

교통 빅데이터를 이용한 전국 도로 안전성능함수 및 안전등급 개발 연구

Development of Safety Performance Functions and Level of Service of Safety on National Roads Using Traffic Big Data

권 건 안* · 박 상 민** · 정 하 림** · 권 철 우*** · 윤 일 수****

* 주저자 : 한국교통안전공단 자동차안전연구원 인증검사처 연구위원
 ** 공저자 : 아주대학교 건설교통공학과 박사과정
 *** 공저자 : 인천연구원 교통물류연구실 연구원
 **** 교신저자 : 아주대학교 교통시스템공학과 부교수

Kenan Kwon* · Sangmin Park** · Harim Jeong** · Cheolwoo Kwon*** · Ilsoo Yun****

* Vehicle Certification Office, Korea Automobile Testing & Research Institute
 ** Dept. of Transportation Eng., Ajou University
 *** Transportation & Logistics Researcher Division The Incheon Institute
 **** Dept. of Transportation Eng., Ajou University
 † Corresponding author : Ilsoo Yun, ilsooyun@ajou.ac.kr

Vol.18 No.5(2019)

October, 2019

pp.34~48

pISSN 1738-0774

eISSN 2384-1729

<https://doi.org/10.12815/kits>

2019.18.5.34

Received 30 April 2019

Revised 23 May 2019

Accepted 2 September 2019

© 2019. The Korea Institute of Intelligent Transport Systems. All rights reserved.

요 약

본 연구는 우리나라 전체 도로 유형을 대상으로 빅데이터를 이용하여 안전성능함수(safety performance function, SPF)를 개발하고, 그것을 바탕으로 다양한 도로에 대한 안전등급을 평가함으로써 상대적으로 위험한 도로에 대한 대책을 수립할 수 있는 기초 정보를 제공하고자 하였다. 교통사고 자료를 국가표준 노드 및 링크 체계를 기반으로 전국의 도로에 매칭 하여 종속 변수로 활용하였으며, 독립변수로 링크 길이, 차로 수 등 기하구조 자료 및 한국교통연구원의 ViewT 교통량 자료, 그리고 사업용 차량에 장착된 운행기록계를 통한 위험운전행동 건수를 활용한 4개 시스템의 교통 빅데이터를 활용하여 연구를 진행하여 7개 도로 유형별 안전성능함수를 개발하고, 개발된 안전성능함수를 활용하여 도로유형별로 A, B, C, D 네 개의 안전등급을 평가하였다. 본 연구에서 사용한 방법론과 분석 결과를 토대로 우리나라의 위험도로를 선정하였다. 도출된 결과를 바탕으로 교통안전 개선사업의 대상을 선정하고 그에 따른 효과 등을 모니터링하고 계량화할 수 있을 것으로 기대된다.

핵심어 : 빅데이터, 안전성능함수, 안전등급, 도로, 교통안전

ABSTRACT

The purpose of this study was two-fold; first, to develop safety performance functions (SPF) using transportation-related big data for all types of roads in Korea were developed, Second, to provide basic information to develop measures for relatively dangerous roads by evaluating the safety grade for various roads based on it. The coordinates of traffic accident data are used to match roads across the country based on the national standard node and link system. As independent variables, this study effort uses link length, the number of traffic volume data from ViewT established by the Korea Transport Research Institute, and the number of dangerous driving behaviors based on the digital

tachograph system installed on commercial vehicles. Based on the methodology and result of analysis used in this study, it is expected that the transportation safety improvement projects can be properly selected, and the effects can be clearly monitored and quantified.

Key words : Big data, Safety performance function, Level of service of safety, Road, traffic safety

I. 서론

1. 연구의 배경 및 목적

2017년 교통사고 사고건수는 216,335건, 사망자수는 4,185명, 부상자수는 322,829명으로 이는 전년에 비해 각각 2.1%, 2.5%, 2.7%가 감소하였다. 최근 10년간 연평균 교통사고 사고건수, 사망자수 및 부상자수는 각각 0.0%, 3.7%, 0.5%로 사고건수를 제외하고 모두 감소 추세를 나타내고 있다(Korea Road Traffic Authority, 2018). 이러한 성과는 최근 정부가 수립한 ‘교통사고 사망자 수를 2022년까지 절반으로 감축한다.’는 목표에 따라 교통 인프라를 개선하고, 단속 및 교육을 강화하는 등 다각적인 노력을 지속한 결과라고 할 수 있다. 이러한 노력의 일환으로 “제8차 국가안전 기본계획”에 따라 교통 인프라 강화 및 보행자 중심 패러다임을 통한 보행자 안전강화, 차량의 속도체계 변경 등을 추진하고 있다. 또한 2022년까지 150개 지방지역을 ‘마을주민 보호구간’으로 지정하여 교통사고 취약지역에 대한 개선대책을 마련하고 있으며, 교통약자인 어린이와 고령자 보호구역을 지정하여 운영하고 있다. 이와 함께 교통사고 중점관리 대상 자동차인 사업용자동차의 안전관리 향상을 위해 화물자동차 차령 관리, 위험물 운송모니터링 시스템 시범운영, 고령자 화물자동차 자격관리 강화 등을 추진하고 있다. 그리고 사업용자동차의 운전자에 대한 인센티브 제도로써 실시간 운행기록 제출 시 해당 자동차 보험료의 5~15%를 할인하는 등 운전자의 자발적인 운전행태 개선을 유도하고 있다. 또한 교통사고 예방 강화를 위해 범정부적으로 정보통신기술(Information and Communication Technology, ICT)과 첨단안전장치 등 관련기술이 적용된 대형차량 줄음운전방지장치, 차로이탈경보장치(Lane Departure Warning System, LDWS), 비상자동제동장치(Advanced Emergency Braking System, AEB) 등을 단계적으로 확대하고 있으며 가시적인 성과를 보이고 있다. 이러한 노력에도 불구하고 우리나라 교통안전 수준은 여전히 후진국 수준에 머물러 있다. 교통사고는 국민생활안전에 큰 위협이 되고 있는 사회적인 문제로 당사자와 가족에게는 고통과 슬픔을 야기함은 물론 경제적으로도 연간 23조 5천억 원에 달하는 사회적 비용을 발생시켜 국가경쟁력 약화의 한 원인으로 지목되고 있다(Korea Road Traffic Authority, 2018). 인구 10만 명당 사망자 추세에서 우리나라는 ‘80년 16.9명에서 ‘90년 33.1명으로 증가한 이후 2000년 21.8명, 2014년 9.4명, 2015년 9.1명으로 꾸준히 감소하고 있는 것으로 나타났고, 자동차 1만 대당 사망자 수 또한 ‘80년 59.4명에서 ‘90년 23.9명, 2000년 6.5명, 2014년 2.0명, 2015년 1.9명으로 인구 10만 명당 사망자 추세와 마찬가지로 꾸준히 감소하고 있는 것으로 나타났다. 그러나 우리나라의 교통안전 수준은 선진국과 비교해 열악한 상황이다. 한 예로 자동차 1만 대당 사망자수는 13.8명으로 OECD 회원국 평균인 6.0명 보다 약 2.3배 높은 수준인 것으로 나타났다(Korea Road Traffic Authority, 2018).

최근 제4차 산업혁명시대의 도래와 함께 ICT의 발달로 인하여 교통 분야의 빅데이터 생산량 및 다양성이 증가하고 있다. 또한 최근 빅데이터 분석기법의 발달로 인하여 교통 빅데이터를 이용한 다양한 교통 서비스 개발뿐만 아니라, 교통안전 분야에서도 교통 빅데이터를 이용한 분석의 고도화를 수행하고자 하는 노력이 진행되고 있다. 이러한 다양한 교통 빅데이터는 교통안전과 관련된 설명변수로서 많이 활용될 것으로 기대된다. 따라서 본 연구의 목적은 우리나라 도로에 대하여 가용할 수 있는 다양한 정보를 이용하여 안전성능함

수(safety performance function, SPF)를 개발하고, 개발된 안전성능함수를 바탕으로 다양한 도로에 대한 안전등급을 산출함으로써 상대적으로 위험한 도로에 대한 대책수립을 위한 기초 정보를 제공하는 것이다.

본 연구를 통해 개발된 안전성능함수와 도로안전등급을 활용하여 선정된 위험 도로에 대한 적절한 개선안과 개선안의 적용 효과를 계량화한 결과가 정책에 활용되기를 기대한다.

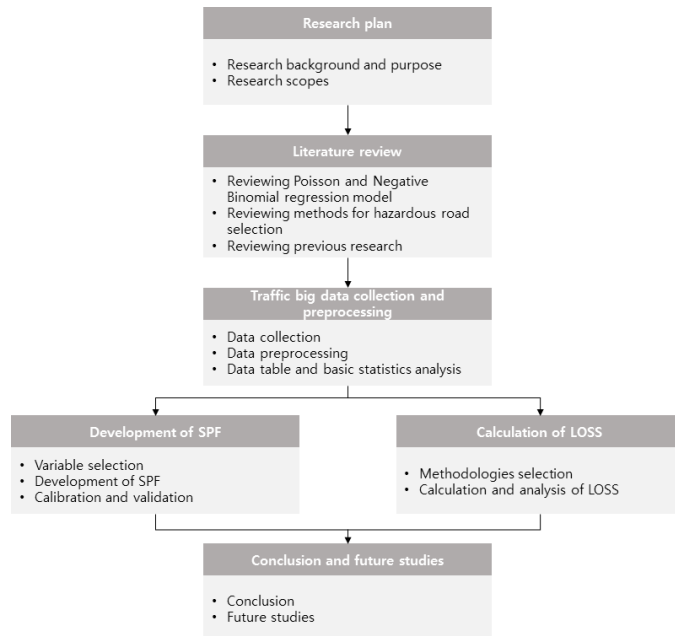
2. 연구의 범위 및 방법

1) 연구의 범위

본 연구의 공간적 범위는 전국의 도로를 1차 대상으로 한다. 안전성능함수와 도로안전등급 개발을 위한 본 연구의 시간적 범위는 교통사고 자료, 디지털운행기록계(digital tachograph, DTG) 자료, 교통량 자료 등이 확보 가능한 2015년으로 한다.

2) 연구의 방법

본 연구에서는 안전성능함수와 도로안전등급과 관련된 적용이론 및 선행연구 논문에 대하여 고찰하였으며, 특히 안전성능함수와 도로안전등급의 성능 향상을 위해 사용된 교통 빅데이터의 종류 및 특징을 파악하고자 하였다. 안전성능함수 개발을 위하여 필요한 종속변수와 독립변수로 사용 가능한 자료를 수집 및 가공하였다. 종속변수로는 교통사고 건수를 사용하였으며, 독립변수로는 차로수, 기하구조, 제한속도, 교통량, 위험운전행동 등의 다양한 자료를 수집하였으며, 수집된 자료에 대한 기초 통계분석을 실시하였다(Kwon, 2019). 특히, 안전성능함수의 성능을 높이기 위해 본 연구에서는 교통 빅데이터를 기반으로 구축된 교통량과 위험운전행동 건수를 독립변수에 포함시킨 것이 특징이라고 할 수 있다. 도로안전등급을 산출하기 위하여 표준편차를 이용하여 네 개의 등급으로 나누는 방법을 적용하였다.



<Fig. 1> Research process

II. 관련 이론 및 연구 고찰

1. 관련 이론 고찰

1) 안전성능함수

안전성능함수는 교통사고 건수 또는 교통사고 심각도와 같은 종속변수와 도로조건, 교통조건, 운영조건, 그리고 환경조건 등 관련 독립변수들 간의 관계를 수학적으로 표현한 것이다. 이러한 안전성능함수를 이용하여 교통사고 건수의 예측하거나, 도로안전등급 평가 또는 교통안전시설물의 설치효과를 추정하고 있다. 교통사고는 임의적(random event)이고 발생 빈도가 낮으며(rare event), 이산적(discrete event)이라는 특성을 가지고 있기 때문에 포아송 분포를 잘 따르고 있다고 할 수 있다. 일반적으로 교통사고 건수 y_i 가 포아송 분포를 따른다고 할 때 도로구간 i 에서에서 평균적으로 λ_i 건의 교통사고가 발생한다면, 해당 도로에서 발생할 교통사고의 확률은 식(1)과 같다. 또한 식(1)을 기반으로 도로구간 i 에서 일어나는 평균 사고건수(기댓값) λ_i 는 식(2)와 같은 포아송 회귀모형식으로 추정할 수 있다(Ryu, 2018).

$$p(y_i) = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!}, y = 0, 1, 2, \dots \dots \dots (1)$$

여기서, $p(y_i)$: 도로구간 i 에서 y_i 번의 사고가 발생하는 확률
 λ_i : 도로구간 i 에서 발생할 수 있는 평균 사고건수(기댓값)

$$\lambda_i = \exp(\beta_0 X_0 + \beta_1 X_1 + \dots + \beta_m X_m) = \exp\left(\sum_{j=0}^m \beta_j X_j\right) \dots \dots \dots (2)$$

여기서, β_j : 회귀추정계수

포아송 회귀모형은 평균과 분산이 같다는 가정 하에 적용할 수 있다. 그러나 실측된 교통사고 건수는 분산이 기댓값(평균)보다 큰 과분산 현상(over dispersion)을 보이는 경우가 많다. 교통사고 건수가 과분산 현상을 보일 때는 평균과 분산이 같다는 조건을 완화시킨 음이항 회귀모형을 사용하여 교통사고 건수를 추정할 수 있다.

음이항 회귀모형(negative binomial regression model)은 평균과 분산이 같다는 포아송 회귀모형의 가정을 일반화한 모형이다. 따라서 교통사고 건수와 같은 종속변수가 과분산 현상을 보일 때 사용할 수 있다. 음이항 회귀모형을 이용하여 도로구간 i 에서 일어나는 평균 사고건수(기댓값) λ_i 를 추정하면 식(3)과 같다(Washington et al., 2003). 식(3)은 포아송 회귀모형에서 설명한 식(2)와 모습은 매우 유사하지만 음이항 회귀모형의 식(3)은 포아송 회귀모형의 식(2)에 오차항(ϵ_i)이 추가되었음을 알 수 있다. 오차항(ϵ_i)은 평균이 1이고 분산이 α 인 감마분포를 따른다.

$$\lambda_i = \exp(\beta_0 X_0 + \beta_1 X_1 + \dots + \beta_m X_m + \epsilon_i) = \exp\left(\sum_{j=0}^m \beta_j X_j\right) \dots \dots \dots (3)$$

여기서, β_j : 회귀추정계수
 ϵ_i : 오차항

또한, 음이항분포의 분산은 $\kappa \lambda_i^2$ 이 과분산 현상으로 발생하는 값이며 k 가 0에 수렴할 때 포아송 분포를 따르며, k 가 1보다 크면 과분산이 발생하여 음이항 분포(Negative binomial distribution)를 따른다(Ryu, 2018).

$$Var[y_i] = \lambda_i(1 + \kappa \lambda_i) = \lambda_i + \kappa \lambda_i^2 \dots\dots\dots (4)$$

여기서, $Var[y_i]$: 도로구간 i 에서 y_i 번의 사고가 발생하였을 때 분산
 λ_i : 도로구간 i 에서 일어날 평균 사고건수(기댓값)

2) 안전성능함수를 이용한 교통안전도 평가기법

현재 가장 과학적이고 신뢰성 있는 방법으로 계량화된 모형인 교통안전성 분석모형은 미국 도로안전편람(Highway Safety Manual, HCM)¹⁾의 모형이라 할 수 있다(AASHTO, 2010). HSM에서 제시하고 있는 분석모형은 SPF 및 사고보정계수(crash modification factor, CMF), 지역보정계수(C)를 적용하여, 도로상 발생 가능한 사고건수를 예측한다. 여기서, CMF는 도로조건에 따른 사고 영향계수이다. 안전성능함수는 이상적인 상태의 도로구간 사고건수를 예측하는 데 활용할 수 있으며, 일반적인 경우 일평균교통량과 도로구간연장을 이용하여 산출한다. 교통사고와 상관성이 높은 변수를 도출하고 계수를 산정하는 등 평가항목을 개발하여 적용성을 검증했다. 평가항목 개발을 위한 모형식은 다음과 같다.

$$N_{predicted} = SPF(CMF_{1x} \times CMF_{2x} \times \dots \times CMF_{yx}) \times C_x \dots\dots\dots (5)$$

여기서, $N_{predicted}$: 사고예측건수
 SPF : 기본사고 예측건수
 CMF_{yx} : 사고보정계수
 C_x : 지역보정계수
 x : 특정지점 또는 구간
 y : CMF의 총 갯수

2. 선행 연구 고찰

Martinelli(2009)는 HSM accident predictive model의 전이성을 살펴보기 위하여 서로 다른 환경과 도로여건, 운전행동 및 사고데이터시스템 등을 이용하였다. 또한, 보정계수(calibration factors)를 위해 서로 다른 네 가지 방법론을 검토하였다. 2002년부터 2004년까지 3개년을 시간적 범위로 설정하고 Italia Arezzo province의 rural two lane highways (1,300km)를 공간적 범위로 적용하였다. 변수로는 기하구조, 교통량, 교통사고, 도로유형 데이터를 사용하였다. 전체도로 938km를 대상으로 8,379개의 동일구간을 구분하였다. 차로별 교통량, 기하구조의 특성에 따른 차로 폭, 노면 폭 등으로 구분하여 총 80개의 동질구간을 분리하였다. 동질구간을 세분화한 결과, 평균길이 112m, 총 구간 8,379개 중 8,106개 구간에서 사고건수가 0건이 기록되는 문제가 발생하는 것을 확인하였다.

Kwak et al.(2010)는 고속도로 교통사고 건수 예측모형을 교통사고에 영향을 미치는 환경적 요인 및 기하

1) HSM 1990 10 , 2010 FHWA(Federal Highway Administration) TRB(Transportation Research Board) . HSM Part A D 4 , ,

구조를 반영하였다. 2006년부터 2008년까지 3년간 경부고속도로 양방향을 활용하였다. 변수는 교통량 자료 및 기하구조를 사용하여 분석하였다. 고속도로 교통사고는 발생확률이 낮은 사건(rare and random event)으로 사고의 대부분이 0건에서 6건으로 분포한 것으로 나타났다. 또한 사고가 발생하지 않는 구간의 분포도 매우 높게 나타나고 있다. 이런 현상을 극복하기 위하여 포아송 및 음이항 회귀모형과 영과잉 포아송과 영과잉 음이항 회귀모형을 활용하였다. 안전성능합수는 포아송 회귀모형을 통해 구축된 모형보다 음이항 회귀모형이 통계적으로 적합한 것으로 분석하였다. 또한 Vuong 검정을 통해 영과잉 포아송 회귀모형이 포아송 회귀모형보다 통계적으로 우수한 것으로 나타났다.

Banihashemi(2012)는 데이터에 따른 보정계수 변화에 대한 민감도 분석을 수행하였다. 이 연구는 2006-2008년까지 3년간의 시간적 범위를 구성하고, 워싱턴 주 의 유형별 고속도로를 공간적 범위로 사용하였다. 변수로는 도로유형, 곡선반경, 경사율, 차로 수, 사고 자료를 이용하였다. 민감도분석 분석 결과 Rural Two lane - Undivided 도로의 이상적 보정계수는 1.472로 나타났다. 유의수준 5%, 전체 구간길이 50% 이상의 보정계수를 산출하였고 통계적으로 동일한 값이 산출되었다. 또한 유의수준 10%, 전체 구간길이 20% 이상에서도 동일한 값이 산출되었다. 이 연구는 지방부 2차로 도로 - 분리되지 않은 고속도로의 경우, 유의수준 10%, 전체 구간길이 5~10%의 도로를 대상으로 보정계수를 산출할 경우, 보정계수의 참값과 통계적인 동일 확률이 65~80%로 도출되었다. 전체 구간길이의 5~10%는 약 240~480mile의 수준이며, 사고건수는 약 330~660건/year 수준으로 분석되었다.

Cafiso et al.(2013)는 고속도로 구간 길이를 분할할 때 발생하는 문제점을 통하여 최적의 분할 기준에 대한 연구를 진행하였다. 이 연구는 이탈리아 지방부 고속도로를 대상으로 2002년부터 2009년까지 8년간의 데이터를 활용하였다. 활용변수는 사망사고 및 부상사고 합, AADT, 구간길이, 경사율, 곡선반경 등을 독립변수로 활용하였다. 이 연구는 세 가지 모형의 형태로 개발되었는데, 안전성능합수를 구하기 위해 독립변수로 모든 변수를 반영한 모형, AADT 및 곡선반경 두 가지 변수를 반영한 모형과 AADT만 단일변수만 반영한 기본 모형이었다. 개발된 세 가지 모형을 통하여 QIC₂를 이용하여 적합도가 가장 높은 구간길이를 결정하였다. 분석 결과, QIC가 가장 작은 2curves와 2tangents 기반의 분할한 모형인 것으로 분석되었다.

III. 자료 수집 및 구축

1. 변수 선정

안전성능합수 개발을 위하여 필요한 종속변수와 독립변수로 사용가능한 자료를 수집 및 가공하였다. 종속 변수로는 교통사고 건수를 사용하였으며, 독립변수로는 도로의 규모(차로수 등), 기하구조, 운영여건(제한속도 등), 수요(교통량), 운행행태(위험운전행동 등)의 다양한 자료를 수집하였으며, 수집된 자료에 대한 기초 통계분석을 실시하였다. 특히, 안전성능합수의 성능을 높이기 위해 본 연구에서는 교통 빅데이터를 기반으로 구축된 교통량과 위험운전행동 건수를 독립변수에 포함시킨 것이 특징이라고 할 수 있다(Kwon, 2019).

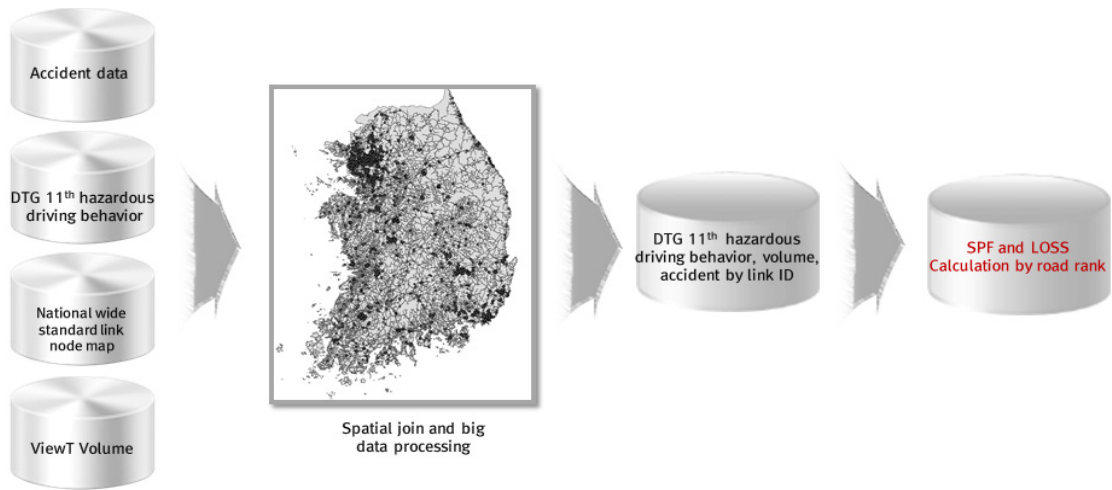
2. 자료 수집 및 가공

본 연구에서는 교통 빅데이터를 이용한 도로안전성능합수 개발을 위해 교통 관련 기관별로 보유하고 있는 교통 빅데이터를 수집하였다. 수집 자료는 국가표준노드링크, 교통사고자료, 11대 위험운전행동 건수,

2) Quasilikelihood under the Independence model Criterion

ViewT 교통량 자료이다. 우선 교통 빅데이터를 이용하기 위해 자료 수집 단위를 국가표준 노드 링크로 선정하였다. 교통사고자료는 경찰청 교통사고 자료를 이용하였다. 경찰청 교통사고 자료는 총 90개의 항목으로 구성되어 있으며, 교통사고 위치가 좌표로 되어있어 국가표준 노드 링크에 공간연산을 통해 쉽게 매칭이 가능하다. 한국교통안전공단이 수집하여 관리하고 있는 DTG 11대 위험운전행동 자료는 사업용자동차의 실시간 운행정보를 저장하여 시간별 변화하는 운행상황을 자동으로 기록하는 운행기록분석시스템에서 제공하는 자료이다. 이를 통해 운전자의 운전습관을 파악하고 분석하여 과학적이고 실증적인 운전자 안전관리를 지원하는 자료이다. 과속유형, 급가속유형, 급차로 변경유형, 급회전유형 등으로 구성되어 있다. 한국교통연구원에서 개발한 ViewT 1.0은 전국단위의 교통 데이터를 제공하며, 개별 차량 단위의 경로데이터 기반의 분석기능을 제공하는 교통 빅데이터 플랫폼이다. 분석지표로 교통량, 속도, 교통혼잡지표 등을 제공하고 있으며, 특히 교통량의 경우 관측교통량, 추정교통량 등을 전국단위로 제공하고 있다.

수집된 자료들은 국가표준 노드 링크 단위로 가공되어 있지 않아 빅데이터 처리 및 공간 연산을 통해 국가표준 노드 링크 단위로 가공하는 것이 필요하다. 특히 DTG 11대 위험운전행동 건수와 교통사고 자료는 빅데이터 처리 및 공간 연산을 통해 국가표준 노드 링크 단위로 가공하였다. 또한 ViewT 교통량의 경우는 국가표준 노드 링크 단위로 가공되어 있기 때문에 링크 ID로 매칭 하여 사용하였다. 다음은 빅데이터 처리 및 공간연산을 통해 자료 가공하는 과정을 나타내는 그림이다(Kwon, 2019).



<Fig. 2> Traffic Big Data source processing

IV. 분석테이블 구축 및 기초 통계 분석

1. 분석테이블 구축

교통 빅데이터의 분석을 위한 자료가공을 통해 링크 ID를 기준으로 매칭하여 분석테이블을 구축하였다. 분석테이블을 구축한 결과 총 6개 변수가 구성되었으며, 136,525개의 데이터가 구축되었다. 분석테이블은 링크 ID별로 자료가 구축되었으며, 도로유형을 구분할 수 있도록 코드를 입력해 놓았다. 아래 그림은 분석 테이블을 구축한 결과를 나타낸 것이다.

Link ID	Type of road	Section length	11th hazardous driving behavior	Number of lane	Speed limit	ViewT traffic volume	Number of accident
2510029302	103	116.1462743	5690	3	60	8715	2
2510029402	103	117.8485784	1925	3	60	8529	0
2510116100	107	359.2912639	9441	2	60	1700	1
2510116200	107	363.3037969	9676	2	60	2060	2
2510115500	107	59.74358118	733	2	60	1700	0
2510115600	107	59.74793111	522	2	60	2060	0
2510114700	107	47.34746862	1564	2	60	1700	0
2510095300	107	25.41282058	212	2	30	1799	0
2510095400	107	25.41285599	85	2	30	1508	0
2510093300	107	78.0740649	1000	2	30	1799	2
2510093400	107	78.07418111	586	2	30	1508	0
2510112300	107	97.61786733	1831	2	30	2723	1
2510112400	107	97.6220429	4997	2	30	2322	1
2510109700	107	92.63621923	1237	2	30	2723	0
2510109800	107	93.00874496	2755	2	30	2322	0
2510106900	107	52.69927542	2226	2	30	2723	0
2510107000	107	52.69829772	2714	2	30	2322	0
2510128200	107	144.8308102	5795	3	60	6162	0
2510133500	107	80.07821234	5445	3	60	6214	0

<Fig. 3> Result of Coding

2. 기초 통계분석

구축된 분석 자료의 특성 파악을 위해 도로유형별로 기초 통계분석을 실시하였고 도로 유형별로 사고건수, 구간길이, 11대 위험운전행동건수, 차로 수, 제한속도 및 ViewT 교통량 등 독립변수를 기반으로 하여 각 변수의 평균값, 표준편차, 중앙값 및 최소·최대 등 결과를 도출하였다. 도로유형별로 종속변수인 사고건수와 독립변수의 상관성을 분석한 결과 대부분 양의 상관관계를 보이는 것으로 나타났다. 특히 11대 위험운전행동의 경우 지방도를 제외하고 다른 도로유형에서 높은 양의 상관관계를 보이는 것으로 나타났다. 다만, 제한속도의 경우 다른 독립변수들에 비해 사고건수와 상관관계가 낮은 것으로 나타났다.

종속변수인 교통사고를 기반으로 기초통계분석을 분석해 보면, 사고건수 최댓값과 평균값의 순위가 다른 것으로 나타나고 있으며, 이는 도로 유형별로 다양한 변수들의 영향이 복합적으로 나타나고 있는 것으로 추정된다(Kwon, 2019).

<Table 1> Result of Relation Variable

Type of Road	Section Length	11th hazardous Driving Behavior	Number of Lane	Limit Speed	ViewT Traffic Volume
Highway	0.40	0.61	0.39	0.21	0.51
Urban highway	0.49	0.40	0.20	0.18	0.36
General national road	0.07	0.51	0.30	-0.03	0.36
Special metropolitan road	0.19	0.56	0.18	0.08	0.30
National local road	0.12	0.41	0.29	0.07	0.31
Provincial road	0.10	0.09	0.29	0.05	0.30
City road	0.09	0.43	0.10	-0.10	0.18

V. 안전성능합수 개발 및 도로안전등급 평가

1. 안전성능합수 개발

안전성능합수 개발을 위해서는 변수 선정이 필요하다. 본 연구에서는 앞서 수행된 문헌고찰과 연구고찰을 통해 안전성능합수 구축을 위한 종속변수와 독립변수를 선정하였다. 대부분의 관련 연구에서는 교통사고 건수를 안전성능합수의 종속변수로 사용하고 있었다. 본 연구에서도 안전성능합수의 종속변수로 국가표준노드 링크에 매칭된 교통사고 건수를 사용하였다. 또한 독립변수로는 국가표준 노드 링크별로 이용 가능한 ViewT 교통량, 구간길이, 차로수, DTG 11대 위험운전행동건수 합계를 사용하였다. 특히 ViewT 교통량과 DTG 11대 위험운전행동 건수는 본 연구에서 사용된 교통 빅데이터이다(Kwon, 2019).

<Table 2> Opinion of Variable

Category	Name of Variable	Unit
Dependant variable	Number of Accident	number
Independent variable	ViewT Traffic Volume	car/day
	Section Length	m
	Number of Lane	number
	DTG 11th total hazardous Driving Behavior	count

도로유형별 안전성능합수를 구축하기 위해 전통적으로 교통사고 예측에 이용되는 음이항회귀모형을 사용하였다. 문헌 고찰 결과 안전성능합수를 구축하는데 있어서 두 가지 모형식을 사용하고 있는 것으로 나타났다. 첫째는 노출계수에 대한 고려 없이 모든 독립변수들이 지수 형태로 반영되고 있는 방법과 교통량과 구간길이 같은 노출 계수를 로그-변환하여 사용하고 있는 방법이다. 본 연구에서는 노출계수로 ViewT 교통량과 구간길이를 선정하고 노출계수를 적용한 모형을 사용하였다. 다음 수식은 노출계수를 적용한 모형식이다(Kwon, 2019).

$$Y = Volume^{\beta_1} \times l^{\beta_2} \times e^{(\alpha + \beta_3 \dots)} \dots \dots \dots (6)$$

여기서, Y : 교통사고 예측 건수
 $Volume$: 구간의 연평균일평균 교통량(대/일)
 l : 구간길이(m)
 α, β : 회귀 계수

선정된 모형식과 구축된 자료를 통계계산과 그래픽을 위한 오픈소스 프로그래밍 언어이자 소프트웨어인 R 프로그래밍 언어를 이용하여 안전성능합수를 구축하였다. 특히 R 프로그래밍 언어는 다양한 통계기법과 수치해석 기법을 지원하며, 통계계산과 소프트웨어 개발을 위한 환경이 필요한 통계학자나 연구자들이 많이 사용하고 있는 프로그래밍 언어이다. 안전성능합수 구축시 모든 변수의 유의수준을 만족하는지 확인하였다. 또한 우도비 검증을 통해 모형의 유의성을 확인하여 각 도로유형별로 안전성능합수를 최종적으로 결정하였다. 구축된 모형 적합도를 나타내는 McFadden Pseudo R-Squared(ρ^2)는 고속도로의 경우 0.33이며, 도시고속도로의 경우 0.11, 일반국도, 특별광역시도, 국가지원지방도, 지방도는 0.14, 시군도는 0.10으로 나타났다. ρ^2 가

0.2~0.4 정도면 모형 적합도가 우수한 것으로 알려져 있다 (Yun, 2001).

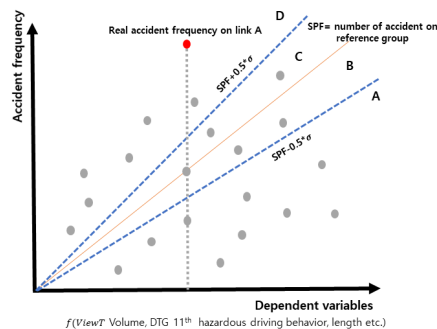
본 연구에서 구축된 안전성능함수는 특별광역시도의 차로수를 제외한 모든 변수에서 유의수준 0.01에서 유의한 것으로 나타났으며, 각 변수의 부호가 양의 부호로 나타나 교통사고에 영향을 주는 것으로 확인 되었다(Kwon, 2019). 이는 위험운전행동건수가 운전자의 운전행태에 기인한 결과이지 도로의 기하구조에 따른 영향의 결과로 나타나지 않음을 반증하고 있다. 다음 표는 구축된 모든 도로유형의 변수 및 유의수준을 정리한 결과이다.

<Table 3> Result of Safety Performance Function

Type of Road	Constant	viewT Traffic Volume	Section Length	Number of Lane	11th hazardous Driving Behavior	Pseudo R-Squared
Highway	-10.7500***	0.6550***	0.5280***	0.0680*	0.00002***	0.33
Urban highway	-4.6620***	0.1260***	0.5900***	0.1170***	0.00001***	0.11
General national road	-6.9130***	0.3790***	0.4650***	0.2390***	0.00002***	0.14
Special metropolitan road	-6.4450***	0.3800***	0.5680***	0.0130	0.00002***	0.14
National local road	-7.0040***	0.3710***	0.4550***	0.2590***	0.00002***	0.14
Provincial road	-7.8250***	0.3720***	0.4900***	0.6960***	0.00001***	0.14
City road	-5.0840***	0.2410***	0.4980***	0.1510***	0.00002***	0.10

2. 도로안전등급 평가

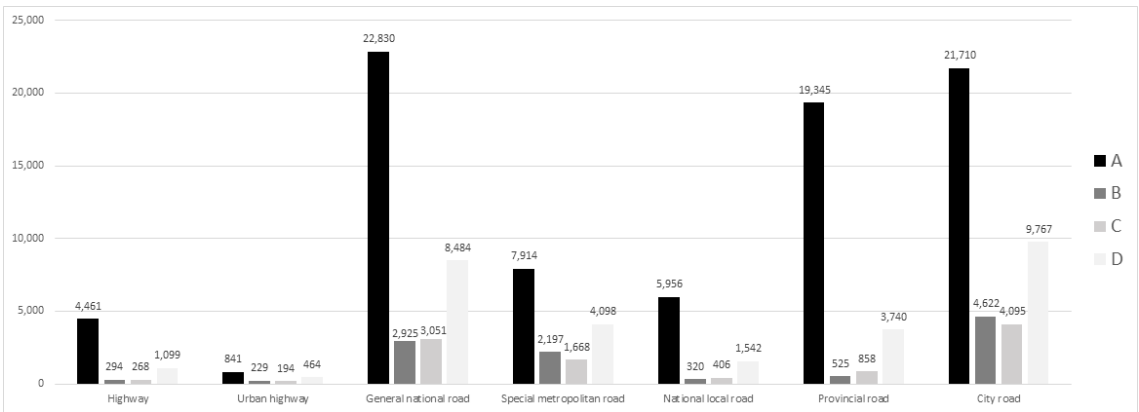
본 연구에서는 도로유형을 산출하기 위해 안전성능함수에 일정한 표준편차를 적용하여 안전수준을 계층화 하는 방법론을 사용하였다(Kononov and Allery, 2003; Kononov and Allery, 2015). 이 방법은 참조집단의 사고건수를 추정하는 것으로 동일 수준의 운전애 노출되면 위험요인이 없을 때, 참조집단만큼만 사고가 발생하며, 참조집단 이상의 사고 발생은 위험한 도로 기하구조(경사, 곡선반경 등), 교통상황(높은 속도 편차 등)에 의한 사고이므로 원인 파악 및 개선이 필요하다는 것을 의미한다. 또한, 안전성능함수를 이용한 안전등급(level of service of safety, LOSS) 추정 기법의 장점은 동일한 모형 설계 방식을 도입하여, 통일성을 확보 가능하고 체계적이고 통계적으로 객관적인 결과를 제시하는 것이 가능하다(Kwon, 2019). 다음 그림은 안전성능함수를 이용한 도로 안전등급 추정 개요를 나타낸다.



<Fig. 4> Concept of LOSS Rating

3. 최종 LOSS 검증 결과 및 분석

도로의 LOSS 산출은 음수이항에 계수를 곱하여 산출한다(Kononov and Allery, 2004; Kononov and Allery, 2015). 특히, 계수가 작거나 큰 경우 LOSS가 산출되지 않는 문제가 발생할 수 있어 도로의 안전관리를 위해서는 균일하게 등급을 산출하는 것이 바람직하다(Kwon, 2019). 본 연구에서는 도로 LOSS를 가능한 균등하게 산출하기 위해 계수에 따른 민감도 분석을 수행하였다. 특히 계수에 따라 LOSS가 산출되지 않는 경우를 방지하기 위하여 적절한 계수값을 선정하는 것이 필요하기 때문이다. 본 연구에서는 계수 값을 산정하기 위해 계수 값을 0.5~1.0까지 조정해가며 도로 LOSS 산출 결과를 평가하였다. 계수별 민감도 분석을 수행한 결과 계수를 0.5로 선정하여 도로 LOSS를 산출을 하였을 경우 LOSS A~D가 다른 계수에 비해 균등하게 나뉘지는 것을 확인 할 수 있었다. 특히 모든 경우에서 LOSS A가 LOSS B~D에 비해 높은 빈도를 보이고 있으며, LOSS B~C에 비해서는 LOSS D인 구간이 높은 것으로 분석 되었다(Kwon, 2019). 본 연구에서는 계수를 0.5로 한 경우를 최종 도로 LOSS로 산출하고 분석을 수행하였다. <Fig. 5>는 최종 도로 LOSS의 도로 유형별 분포이다.



<Fig. 5> Distribution of Road Type of LOSS

산출된 도로별 LOSS의 비율을 분석한 결과, LOSS A의 경우 가장 높은 빈도를 보이고 있으며, 교통사고의 특성상 교통사고가 발생하지 않은 구간이 많기 때문인 것으로 판단된다. 그 다음으로는 D등급이 높은 비율을 보이고 있어 교통사고가 많이 발생한 구간인 것으로 분석되었다.

<Table 4> Ratio of LOSS by road types

Type of Road	LOSS A	LOSS B	LOSS C	LOSS D	Total
Highway	72.9%	4.8%	4.4%	18.0%	100.0%
Urban highway	48.7%	13.3%	11.2%	26.9%	100.0%
General national road	61.2%	7.8%	8.2%	22.8%	100.0%
Special metropolitan road	49.8%	13.8%	10.5%	25.8%	100.0%
National local road	72.4%	3.9%	4.9%	18.8%	100.0%
Provincial road	79.1%	2.1%	3.5%	15.3%	100.0%
City road	54.0%	11.5%	10.2%	24.3%	100.0%
Average	62.6%	8.2%	7.6%	21.7%	100.0%

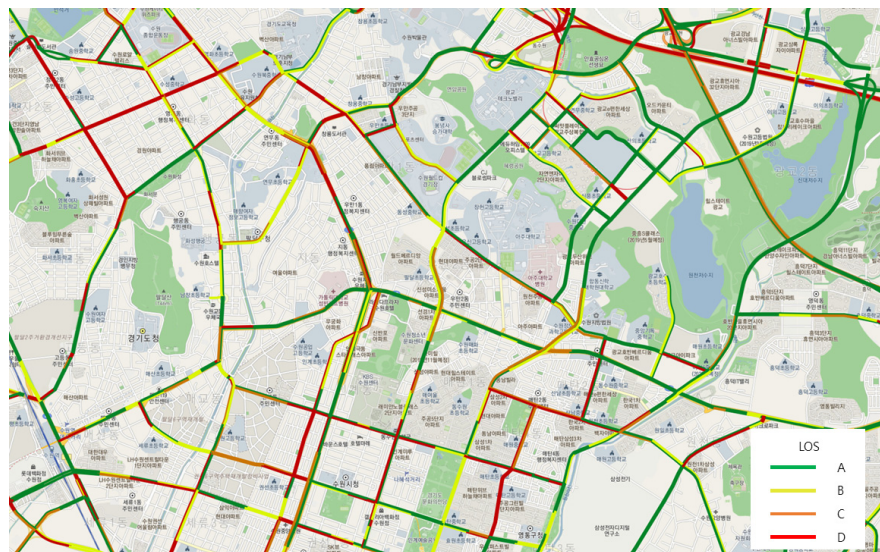
산출된 도로별 LOSS의 적합성 검정을 위해 도로 유형별 및 LOSS별로 평균 교통사고 건수와 표준편차를 분석한 결과, LOSS A에서 D로 갈수록 평균 사고 건수가 높아짐을 확인할 수 있었다. 하지만 국가지원지방도와 지방도의 경우 LOSS B와 C의 등급이 역전되는 현상이 일어났는데, 이는 B 등급에 위치한 교통사고들의 표준편차가 높기 때문인 것으로 분석된다. 전체적으로 LOSS 별로 평균 교통사고건수가 높아짐에 따라 LOSS를 이용하여 교통안전관리에 활용할 수 있을 것으로 판단된다.

<Table 5> Average crash by LOSS and road types

() : Standard deviation

Type of Road	LOSS A	LOSS B	LOSS C	LOSS D
Highway	0.10 (1.14)	2.43 (3.92)	2.60 (2.48)	3.01 (3.48)
Urban highway	0.34 (1.16)	1.98 (2.26)	2.42 (1.77)	4.48 (3.77)
General national road	0.11 (0.84)	1.73 (2.12)	1.90 (1.88)	3.13 (3.01)
Special metropolitan road	0.45 (1.73)	2.32 (2.52)	2.97 (2.69)	4.94 (3.96)
National local road	0.04 (0.41)	1.86 (2.18)	1.77 (1.74)	2.34 (2.30)
Provincial road	0.01 (0.18)	1.61 (1.73)	1.56 (1.60)	2.14 (2.31)
City road	0.14 (0.88)	1.53 (1.45)	1.68 (1.68)	3.64 (2.96)

산출된 도로 LOSS 분포를 확인하고, 도로의 안전관리에 활용할 수 있도록 시각화하기 위해서 Q-GIS를 이용하여 국가표준노드링크에 표출하였다. <Fig. 6>은 GIS에 표출된 도로 LOSS 예시이다.



<Fig. 6> Example of LOSS

VI. 결론 및 향후 연구과제

1. 결론

본 연구에서는 우리나라 전체 도로 유형을 대상으로 가용할 수 있는 다양한 정보인 빅데이터를 기반으로 안전성능함수를 개발하고, 개발된 안전성능함수를 바탕으로 다양한 종류의 도로에 대한 LOSS를 산출함으로써 상대적으로 위험한 도로에 대한 안전대책을 수립할 수 있는 기초 정보를 제공하고자 하였다. 이를 위하여 교통사고 자료의 좌표정보를 이용하여 교통사고를 국가표준 노드 및 링크 체계를 기반으로 전국의 도로에 매칭시켜서 종속변수로 활용하였으며, 독립변수로 국가표준 노드 및 링크 체계상에 활용 가능한 링크 길이, 차로 수 등 기하구조 자료 및 한국교통연구원에서 구축한 ViewT 교통량 자료, 그리고 사업용 차량에 장착된 DTG를 활용한 위험운전행동건수를 기반으로 4개 시스템의 교통 빅데이터를 활용하여 본 연구를 진행하였다. 이러한 종속변수와 독립변수를 활용하여 7개의 도로유형별 안전성능함수를 개발하였고, 개발된 안전성능함수를 활용하여 도로 유형별로 표준편차 0.5를 적용한 A, B, C, D 네 가지 LOSS를 산출하였다. 이렇게 산출된 LOSS는 도로교통사고 예방을 위한 관리 및 개선을 위해 활용할 수 있도록 준거로서 제시하였다 (Kwon, 2019). 본 연구에서는 도로 기하구조는 물론 실제 교통량 및 운전자 운전행태분석 데이터 등 시스템에서 빅데이터를 활용한 것이 가장 큰 특징이다. 본 연구에서 사용한 방법론과 분석 결과를 토대로 우리나라 전국 도로에 대하여 위험도로를 구분할 수 있는 기준을 설정할 수 있고 또한 교통안전 개선사업의 효과 등을 모니터링하고 계량화할 수 있을 것으로 판단된다. 또한 본 연구에서 도출한 위험도로 선정 방법을 통해 교통안전 개선사업의 대상을 선정하는 것도 가능할 것으로 판단된다. 전국 단위의 위험도로를 도출할 수 있는 방법론을 제시함으로써 궁극적으로는 우리나라 교통사고 건수, 사망자 수 및 부상자 수를 줄이는 국정 목표 달성에 기여할 수 있을 것으로 기대된다.

2. 향후 연구과제

본 연구는 국가표준노드·링크, 경찰청 교통사고데이터, ViewT 교통량 및 DTG 위험운전행동건수 등 교통 빅데이터를 추출하여 교통사고건수를 종속변수로 구간길이, 교통량, 차로수 등을 독립변수로 사용한 음이항 회귀모형을 활용하여 도로안전성능함수를 개발하였다. 안전성능함수 도출을 위해 사용한 독립변수를 도로의 종류에 구분 없이 동일하게 사용하였고 도로유형별 특징을 고려한 독립변수를 사용하지 못한 한계가 있어 향후 도로유형을 세분화하여 도로유형별로 교통사고와 연관성이 높은 독립변수를 적용하여 산출할 경우, 보다 정밀한 안전성능함수가 개발될 수 있을 것으로 보인다. 이를 위해 도로유형에 따라 교통사고건수와의 상관분석을 통해 유의한 변수를 찾아내는 추가 연구가 수행된다면 보다 더 정확한 안전성능함수를 도출할 수 있을 것으로 사료된다. 또한, 8개의 도로유형으로 분류하였으나, 기타도로의 경우 국가표준 노드 링크 상의 도로유형 코드별 매칭 데이터 수 비율이 11.42%로 낮아 안전성능함수 개발에서 제외 하였지만, 향후 도로별 세분화 연구를 진행할 경우 함께 포함한다면 더욱 정확한 안전성능함수를 개발할 수 있을 것으로 사료된다 (Kwon, 2019). 전국 단위의 도로에 대하여 가용할 수 있는 다양한 자료를 활용하여 도로유형별 안전성능함수를 개발하였으나 아직도 많은 자료들이 전국 단위의 도로에 구축되어 있지 못한 것이 사실이다. 특히, 지방도의 경우 다양한 기하구조를 가지고 있기 때문에 향후 연구를 통해 지방도의 특성을 고려한 변수를 찾아보고 이를 범주화한 연구가 필요하다 사료된다. 본 연구에서는 사업용자동차로부터 수집되는 운행기록계 자료를 이용하여 계산된 위험운전행동 건수를 모형 개발에 활용하였다. 빅데이터 중 운전자의 위험운전행동 건

수는 인적요인으로 기하구조와 관련성은 낮으나 특정 지역에 반복적으로 나타나는 위험운전행동의 경우 주변 환경, 도로의 수요, 기하구조 여건 등과 연관성이 있어 향후 연구과제로 다룰 수 있다. 향후 C-ITS 사업으로부터 생산되는 승용차를 포함한 다양한 차량들에서 수집되는 위험운전행동 건수 등의 자료를 활용할 필요가 있다. 본 연구에서는 포아송 회귀모형 또는 음이항 회귀모형과 같은 전통적인 통계적 모형을 이용하여 안전성능합수를 개발하였으나 향후 빅데이터의 증가를 토대로 전통적인 통계적 모형이 아닌 딥러닝과 같은 빅데이터 분석 기법을 활용하여 안전성능합수와 같은 모형의 개발이 가능할 것으로 기대되므로 빅데이터 분석 기법을 활용하는 방안을 지속적으로 연구할 필요가 있다고 판단된다. 마지막으로 본 연구에서 개발한 빅데이터를 활용한 전국 단위 도로유형별 안전성능합수 및 안전 등급은 우리나라 교통사고 건수, 사망자 수 및 부상자 수를 줄이기 위한 아주 미약한 노력에 지나지 않는다. 실질적으로 도로 교통안전성을 향상시키기 위해서는 이러한 기법을 바탕으로 전국 단위 도로유형별 안전등급을 생성, 관리 및 모니터링 할 수 있는 체계를 구축하고, 또한 개별 교통안전 개선사업의 효과뿐만 아니라 전국 단위의 교통안전 정책 및 사업의 효과를 지속적으로 관리할 수 있는 체계와 조직을 갖추는 것이 필요할 것으로 판단된다(Kwon, 2019). 따라서 안전성능합수 및 안전 등급 개발 이후에 개발된 합수 및 등급을 활용하기 위한 정책적인 부분의 연구가 추가적으로 필요하다고 판단된다.

ACKNOWLEDGEMENTS

이 논문은 2019년 아주대학교 대학원 건설교통공학 박사학위 논문을 토대로 작성하였고, 한국교통안전공단에서 구축·운영 중인 운수안전 컨설팅 지원시스템과 한국교통연구원에서 구축·운영 중인 ViewT 1.0의 자료를 활용하였음.

REFERENCES

- AASHTO(2010), *Highway Safety Manual*, First Edition, American Association of State and Highway Transportation Officials, Washington DC.
- AASHTO(2014), *Highway Safety Manual*, First Edition, 2014 Supplement, American Association of State and Highway Transportation Officials, Washington DC.
- Banihashemi M.(2012), "Sensitivity Analysis of Data Set Sizes for Highway Safety Manual Calibration Factors," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2279, pp.75-81.
- Bonneson J., Lord D., Zimmerman K., Fitzpatrick K. and Pratt M.(2007), *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions*, FHWA/TX-07/0-4703-4, Texas Transportation Institute, College Station.
- Bonneson J., Zimmerman K. and Fitzpatrick K.(2005), *Roadway Safety Design Synthesis*, FHWA/TX-05/0-4703-P1, Texas Transportation Institute, College Station.
- Cafiso S. and Di Silvestro G.(2011), "Performance of safety indicators in identification of black spots on two-lane rural roads," *Transportation Research Record 2237*, pp.78-87.
- Cafiso S., D'Agostino C. and Persaud B.(2013), "Investigating the influence of segmentation in

- estimating safety performance functions for roadway sections,” *Journal of Traffic and Transportation Engineering*, pp.1-8.
- Cafiso S., Di Graziano A., Di Silvestro G. et al.(2010), “Development of comprehensive accident models for twolane rural highways using exposure, geometry,” *consistency and context variables. Accident Analysis & Prevention*, vol. 42, pp.1072-1079.
- Caliendo C., Guida M. and Parisi A.(2007), “A Crash-Prediction Model for Multilane Roads,” *Accident Analysis & Prevention*, vol. 39, pp.657-670.
- Do C. W.(2013), *Traffic Safety Engineering*, Chungmungak.
- Jinyan L., Kirolos H., Priyanka A. and Albert G.(2013), “Full versus simple safety performance functions: comparison based on urban four-lane freeway interchange influence areas in Florida,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2398, pp.83-92.
- Kim D. H., Lee D. M. and Sung N. M.(2010), “A Development of Traffic Crash Frequency Prediction Models for Rural 3-legged Intersections,” *Journal of Transport Research*, vol. 17, no. 1, pp.37-48.
- Kim K. H.(1991), *Traffic safety engineering*, Thalim.
- Kononov J. and Allery B. K.(2004), “Level of service of safety: Conceptual blueprint and analytical framework,” *Transportation Research Record: Journal of the Transportation Research Board 1840*, (Statistical Methods and Modeling and Safety Data, Analysis, and Evaluation), pp.57-66.
- Kononov J. and Allery B. K.(2015), “Level of Service of Safety Revisited,” *Transportation Research Record: Journal of the Transportation Research Board*, No. 2514, pp.10-20.
- Korea Road Traffic Authority(2018), *Comparison of Traffic Accident of OECD Member States*.
- Korea Road Traffic Authority(2018), *Traffic Accident Statics Analysis*.
- Kwak H. C., Su Y. J., Kho S. Y. and Rhee S. M.(2010), “Establishment of Influence Area using Select Link Analysis for Highway Investment Projects,” *Journal of the Eastern Asia Society for Transportation Studies*, vol. 8, pp.593-602.
- Kwon K. A.(2019), *Development of national road Safety Performance Function and Level of Safety using Traffic Big Data*, doctoral dissertation, Ajou University.
- Lee D. M., Kim E. C., Choi E. J. and Kim D. H.(2010), “Development validation of accident modifications factors of 2-lane rural roadways,” *Journal of Korean Society of Transportation*, vol. 28, no. 3, pp.131-143.
- Ministry of Land, Infrastructure and Transport(2016), *The 8th National Traffic Safety Basic Subject(2017~2021)*.
- Ministry of Land, Infrastructure and Transport(2016), *Traffic Accident Statics Analysis*.
- National Police Agency(2016), *2015 Traffic Accident Statics*.
- Park J. T., Lee S. B. and Lee D. M.(2011), “Development of traffic prediction for rural road section,” *Journal of Transport Research*, vol. 18, no. 1, pp.15-26.
- Ryu J. D., Park S. M., Park S. H., Kwon C. W. and Yun I. S.(2018), “A study for development traffic accident prediction model using deep learning,” *The Korea Institute of Intelligent Transport Systems*, vol. 17, no. 4, pp.14-25.
- Yun D. S.(2016), *Traffic Demand Analysis*, Parkyoungsa, p.259.