DEVELOPMENT OF FREEWAY INCIDENT DURATION PREDICTION MODELS

Chi-Hyun SHIN
Associate Professor
Dept. of Transportation Engineering
KYONGGI University
San-94 Yiui-Dong Paldal-Gu, Suwon
442-760, Korea
Fax : +82-31-249-9775
E-mail : chshin@kyonggi.ac.kr

Abstract: The real-time decisions regarding the resources needed to clear and manage the incident and the information disseminated to travelers all depend on the knowledge of incident duration. An accurate and reliable estimate of the duration can be the main difference between an effective incident management operation and an unacceptable one. With a partial opening of the Seoul Urban Expressway Traffic Management Center, it was imperative to have incident duration models based on the incident data sets collected in the region to reflect local conditions such as traffic, roadway and incident management resources. Two main sources of various incident cases were from both the Traffic Incident Reports of the Korea Highway Corporation and the at-field accident investigation effort. Total of 506 freeway mainline incident data were collected and reduced to a set of 270 cases. Several multiple regression models have been developed and presented herewith.

Key Words: Freeway Incident, Incident Duration, Incident Management, Regression Models

1. INTRODUCTION

1.1 Background and Objectives

Incident is non-recurrent events on roadway such as traffic accidents, vehicle disablement, load spills and flat tires that causes traffic congestion. It is the major cause of traffic delay and air quality problems in urban area let alone personal safety and property loss.

Time between incident occurrence and incident clearance, referred as incident duration consists of several components as shown below. The major portion of the duration is occupied by incident clearance time.

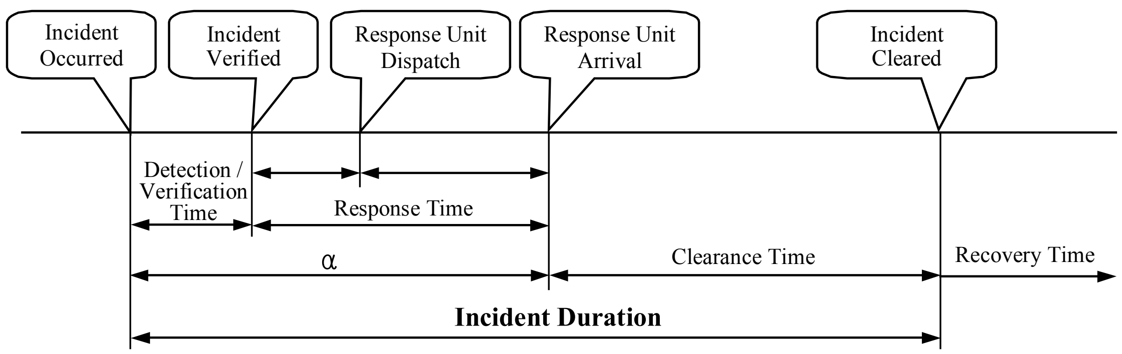

Figure 1. Time Portions of Incident Duration

Duration prediction is one of the important steps of the incident management process. The real-time decisions regarding the resources needed to clear and manage the incident and the information given to travelers all depend on the knowledge of incident duration.

For many freeway traffic management centers including that of Seoul, it was most recognized to have incident duration models if and when an incident management system is put into operation. With sponsored by the Ministry of Construction and Transportation, this study set up following objectives. Firstly, develop reliable estimation models of incident duration based on a local database, that are simple to understand and less troublesome when modifications are needed. Secondly, develop an incident duration decision tree in order to immediately and easily determine or verify an average incident duration an its range.

1.2 Study Scope and Procedures

The research entails literature reviews, visitation of incident response agencies, interviews, data gathering from agencies and field surveys like wrecker truck on-board data acquisition as shown in Figure 2. These efforts are mainly focused on the setting up of a small but adequate database upon which incident duration models are developed. Data collection effort was geographically limited to the freeways in Seoul metropolitan region.

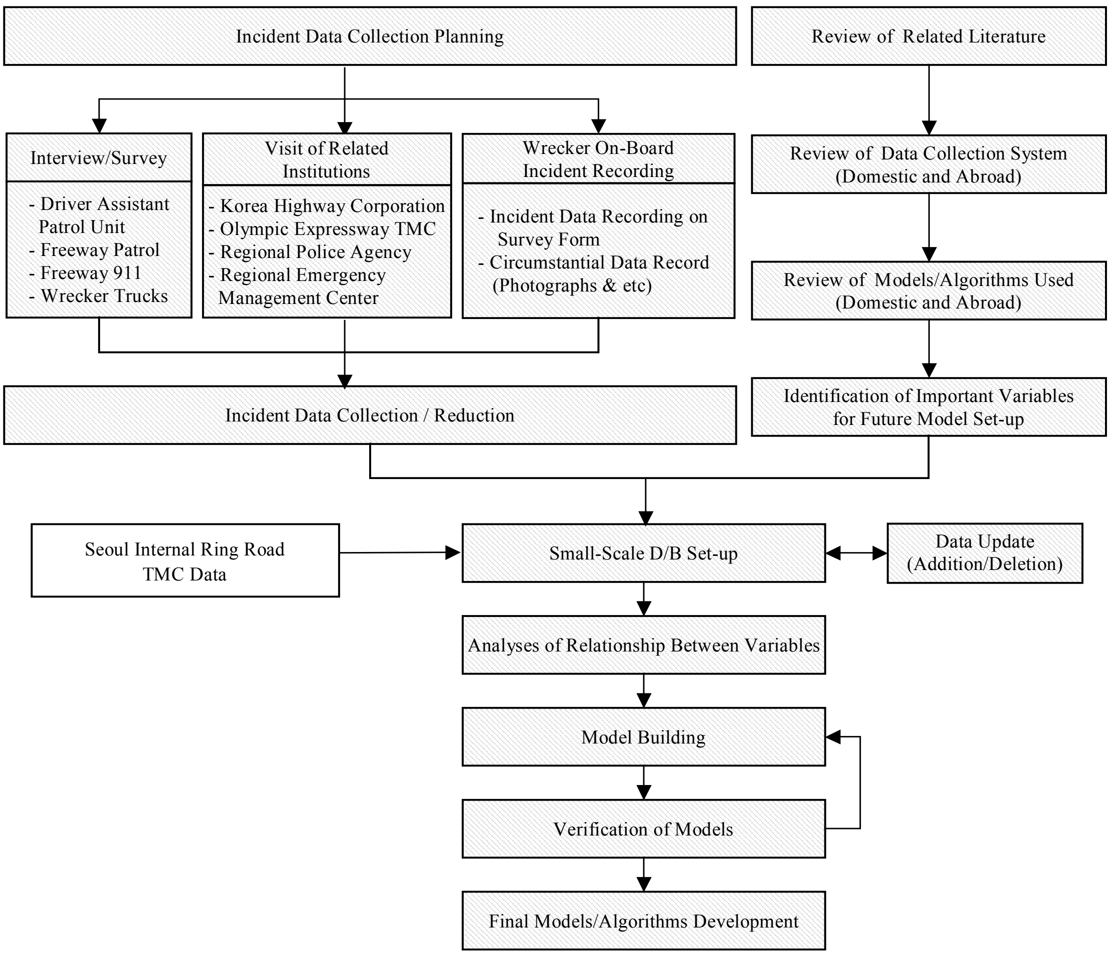

Figure 2. Research Procedures

2. LITERATURE REVIEW AND ISSUE STATEMENT

The one of major research efforts for modeling freeway incident clearance time has been carried out at Northwestern University using operational, incident-type and environmental factors. Another study was also conducted in Northwestern University as part of the ADVANCE project that has produced a decision tree.

Garib(1997) has proposed an incident duration model using the I-880 data set that has a regression formula form. Recently FHWA sponsored a research to develop a model, called IMPACT, to predict freeway incident and delay. Mannering(2000) has rigorously proposed hazard- function-based duration models that are somewhat cumbersome for TMC to adopt.

In Korea, Han(2000) proposed an incident duration model using the data set of a TMC in Washington State. However, one must recognize that Korean road environment, traffic mix, deployment of wrecker trucks, accident investigation procedures and like are too different to adopt the models as above. Therefore, it is imperative to develop such models using data sets prepared in Korea.

3. DATA COLLECTION AND REDUCTION

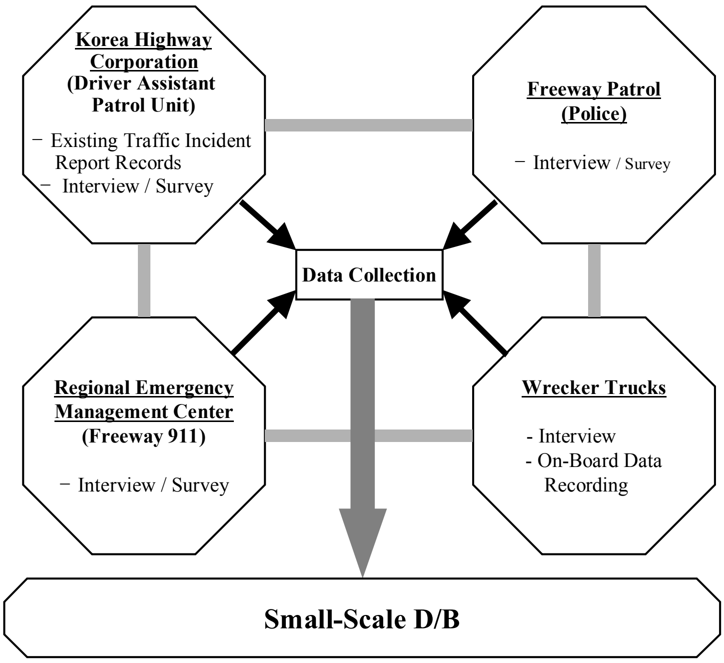

Figure 3. Data Collection Steps

As a preliminary effort, research team has visited many institutions such as the Korea Highway Corporation, the Olympic Expressway TMC, the Seoul Police Agency and the Traffic Radio Service to investigate incident management processes and to collect available data sets. Many interviews and surveys have been also conducted to obtain the hands-on experience and knowledge of incident response/report personnel at fields.

Specially, the members of research team have boarded wrecker trucks for 3 months to acquire more reliable incident related data and to precisely isolate important explanatory variables. A portion of data has been also collected from wrecker truck drivers on payment-basis.

In brief, data has been collected from two major data gathering sources. One major source is the Traffic Incident Reports of the KHC and the other data acquired from the wrecker-on-board effort (field survey) as cited. These two data sets have been reduced to one final integrated database.

3.1 Data Collection from Traffic Incident Reports of the KHC

Total of 384 incident data was obtained from the Korea Highway Corporation. From the data set, 170 mainline incident cases have been prepared for analysis since incidents on roadside and toll plaza area were eliminated.

Table 1. Description of the KHC's Data

Dataset	Collection Period	Area Coverage and Reduction Criteria	# of Data
Initial Dataset	2000. 1.~ 2001. 9. (21 months)	Seoul Han-nam bridge(0.0km)～ Kyonggi An-sung(63.2km)	384
Reduced dataset		Only main-lane accidents selected Small scale accidents eliminated Pedestrian accidents eliminated	170

3.2 Data Collection by Boarding Wrecker Trucks (Field Survey)

Besides the recording items appeared on the KHC's incident report log sheets, the number of response vehicles (police cars, ambulances, fire trucks, wreckers) was acquired from fields to take operational factors into account. The rest of information on incidents was collected in accordance with the KHC's traffic incident report format.

Total of 122 incident data was obtained during a 3-month period. From the data set, 100 incident cases have been prepared for analysis.

3.3 Integration of the Two Data Sets

The KHC's traffic incident report data and the field survey data obtained from wrecker-on-board effort were integrated into one database.

Table 2. Description of Integrated Dataset

Incident Types	Passenger Car		Small Truck		Medium Truck		Large Truck		Total
	Stop	Tumbled / Overturned	Stop	Tumbled / Overturned	Stop	Tumbled / Overturned	Stop	Tumbled / Overturned	
Single	31	20	3	7	4	6	6	6	83
Double	29	4	14	6	7	5	18	4	87
Multiple	36	1	12	3	5	2	16	0	75
Total	96	25	29	16	16	13	40	10	245

* Another dataset of 25 incident records was used for model validation

Trucks with its weight of 5 tons or more were classified into heavy vehicles (Truck), and vehicle overturn and vehicle tumbledown have been reduced to one incident-type category due to limited amount of incidents of these types. Data were further classified to detailed categories according to the time of incident occurrence (day or night), weather conditions (snow or rain), vehicle fire and load spills.

Total of 245 incident cases includes 36 spillages, 20 fire cases, 18 accidents of fatality and 27 adverse weather conditions. The time length of incident duration ranges from 10 minutes to 280 minutes. In particular, incidents involving the overturn and tumbledown of medium to full size truck often appeared to create load spills resulting in excessively long incident duration time.

4. DATA ANALYSES AND MODEL BUILDING

4.1 Models for Traffic Incident Report Records of The KHC

Total of 384 incidents data collected with help of Korea Highway Corporation. The dataset was reduced to 170 incident cases that clearly affected traffic flows. 145 incident data was used for modeling and the rest for model validation.

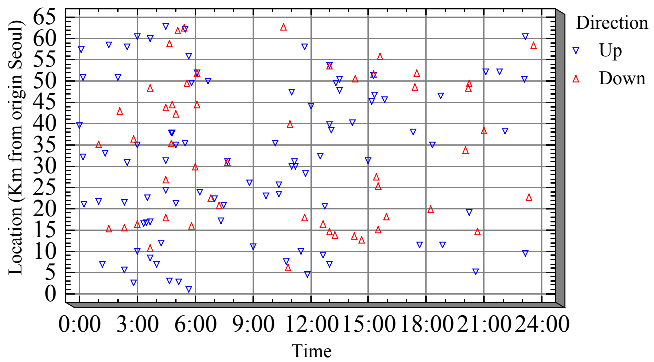

Figure 4. Time-space Diagram for Incidents

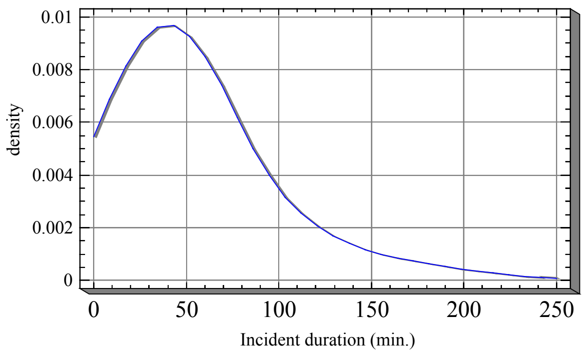

Figure 5. Incident Duration Density Trace

Analysis results show that P-values calculated for independent variables were lower than 0.1 indicating their significance at the 90% confidence level. In particular, two variables, the number of heavy vehicles and cargo spills, were found significant and important for models showing correlation coefficients of 0.5 or higher. And regression method was chosen for model development since, from a viewpoint of TMC operators, the models needed to be simple to understand and less complex to modify when necessary.

Multiple linear regression analysis was employed, and specially several dummy variables were introduced to describe weather condition, environmental status, accident types, and so on. The regression models have following forms.

$$Y = 29.01 + 6.35X_1 + 23.85X_2 + 17.82X_3 + 17.15X_4 + 26.27X_5 + 74.16X_6 + 11.20X_7 + 1.19X_8 \quad (1)$$

$$Y = 29.36 + 6.19X_1 + 23.41X_2 + 18.00X_3 + 16.87X_4 + 26.00X_5 + 73.73X_6 + 11.01X_7 + 0.58X_8X_9 \quad (2)$$

where,

Y = Incident duration,
X_1 = 1 if night or rain, otherwise 0,
X_2 = 1 if snow, otherwise 0,
X_3 = 1 if heavy vehicle turnover, otherwise 0,
X_4 = 1 if vehicle fire, otherwise 0,
X_5 = 1 if small cargo spill, otherwise 0
X_6 = 1 if large cargo spill, otherwise 0
X_7 = Number of heavy vehicles
X_8 = Number of injuries & fatalities
X_9 = Number of lanes closed

The R-Squared statistic indicates that the models as fitted explain 58.6% of the variability in incident duration. The standard error of the estimate shows the standard deviation of the residuals to be 27.46. It is believed that this somewhat low fitting result is produced from the weakness in the database where 20 to 60 minute long incidents cover 72% of all cases.

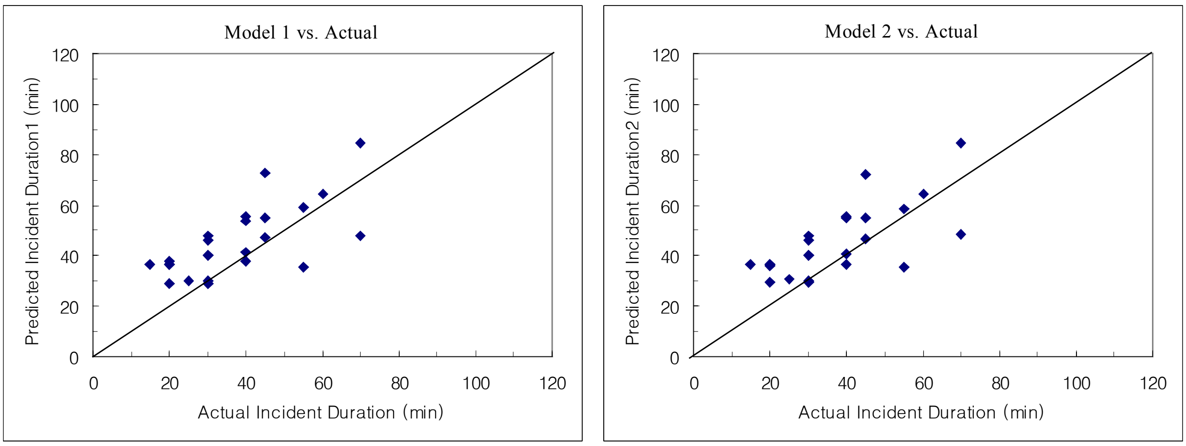

Figure 6. Verification of Model 1&2 Developed Using KHC Data Set

As shown above, models tend to overestimate the durations due to the skewed dataset with high frequency in 20 to 60 minute incidents. A variable, X_8X_9, was used in the equation (2) and represents severity of incident in terms of both human casualty and spatial blockage made by involved vehicles.

4.2 Models for Wrecker-On-Board Data Records

Total of 122 incidents data was collected from 3-month field surveys. Data was reduced to100 incident cases that impacted directly traffic flows.

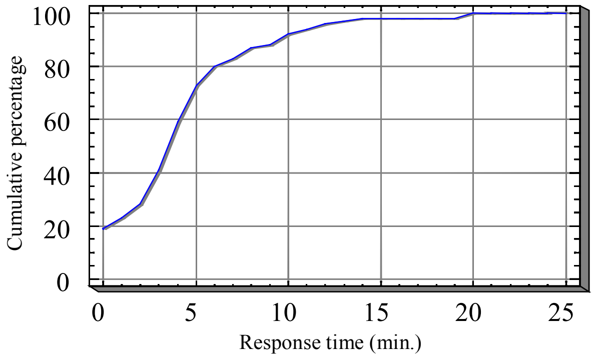

Figure 7. Wrecker Response Time

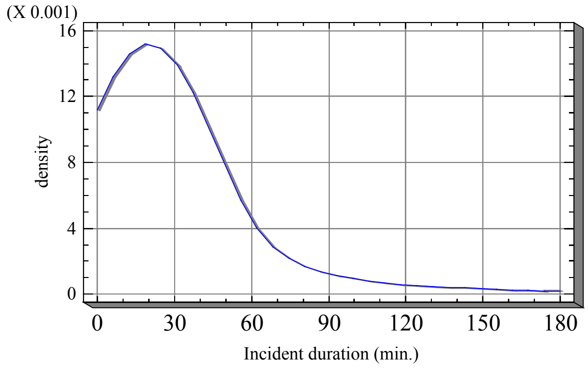

Figure 8. Incident Duration Density Trace

For this data set, an operational variable, the number of response vehicles (police cars, ambulance, wrecker truck, 911 vehicles, special vehicles and cranes) was introduced to the model and found relatively significant. The regression model has a following form.

$$Y = 8.08 + 7.78X_1 + 34.33X_2 + 9.54X_3 + 27.75X_4 + 11.13X_5 + 2.37X_6 \qquad (3)$$

where,

X_1 = 1 if night or rain, otherwise 0,
X_2 = 1 if heavy vehicle turnover, otherwise 0,
X_3 = 1 if vehicle fire, otherwise 0,
X_4 = 1 if cargo spill, otherwise 0
X_5 = Number of heavy vehicles
X_6 = Number of response vehicles

R-square value for Model 3 or equation (3) was as high as 0.794 because the variable, the number of response vehicles, was used in the model and well explains the severity of incidents that directly relates to incident duration. The standard error of the estimate shows the standard deviation of the residuals to be 12.61.

With increasing number of surveillance devices like CCTV along congested urban freeways, the operational variable is easily obtainable without much delay through monitors in traffic management center. For the reference, Seoul urban expressway traffic management system has installed CCTVs for every 4 kilometers. The CCTV has powerful zooming capability and even the license plate number of a vehicle on roadway is legible through monitor at TMC.

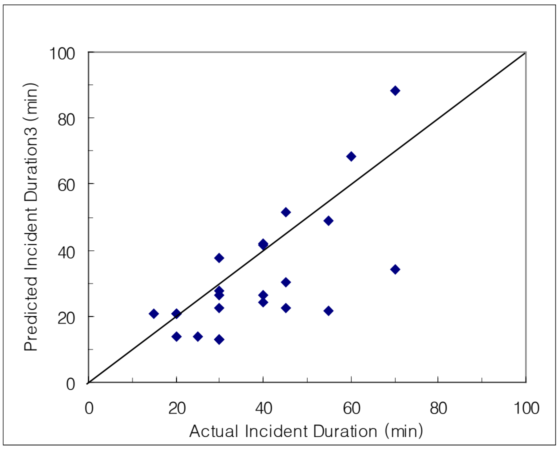

Figure 9. Verification of Model 3

Model 3 marginally tends to underestimate durations because the dataset used has relatively small number of severe incidents and produced a small value of 8.08 for an intercept in the resulting regression model.

4.3 Model Development Using Integrated Database

Two data sets previously described were integrated into one small database, and final and more general regression models were developed as follows. For this time, the operational variable was not used in the model since rural freeway conditions usually does not allow the variable to be easily obtainable. Refer to equation (1) and (2) for variable descriptions.

$$Y = 21.84 + 10.50X_1 + 29.89X_2 + 23.35X_3 + 21.10X_4 + 26.13X_5 + 70.16X_6 + 12.18X_7 + 1.18X_8 \quad (4)$$

$$Y = 22.19 + 10.39X_1 + 29.38X_2 + 23.44X_3 + 20.76X_4 + 25.90X_5 + 69.82X_6 + 11.98X_7 + 0.58X_8X_9 \quad (5)$$

R-square values for Model 4 and 5 were both 0.649. The standard error of the estimate shows the standard deviation of the residuals to be 18.45. This result could lead us to a conclusion that, with extended data collection effort on severe incident cases, there is great possibility for model improvement later on.

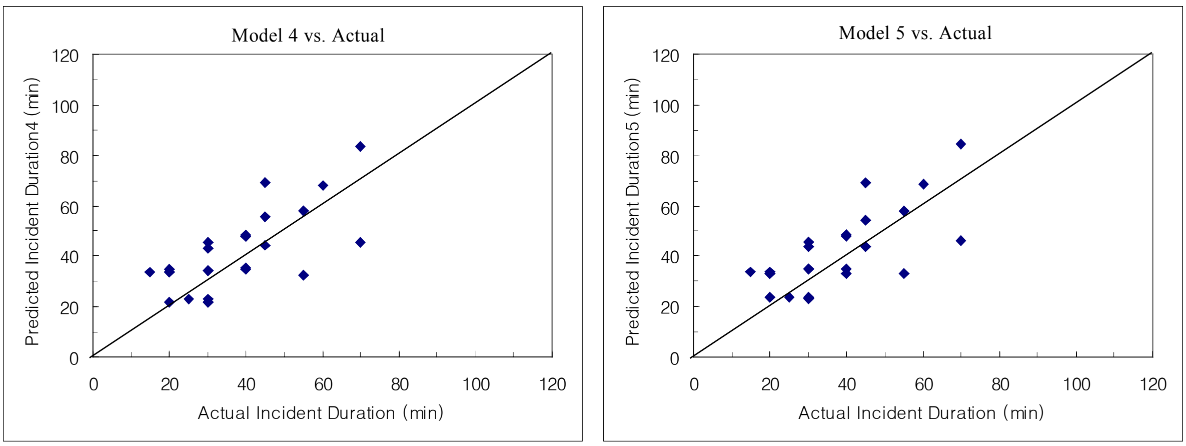

Figure 10. Verification of Model 4&5 Developed Using an Integrated Data set

4.4 Development of Decision Tree

Besides multiple linear regression models proposed, an incident duration decision tree was developed, as shown in Figure 11. The purpose of this type of decision tree is to immediately and easily determine an average incident duration and its inter-quartile range based on incident types and characteristics identified.

The structure of the tree consists of two levels. The number of vehicles involved is identified in the first level, followed by the next level within which detailed information such as heavy vehicle involvement, overturn of vehicle and cargo spill is identified as available.

The decision tree provides TMC's incident management staffs with a convenient and ready-to-use tool for incident duration prediction based on incident types and characteristics identified and the knowledge of incident on the scenes. Due to paper length restriction, trees for incidents with multiple vehicle involvement were not presented here.

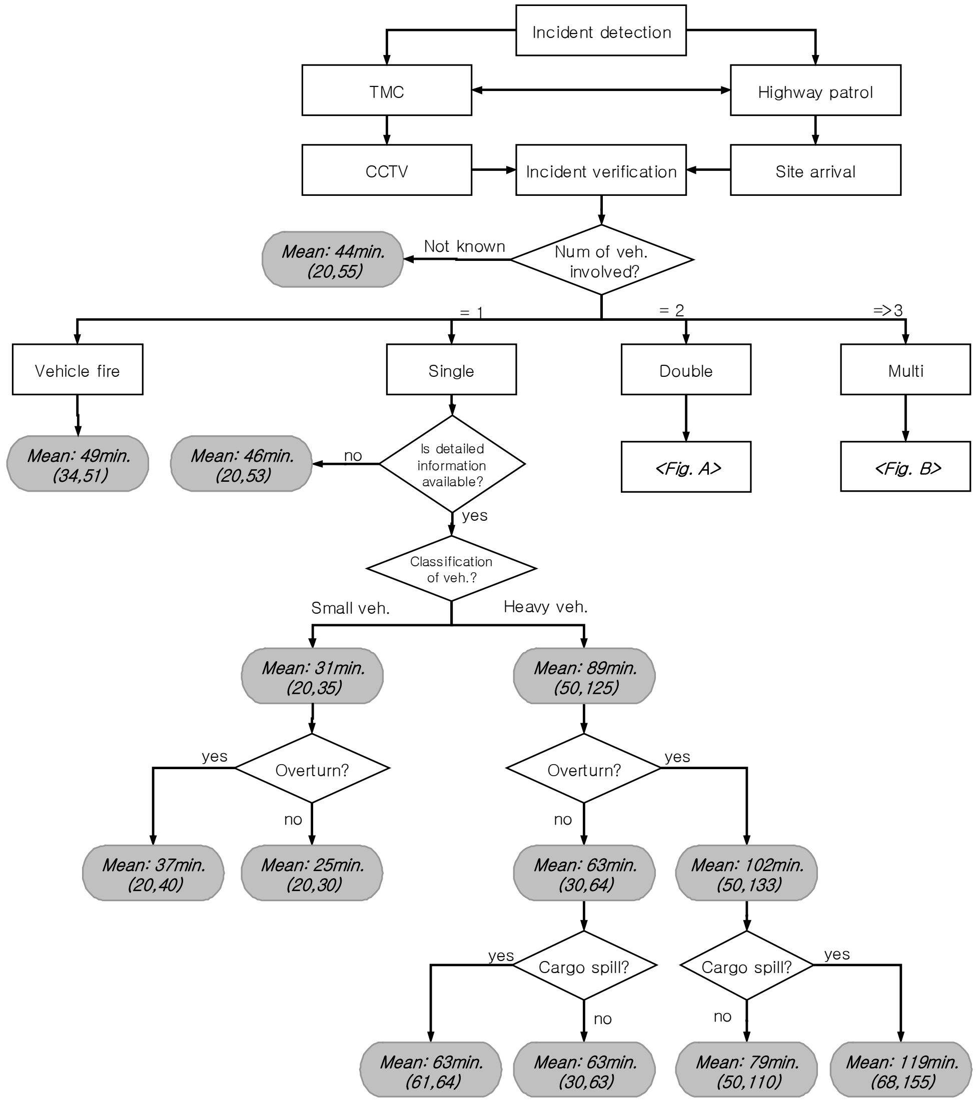

Figure 11. Decision Tree for Freeway Incident Duration Estimation

5. CONCLUSIONS AND FUTURE RESEARCH

5.1 Conclusions

Total of 506 incident data from the traffic incident reports of the KHC and the field surveys was collected. Data went under a screening process and was finally reduced to 245 incident cases. With the dataset reduced, statistical analyses were performed, and regression models and decision trees for incident duration prediction have been developed and proposed.

It was found the majority of minor incidents were cleared instantly after the arrival of wrecker trucks that are in competitive business. The variables that significantly affect incident duration times were found to be the number of heavy vehicles involved, vehicle tumble or turnover, cargo spills, time of day and weather conditions.

Regression models are separately proposed for two different data sources even though the models are similar in the choice of independent variables. The variables showing strong relationships with the dependent variable were those cited as above. The proposed model such as equation (3) shows that the number of response vehicles on the scene obtainable with proper surveillance devices plays an important role in model precision.

However, the knowledge on the number of response vehicles on the scene requires small time lag after incident occurrence and might need the full coverage of scenes using traffic surveillance systems. Therefore, proposed models should be divided into 2 groups (Model I, II, IV, V and Model III) and the decision of model choice is dependent upon the surveillance capability of TMCs.

It was also found models tend to overestimate the duration times when a dummy variable, load spill, is used. It was simply because several of incidents with load spill data had excessively long clearance times and a variety of load spills was not incorporated in the models.

5.2 Future Research

For this research, large portion of incident cases was related to minor incidents. To improve model precision, more extensive collection of severe incident data such as vehicle fire, load spills, vehicle turnover is needed. Specially, more of severe incident cases with duration time of 60 minutes or more should be incorporated into the existing database. Cases with many different environmental conditions also need to be collected.

Since this research used the incident data obtained mainly within Seoul metropolitan region, it is possible that the incident duration models proposed have limited prediction capabilities when used for totally different geographical settings. It is therefore most necessary to extend the data collection effort to the freeways around the nation.

In addition, one might wish to rest on more complex model development efforts utilizing fuzzy logic or neural network theory. But it must be recognized that the decision of what kind to use is a matter of trade-off between accuracy and economy.

REFERENCES

Ozbay, K. and Kachroo, P. (1999) **Incident Management in ITS.** Artech House, Boston, MA.

Garib, A. E. Radwan and H. Al-Deek. (1997) Estimating Magnitude and Duration of incident Delays. **Journal of Transportation Engineering, Vol.123, N0. 6**, 459-466.

Khattak, A., Schofer, J. and Wang, M.H. (1995) A Simple Time Sequential Procedure for Predicting Freeway Incident Duration. **IVHS Journal, Vol. 2, No. 2**, 113-138.

Nam, D.H. and Mannering, F. (2000) An Exploratory Hazard-Based Analysis of Highway Incident Duration. **Transportation Research Part A, Vol. 34A, No. 2**, 85-102.

Manugistics (2000) **STATGRAPHICS 5 Plus.** Manugistics Inc., Maryland.

FHWA (2000) **Traffic Incident Management Handbook.** Federal Highway Administration Office of Travel Management, PB Farradyne.

Han, U.G. (2000) A model for Estimation Incident Duration. Master thesis, Ajou University

Vaneet Sethi., *et al*. (1994) Duration and Travel time Impacts of Incident. ADVANCE Project Technical Report TRF-ID-202, Northwestern University Transportation Center.