.7	14.9
Group 3	48.7	18.3	13.2	NA	NA	NA	26.7	51.4	28.6	28.7
Group 4	82.5	50.4	19.0	10.9	NA	NA	40.7	83.0	41.0	41.5
Group 5	95.6	68.8	41.9	16.9	8.9	NA	46.4	98.3	47.2	47.2
Group 6	115.5	90.5	65.5	40.5	15.5	10.5	56.3	112.5	61.2	60.8

Note: ANN = artificial neural network; RF = random forest; MAE = mean absolute error; NA = not available.

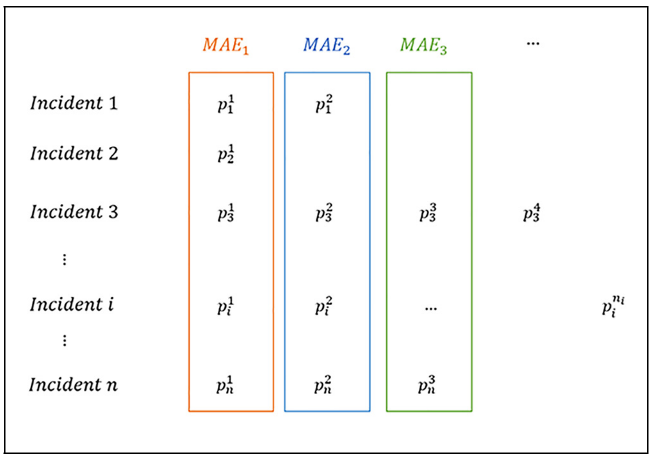

Figure 7. Illustration of definitions of mean absolute errors (MAEs).

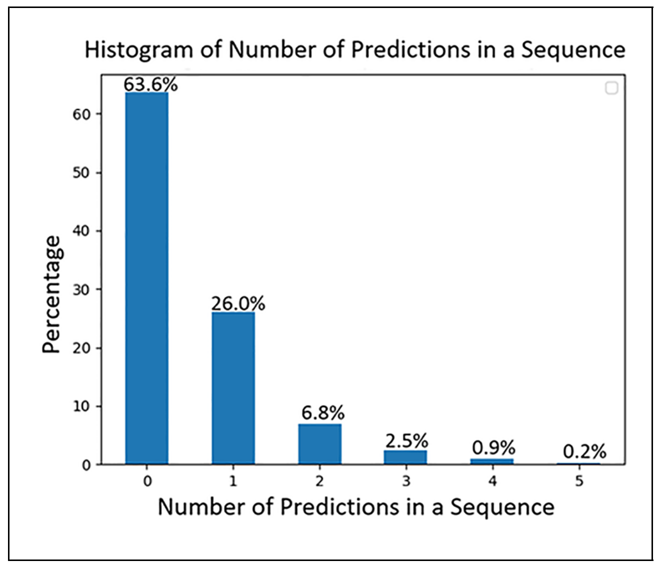

Figure 8. The number of predictions in the reinforcement learning predictions sequence.

agent over 8000 episodes, with a total running time of approximately 25 min. To assess the effectiveness of the agent, we first compared the performance of our RL agent's sequential predictions to that of the most commonly used one-time prediction model, the CPH-based model (*21*). We generated sequential predictions for incidents in the testing dataset. For each incident i, the agent produced a sequence $P_i = \left(p_i^1, p_i^2, \ldots, p_i^{n_i}\right)$, where $n_i = |P_i|$. Then, we grouped the incidents into six groups based on the length of the RL agent's predicted sequence. Specifically, Group 1 includes incidents with a predicted sequence of length 1, Group 2 includes incidents with a predicted sequence of length 2, and so on up to a maximum sequence length of 6. We calculated the mean absolute error (MAE) for each prediction in the sequence, with MAE_1 representing the MAE for the first prediction, MAE_2 representing the MAE for the second prediction, and so on, as illustrated in Figure 7. The group with the larger predicted sequence length often includes incidents with durations that are outliers from the distribution, representing more extreme cases that experience uncommonly long durations. This grouping allows us to evaluate our model's performance in predicting such extreme cases. By dividing the incidents into six groups, we were able to assess the RL agent's performance in predicting incidents of different levels of complexity, and examine how the accuracy of the agent's predictions changed as the length of the sequence increased. Table 4 displays the results of our evaluation. We observe that, for incidents in Group 1, the RL agent and CPH model have very similar MAEs. However, starting from Group 2, the RL agent's predictions have smaller MAEs compared to those of the CPH model. The most notable advantage of the RL agent's sequential predictions is that the MAEs decrease as new predictions are produced. This compensates for the limitations of one-time predictions, which only produce a single prediction and cannot improve over the duration of an incident.

Figure 8 displays the distribution of the number of predictions generated by the RL agent, that is, the number of incidents in each group. The results reveal that for

63.6% of incidents, the RL agent only needs to produce one prediction. For 26.0% of incidents, two consecutive predictions are required, while for 6.8% of incidents, three consecutive predictions are needed. More than three sequential predictions are required for only 3.5% of incidents. Therefore, the agent primarily operates as a one-time prediction method. However, in cases where a one-time prediction has a considerable error and underestimates, additional predictions can be generated to compensate for the limitations of one-time prediction approaches.

To further assess the effectiveness of the RL agent, we also compare its performance to that of other models constructed for the same problem and application. Specifically, we trained an ANN model (two hidden layers with size 8) and a random forest (RF) model, using the same training and testing datasets.

Recognizing that both the ANN and RF models are fundamentally designed for single predictions rather than sequential updates, we implemented a strategy to adapt them for our specific needs. To account for the evolving nature of incident durations, we included "time elapsed" as a covariate in both models during training. This approach allows the models to incorporate the elapsed time since the incident occurred, which is crucial for making more accurate duration predictions.

To further simulate scenarios where new predictions are necessary, we artificially interpolated samples into the training dataset. For instance, consider an incident characterized by a set of attributes $(x_1, x_2, \cdots, x_n)$ and ground truth duration p. The first training sample is $(x_1, x_2, \cdots, x_n, 0)$ with an output of p (here 0 indicates that no time has elapsed since the incident's occurrence). In contrast, we synthesized additional training samples by varying the elapsed time; for example, $(x_1, x_2, \cdots, x_n, t)$ with output an output of $p - t$ (here t refers to the time elapsed), where t is a random value ranging from 0 to p. This method allowed us to effectively train the ANN and RF models to handle situations where updated predictions might be required, despite their original design constraints.

After comparing the MAE of all the predictions in the sequences, we found that the RL agent achieved a MAE of 16.9 min, while the ANN and RF models achieved respective MAEs of 18.2 and 18.4 min. Furthermore, as indicated in Table 4, the MAEs for the ANN and RF models for each incident group were consistently larger than those of the RL agent. This result suggests that the RL agent performs better in making sequential predictions, which can be attributed to the nature of RL as a sequential decision-making process. In contrast, other models, such as the ANN and RF, may not be explicitly designed to make sequential predictions and may be trained on isolated samples without considering the sequential nature of the problem. Therefore, the RL agent's advantage in making sequential predictions may be partly because of its training procedure, which emphasizes the importance of the sequence of actions taken by the agent.

Overall, compared to the one-time prediction methods, the RL agent's predictions have smaller MAEs, and the MAEs decrease as new predictions are produced, which compensates for the limitations of one-time predictions; compared to other models applied in sequential prediction such as the ANN and RF, the RL agent also achieves a lower MAE, which can be attributed to the nature of RL as a sequential decision-making process.

Discussion

The results of this study underscore the significance of utilizing RL for improving incident duration predictions in TIM. Our findings illustrate that the RL agent's ability to adjust its predictions based on real-time data enhances the accuracy of estimated incident durations, particularly in situations where initial predictions have proven inaccurate. This adaptability is crucial in dynamic traffic environments where incidents can evolve, leading to prolonged durations that may not be captured by static prediction models.

By incorporating our RL framework, traffic operators can provide timely updates to travelers with respect to incident durations. This capability not only aids in more efficient route planning but also minimizes the negative impact of incidents on overall traffic flow. Improved accuracy in predictions can significantly reduce congestion, leading to better resource allocation and incident response strategies.

Our study contributes to the existing body of knowledge by highlighting the limitations of one-time prediction models. The RL agent's ability to produce a sequence of predictions, with decreasing MAEs over time, demonstrates a more nuanced approach to incident duration estimation. This advancement can inspire further research and development in predictive modeling within the field of traffic management.

To build on our findings, future studies should focus on integrating a more extensive dataset that includes a wider variety of incident types and conditions. In addition, incorporating external factors such as weather conditions, time of day, and historical traffic patterns may further improve the accuracy of predictions. Exploring the use of hybrid models that combine RL with traditional machine learning techniques could also provide deeper insights into traffic incident dynamics.

Conclusions

The growth and economic development of a society depend significantly on the efficiency of its road and highway system. However, traffic congestion often poses challenges to the effective use of the transportation system. To mitigate the harmful effects of incidents on travelers, collecting incident-related information through a TIM system and disseminating accurate incident duration predictions to travelers through an ATIS is essential. This allows travelers to make informed decisions about the best routes and times to depart, mitigating the negative impacts of incidents on traffic flow. There is a situation that requires sequential prediction for incident duration but has not been extensively explored in the literature or practice, which is when an incident has not been cleared even though the estimated incident duration time has elapsed. This paper introduces a methodology for training an RL agent to generate sequential predictions under this circumstance.

We applied our methodology to the Houston TranStar incident data and trained an RL agent to produce new predictions when earlier ones underestimated, resulting in a sequence of predictions. The RL agent's average cumulative reward for each episode during training significantly improved as the training progressed, from −80 to 60, indicating the agent's improving ability to make accurate duration predictions over time. To assess the effectiveness of the agent, we first compared the performance of our RL agent's sequential predictions to that of the most commonly used one-time prediction model, the CPH-based model. Compared to the CPH model, the RL agent's predictions have smaller MAEs, and the MAEs decrease as new predictions are produced, which compensates for the limitations of one-time predictions. In addition, we applied other models in sequential prediction, such as the ANN and RF: the RL agent also achieves a lower MAE, which can be attributed to the nature of RL as a sequential decision-making process.

Furthermore, we have found that in 63.6% of incidents, the agent only needs to make a single prediction, indicating that the agent primarily serves as a one-time prediction method. However, in cases where a one-time prediction incurs a significant error, the agent can generate additional predictions to address the limitations of the one-time prediction approach.

To the best of our knowledge, this is one of the first studies that addresses the issue of accounting for inaccurate earlier incident duration estimates when an incident remains uncleared even though the estimated incident duration time has elapsed. Our proposed framework facilitates traffic operators to generate updated predictions and provide travelers with updated information, empowering them to make informed decisions. Future research should focus on gathering more comprehensive incident data and incorporating additional state information to enhance the accuracy of the incident duration predictions.

Acknowledgments

The authors would like to thank TranStar and the Texas Department of Transportation (TxDOT) Houston District for sharing the incident data, and are grateful for Dr. Brenda Bustillo's review of an earlier version of this study.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Z. (R.) Huang, X. Li, Y.-C. Chiu; data collection: Z. (R.) Huang; analysis and interpretation of results: Z. (R.) Huang, X. Li, Y.-C. Chiu; draft manuscript preparation: Z. (R.) Huang, X. Li, Y.-C. Chiu, Y.-J. Wu. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Ray (Zirui) Huang https://orcid.org/0000-0001-6640-6271
Xiaofeng Li https://orcid.org/0000-0001-5526-9961
Yi-Chang Chiu https://orcid.org/0000-0002-1993-9478
Yao-Jan Wu https://orcid.org/0000-0002-0456-7915

Data Accessibility Statement

The data that support the findings of this study cannot be publicly shared.

References

1. Owens, N. D., A. H. Armstrong, C. Mitchell, and R. Brewster. *Federal Highway Administration Focus States Initiative: Traffic Incident Management Performance Measures Final Report*. United States. Federal Highway Administration, Washington, D.C., 2009.
2. Clemons, J., L. Aultman-Hall, and S. T. Bowling. *ARTIMIS Telephone Travel Information Service: Current Use Patterns and User Satisfaction*. Report No. KTC-99-24. Kentucky Transportation Center, Lexington, 1999.
3. Lerner, N., R. Huey, P. Zador, J. Harpster, and D. Duncan. User Preferences for Information Types in Advanced Traveler Information System Applications. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 42, No. 17, 1998, pp. 1200–1204.
4. 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.

5. 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.
6. Javid, R. J., and R. J. Javid. A Framework for Travel Time Variability Analysis Using Urban Traffic Incident Data. *IATSS Research*, Vol. 42, No. 1, 2018, pp. 30–38.
7. 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.
8. Lin, L., Q. Wang, and A. W. Sadek. A Combined M5P Tree and Hazard-Based Duration Model for Predicting Urban Freeway Traffic Accident Durations. *Accident Analysis & Prevention*, Vol. 91, 2016, pp. 114–126.
9. 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.
10. 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*, 2016. 2554:139–148.
11. 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* (S. V. Ukkusuri, and K. Ozbay, eds.), Springer, New York, 2013, pp. 287–305.
12. 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.
13. 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.
14. Zhu, W., J. Wu, T. Fu, J. Wang, J. Zhang, and Q. Shangguan. Dynamic Prediction of Traffic Incident Duration on Urban Expressways: A Deep Learning Approach Based on LSTM and MLP. *Journal of Intelligent and Connected Vehicles*, Vol. 4, No. 2, 2021, pp. 80–91.
15. 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.
16. Li, R., F. C. Pereira, and M. E. Ben-Akiva. Overview of Traffic Incident Duration Analysis and Prediction. *European Transport Research Review*, Vol. 10, No. 2, 2018, pp. 1–13.
17. 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.
18. 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.
19. Pereira, F. C., F. Rodrigues, and M. Ben-Akiva. Text Analysis in Incident Duration Prediction. *Transportation Research Part C: Emerging Technologies*, Vol. 37, 2013, pp. 177–192.
20. Dezfouli, A., and B. W. Balleine. Habits, Action Sequences and Reinforcement Learning. *European Journal of Neuroscience*, Vol. 35, No. 7, 2012, pp. 1036–1051
21. Hensher, D. A., and F. L. Mannering. Hazard-Based Duration Models and Their Application to Transport Analysis. *Transport Reviews*, Vol. 14, No. 1, 1994, pp. 63–82.

The findings expressed in this paper are solely from the authors and do not represent the viewpoints of Houston TranStar or TxDOT.