

의사결정나무를 이용한 고속도로 공사구간 사고 심각도에 관한 연구

박용우 · 백세흠 · 박신행* · 권오훈
계명대학교 도시학부 교통공학전공

A Study on the Crash Severity of Expressway Work Zones Using Decision Tree

PARK, Yong Woo · BACK, Sehum · PARK, Shin Hyoung* · KWON, Oh Hoon

Department of Transportation Engineering, Keimyung University, Daegu 42601, Korea

*Corresponding author: shpark@kmu.ac.kr

Abstract

This study aims to identify factors that affect the degree of injury severity sustained in traffic crashes on work zone of Korean expressways. To this end, decision tree method was applied to identify influential factors on injury severity and compare characteristics of those factors between work zone and non-work zone. The results from the comparison show that the risk of severity was low when traffic volume and heavy vehicle ratio are high because the factors lower the overall section speed. On the other hand, when the traffic volume and the heavy vehicle ratio are low, the section speed increased and the tendency for high injury severity was confirmed. These findings are expected to help transportation planners and engineers understand which risk factors contribute more to severe injury in the work zones such that they can effectively prepare and implement safety countermeasures.

Keywords: crash severity, data mining, decision tree, expressway, work zone

초록

본 연구는 고속도로의 공사구간의 사고 심각도에 영향을 미치는 주요인들을 파악하고 사고 심각도와 관계를 규명하는 것을 목적으로 한다. 2011년부터 2015년까지 국내 고속도로에서 발생한 교통사고 자료를 공사구간 사고와 일반구간(비공사구간) 사고로 구분하였으며, 이 중 물피사고를 제외하고 분석에 활용하였다. 의사결정나무(Decision Tree)기법을 적용하여 공사구간과 일반구간의 사고 심각도에 영향을 끼치는 요인들을 파악한 뒤 사고 특성 및 영향요인의 차이점을 비교하였다. 특히 공사구간에서는 교통량이 많고 중차량 비율이 높을 경우 전반적인 구간속도 저하로 사고 심각도가 낮아지는 반면, 교통량이 적고 중차량 비율이 낮을수록 구간속도가 증가하여 심각도가 높은 사고로 이어지는 경향을 확인할 수 있었다. 본 연구를 통해 도출한 공사구간 사고심각도 영향 요인들은 공사구간 안전 관리 대책 마련에 도움을 줄 수 있을 것으로 기대한다.

주요어: 사고 심각도, 데이터마이닝, 의사결정나무, 고속도로, 공사구간

J. Korean Soc. Transp.
Vol.34, No.6, pp.535-547, December 2016
<https://doi.org/10.7470/jkst.2016.34.6.535>

pISSN : 1229-1366
eISSN : 2234-4217

Received: 20 October 2016

Revised: 12 December 2016

Accepted: 14 December 2016

Copyright ©
Korean Society of Transportation

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

서론

1. 연구배경 및 목적

최근 고속도로의 노후화에 따라 도로의 보수나 확장뿐만 아니라 각종 시설물의 유지 보수 등과 같은 공사가 지속적으로 증가하면서 도로 이용자를 위한 공사구간 안전관리의 중요성이 더 커지고 있다. 고속도로의 공사작업은 일부 차로를 점유하거나 차단하여 진행하게 되는데, 이 때 차량의 속도감소, 차로변경 등에 따른 차량의 상충이 증가하여, 특히 빠른 속도로 주행하는 고속도로 상에서 차로가 감소되는 구간에서는 차량의 안전뿐만 아니라 작업자의 안전까지도 위협받게 된다. 교통사고에 관한 기존 연구들은 주로 사고와 관련 있는 다양한 요인들을 밝혀내어 각각의 요인들이 교통사고에 어떤 영향을 미치는지 분석하였다. 그러나 교통사고의 발생은 하나의 요인으로 인해 발생하는 것이 아니라 여러 가지 요인이 복합적으로 작용하여 발생하므로, 이러한 관계를 면밀하게 밝히는 것이 필요하다.(Lee J. Y. et al., 2008)

해외에서는 고속도로 공사구간을 사고 위험 구역으로 보고 공사구간의 교통사고의 감소를 위해 사고의 특성 및 발생요인에 대한 연구가 많이 진행되어 왔다. 하지만 국내의 교통사고 관련 연구는 주로 고속도로 일반구간의 교통사고에 대한 연구가 주를 이루고 있고, 공사구간의 교통사고에 대한 연구는 아직 미진한 실정이다. 따라서 공사구간의 사고 심각도에 관한 주요인들을 파악하는 것은 공사구역에서 발생하는 사고특성을 이해하고, 사고가 발생하는데 영향을 주는 주요 원인을 제거함으로써 사고를 예방하는데 도움을 줄 수 있다. 이에 본 연구에서는 고속도로 교통사고 자료를 바탕으로 사고 심각도 분석을 수행하여 심각도에 영향을 주는 주요인들을 파악하고, 전체 고속도로 상에서 일반구간과 공사구간을 비교하여 사고 심각도와와의 관계를 규명하고자 한다.

2. 연구방법 및 범위

본 연구의 시간적 범위는 2011년부터 2015년까지의 5년 동안이며, 공간적 범위로는 한국도로공사 관리 노선을 대상으로 진행하였다. 한국도로공사 고속도로 사고 자료와 국토교통부 교통량 정보제공 시스템(<http://www.road.re.kr>)에서 제공하는 교통량 자료를 활용해 AADT, 트럭 및 버스 비율 자료를 수집하였다.

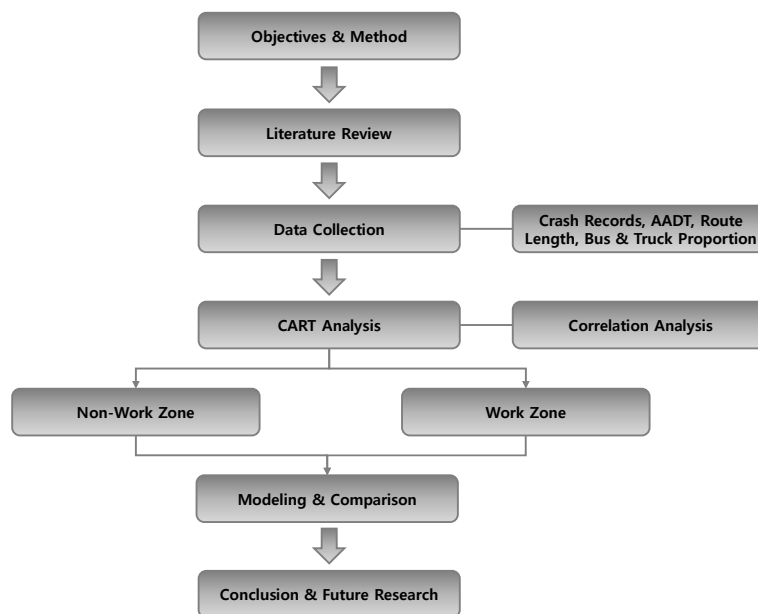


Figure 1. Study flow

사고 심각도 분석은 데이터마이닝기법의 하나인 의사결정나무법(Classification And Regression Tree, CART)을 적용하였으며, 상관분석을 통해 적절한 독립변수를 설정하였다. 공사구간과 일반구간에 대해 각각 의사결정나무를 구축하여 영향요인을 파악하였으며, 이를 서로 비교하여 시사점을 도출하였다. 본 연구의 수행절차는 Figure 1과 같다.

문헌고찰

1. 사고 심각도에 관한 연구

Ha et al.(2005)은 교차로에서 발생하는 교통사고의 심각도를 분석하기 위해 충청남·북도의 77개 교차로에서 발생한 교통사고에 대해 순서형 프로빗 모형을 이용하여 사고 심각도를 분석하였다. 그 결과 부도로 교통량, 주도로 중차량 비율, 주·부도로 우회전 비율 등이 유의미한 요인으로 도출되었다.

Lee J. Y. et al.(2008)은 구조방정식 모형을 이용하여 고속도로 사고 심각도를 분석하였다. 사고 심각도에 영향을 미치는 요인들을 크게 ‘도로 요인’ 및 ‘운전자 요인’, ‘환경 요인’ 등으로 구분하고, 총 2,880개의 사고데이터를 이용하여 구조방정식 모형을 구축하였는데, 도로 및 환경 요인은 통계적으로 유의한 수준에서 사고 심각도와 강한 관계를 가지는 것으로 나타났다.

Mun et al.(2011)은 2005년부터 2008년까지 국내 전 고속도로를 대상으로 이항 로지스틱 모형을 적용하여 사고 심각도에 미치는 요인을 사망사고 모형과 차량완파사고모형으로 구분하여 분석하였다. 그 결과 사고발생지점, 사고 유형, 노면상태 등의 변수가 심각도에 영향을 끼치는 주요인으로 분석되었다.

Park et al.(2012)은 2008년 고속도로에서 발생한 사고에 대해 순서형 프로빗과 순서형 로짓, 다항로짓 모형을 적용하여 고속도로 사고 심각도 영향요인을 분석하였다. 각각의 범주별로 기준이 되는 변수값과 비교하였을 때 사고 심각도를 증가시키는 것으로는 심야시간대(자정-오전 6시), 졸음운전, 중앙선 침범, 차대차 사고, 차대사람 사고, 차량단독사고, 2대 이상의 차량이 관련된 사고, 승합차가 포함된 사고이며, 심각도를 감소시키는 요인으로는 오전 시간대(오전 6시-정오), 램프, 톨게이트, 차량결합, 장애물과 열악한 도로환경, 비 또는 눈 등으로 파악되었다.

2. 공사구간 사고 심각도에 관한 연구

Park et al.(2008)은 고속도로 전 구간을 대상으로 2003년부터 2005년까지 3년 동안의 교통통제가 이루어진 공사구간에서 발생한 교통사고를 조사하고 유형별로 정리하여 일반구간과 공사구간을 비교·분석하였다. 그 결과 일반구간보다 공사구간에서 충돌사고가 더 빈번하게 발생하고, 사고 발생 시 심각도가 더욱 높은 것으로 나타났다.

Hong et al.(2011)은 2006년부터 2007년까지 서울 및 경기도 지역의 공사구간에서 발생한 교통사고를 대상으로 순서형 프로빗 모형을 적용하여 사고 심각도 분석을 진행하였다. 그 결과 노면상태, 사고 지점의 도로선형, 공사차량 포함여부, 공사장 진출입 여부 등이 공사구간 교통사고 심각도를 설명하는 주요인으로 분석되었다.

Yoon et al.(2016)은 공사구간 교통관리특성을 반영하여 사고 심각도 영향요인을 순서형 프로빗 모형으로 분석하였다. 분석 결과, 중분대작업, 도로재포장, 주의구간이 시작되는 지점부터 공사구간 방향으로 1.5-1km 구간, 차사람, 용량 67%감소, 용량 33%감소, 사고 지점 상류부로부터 가장 인접한 검지기에서 측정된 사고 5분전 속도, 사고 지점 하류부로부터 가장 인접한 검지기에서 측정된 사고 5분전 속도 등이 주요 영향요인으로 도출되었다.

Li et al.(2008)은 캔자스 주 고속도로 공사구간에서 발생한 사고의 특성을 분석하였다. 그 결과 충돌사고에서 사망사고의 비율이 가장 높았으며, 추돌사고에서 부상사고의 비율이 가장 높았다. 그리고 중차량 비율이 높아짐에 따라 사망사고의 비율이 높아지는 것을 확인하였다.

3. 의사결정나무(Decision Tree)를 활용한 연구

Park et al.(2011)은 도로교통공단의 ‘사고 잦은 곳 개선사업’을 참조하여 1998년부터 2007년까지 개선사업이 완료된 1,044개의 지점자료를 구축하여 교통사고율 감소에 영향을 미치는 개선내역에 대해 CART 분석을 수행하였다. 수도권과 서울시의 경우 교차로의 형태(3지 교차로, 4지 교차로)에 따라 시설물 개선효과가 다르게 나타났고, 광역시의 경우 신호교차로/비신호교차로에 따라 시설 개선 효과가 다르게 나타나는 것으로 분석하였다.

Han et al.(2014)은 CART 분석기법으로 서울시의 도로교통사고 자료를 이용하여 화물자동차 사고와 비화물자동차 사고를 비교하였다. 분석결과 화물자동차 사고는 보호장구 미착용 시 사망사고 가능성이 높아짐을 증명하였다.

Ha et al.(2010)은 경부고속도로(양재-부산)의 2003년부터 2007년까지 발생한 돌발상황 이력자료를 이용하여 사고등급(A, B, C등급)에 따라 고속도로 돌발상황 처리시간 예측모형을 개발하고 고속도로 운영관리자가 돌발상황 발생 시 신속히 적용할 수 있는 의사결정나무(Decision Tree)를 구축하였다.

Lee J. M. et al.(2008)은 전국의 4차로 국도를 대상으로 교통사고율 및 도로환경요인 관련 자료를 수집하여 CART분석을 이용한 교통사고예측모형을 개발하고 기존의 교통사고예측 모형과 CART 분석모형을 비교하였다. 교통사고발생률에 큰 영향을 미치는 도로기하구조 요인으로는 구간거리(km), 횡단보도 폭(m), 횡단길어깨(m), 교통량 순으로 나타났다.

4. 의사결정나무(Decision Tree)의 기본 이론

의사결정나무(Decision Tree)는 데이터마이닝 분석의 대표적인 기법중 하나로 주어진 데이터를 특정 기준에 따라 분류(Classification)하거나 예측(Prediction)하는 분석방법이다. 의사결정나무분석은 나무구조를 이루고 있어 연구자가 분석과정을 쉽게 이해하고, 설명할 수 있으며 의사결정을 하는데 직접적으로 사용할 수 있는 장점이 있다. (Park et al., 2011)

이 알고리즘은 분리기준과 정지규칙, 가지치기에서 서로 다른 형성기준이 있고 이들이 어떻게 결합하느냐에 따라 서로 다른 알고리즘으로 나뉜다. 대표적인 알고리즘으로는 CHAID, QUEST, CART 등이 있는데, CHAID 알고리즘은 목표변수가 이산형일 경우 카이제곱 검정-통계량을 사용하고, 목표변수가 연속형일 경우 F-검정을 이용하여 다지 분리(Multiway Split)를 수행하는 알고리즘이다. QUEST 알고리즘은 명목형 목표변수에 대해서만 분석할 수 있으며 카이제곱 검정-통계량을 이용하여 예측변수의 측도에 따라 서로 다른 분리기준을 사용하여 이진분리(Binary Split)를 하는 알고리즘이다.

본 연구에서 사용할 CART 알고리즘은 그룹(Group) 내 동질성이 비슷한 그룹을 판별하기 위해 불순도함수(Impurity Function)를 사용하여 이진분리를 한다. 목표변수가 이산형일 경우 지니지수(Gini Index) 혹은 투잉지수(Toving Index)를 분리기준으로 사용하고 목표변수가 연속형인 경우 분리기준으로 분산의 감소량을 사용한다. 본 연구에서는 동질성이 높은 변수를 그룹화 할 때 가장 보편적으로 사용하는 지니지수를 활용하였으며, 산출식은 Equation 1과 같다.

$$G = \sum_{j=1}^c P(j)(1 - P(j)) = 1 - \sum_{j=1}^c P(j)^2 = 1 - \sum_{j=1}^c (n_j/n)^2 \quad (1)$$

여기서 c 는 목표변수의 범주의 수, $P(j)$ 는 j 범주에 분류될 확률이며, n 은 마디에 포함되어 있는 관찰치의 수, n_j 는 목표변수의 j 번째 범주에 속하는 관찰치의 수를 나타낸다.(Lee J. M et al., 2008)

자료 수집 및 일반 특성 분석

1. 자료 수집 및 사고 특성 분석

본 연구에서는 한국도로공사에서 제공하는 고속도로 사고 자료를 이용하여 분석을 진행하였다. 2011년부터 2015년까지 5년 동안 고속도로 전 노선에서 발생한 사고 수는 40,684건이며 이 중 일반구간에서 발생한 사고 수는 39,979건, 공사구간에서 발생한 사고 수는 735건이다. 공사구간의 사고 빈도수는 일반구간과 비교하면 현저히 적은 하나, 사고 심각도에 따른 사고 특성을 비교하면 많은 차이가 있다. 사고 시 사망자가 발생한 경우 사망사고, 사망자 없이 중상자만 있을 경우 중상사고, 경상자만 있을 경우 경상사고라고 정의하며, 사망자 또는 부상자가 없이 재산적 손해만 입은 경우 물피사고로 구분한다. 수집한 교통사고 자료를 사고 심각도에 따라 구별하면 Table 1과 같다. 공사구간의 경우 물피사고를 제외한 나머지 사망, 중상, 경상사고의 비율이 일반구간보다 높으며, 특히 사망사고의 경우 전체 8%로 일반구간 2.6%와 비교해 약 3배 이상 많이 발생함을 알 수 있다.

Table 1. Crash frequency and proportion by crash severity

Section	Crash Severity	Frequency	Proportion(%)
Non-work zone	Fatal	1,032	2.6
	Severe Injury	996	2.5
	Light Injury	2,763	6.9
	Property Damage Only	35,158	88.0
	Total	39,949	100
Work zone	Fatal	59	8.0
	Severe Injury	33	4.5
	Light Injury	74	10.1
	Property Damage Only	569	77.4
	Total	735	100

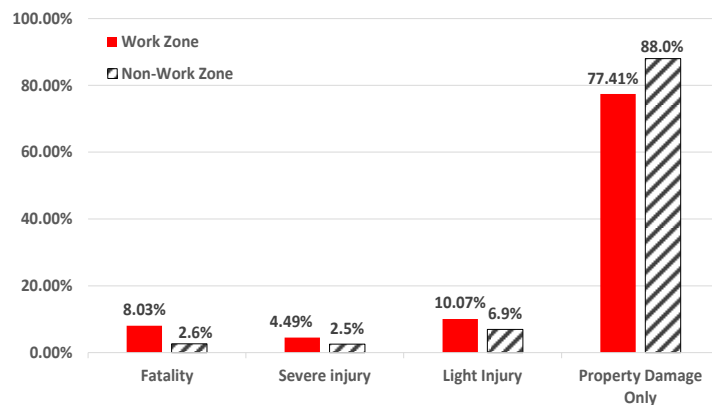


Figure 2. Crash configuration by severity

또한, 고속도로에서 사고가 발생한 경우 인명피해 및 차량피해 유형에 따라 A-D등급으로 분류하는데 그 기준은 Table 2와 같다(물피사고인 D등급은 제외). 사고등급별 사고현황을 살펴보면 Table 3과 같이 공사구간의 경우 일반구간에 비해 A등급, B등급, C등급 사고의 발생 비율이 높은 것을 알 수 있다.

Table 2. Classification criteria

Type	Level A	Level B	Level C
Casualty	Fatalities 3 & more Casualties 10 & more Injuries 20 & more	Fatalities 1 & more Injuries 5 & more	Injuries 1 & more
# of vehicles involved	10 & more (Fatal Crash)	10 & more (Injured Crash)	3 & more

Table 3. Crash frequency and proportion by crash class

Section	Crash Class	Frequency	Proportion(%)
Non-work zone	Class A	22	0.06
	Class B	1,095	2.74
	Class C	9,135	22.87
	Class D	29,697	74.34
	Total	39,949	100
Work zone	Class A	3	0.41
	Class B	62	8.44
	Class C	183	24.90
	Class D	487	66.26
	Total	735	100

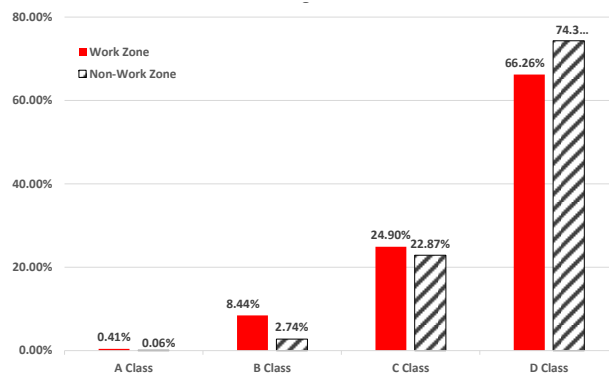


Figure 3. Crash configuration by class

특히, B등급사고의 경우 일반구간과 비교해 약 2배 이상 높으며, A등급사고의 경우 약 6배 이상 높은 것으로 나타나 사고 발생 시 일반구간과 비교해 더 심각한 사고가 발생할 수 있음을 다시 한 번 확인할 수 있다.

결론적으로 공사구간은 일반구간에 비해 사망, 중상사고의 비중이 높아 사고 심각도가 높다고 볼 수 있으며, 사고 등급별로도 더 높은 등급의 사고 발생 비중이 높으므로 사고요인에 대한 분석이 필요함을 알 수 있다.

2. 변수 설정

본 연구의 주목적은 공사구간에서 발생한 사고를 사망사고, 중상사고, 경상사고로 구분하고 각 사고에 영향을 끼치는 주요인을 의사결정나무를 통해 분석하는 것이다. 이를 위해 고속도로 사고 자료에서 사고 발생에 영향을 미칠 것으로 판단되는 요인들을 Table 4와 같이 독립변수로 정리하였다.

Table 4. Variables

Variables	Possible values	Statistics	
		Non-work zone	Work zone
Crash severity	1:Fatal	1,032(34.1%)	59(35.5%)
	2:Severe injury	996(32.9%)	33(19.9%)
	3:Light injury	1,000(33.0%)	74(44.6%)
Location	1:TG(TCS)	33(1.1%)	0(0.0%)
	2:TG(Hipass)	15(0.5%)	0(0.0%)
	3:Ramp	179(5.9%)	1(0.6%)
	4:Mainline	2,651(87.5%)	155(93.4%)
	5:Tunnel	100(3.3%)	10(6.0%)
	6:Service area	50(1.7%)	0(0.0%)
Causal factor	1:Driver	2,727(90.1%)	5(3.0%)
	2:Vehicle	142(4.7%)	156(94.0%)
	3:Others	159(5.3%)	5(3.0%)
Day of the week	1:Weekday	2,082(68.8%)	146(88.0%)
	2:Weekend	946(31.2%)	20(12.0%)
Weather	1:Sunny	1,974(65.2%)	125(75.3%)
	2:Cloudy	517(17.1%)	27(16.3%)
	3:Rainy	471(15.6%)	14(8.4%)
	4:Snowy	66(2.2%)	0(0.0%)
Day & night	1:Day	1,599(52.8%)	118(71.1%)
	2:Night	1,429(47.2%)	48(28.9%)
Horizontal alignment	1:Straight	2,191(72.4%)	120(72.3%)
	2:Right curve	413(13.6%)	21(12.7%)
	3:Left curve	424(14.0%)	25(15.1%)
Vertical grade	1:Flat	1,625(53.7%)	92(55.4%)
	2:Uphill	679(22.4%)	31(18.7%)
	3:Downhill	724(23.9%)	43(25.9%)
Cut & fill	1:None	1,547(51.1%)	81(48.8%)
	2:Fill	1,034(34.1%)	65(39.2%)
	3:Cut	447(14.8%)	20(12.0%)
Median barrier	1:Yes	2,655(87.7%)	153(92.2%)
	2:No	373(12.3%)	13(7.8%)
Roadside barrier	1:Yes	2,373(78.4%)	127(76.5%)
	2:No	655(21.6%)	39(23.5%)
At fault	1:Car	1,891(62.5%)	73(44.0%)
	2:Truck	1,123(37.1%)	92(55.4%)
	3:Others	14(0.5%)	1(0.5%)
AADT		1,191–249,256(veh.) avg.:68,026(veh.)	7,474–243,868(veh.) avg.:63,952(veh.)
Bus Proportion		0.3–11.4(%) avg.:4.10(%)	0.7–8.5(%) avg.:3.76(%)
Truck Proportion		8.3–58.8(%) avg.:29.10(%)	11.2–60.4(%) avg.:30.74(%)

본 연구는 고속도로 공사 시행 시 사고 방지를 위한 안전 대책을 마련하는 데에 초점을 두고 있으므로 사고 심각도에 영향을 끼치는 요인을 도로환경적인 요인으로 제한하였다. 또한, 교통량, 버스비율, 트럭비율은 사고 발생 시 사고 심각도에 영향을 미칠 것이라 판단해 사고가 발생한 지점에 대한 노선 및 이점 자료를 토대로 교통량 정보제공 시스템에서 제공하는 각 노선의 구간별 AADT, 버스비율, 트럭비율 자료를 변수로 추가하여 활용하였다.

의사결정나무를 구축할 때 노드 분리 시 변수의 세부 분류수가 많아지게 되면 정확히 분류되지 않는 경향이 있으므로 변수의 세부 분류를 최대한 축소하였으며, 총 14개의 독립변수를 사용하여 분석을 진행하였다.

모형 구축 및 비교분석

1. 모형 구축

본 연구에서는 의사결정나무를 이용해 사고 자료를 일반구간과 공사구간으로 구분해 모형을 구축하였고, 분석 결과를 서로 비교하여 공사구간의 사고 심각도에 영향을 끼치는 주요인을 알아보려고 하였다. 모형 구축 시 물피사고까지 고려하게 될 경우 물피사고의 빈도가 너무 높아 물피사고로 치우친 예측값이 나올 수 있으므로 종속변수는 사망사고, 중상사고, 경상사고 3가지로 설정하였다. 이에 따라 물피사고를 제외한 3,028건의 일반구간의 사고와 166건의 공사구간 사고자료가 활용되었다. 의사결정나무 구축을 위한 CART분석은 SPSS를 통해 수행하였으며 노드 분리는 지니계수(Gini Index)를 이용해 이지분리를 시행하였다.

의사결정나무를 구축한 결과, 일반구간과 공사구간 모두 23개의 노드와 12개의 터미널 노드, 5개의 계층으로 형성되었다. 모형의 예측력으로 설명될 수 있는 재현율(recall)은 Table 5와 같다. 공사구간의 재현율은 63.3%로 모형으로 사고 심각도에 따른 사고 예측 시 오분류 가능성이 36.7%가 됨을 의미한다. 일반구간과 비교했을 때 전반적으로 예측력이 높고 비율의 편차가 크지 않아 보다 안정적이라 볼 수 있다.

Table 5. Comparison of predictive power

Crash	Recall	
	Non-work zone	Work zone
Fatal	41.4%	64.4%
Severe Injury	59.3%	57.6%
Light Injury	32.2%	64.9%
Total	44.3%	63.3%

전반적으로 예측력이 높지 않은 원인으로는 교통사고의 특성 상 임의성이 크기 때문에 특정 요인에 무관하게 발생하는 사고가 많고, 사용된 변수 이외에 사고에 영향을 주는 요인들이 존재하기 때문인 것으로 판단된다. 교통사고에 영향을 줄 수 있는 모든 요인들을 적용하는 것은 한계가 있으므로, 본 연구에서는 도로, 환경적 요인에 국한하여 분석을 실시하였다. 또한 이지분리를 시행하는 CART 분석 특성상 세부 분류 수가 많으면 효과적으로 분류되지 않을 수 있다. 추후 연구 시에는 이를 반영하여 사고 발생에 영향을 미치는 독립변수를 추가하고 세부 분류를 이원화시키는 한편, 표본을 보다 많이 확보하여 분석을 진행한다면 예측력이 좀 더 향상될 것이라 생각된다.

2. 일반구간과 공사구간 비교분석

일반구간과 공사구간 의사결정나무 구축 결과는 각각 Figure 4와 Figure 5로 나타내었다. 분류 기준 및 독립변수의 중요도에서 차이점이 나타났다. Table 6과 같이 일반구간에서 가장 중요한 독립변수는 버스비율로 나타났다. 버스비율의 중요도는 1.3%로 다소 낮은 편이지만 의사결정나무 구축 시 가장 큰 영향을 미치는 변수이며 그 다음으로 AADT, 절성토구분, 트럭비율, 주야 순으로 나타났다. 공사구간에서 가장 중요한 독립변수는 트럭비율이고 그 뒤로 절성토구분, AADT, 날씨, 종단구배 순으로 나타나 일반구간과 다소 차이가 있다. 특히 트럭 비율은 공사구간에서 발생하는 사망사고, 중상사고, 경상사고를 구분하는 주요 요인으로 나타났으므로 고속도로 공사 시 트럭 비율에 따른 사고 방지 대책을 마련한다면 공사구간에서의 사고 심각도를 낮추는 데 도움이 될 것이라 판단된다.

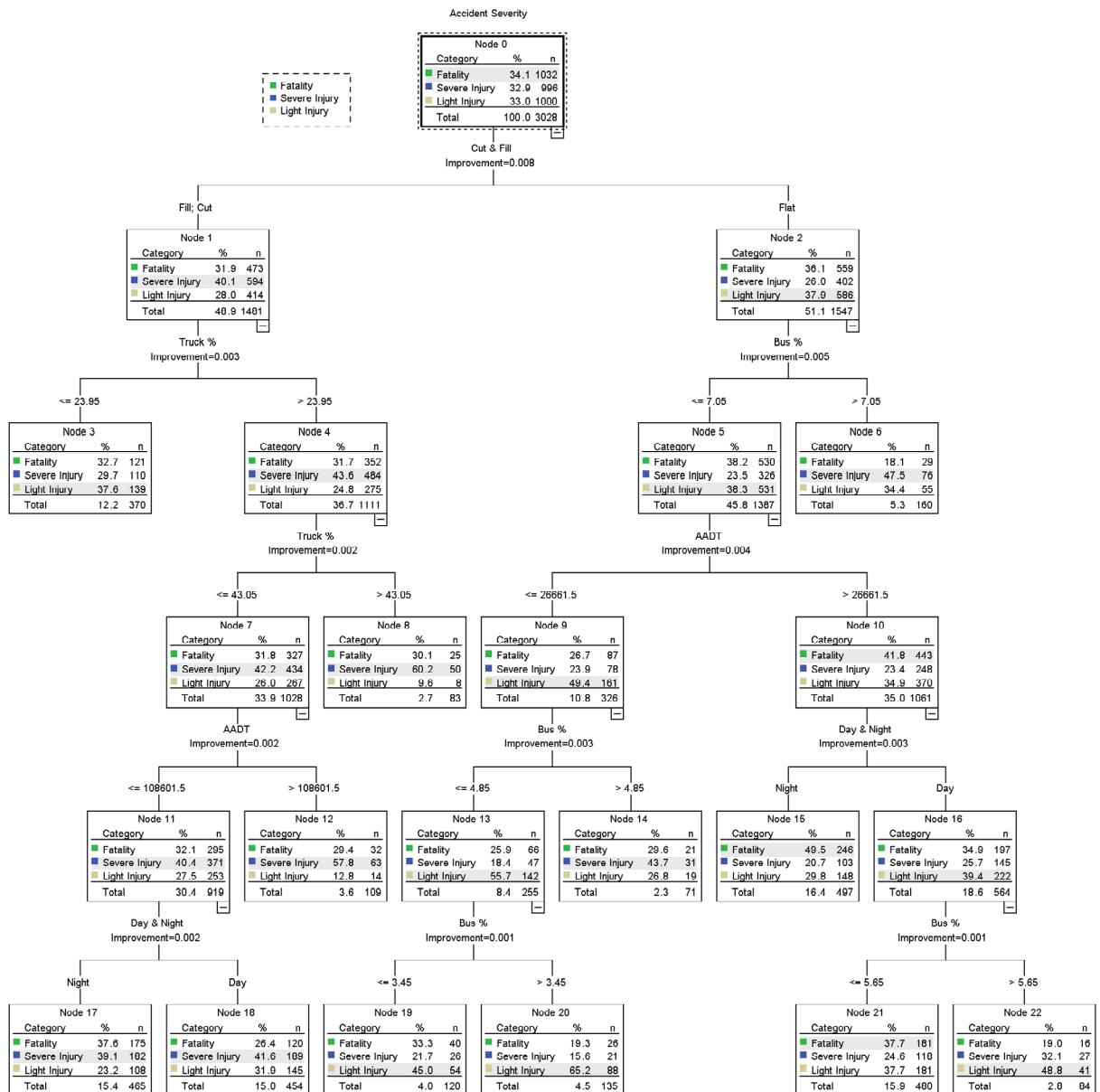


Figure 4. Decision tree - non-work zone

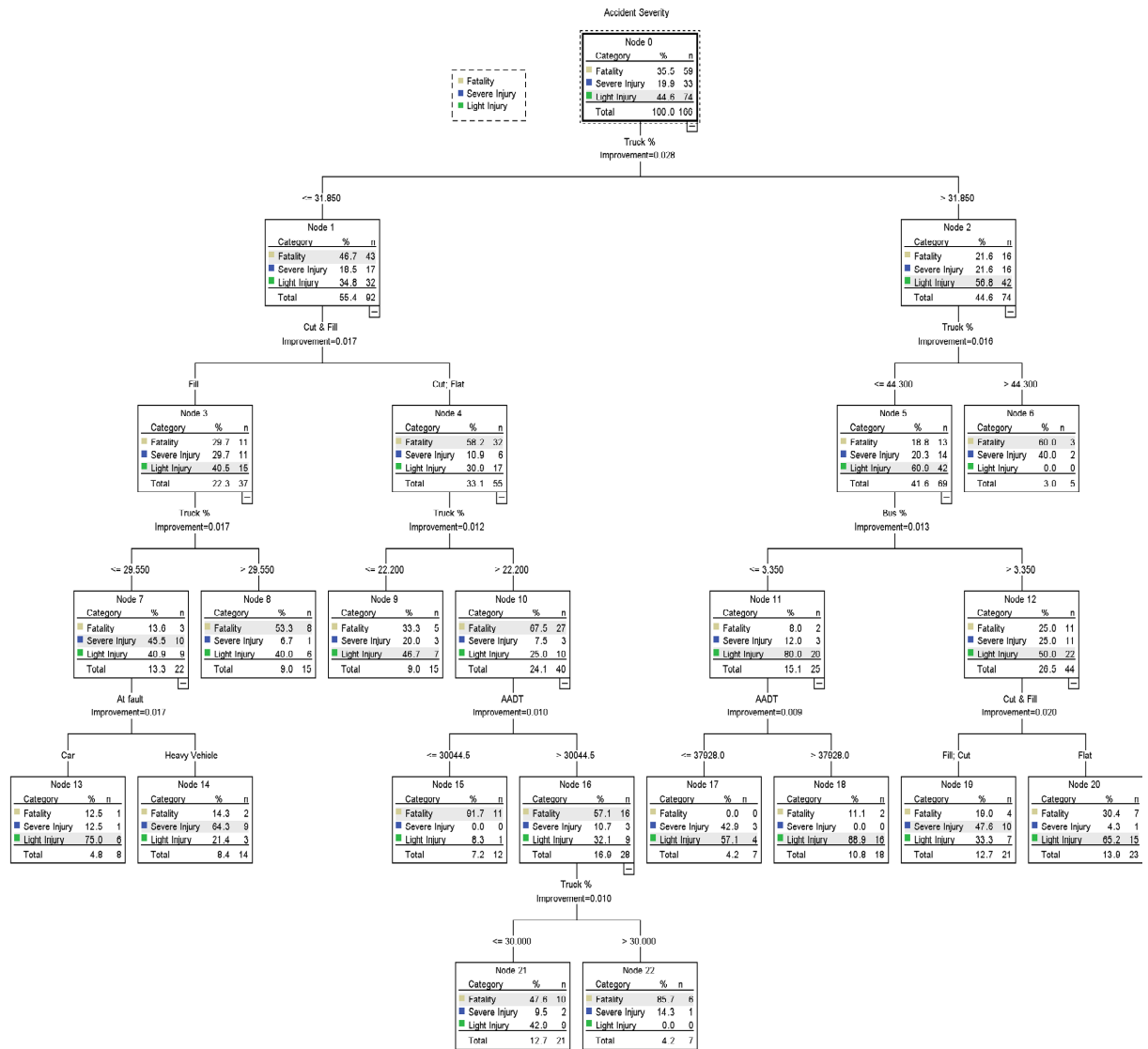


Figure 5. Decision tree - work zone

Table 6. Significance of explanatory variables

Ranking	Non-work zone		Work zone	
	Explanatory variables	Significance	Explanatory variables	Significance
1	Bus proportion	0.013	Truck proportion	0.095
2	AADT	0.008	Cut & fill	0.042
3	Cut & fill	0.008	AADT	0.034
4	Truck proportion	0.008	Weather	0.032
5	Day & night	0.005	Vertical grade	0.024
6	Roadside barrier	0.003	Bus proportion	0.019
7	Location	0.002	Causal factor	0.018

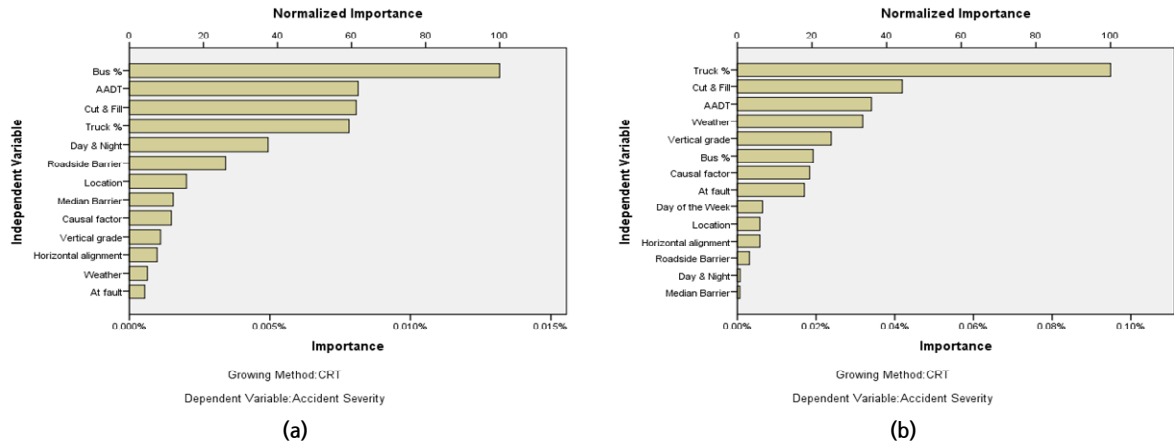


Figure 6. Significance of explanatory variables: (a) non-work zone, (b) work zone

의사결정나무를 통해 구축된 노드를 기준으로 일반구간 및 공사구간에서의 사망사고, 중상사고, 경상사고에 가장 큰 영향을 미치는 주요 요인을 Table 7과 같이 정리하였다.

Table 7. Influential factors on crash severity and their criteria

Section	Crash	Node number	Criteria
Non-work zone	Fatal	15	Flat Bus% ≤ 7% AADT > 26,661(veh.) night
	Severe Injury	8	Cut or fill 43% < Truck%
	Light Injury	20	Flat AADT ≤ 26,661(veh.) 3.4% < Bus% ≤ 4.8%
Work zone	Fatal	15	Cut or flat 22.2% < Bus% ≤ 31.8% AADT ≤ 30,044(veh.)
	Severe Injury	14	Fill Bus% ≤ 29.6% Heavy vehicle
	Light Injury	18	31.8% < Truck% ≤ 44.3% Bus% ≤ 3.4% 37,928(veh.) < AADT

공사구간의 경상사고는 트럭의 비율이 31.8% 이상이고 AADT가 37,928대 이상일 시 발생하는 것으로 나타났는데 이는 교통량이 많고 트럭의 비율이 높을 때 주로 발생한다는 것을 의미한다. 이러한 구간의 경우 교통량에 영향을 받아 운전자들이 방어 운전을 하는 경향이 있어 차량의 평균 속도가 감소하게 되고 이러한 속도로 공사구간에서 사고가 발생할 때는 심각한 사고로 이어지지 않기 때문인 것으로 판단된다. 이와 반대로 사망사고의 경우 AADT가 30,044대 이하이고, 트럭의 비율이 22.2-31.8% 사이 일 때 많이 발생한다. 교통량이 적고 트럭의 비율이 낮아지게 되면 교통흐름이 원활해지고 차량들의 평균 속도는 높아진다. 이러한 속도로 공사구간 내에서 사고가 발생하게 되면 사망사고로 이어져 사고 심각도가 높아지는 것으로 판단된다. 또한, 공사구간 내에서는 사고 원인 차종이 중차량

일 때 중상사고의 요인으로 작용하는 것으로 나타났으므로 공사 진행 시 중차량이 안전하게 통행할 수 있는 방안을 마련한다면 사고 심각도를 낮출 수 있을 것이라 생각된다. 고속도로의 공사구간의 사고 심각도는 일반구간과 비교하여 트럭비율 및 AADT에 주로 영향을 받는 것으로 나타났는데, 이는 차량들의 평균 속도와 연관이 있는 것으로 판단된다. 추후에는 이를 규명하기 위해 속도 자료를 추가로 활용하여 연구를 진행할 필요가 있다.

결론 및 향후 연구과제

본 연구에서는 의사결정나무의 분석방법인 CART분석을 이용하여 일반구간과 공사구간의 사고 심각도에 영향을 미치는 주요인을 규명하고자 하였다. 2011년부터 2015년까지 5년간의 고속도로 사고자료를 이용하여 사망사고, 중상사고, 경상사고를 종속변수로 설정하였으며 14개의 독립변수를 선정하여 의사결정나무를 구축하였다. 구축된 의사결정나무를 통해 다음과 같은 결과를 도출하였다.

첫째, 일반구간과 공사구간은 사망사고, 중상사고, 경상사고에 영향을 미치는 주요 변수에 차이가 있다. 일반구간에서는 구간의 버스비율, AADT, 절성토구분 순으로 중요한 변수이며 공사구간의 경우 구간의 트럭비율, 절성토구분, AADT 순으로 나타났다.

둘째, 공사구간의 경우 사망사고의 주요인이 트럭의 비율로 나타났으나, 트럭 비율이 높을수록 사망사고가 발생하기보다 일정 수준 이하일 때 오히려 사망사고의 비율이 증가한다는 것이 발견되었다. 또한 사고 원인 차종이 중차량일 시 공사구간에서의 중상사고에 영향을 미치는 것으로 나타났다.

결과적으로 트럭비율이 공사구간 사고 심각도에 가장 큰 요인으로 작용한다는 것을 감안하면, 공사 진행 시 구간을 통과하는 트럭 비율에 따라 안전대책을 수립함으로써 사고의 심각도를 낮출 수 있을 것이라 판단된다.

본 연구에서 구축한 의사결정나무는 일반구간과 공사구간의 사고 심각도에 영향을 미치는 요인을 다르게 분석해 냈지만 예측력이 떨어진다는 단점이 있다. 예측력이 높지 않은 데에는 교통사고가 특정 요인에 무관하게 발생하는 사고가 많고, 사용된 변수 외에 사고에 영향을 주는 요인이 존재하기 때문인 것으로 판단된다. 본 연구에서는 독립변수를 도로, 환경적인 요인으로 제한하여 분석했기 때문에 사고원인 및 운전자와 관련된 변수는 고려하지 않았다. 따라서 추후에는 이번 연구에서 반영하지 못한 다양한 변수를 활용하여 모형을 구축해볼 필요가 있을 것으로 생각된다. 또한 단순히 의사결정나무로만 사고 심각도를 분석하기보다는 다항로지스틱모형, 인공신경망과 같은 다양한 통계적 분석 모형을 함께 적용하여 종합적으로 사고 심각도를 분석하는 연구도 진행되어야 할 것이다.

ACKNOWLEDGEMENT

This research was supported by a grant(16RDRP-B115529-01) from Regional Development Research Program (RDRP) funded by Ministry of Land, Infrastructure and Transport of Korean Government.

This work was also supported by a research project of the Korea Expressway Corporation.

알림: 본 논문은 대한교통학회 제75회 학술발표회(2016.09.23)에서 발표된 내용을 수정·보완하여 작성된 것입니다.

REFERENCES

- Ha O. K., Oh J. T., Won J. M., Sung N. M. (2005), The Study on the Accident Injury Severity Using Ordered Probit Model, J. Korean Soc. Transp., 23(4), Korean Society of Transportation, 47-55.
- Ha O. K., Park D. J., Won J. M., Jung C. H. (2010), The Prediction Models for Clearance Times for the Unexpected Incidences According to Traffic Accident Classifications in Highway, J. Korea Inst. Intell. Transp. Syst., 9(1), The

- Korea Institute of Intelligent Transportation Systems, 101-110.
- Han S. J., Jo W. B., Chang J. S. (2014), Analysis on Truck Accidents Using Classification and Regression Trees, *Journal of Transport Research*, 21(4), The Korea Transport Institute, 87-103.
- Hong J. Y., Kim K. T., Lee S. B. (2011), Traffic Safety: Developing the Accident Injury Severity on a Field of Construction Work Using Ordered Probit Model, *Journal of the Korean Society of Safety*, 26(2), The Korean Society of Safety, 89-98.
- Lee J. M., Kim T. H., Lee Y. T., Won J. M. (2008), Developing the Traffic Accident Prediction Model Using Classification and Regression Tree Analysis, *Journal of the Korean Society of Road Engineers*, 10(1), Korean Society of Road Engineers, 31-39.
- Lee J. Y., Chung J. H., Son B. S. (2008), Analysis of Traffic Accident Severity for Korean Highway Using Structural Equations Model, *J. Korean Soc. Transp.*, 26(2), Korean Society of Transportation, 17-24.
- Li Y., Bai Y. (2008), Comparison of Characteristics Between Fatal and Injury Accidents in the Highway Construction Zones, *Safety Science*, 46(4), Elsevier, 646-660.
- Mun S. R., Lee Y. I. (2011), Analysis of Traffic Crash Severity on Freeway Using Hierarchical Binomial Logistic Model, *Journal of the Korean Society of Road Engineers*, 13(4), Korean Society of Road Engineers, 199-209.
- Park J. T., Lee S. B. (2011), Effects Analysis of Traffic Safety Improvement Program Using Data Mining: Focusing on Urban Area, *Journal of Transport Research*, 18(2), The Korea Transport Institute, 77-91.
- Park S., Jang K., Park S. H., Kim D. K., Chon K. S. (2012), Analysis of Injury Severity in Traffic Crashes: A Case Study of Korean Expressways, *KSCE J Civ Eng*, 16(7), Korean Society of Civil Engineers, 1280-1288.
- Park T. H., Park J. J., Yoon P. (2008), A Study on Traffic Accident Characteristics of Freeway Work Zones, *J. Korea Inst. Intell. Transp. Syst.*, 7(1), The Korea Institute of Intelligent Transportation Systems, 127-136.
- Yoon S. M., Oh C., Park H. J., Chung B. J. (2016), Identification of Factors Affecting the Crash Severity and Safety Countermeasures Toward Safer Work Zone Traffic Management, *J. Korean Soc. Transp.*, 34(4), Korean Society of Transportation, 354-372.