공학석사학위논문

An Exploratory Hazard-based Analysis of Traffic Accident Impact on Freeway Traffic Flow

고속도로 교통류에 미치는 교통사고 영향의 위험 기반 분석

2017년 2월

서울대학교 대학원
건설환경공학부
이 상 협

Abstract

An Exploratory Hazard-based Analysis of Traffic Accident Impact on Freeway Traffic Flow

Lee, Sanghyeop
Department of Civil and Environmental Engineering
The Graduate School
Seoul National University

The traffic congestion can be categorized into two groups regards to its causes; Recurring congestion and Non-Recurring congestion. Of these two types, Non-Recurring congestion occurs more frequently that it occupies 60% of traffic congestion. The main causation of the Non-Recurring congestion is traffic accident.

For making efficient plan for preventing congestion caused by the accidents on highway, this research aimed to analyzing impacts of traffic accident on traffic flow. For achieving this goal, cox proportional hazard model has been adopted. Four years of traffic accident records and Vehicle Detectors (VDs) data on the Gyeongbu line, the Seohaean line, the Yeougdong line in Republic of Korea (Korea) are collected for analysis. By using the hazard model, various factors of traffic accident that affects the traffic flow have been identified. If the traffic flow before the accident is stable, the hazard of getting worse level of service(LOS) is lower.

Moreover, hazard is increased proportionally with the degree of vehicle damages and volume at upstream. Also, location of lane closure and involvement of heavy vehicle, alignment also have an effect on occurring congestion.

Meanwhile, factor influence is different according to traffic flow characteristics before accident, and management goal. The cases of Model 1 (decrease in LOS from LOS A and B to LOS C), Model 3 (from LOS C and D to LOS E) are influenced by accident severity more than Model 2(from LOS A and B to LOS E). But, the cases of Model 2 are mostly affected by capacity decrease and demand. It means severely accumulated queue is needed for incurring congestion in Model 2. It is because Model 2 is the case representing conversion of the state (from free flow to severe congested flow).

From this results, the research suggested what type of operations is needed to prevent congestion and what extent of operations is needed.

Keyword : Traffic accident, Traffic flow, Traffic operations, Cox proportional hazard model

Student Number : 2015-21303

<Table of Contents>

<List of Tables>

<List of Figures>

Chapter 1. Introduction

1.1. Research Background and Purpose

1.1.1 Research Motivation

Traffic demand has been increasing explosively for past decades because of population and economic growth (Giuliano, 1989). Increasing traffic demand has caused traffic congestion under restricted infrastructure condition (Giuliano, 1989). Economic cost caused by traffic congestion is assumed to be about $ 37.5 billion annually at 50 cities in the USA (Lee, 2010). Under these circumstances, various traffic operation methods such as ramp metering has become critical to address the problem of traffic congestion.

The congestion is generated by numerous factors. The congestions are divided largely into two sections according to causes: Recurring congestion and Non-Recurring congestion (McGroarty, 2010; Giuliano, 1989). Recurring congestion is occurred due to the surplus of traffic demand to road capacity during specific time. Whereas Non-Recurring congestion is caused by unexpected events like traffic accident, work zones, bad weather etc.

Among these two types of congestion, Non-Recurring congestion occupies 60% of the total congestion (Lindley, 1987). Also, in the case where the Non-Recurring congestion is occurred, road users cannot respond to the situation in advance. This makes the heavier damage to traffic flow and road user more than Recurring congestion. So, traffic operation authorities should concentrate on managing Non-recurring congestion on the road (Lee, 2010). However, strategies for the Non-recurrent congestion are hard to be developed because its occurrence and effects are difficult to be expected (Baek, 2004).

Meanwhile, the existing manual for traffic accident on Korean highway has many limitations. The countermeasure is composed of two steps. First, classification of traffic accidents into three levels according to severity of crashes. Second, determination of what kind of operation strategy is needed based on severity level (Kang, 2008).

However, the two-steps-plan is improper to manage traffic accident and traffic flow on highway efficiently. The reason is because there is no prearranged plan for general crashes on the highways. The two steps mentioned above are only applied to the accidents occurred on the bridge or tunnel (Kang, 2008). Moreover, the effects of traffic accidents on traffic flow are analyzed based on accident severity like death rate, relevant vehicles, traffic blockage only. On the other hand, deterioration of traffic flow (decrease of speed and increase of density) is affected by more diverse factors: Traffic flow characteristics, road geometry, accident severity etc. (Giuliano, 1989) Thus, expectation of traffic flow status after the accident based on existing plan cannot be accurate. This in turn, the operation to recover the traffic flow becomes inefficient.

The countermeasure established by the two-step-plan is also not focused on demand control such as detour and lane blockade, but only on the clearance of accidents. By doing so, traffic accident operation for preventing congestion is restrictedly done. Therefore, in order to be more efficient, demand control is indispensable.

For resolving these problems, this research will be focused on traffic accident accounting for 25% of total congestion (Tavassoli Hojati, 2010). The traffic accident data and Vehicle Detectors (VDs) data are analyzed for modeling to expect the effects of traffic accident on traffic flow. The analysis is carried out to identify significant factors and influence of them to traffic flow and determine the extent of demand control for efficient traffic accident management.

1.1.2 Research Purpose

In order to maintain traffic flow in stable condition after a traffic accident, we should constitute analytical plans handling traffic accidents. The plans firstly expect the impacts of traffic crashes on traffic flow characteristics and then decide the extent of demand control according to accident types, management criteria.

The main purpose of this research is to provide rationale base for the plan using statistical method. Then, there are three sub-purpose to accomplish the main goal efficiently.

First, this research will analyze effects of various factors related with the traffic accidents on traffic flow. It provides highway operators with knowledge that what kinds of management skills should be done in the specific situation.

Second, this research will analyze hazard of traffic accident to deteriorate traffic flow characteristics like increase in density according to types of traffic accidents. Using hazard of traffic accidents, traffic operators can determine what type of accidents should be handled first for efficient management under restricted resources.

Third, this research includes an analysis of to what extent demand control of each type of traffic accident is needed to prevent congestion. If traffic operator knows this information at the occurrence, traffic operator can control inflow properly and prevent further congestion caused by crashes.

Through achieving these purposes, traffic operator could make timely treatment for general accidents. This guarantees road users and traffic operators to prevent severe traffic turbulent, additional accidents (Fang, 2013; Baykal-Gursoy 2006). Also, it can contribute to a quick recovery of traffic flow to normal status.

1.2. Research Composition

In Chapter 2, four-parts of literature review is done. 1) Analyzing impacts of traffic flow on traffic accidents 2) Analyzing impacts of traffic accidents on traffic flow using simulation 3) Analyzing impacts of traffic accidents on the traffic flow using statistical methods 4) Analyzing impacts of traffic accident on traffic flow using queueing model

In Chapter 3, data collected from Korea-Expressway Corporation for research is outlined. And the research suggests how to filter data and how to merge data for analysis.

In Chapter 4, methodology is outlined: What kind of statistical model and variables are used, how model is constituted for analysis.

In Chapter 5, estimation results are suggested. The effects of various factors on traffic flow and hazard model which can be used to draw relative hazard are proposed.

In Chapter 6 includes conclusions and limitation of this research

Chapter 2. Literature Review

This research reviews existing papers which dealt a relationship of traffic accidents and traffic flow. Existing papers can be partitioned into four parts.

First, analyzing impacts of traffic flow on traffic accidents (Oppe, 1979; Ivey, 1980; Ceder, 1980; Accid, 1981; Jovanis, 1989; Hiselius, 2004). Second, analyzing impacts of traffic accidents on traffic flow using simulation tools (Nakatani, 1993; Kurata, 2002; Zhu, 2008; Fang, 2013; Chen, 2014). Third, analyzing impacts of traffic accidents on traffic flow using statistical methods (Golob, 1986; Giuliano, 1989; Nam, 2000; Golob, 2003; Chung, 2007; Lee, 2010; Dimitrious, 2015). Fourth, analyzing impacts of traffic accidents on traffic flow using queueing model (Morale, 1987; Olmstead, 1999; Baykal-Gursoy, 2006)

2.1. Analyzing impacts of traffic flow on traffic accidents

The research in 20th Century mainly focused on expecting traffic accidents frequency and crash risk based on traffic flow characteristics. These research are classified by what kind of variables and methodologies are used for explaining traffic accident risk. Oppe (1979) explained accident ratios by surface skidding resistance and hourly traffic volume using Conjoint Measurement model. Ceder and Livneh (1980) used accident measures considering exposure time like accident density per an hour exposure of traffic flows. This research predicted the accident measures by hourly traffic volume using regression model. Ivey et al. (1980) expected accident rate in wet weather by traffic flow characteristics, highway geometric, pavement surface and rainfall exposure characteristics using regression model. Jovanis (1989) used accident measure reflecting exposure to accident risk. For explaining the accident measure, the researcher took variables like accident time whether accidents occurred at winter and accidents occurred at night, the years of driving experiences, cargo weight, number of hours on-duty, number of consecutive hours off-duty. This paper predicted accident measure for each accident severity and type of collision targeting truck accidents using survival theory. Hiselius (2004) estimated accident frequency considering hourly traffic volume, and heterogeneous vehicles types using regression model.

These research showed effects of various factors like traffic flow on accident occurrence and risk using mainly regression model. Based on these results, traffic operators could prevent accidents. However, traffic operators also have to be aware of how to deal with the occurred accidents to prevent congestion.

<Table 1> Research on impacts of traffic flow on traffic accidents

Author(s)	Year	Variables	Methodology
Oppe.	1979	Surface skidding resistance Hourly traffic volume	Conjoint Measurement Model
Ceder and Livneh	1982	Hourly traffic volume	Regression model
Ivey et al.	1980	Traffic flow characteristics Highway geometric Pavement surface Rainfall exposure characteristics	Regression model
Jovanis	1989	Accident time The years of driving experiences Cargo weight The number of hours on-duty The number of consecutive hours off duty	Survival theory
Hiselius	2004	Hourly traffic volume Vehicle type	Regression model

2.2. Analyzing impacts of traffic accidents on traffic flow using simulation

There are some research analyzing effects of traffic accidents on traffic flow by employing simulation tools. Research taking simulation as analysis tool has been developed to improve vehicle behavioral model for reflecting realistic vehicle movement and to extend lanes of targeted highway set for analysis. These researches are classified by what kind of methodologies is used which depicts vehicle movement when occurs accident on the highway. Also, they can be partitioned by how many lanes are assumed for simulation analysis.

Nakatani (1993) simulated traffic congestion caused by accident in the two-lane highway. Car behavior was regulated by deterministic cellular automation model. Kurata (2002) simulated traffic jam induced by accident in the two-lane highway also. In this paper, car movement was depicted by extension model of the optimal velocity model considering lane changing on the highway. Zhu (2008) simulated jamming transition and reduction of capacity caused by accident taking into account of blockage of specific lane on the two-lane highway. In this article, the physics of congestion on the highway was studied with cellular automaton model considering asymmetric lane changing rules other than symmetric lane changing rules. Fang (2013) simulated traffic flow induced by typical road accident on the two-lane highway using ANSYS software. In this research, traffic flow model was analyzed fluid mechanics theory. Chen (2014) simulated car accident on the three-lane highway. The model used in this paper was extension of full velocity difference model reflecting lane changing behavior on the highway.

Simulation method has many advantages but also many definite disadvantages in the case of analyzing effect of accident on traffic flow. First, it is difficult to define meaningful factors contributing to outbreak of congestion and analyze the extent of contribution of the factors. Therefore, simulation results cannot help traffic operators to determine what type of strategies needed for preventing congestion because they cannot find the factors that bring about congestion. Second, the analysis using simulation has not been done to four-lane, five-lane highway which is common in Korea. Third, simulation takes a long time to get reliable results because of many tries needed to decrease variation of results (Baykal-Gursoy, 2006)

<Table 2> Research on impacts of traffic accidents on traffic flow using simulation

Author(s)	Year	Traffic flow model	Number of lane
Nakatani.	1993	Deterministic cellular automation model	2
Kurata	2002	Extension model of the optimal velocity model	2
Zhu	2008	Cellular automaton model	2
Fang	2013	Fluid mechanics theory	2
Chen	2014	Extension of full velocity difference model	3

2.3. Analyzing impacts of traffic accidents on traffic flow using statistical methods

The part of research studying impacts of traffic accidents on traffic flow uses various statistical methods. Most of the studies represented effects of accidents on traffic flow as accident duration because the longer accident duration usually means accident's influence to traffic flow also be longer and severe (Khattak et al, 2009)

These research are divided by what kind of variables and methodologies are employed for explaining influence of traffic accidents. Golob (1986) classified truck-related accident type according to collision type, lane closure, the number of fatalities and injuries. And the research draws mean and variation of accident duration each accident type. Giuliano (1989) divided accidents by accident type, time of day, whether trucks are involved, lane blockage. It carried out an analysis of variance for selecting main factors among variables mentioned above. The research also calculated the means and the variations of accident duration for each accident set categorized by main factors. Nam and Mannering (2000) used temporal characteristics, environmental characteristics, geographic information, incident characteristics to expect incident detection time, reporting time, response time, clearance time. This paper with hazard-based analysis found that several variables affect each accident duration times significantly.

Golob (2003) characterized the accident by accident types and locations, the number of vehicles, the behavior of vehicle before occurring traffic accident, accident severity considering weather and lighting environment on the highway. It analyzed relationships between each traffic accident type and traffic flow represented by speed, volume different from other researches targeting accident duration. In order to draw results, he performed principal components analysis and nonlinear canonical correlation analysis. Lee (2010) selected significant factors among accident characteristics, traffic flow characteristics, accident time, distance between accident spot and detector, road geometry by genetic algorithm. This research made two artificial neural network-based models based on significant factors selected above for sequent prediction of accident duration. Dimitrious (2015) collected data of variables: whether accident type is primary accident or secondary, collision severity, number of lanes, number of vehicles involved, involvement of heavy vehicle, traffic flow characteristics, rainfall, alignment, geometry for predicting accident duration. The research took fuzzy rule-base system as methodologies and compared results with conclusion from other model like survival analysis.

However, these research did not analyze the impacts of traffic accidents on traffic flow directly using variation of traffic volume, density, speed but indirectly investigated the effects by the predicting time taken for handling the accident. Since the influence of traffic accident is not determined by accident duration only, more research which directly analyze the influence considering various significant factors are needed for accurate prediction.

<Table 3> Research on impacts of traffic accidents on traffic flow using statistical methods

Author(s)	Year	Variables	Methodology
Golob.	1986	Collision type Lane closure The number of fatalities and injuries	Calculating Mean and variation
Giuliano	1989	Accident type Time of day Involvement of truck Lane closure	Calculating Mean and variation
Nam and Mannering.	2000	Temporal characteristics (Seasons, Time of day) Weather Geographic characteristics (District, HOV） The number of vehicles involved Fatal and injury	Hazard based analysis
Golob	2003	Accident type Geographic characteristics (Accident location) The number of vehicles involved The behaviour of vehicle before occurring traffic accident Accident severity Weather Lighting environment	Principle components analysis Nonlinear canonical correlation analysis
Lee	2010	Accident characteristics Traffic flow characteristics Accident time Distance between accident spot and detector Geometry	Genetic algorithm Artificial neural network based model
Dimitrious	2015	Accident type (Primary accident, secondary accident) Collision severity The number of lane The number of lane involved Involvement of heavy vehicles	Survival analysis

	Traffic flow characteristics Rainfall Geometric characteristics (Alignment)

2.4. Analyzing impacts of traffic accidents on traffic flow using queuing model

There are several research analyzing impacts of traffic accidents on traffic flow using queuing model to calculate accident-induced delay. Morale (1987) presented the analytical procedure and micro-computer model by means of queuing model. Olmstead (1999) posited new method using deterministic queuing model that can draw incident-induced total delay more accurately by changing the notion of delay “delay due an average incident” to “average delay due to incidents”. Baykal-Gursoy (2006) calculated the delay time induced by traffic accident using the analytical queuing model presented by Baykal-Gursoy and Xiao (2004) and compared the result with a simulation model.

However, queuing model that calculates total delay caused by accidents has the same defects as the simulation method. The model cannot define significant variables related to the congestion and also fails to analyze the level of influence.

Chapter 3. Data Collection

In order to analyze impacts of traffic accidents on traffic flow, the data about traffic accidents, traffic flow is needed. For this research, the accident record data is collected from Korea-Expressway Corporation. The accident record is ranged from 2010 to 2013 and it is collected from the Gyeongbu line (to Seoul), the Yeongdong line (to Gangneung), the Seohaean line (to Seoul). Also, VDs data recorded every 5 minute on the highway is acquired from Korea-Expressway Corporation for getting traffic flow characteristics information. VDs data is also ranged from 2010 to 2013.

In this paper, accident cases and VDs data are combined according to procedures written below for making data set. The accident cases are combined with the two different types of VDs data; pre-accident data and post-accident data. The pre-accident VDs data is aggregated before accident 5 minutes ago for representing traffic flow status prior to occurring accidents. The post-accident VDs data is collected during accident duration for representing traffic flow status under influence of accident. The duration represents the time between accident occurrence and accident clearance.

<Table 4> List of variables on accident record

Category	Variables
Accident characteristics	Cause of accident, Depiction of accident, Collision type
Accident time	Accident occurrence time
Accident location	Distance to accident site, Name of lane, Direction of lane, Lane accident occurred
Accident severity	The number of fatalities and injuries. The number of vehicles involved, Damage of vehicles
Geometric Characteristics	Alignment, vertical curve
Accident duration	Reporting duration, Response duration, Clearance duration Traffic flow control duration
Weather	Weather
Vehicle type	Vehicle type involved accident
Construction	Whether construction is done or not
Traffic flow characteristics	Volume, Occupancy, Speed

Data filtering is done only to study the influence of traffic accident on traffic flow among factors causing non-recurrent congestion. Filtering also is carried out to delete data which has not essential information for analysis. The data filtering has been conducted by following procedures.

Firstly, all data except from clear days are removed to exclude weather factor. Secondly, all data except from non-construction data are ruled out to exclude construction factor. Thirdly, all data, which has not necessary information for study such as lane blockage, are removed. Fourthly, because this research takes into account of influence of demand, the study removed accident data occurred near ramp. This filtering is done to complement absence of volume data entering from ramp. If we use only the accident data at some distance from ramp, researcher could assume volume from second upstream VD as demand at this period. Fifthly, the accidents which are blocked several lanes due to lack of cases and difficulty of analysis. As a result, 1413 data has been selected for this research.

Chapter 4. Methodology

4.1 Statistical Model

The analytical models are considered to analyze the impacts of accidents on the traffic flow. Among them, simulation model and queuing model have disadvantages that the methods cannot define what are significant factors to causing congestion and their level of influence. This defects are critical to traffic operators because they cannot know what kind of operation are needed to prevent congestion.

For overcoming the problems, this research will employ multi-variates survival analysis which can contribute to find causes of congestion and their effects on traffic flow. Moreover, Survival analysis is appropriate method when researcher uses data which is combined discrete data with aggregate data for analysis. If these data are combined improperly, relationship between discrete data and aggregate data can be blurred. Survival analysis can merge them into a mathematically consistent disaggregate level (Jovanis, 1989). So, survival analysis is suitable method for this research using discrete accident data and aggregated VDs data. Also, survival model analyzes censored data which means not-occurring-event data for estimating the effect of covariates. If study is done without censored data, the coefficient of factors could be overestimated or underestimated (Lee, 2016).

Survival analysis is statistical method which analyzes duration time until the event occurs for various purposes (Cox, 1972). Among models of survival analysis, this research uses cox proportional hazard model because of several reasons. Firstly, cox model is multi-variates survival model which is an essential characteristic for analyzing the effect of various factors. Secondly, the cox model is semi-parametric model that is not supposing distribution of targeted data. Therefore, there is not a large bias which is occurred when researcher assume inaccurately the distribution of data. Since the duration time has skewed-distribution, it is important to select appropriate model for the distribution (Han, 2013).

The cox proportional hazard model is outlined as function (1)

$$h(t, x) = h_0(t) * \exp(\boldsymbol{\beta x}) \quad (1)$$

Where,

$h_0(t)$: Baseline hazard function

$\boldsymbol{\beta}$: Vector of coefficient

$\mathbf{x}$: Vector of factor

The Baseline hazard function is hazard when all covariates are zero (Han, 2013). By using cox proportional hazard model, research can get hazard ratio of occurring the event between various cases at specific time *t*.

4.2 Variables

Using all usable data for revealing the effect of accident is not desirable. Too many variables can increase complexity of correlation between various factors. The increase in the complexity may generate noise in the model that caused inaccurate expectation (Lee, 2010).

Thus, this research chooses the variables representing accident (independent variables) and traffic flow (event variable). For the first step, this research reviewed the related papers to see what they used at <table 4> for representing accident. Most of papers used accident type, accident severity, lane closure, temporal characteristics, environmental characteristics, geographic characteristics, accident location, lighting environment, vehicle behavior before accident occur, traffic flow characteristics. Then, this research selects variables can be observed immediately after accident occurs. This is because the purpose of research is to expect traffic flow status after traffic accident in order to conduct appropriate operation strategy at the initiatory stage. It is also important to remove highly correlated variables. Therefore, the research also conducted correlation analysis.

<Table 4> List of variables on this research

Category	Variables
Accident time	Day:0, Night: dummy
Accident severity	The number of fatalities and injuries. The number of vehicles involved, Damage of vehicles
Geometric characteristics	Alignment (Straight lane: 0, Left curve: dummy, Right curve: dummy) Vertical curve(Flat lane: 0, Ascent: dummy, Descent: dummy
Demand	Volume from second upstream VD
Vehicle type	Vehicle type involved accident(Passenger car: 0, Trailer: dummy, Truck: dummy
Lane blockage	Location of lane blockage(Shoulder: 0, Leftmost lane: dummy, Middle lane: dummy, Rightmost lane: dummy)
Traffic flow characteristics	Volume, Speed, Density from first upstream VD

Meanwhile, variation of density calculated by green-shield method is selected as a representative of effects on traffic flow characteristics. There are many reasons that the density is chosen as an index among other characteristics. First, the index has to be proportional to congestion incurred by accidents. Density and speed can change consecutively with growth of congestion. However, the volume increases until the saturated flow level and decreases after. Therefore, it is adequate to use density or speed as the index. Second, it is more convenient to use density as index when traffic operator conducts strategies such as ramp-metering. It is because he wants to know how many vehicles can be entered highway mainline without incurring congestion.

<Table 5> Descriptive Statistics of the research's data

Index	Frequency (case)	Proportion (%)
Collision type		
Vehicle-facility	734	51.8
Vehicle-vehicle	328	23.1
Vehicle-human	15	1.1
Etc	341	24.0
Accident type		
Day	818	56.8
Night	621	43.2
The number of fatalities and injuries	1439	100.0
Fatality	62	4.3
Injury	255	17.7
The number of vehicles involved		
1	813	56.5
2	436	30.3
3	114	7.9
4	36	2.5
≥5	40	2.7
Damage of vehicles	1439	100.0
Vehicle crushed totally	77	5.4
Vehicle crushed half	235	16.3
Vehicle crushed 1/3	797	55.4
Vehicle crushed slightly	709	49.3
Alignment		
Straight lane	1084	75.3
Left lane	182	12.6
Right lane	173	12.0
Vertical curve		
Flat lane	840	58.4
Ascent	310	21.5
Descent	289	20.1
Vehicle type involved accident	1439	
Passenger car	433	30.1
Trailer	113	7.9
Truck	87	6.1
Location of lane blockage		
Shoulder	948	54.5
Leftmost lane	300	17.2
Middle lane	234	13.4
Rightmost lane	259	14.9

4.3 Model constitution

Survival analysis has to set an event for analysis. Level of Service (LOS) presented by Ministry of Land, Infrastructure and Transport is selected as event criteria.

According to LOS, the model is divided into three parts. Model 1 consists of status changing from LOS A and LOS B (before accident) to LOS C (after accident), Model 2 represents status changing from LOS A and LOS B to LOS E, and Model 3 observed status changing from LOS C and LOS D to LOS E. The reason for rearranged six level-LOS into three levels in this analysis is that difference of traffic flow characteristics among six levels is not that large (Cho, 2013).

Model 1 and Model 2 are used to compare the effects of factors according to level of congestion incurred by the accidents. Model 2 and Model 3 are for comparing effect of the covariates according to status of traffic flow before accidents because impacts of traffic accident on traffic flow is different in free-flow and congested-flow (Ceder. 1980)

The expected results from this model is as following; the types of accident that has a greater potential on congestion, the strategies needed to prevent congestion at initial stages, and the extent of operations needed to stave off the traffic jam.

Chapter 5. Estimation Results

5.1. Model 1 (From LOS A and B to LOS C)

Cox proportional hazard model for Model 1 is derived from 1038 cases using the independent variables selected at the Chapter 4. The significant variables are selected by forward selection method which is provided from SPSS. Addition to these selected variables, some variables which are needed to analysis are added. The form of the model is suggested in Equation 2:

$$h(t,x) = h_0(t) * \exp(\beta_1 x_{total} + \beta_2 x_{half} + \beta_3 x_{\frac{1}{3}} + \beta_4 x_{Speed} + \beta_5 x_{vol}$$
$$+\beta_6 x_{demand} + \beta_7 x_{trail} + \beta_8 x_{ri-cur} + \beta_9 x_{le-lane} + \beta_{10} x_{ri-lane}$$
$$+\beta_{11} x_{mi-lane} + \beta_{12} x_{demand} * x_{le-lane} + \beta_{13} x_{demand,ri-lane} \quad (2)$$
$$+\beta_{15} x_{demand} * x_{mi-lane}$$

Where:

x_{total}	: The number of crushed totally(veh)	
x_{half}	: The number of crushed half(veh)	
$x_{1/3}$	: The number of crushed 1/3(veh)	
x_{speed}	: Average speed before accident(km/h)	
x_{vol}	: Volume at 1st upstream VD(veh/min/lane)	
x_{demand}	: Volume at 2nd upstream VD(veh/min/lane)	
x_{trail}	: 1, if involvement of trailer	; 0, otherwise
x_{ri-cur}	: 1, if alignment at accident location is right curve	; 0, otherwise
$x_{le-lane}$	: 1, if vehicle blocked the leftmost lane	; 0, otherwise
$x_{ri-lane}$	: 1, if vehicle blocked the rightmost lane	; 0, otherwise
$x_{mi-lane}$	: 1, if vehicle blocked lanes except leftmost, rightmost lane	; 0, otherwise

When comparing with the constant only model, the hazard function drawn using cox method is statistically significant ($\chi^2(14) = 443.613,\ p < .01$). Moreover, at the confidence level of 99%, the majority of coefficients are significant.

<Table 6> Estimation results: From LOS A and B to LOS C

Variable	β	Std. err	exp(β)
Traffic flow characteristics			
Average Speed	-.044***	.006	.957
Volume	.209***	.027	1.233
Accident severity			
Vehicle crushed totally	1.207***	.264	3.345
Vehicle crushed half	.571***	.170	1.771
Vehicle crushed 1/3	.355***	.118	1.426
Demand			
Volume at 2nd VD upstream	.127***	.022	1.135
Volume * leftmost lane	.044	.039	1.045
Volume * middle lane	.042	.049	1.043
Volume * rightmost lane	.019	.038	1.019
Blocked lane			
Shoulder	Reference		1
Leftmost lane	.545	.492	1.725
Middle lane	.888	.610	2.431
Rightmost lane	.681	.510	1.976
Type of involved vehicle			
Car	Reference		1
Trailer	.649***	.222	1.914
Alignment			
Straight lane	Reference		1
Right curve	-.476**	.212	.621

$*\ p < .1,\ **\ p < .05,\ ***\ p < .01$

The results are drawn on the influences of factors which occur congestion.

1) Traffic characteristics before accidents can affect highway LOS. For example, average speed has a coefficient of β= -0.044. It means that velocity increase of 1km/h decreases 4.3% ($\exp(\beta) = 0.957$) hazard of occurring congestion. The volume has β= 0.209. That is, volume increment of 1 veh/min/lane increases 23.3% hazard. Analysis based on speed and volume factors concludes that the more stable traffic flow before accident, the less hazard of resulting worse LOS.

2) If accident is severe in terms of degree of vehicle damage, the hazard increased proportionally. The case of vehicle crushed totally has the highest level of hazard (β= 1.207).

3) The high demand at upstream has adverse effect on the hazard. The effect of demand is different according to location of lane blockage by vehicles. When leftmost lane is blocked, the influence of demand is most severe (β= 0.127+0.44 = 0.171).

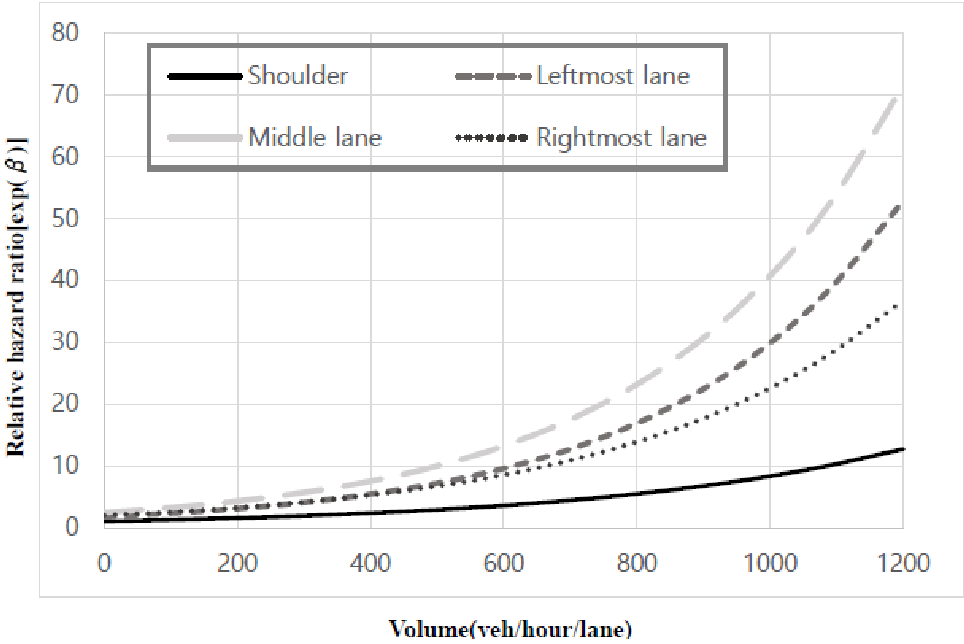
4) Location of blocked lane itself has adverse influence on the hazard. Among three locations of lane, the case of middle lane blockage is risky the most(β= 0.888).

5) The type of involved vehicle has the influence to LOS. The heavy vehicle like trailer has worse effect on traffic flow relative to the passenger car.

6) Alignment of accident location also need to be considered when calculates the hazard of traffic accident. In this analysis, the accident occurred on right curve has more hazard than straight lane.

Besides, the hazard model can be used when traffic operator compares hazard of various accidents. For example, hazard ratio of Case 1 (Average speed: 90km/h; volume: 10veh; vehicle crushed totally: 1veh; vehicle crushed half: 1veh; demand: 10veh; blocked lane: middle lane; type of vehicle involved: truck; Alignment: straight lane) and Case 2 (Average speed: 80km/h; volume: 12veh; vehicle crushed half: 2veh; demand: 13veh; blocked lane: shoulder; type of vehicle involved: passenger car; Alignment: straight lane) is about 3.87. It means that hazard of Case 1 accident is 3.87 times the hazard of Case 2. So, traffic operator should treat Case 1 prior to Case 2 for efficient operation under restrictive condition.

Also, the hazard model suggests the extent of upstream volume control considering type of accidents. In this research, location of lane blockage is taken as classification standard to analyze the volume control. The analysis is carried out under condition that other variables are identical. As a result, <Figure 1> is presented.



<Figure 1> Relative hazard change by volume at upstream from accident (From LOS A and B to LOS C)

According to <Figure 1>, the influence of volume at upstream can be ordered Middle lane - Left lane - Right lane - Shoulder. The reason why the accident (vehicle is stalled at middle lane) has the biggest influence on traffic flow is that the traffic flow is interfered at two points (left and right of accident spot) with increasing number of vehicles from upstream. The other hand, the accidents (vehicles are stalled at leftmost and rightmost lane) interfere traffic flow at the nearest lane with increment of demand. <Figure 1> suggested that the extent of demand control should be conducted to secure stable traffic flow status. Using this, the strategies can be optimized for efficient operations.

5.2. Model 2 (From LOS A and B to LOS E)

The Cox model for Model 2 is derived from 1038 cases using the independent variables selected at Chapter 4. The significant variables are selected by forward selection method. And some variables which are needed to analysis are added. The form of the model is suggested in Equation 3:

$$h(t,x) = h_0(t) * \exp(\beta_1 x_{total} + \beta_2 x_{half} + \beta_3 x_{Speed} + \beta_4 x_{vol} + \beta_5 x_{demand} + \beta_6 x_{le-lane} + \beta_7 x_{ri-lane} + \beta_8 x_{mi-lane} + \beta_9 x_{demand} * x_{le-lane} + \beta_{10} x_{demand} * x_{ri-lane} + \beta_{11} x_{demand} * x_{mi-lane}) \quad (3)$$

Where:

x_{total}	: The number of crushed totally(veh)	
x_{half}	: The number of crushed half(veh)	
x_{speed}	: Average speed before accident(km/h)	
x_{vol}	: Volume at 1st upstream VD(veh/min/lane)	
x_{demand}	: Volume at 2nd upstream VD(veh/min/lane)	
$x_{le-lane}$	: 1, if vehicle blocked the leftmost lane	; 0, otherwise
$x_{ri-lane}$	: 1, if vehicle blocked the rightmost lane	; 0, otherwise
$x_{mi-lane}$	: 1, if vehicle blocked lanes except leftmost, rightmost lane	; 0, otherwise

When compare with the constant only model, the hazard function drawn using cox method is statistically significant ($\chi^2(13) = 292.141,\ p < .01$). Moreover, at confidence level of 95%, the part of coefficients are shown to be significant.

<Table 7> Estimation results: From LOS A and B to LOS E

Variable	β	Std. err	exp(β)
Traffic flow characteristics			
Average Speed	-.034***	.009	.966
Volume	.166***	.037	1.181
Accident severity			
Vehicle crushed totally	.718**	.363	2.051
Vehicle crushed half	.678***	.200	1.970
Demand			
Volume at 2nd VD upstream	.131***	.041	1.140
Volume * leftmost lane	.000	.049	1.000
Volume * middle lane	.073	.065	1.075
Volume * rightmost lane	.039	.056	1.040
Blocked lane			
Shoulder	Reference		1
Leftmost lane	1.689**	.662	5.414
Middle lane	1.221	.854	3.391
Rightmost lane	1.216	.757	3.375

$* \ p < .1, ** \ p < .05, *** \ p < .01$

The results are drawn on the influences of factors which occur congestion

1) Traffic characteristics before traffic accidents can affect highway LOS. For example, average speed has coefficient as β= -0.034. It means that velocity increase of 1km/h decreases 3.7% ($\exp(\beta) = 0.966$) hazard of occurring congestion. The volume has β= 0.166. That is, volume increment of 1 veh/min/lane increases 18.1% hazard. Analysis based on speed and volume factors concludes that the more stable traffic flow before accident, the less hazard of resulting worse LOS.

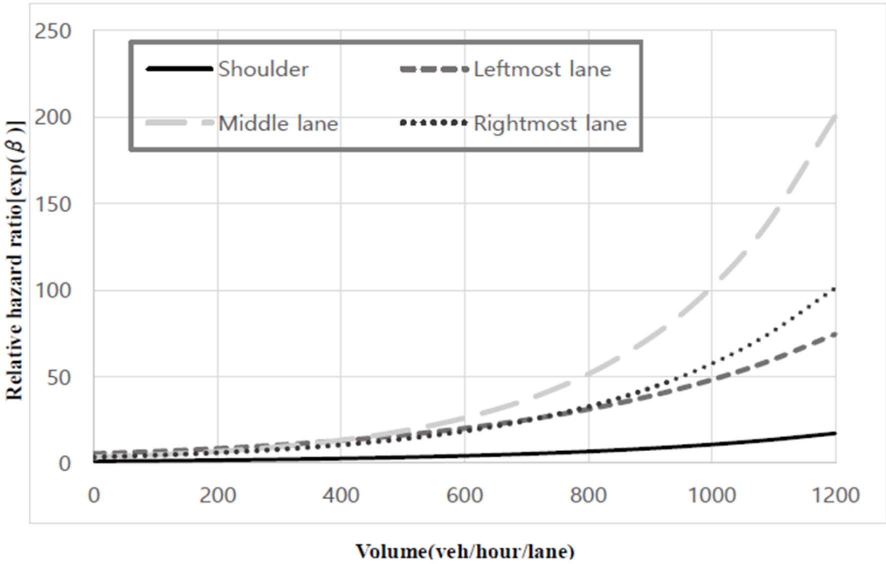
2). If accident is severe in terms of degree of vehicle damage, the hazard increased proportionally. The case of vehicle crushed totally has the highest level of hazard (β= 0.718).

3) The high demand at upstream has adverse effect on the hazard. The effect of demand is different according to location of lane blockage by vehicles. When middle lane is blocked, the influence of demand is most severe (β= 0.131+0.073 = 0.204).

4) Location of blocked land itself has adverse influence on the hazard. Among three locations of lane, the case of leftmost lane blockage is risky most (β= 1.689).

Besides, the hazard model can be used when traffic operator compares hazard of various accidents like 5.1.

Also, the hazard model suggests the extent of upstream volume control considering type of accidents. In this research, location of lane blockage is taken as classification standard to analyze the volume control. The analysis is carried out under condition that other variables are identical. As a result, <Figure 2> is presented.



<Figure 2> Relative hazard change by volume at upstream from accident (From LOS A and B to LOS E

According to <Figure 2>, the influence of volume at upstream can be ordered Middle lane - Right lane - Left lane - Shoulder. The reason why the accident (vehicle is stalled at middle lane) has the biggest influence on traffic flow is that the traffic flow is interfered at two points (left and right of accident spot) with increasing number of vehicle from upstream. The other hand, the accidents (vehicles are stalled at leftmost and rightmost lane) interfere traffic flow at the nearest lane with increment of demand. As a result, the hazard of middle lane-accident overtakes the hazard of left lane-accident at 400veh. <Figure 2> suggested that the extent of demand control should be conduct to secure stable traffic flow status. Using this, the strategies can be optimized for efficient operations.

5.3. Model 3 (From LOS C and D to LOS E)

Cox proportional hazard model for Model 3 is derived from 152 cases using the independent variables selected at Chapter 4. The significant variables are selected by forward selection method. And some variables which are needed to analysis are added. The form of the model is suggested in Equation 4:

$$h(t,x) = h_0(t) * \exp(\beta_1 x_{den} + \beta_2 x_{demand} + \beta_3 x_{ri-cur} + \beta_4 x_{le-lane} + \beta_5 x_{ri-lane} + \beta_6 x_{mi-lane}) \quad (4)$$

Where:

x_{den}	: Density before accident(veh/km)	
x_{demand}	: Volume at 2nd upstream VD(veh/min/lane)	
x_{ri-cur}	: 1, if alignment at accident location is right curve	; 0, otherwise
$x_{le-lane}$	: 1, if vehicle blocked the leftmost lane	; 0, otherwise
$x_{ri-lane}$	: 1, if vehicle blocked the rightmost lane	; 0, otherwise
$x_{mi-lane}$	: 1, if vehicle blocked lanes except leftmost, rightmost lane	; 0, otherwise

When compare with the constant only model, the hazard function drawn using cox method is statistically significant ($\chi^2(9) = 41.395,\ p < .01$). Moreover, at confidence level of 90%, the majority of coefficients are shown to be significant.

<Table 8> Estimation results: From LOS C and D to LOS E

Variable	β	Std. err	exp(β)
Traffic flow characteristics			
Density	.120***	.046	1.128
Demand			
Volume at 2nd VD upstream	.062**	.029	1.064
Volume * leftmost lane	-.075*	.041	0.927
Volume * middle lane	-.054	.054	0.948
Volume * rightmost lane	.017	.065	1.017
Blocked lane			
Shoulder	Reference		1
Leftmost lane	2.303***	.832	10.007
Middle lane	1.883	1.207	6.572
Rightmost lane	.332	1.271	1.394
Alignment			
Straight lane	Reference		1
Right curve	.601*	.342	1.824

* $p < .1$, ** $p < .05$, *** $p < .01$

The results are drawn on the influences of factors which occur congestion.

1) Traffic characteristics before traffic accidents can affect highway LOS. Demand factor has coefficient as β= 0.120. It means that density increment of 1 veh/km increases 12.8% ($\exp(\beta) = 1.128$) hazard of occurring congestion. Analysis based on density concludes that the more stable traffic flow before accident, the less hazard of resulting worse LOS.

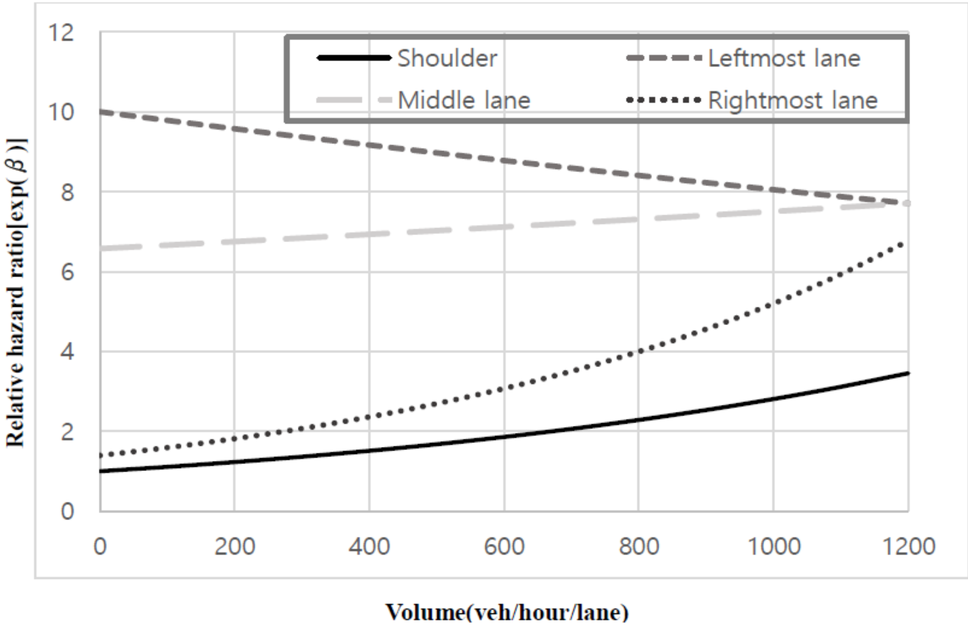
2) The high demand at upstream has adverse effect on the hazard. And the effect of demand is different according to location of lane blockage by vehicles. When rightmost lane is blocked, the influence of demand is most severe (β= 0.062+0.017 = 0.079).

3) Location of blocked land itself has adverse influence on the hazard. Among three locations of lane, the case of leftmost lane blockage is risky most (β= 2.303).

4) Alignment of accident location also needed to be considered when calculates hazard of traffic accident. In this analysis, the accident occurred on right curve has more hazard than straight lane.

Besides, the hazard model can be used when traffic operator compares hazard of various accidents like the models above.

Also, the hazard model suggests the extent of upstream volume control considering type of accidents. In this research, location of lane blockage is taken as classification standard to analyze the volume control. The analysis is carried out under condition that other variables are identical. As a result, <Figure 3> is presented.



<Figure 3> Relative hazard change by volume at upstream from accident (LOS C and D to LOS E

According to <Figure 3>, the influence of volume at upstream can be ordered Right lane - Shoulder - Right lane – Left lane. The covariates of the volume when accident vehicle is stalled on leftmost lane is minus. And overall, the effects of the volume are small in LOS C & D cases.

From this, this research finds that traffic flow is mainly influenced by reduction of capacity such as lane closure when accident is occurred on congested-flow. In other words, congestion is not occurred by accumulation of queue when there are enough vehicles but reduction of capacity

Chapter 6. Conclusion

This three-model analysis drew a result of what kind of operations and extents of operation are required to prevent congestion.

The hazard model showed that various factors of traffic accident affects to traffic flow. If traffic flow before the accident is stable, the hazard of getting worse LOS. Moreover, hazard increased proportionally with the degree of vehicle damages and volume at upstream. The location of lane closure and involvement of heavy vehicle, alignment also have an effect on occurring congestion.

Meanwhile, influence of the factors is different according to traffic flow characteristics before accident and the management goal. The cases of Model 1 and Model 3 are influenced by accident severity more than Model 2. However, the cases of Model 2 are affected by capacity decrease and demand. It may be because Model 2 represents a case, which converges from free flow to severely congested flow. Thus, severely accumulated queue are required for incurring congestion.

Moreover, the research suggests that what extent of volume control should be carried out according to the spot of lane blockage for reducing hazard of congestion. Through applying this knowledge, highway operators can control inflow to mainline moderate level.

In the one hand, the further research should overcome a limitation of using 5 minutes aggregated data by applying more discrete data. By acquiring more data, the further research should include extra variables and analyze the effect of factors more accurately.

References

- Cox, D.R., (1972). Regression models and life tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 187-220.
- Oppe, S. (1979). The use of multiplicative models for analysis of road safety data. Accident Analysis & Prevention, 11(2), 101-115
- Ivey, D. L., Griffin, L. I., Newton, T. M., Lytton, R. L., & Hankins, K. C. (1981). Predicting wet weather accidents. Accident Analysis & Prevention, 13(2), 83-99.
- Ceder, A., & Livneh, M. (1982). Relationships between road accidents and hourly traffic flow: analyses and interpretation. Accident Analysis & Prevention, 14(1), 19-34..
- Golob, T. F., Recker, W. W., & Leonard, J. D. (1987). An analysis of the severity and incident duration of truck-involved freeway accidents. Accident Analysis & Prevention, 19(5), 375-395.
- Morales, J. M. (1987). Analytical procedures for estimating freeway traffic congestion. ITE journal, 57(1), 45-49.
- Jovanis, P. P., & Chang, H. L. (1989). Disaggregate model of highway accident occurrence using survival theory. Accident Analysis & Prevention, 21(5), 445-458.
- Wei, L. J. (1992). The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis. Statistics in medicine, 11, 1871-1879.
- Nagatani, T. (1994). Dynamical jamming transition induced by a car accident in traffic-flow model of a two-lane roadway. Physica A: Statistical Mechanics and its applications, 202(3), 449-458.
- Olmstead, T. (1999). Pitfall to avoid when estimating incident-induced delay by using deterministic queuing
- Dickerson, A., Peirson, J., & Vickerman, R. (2000). Road accidents and traffic flows: An econometric investigation. Economica, 67(265), 101-121.
- Nam, D., & Mannering, F. (2000). An exploratory hazard-based

analysis of highway incident duration. Transportation Research Part A: Policy and Practice, 34(2), 85-102.

- Golob, T. F., & Recker, W. W. (2003). Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. Journal of transportation engineering, 129(4), 342-353.
- Kurata, S., & Nagatani, T. (2003). Spatio-temporal dynamics of jams in two-lane traffic flow with a blockage. Physica A: Statistical Mechanics and its Applications, 318(3), 537-550.
- Hiselius, L. W. (2004). Estimating the relationship between accident frequency and homogeneous and inhomogeneous traffic flows. Accident analysis & prevention, 36(6), 985-992.
- Golob, T. F., Recker, W. W., & Alvarez, V. M. (2004). Freeway safety as a function of traffic flow. Accident Analysis & Prevention, 36(6), 933-946.
- 백승걸, & 박재범. (2004). 고속도로 돌발상황으로 인한 교통영향예측시스템 고찰.
- Baykal-Gursoy, M., Xiao, W., Duan, Z., & Ozbay, K. (2006). Delay estimation for traffic flow interrupted by incidents. In 86th Annual Transportation Research Conf., Transportation Research Board, Washington, DC.
- Jeong, Y. S., Song, S. G., & Choe, G. J. (2007). A Prediction Model on Freeway Accident Duration using AFT Survival Analysis. Journal of Korean Society of Transportation, 25(5), 135-148
- 강연수, 강경표, 최정민, 김범일, & 김현미. (2008). 재난대응 첨단교통관리체계 구축방안 연구.
- Zhu, H. B., Lei, L., & Dai, S. Q. (2009). Two-lane traffic simulations with a blockage induced by an accident car. Physica A: Statistical Mechanics and its Applications, 388(14), 2903-2910
- Khattak, A., Wang, X., & Zhang, H. (2009). Are incident durations and secondary incidents interdependent?. Transportation Research Record: Journal of the Transportation Research Board, (2099), 39-49
- Lee, Y., & Wei, C. H. (2010). A computerized feature selection

method using genetic algorithms to forecast freeway accident duration times. Computer-Aided Civil and Infrastructure Engineering, 25(2), 132-148.

- Cader, H. A., & Leatherman, J. C. (2011). Small business survival and sample selection bias. Small Business Economics, 37(2), 155-165.
- Xie, F., & Feng, Q. (2013, December). Research of effects of accident on traffic flow characteristics. In Mechatronic Sciences, Electric Engineering and Computer (MEC), Proceedings 2013 International Conference on (pp. 1670-1674). IEEE.
- George, B., Seals, S., & Aban, I. (2014). Survival analysis and regression models. Journal of Nuclear Cardiology, 21(4), 686-694.
- Chen, J., Peng, Z., & Fang, Y. (2014). Effects of Car Accidents on Three-Lane Traffic Flow. Mathematical Problems in Engineering, 2014.
- 한바울, & 백준걸. (2014). 생존분석을 이용한 디스플레이 FAB 의 반송시간 예측모형. 대한산업공학회지, 40(3), 283-290.
- Tavassoli Hojati, A. (2014). Modelling the impact of traffic incidents on travel time reliability.
- Shim, J., Jang, K., Chung, S. B., & Park, S. H. (2015). Comparison of section speed enforcement zone and comparison zone on traffic flow characteristics under free-flow conditions in expressways. Journal of Korean Society of Transportation, 33(2), 182-191.
- Dimitriou, L., & Vlahogianni, E. I. (2015). Fuzzy modeling of freeway accident duration with rainfall and traffic flow interactions. Analytic Methods in Accident Research, 5, 59-71.
- 최윤영, 고승영, 명묘희, & 김동규. (2015). 운전자 속성과 처벌경험이 교통법규 준수기간에미치는 영향에 관한 연구. 경찰학연구, 15(1), 193-214.
- Lee (2016), Deterrent Effects of Demerit Points and License Sanctions on Traffic Violation Using Proportional Hazard Model

국문 초록

교통 혼잡은 그 원인에 따라 크게 반복정체와 비반복정체로 구분할 수 있다. 특히, 비반복정체의 경우 전체 혼잡의 60%를 야기하고 있다. 이러한 비반복정체 요인 중에서도 가장 주요한 요인은 교통사고로 지목되고 있다.

이러한 교통 혼잡을 효과적으로 예방하기 위해서는 교통 사고에 대한 대응을 효율적으로 하는 것이 중요하다. 이에 본 연구에서는 교통사고가 교통류에 미치는 영향을 분석하는 것을 목적으로 한다. 이 목적을 달성하기 위해서 본 연구는 콕스의 비례해저드 모형을 사용하였다. 또한, 분석을 위해 한국도로공사로부터 경부선, 서해안선, 영동선의 4년치 사고자료 및 VDS 자료를 구득하였다.

해저드 모형을 도출한 결과 교통사고와 관련된 다양한 요인들이 교통류에 영향을 끼치고 있는 것으로 분석되었다. 사고 발생 전의 교통류가 안정적일수록 LOS가 낮아질 해저드는 감소하였으며, 차량의 손상 정도가 커질수록, 교통사고 발생 지점의 상류부에서의 교통량이 증가할수록 해저드는 증가하는 경향을 보였다. 또한, 사고차량이 정차한 위치, 중차량의 연관, 평면선형이 혼잡 발생에 영향을 끼치는 것으로 분석되었다.

한편, 교통사고 관련 변수들의 영향력은 교통사고 발생 전의 교통류 상태와 운영기관의 목표운영수준에 따라 다른 것으로 나타났다. Model 1과 Model 3 은 Model 2에 비해서 사고 심각도의 영향을 많이 받는 것으로 나타났다. 반면에 Model 2는 사고로 인한 용량 감소와 상류부에서 발생하는 교통량의 영향을 더 많이 받는 것으로 나타났다. 이는 Model 2는 Model 1, Model 3과 달리 자유류 교통상태에서 심각한 혼잡상황으로의 교통류 변동을 설명하는 모델이기 때문이다. 이를 통해 심각한 수준의 혼잡은 사고의 심각도 뿐만 아니라 용량 감소 및 수요 증가가 맞물려야 발생함을 알 수 있었다. 또한 이러한 결과를 통해 혼잡을 예방하기 위해서 어느 수준의 교통량 조절이 필요한지에 대한 지식을 얻을 수 있었다.

주요어 : 교통사고, 교통류, 교통 운영, 비례해저드 모델, 교통량 조절

학번: 2015-21303