

# 도로교통사고 예방 위한 위험도 평가기법 모색

홍상연 정재훈

A Study on Risk Assessment Technique for Prevention of Road Traffic Accident



서울연구원  
The Seoul Institute

**도로교통사고 예방 위한  
위험도 평가기법 모색**

## \ 연구책임

홍상연 서울연구원 교통시스템연구실 부연구위원

## \ 연구진

정재훈 서울연구원 교통시스템연구실 연구원

이 보고서의 내용은 연구진의 견해로서  
서울특별시의 정책과는 다를 수도 있습니다.

## 요약

# 교통사고 위험도 평가 위해 다양한 자료수집 첨단 교통정보 활용·데이터 플랫폼 구축 필요

## 서울시 교통사고 줄이기 위한 예방적 접근방법이 필요한 시점

중앙정부와 서울시 차원에서 교통사고 사망자 줄이기에 노력한 결과, 서울시의 인구 10만 명당 교통사고 사망자수는 2010년 4.1명에서 2017년 3.4명으로 지속적인 감소 추세를 보이고 있다. 이는 국내 특별광역시도 중에서도 가장 낮은 수준으로, 서울시의 교통안전 개선 노력이 성과를 거두고 있는 것으로 볼 수 있으나, 국외 주요 선진도시들에 비해서는 여전히 높은 사망자수 비율을 나타내고 있다(2016년 기준 인구 10만 명당 교통사고 사망자수: 서울 3.5명, 뉴욕 2.9명, 런던 1.5명, 베를린 1.5명 등). 더욱이 최근 들어 서울시의 교통사고 사망자수 감소 추세가 이전에 비해 둔화되고 있어, 지금까지와는 다른 차별화된 접근방법의 검토가 필요한 시점이다.

지금까지 교통안전 개선사업은 교통사고가 잦은 곳을 중심으로 교통안전시설 설치 및 신호운영체계 조정 등의 개선방안을 제시하는 방식으로 추진되었다. 1980년대 후반에 도입된 ‘교통사고 잦은 곳 개선사업’은 국내 교통사고를 획기적으로 줄이는 데에 지대한 공헌을 해왔으나, 이미 발생한 교통사고 통계를 기초로 하는 특성상 결과적, 사후적(retroactive) 대응에 그칠 수밖에 없는 근본적인 한계를 지닌다.

교통사고는 교통체계뿐 아니라 주변 토지이용과 사람들의 행태 간의 복합적인 관계에 영향을 받으므로, 잠재적인 교통사고 위험도를 판단하기 위해서는 교통사고를 사회환경의 맥락 안에서 이해하려는 노력이 필수적이다. 교통사고가 실제 많이 발생하지 않은 지역이더라도 교통여건과 사회환경을 고려할 때 교통사고가 발생할 개연성이 높은 곳에 대해서는 교통사고에 예방적(proactive)으로 대응함으로써 현재의 교통안전 개선체계를 한 단계 발전시키는 계기로 삼아야 할 것이다.

## 교통안전 분야 데이터는 생성단계...활용 가능한 기초자료 미흡

서울시의 기초자료를 조사하여 활용가능한 후보군을 분석한 결과, 최근 다양한 분야에서 언급되고 있는 '빅데이터'로서의 교통안전 분야의 데이터는 이제 막 생성되고 있는 단계에 있는 것으로 판단된다. 즉, 데이터의 양(volume), 데이터의 생성 속도(velocity), 데이터 종류의 다양성(variety), 데이터의 정확성(veracity)을 특징으로 하는 빅데이터로서의 데이터 중 실제 공공 분야에서 저장관리하는 것은 스마트카드 이용내역 등 일부에 국한되며, 통신사 이용내역, 보험사 정보 등은 한시적일회성으로 가공되어 공공에 제공되고 있는 수준이다.

이에 따라 보다 고도화된 빅데이터 활용을 위해서는 공공분야의 자체적인 데이터 수집, 가공뿐 아니라 민간분야의 참여를 유도할 수 있는 협력방안을 구축하는 것이 급선무일 것으로 판단된다. 따라서 보다 활용도 높은 기초자료 확보를 위해서는 지속적으로 1) 정보수집원의 다양화, 2) 정보분석의 고도화, 3) 민관 정보교류 등을 추진해야 할 것이다.

## 교통체계·사회환경·인구활동 연계한 교통사고 위험도 평가법 제시

본 연구에서는 교통공학적 관점의 전통적인 교통사고 분석에서 벗어나 교통체계-사회환경-인구활동을 연계한 교통사고 위험도 평가방법을 제시하고, 평가에 활용할 수 있는 서울시 기초자료를 검토하였다.

교통사고 위험도 평가모형은 종합적인 정책적 시사점 도출을 위하여 교통사고 발생건수 예측과 교통사고 발생원인 분석 모두를 포함하도록 구성하고, 전통적인 통계적 기법 외에 최근 활용도가 높아지고 있는 기계학습 기법을 포함하였다.

본 연구의 목적에 따라 선정한 평가모형은 지리적 이질성을 고려한 회귀분석 기법인 지리가중 회귀분석(Geographically Weighted Regression), 분석대상 간에 존재하는 패턴을 예측 가능한 여러 가지 규칙들의 조합으로 분류하는 의사결정나무(Decision Tree), 자료 간의 상대 거리를 통해 분석대상을 분류하는 K-최근접이웃(K-Nearest Neighbors) 등이다.

## 주차면수·교차로수 등에 따라 행정동별로 적용할 개선대책 달라야

교통사고 발생요인이 교통사고 발생에 미치는 영향은 행정동별로 다르게 나타났다. 특히 노상 주차면수나 교차로수와 같이 교통안전 개선대책에 직접적으로 활용할 수 있는 변수의 경우에도 행정동에 따라 계수값이 양(+) 또는 음(-)의 값을 보이는 것으로 분석되었다. 이는 유사한 교통사고 발생건수에도 불구하고 행정동별로 적용되어야 하는 개선대책은 전혀 반대로 접근할 수 있음을 시사하고 있다.

또한 교통체계-사회환경-인구활동에 관한 각 변수는 교통사고 발생에 독립적으로 영향을 미치지 않고 상호 조건부의 관계를 맺고 있는 점도 밝혀졌다. 예를 들어 노상주차면수는 횡단보도수가 충분한 경우에는 교통사고 위험도에 별다른 영향을 미치지 않지만, 횡단보도수가 적은 곳에서는 교통사고 위험도 판별기준으로 사용되었으며 노상주차면수가 임계수준보다 높을 경우 교통사고 위험도도 증가하는 것으로 분석되었다.

교통사고 심각도별로 영향을 미치는 요인도 다르게 나타났다. 사상사고(Fatal and Injury, FI)의 경우에는 도로시설의 특징과 고령자수에 의해 주로 영향을 받는 것으로 분석된 반면, 물피사고(Property Damage Only, PDO)의 경우에는 노변에서의 인구활동에 의해 주로 영향을 받는 것으로 나타났다.

마지막으로 교통안전 개선대책이 수립되었을 때 장래 개선효과를 전망하기 위한 분석기법을 제시하였으며, 각 행정동별로 독립변수를 임의로 증가시켰을 때의 변화 양상에 대한 결과를 제시하여 모형의 적용성을 검토하였다.

## 교통안전 개선사업별 평가모형 지속 개발 등 4가지 개선방안 제시

본 연구를 통해 살펴본 결과, 향후 보다 체계적이고 과학적인 교통안전 개선대책 수립을 위해서는 다음과 같은 준비가 필요할 것으로 판단된다.

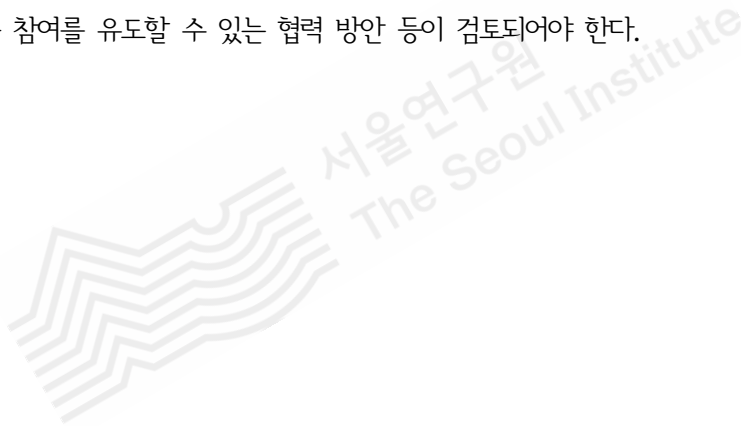
첫째, 다양하고 구체적인 교통사고 자료의 수집이 필요하다. 이를 위해 단기적으로는 각 보행

사에서 구축하고 있는 교통사고 데이터베이스를 수집하여 분석에 활용함으로써 보다 신뢰성 있는 평가체계를 구축할 수 있을 것이다.

둘째, 잠재적인 위험도 평가를 위한 첨단 교통정보의 활용이 필요하다. 현재도 통신사, 카드사 등 민간자료를 결합하고, 상용차량 관리시스템을 활용한다면 일부 자료의 획득이 가능하며, 장기적으로 C-ITS, V2X 등이 보편화되면 다양한 첨단 교통정보의 활용이 확대될 것으로 전망된다.

셋째, 자료 기반의 의사결정체계 구축을 위한 지속적인 노력이 요구된다. 중앙정부와 서울시에 서 운영하는 각종 교통안전 개선사업의 목적과 대상에 적합한 평가모형을 지속적으로 발굴하 고 현장 데이터에 적용하여 검증하는 노력이 필요하다.

넷째, 교통사고 분석체계 운영을 위한 데이터 플랫폼 구축이 시급하다. 공공에서 직접 수집하 는 정보를 다양화하는 방안, 기존에 가공하지 않고 버려지던 정보를 고도화하는 방안, 민간분 야의 플랫폼 참여를 유도할 수 있는 협력 방안 등이 검토되어야 한다.



# 목차

01 연구의 개요	2
1_연구의 배경 및 목적	2
2_연구의 주요내용 및 방법	4
02 서울시 교통사고 현황 분석	6
1_교통사고 특성 분석	6
2_기존 교통안전 개선사업 검토	11
03 교통사고 위험도 평가를 위한 기반 검토	16
1_교통사고 위험도 평가 기법	16
2_교통사고 영향요인 검토	20
3_서울시 기초자료 조사	21
4_시사점 및 연구방향 설정	23
04 교통사고 위험도 평가모형 구축	26
1_평가모형 선정	26
2_분석자료 구축	33
3_사례분석 결과	37
4_평가모형 검토 결과	45



05 교통사고 예방을 위한 정책제언	48
1_결론 및 정책제언	48
2_향후 발전방향	49
참고문헌	53
Abstract	55



## 표

[표 1-1] 서울시 교통사고 발생건수 및 사망자수 추이	2
[표 2-1] 서울시 교통사고 기초통계 (2014년 기준)	7
[표 2-2] 서울시 사고유형별 교통사고 (2014년 기준)	8
[표 2-3] 서울시 연령층별 사상자 (2014년 기준)	9
[표 2-4] 가해운전자 법규위반별 교통사고 (2014년 기준)	10
[표 3-1] 교통안전대체지표의 사례	19
[표 3-2] 분야별 교통사고 영향요인 검토	21
[표 3-3] 서울시 기초자료 검토 결과	22
[표 4-1] 교통사고 위험도 평가모형	26
[표 4-2] 선정된 평가모형	28
[표 4-3] 교통사고분석시스템의 주요 변수	33
[표 4-4] 사상사고 영향요인 변수 선정결과	35
[표 4-5] 물피사고 영향요인 변수 선정결과	36
[표 4-6] 지리가중회귀분석과 일반회귀분석 결과 비교 (사상사고)	38
[표 4-7] 노상주차면수의 계수값 (일부 지역)	38
[표 4-8] 교차로수의 계수값 (일부 지역)	39
[표 4-9] 지리가중회귀분석과 일반회귀분석 결과 비교 (물피사고)	39
[표 4-10] 노상주차면수의 계수값 (일부 지역)	40
[표 4-11] 횡단보도수의 계수값 (일부 지역)	40
[표 4-12] 장래 여건변화에 따른 사상사고 위험도 사례분석	44

[표 4-13] 장래 여건변화에 따른 물피사고 위험도 사례분석	44
[표 5-1] 교통 관련 빅데이터 원자료 주요 내용	51
[표 5-2] 단계별 데이터 플랫폼 구축 방안	52



## 그림

[그림 1-1] 서울시 교통사고 발생건수 및 사망자수 추이	2
[그림 1-2] 주요도시 인구 10만 명당 사망자수	3
[그림 2-1] 서울시 연령층별 사상자 (2014년 기준)	9
[그림 2-2] 교통사고 잦은 곳 개선사업 추진절차	12
[그림 3-1] 빅데이터와 공공데이터의 정의	24
[그림 4-1] 인공지능, 기계학습 및 딥러닝의 관계	27
[그림 4-2] 지리가중회귀분석의 지역별 적합도( $r^2$ ) 분석 사례	29
[그림 4-3] 의사결정나무 분석결과 예시	30
[그림 4-4] K-최근접이웃의 분석 개념도	31
[그림 4-5] 사상사고의 변수별 중요도 분석결과	35
[그림 4-6] 물피사고의 변수별 중요도 분석결과	36
[그림 4-7] 의사결정나무 분석결과 (사상사고)	41
[그림 4-8] 의사결정나무 분석결과 (물피사고)	42

# 01

---

## 연구의 개요

1\_연구의 배경 및 목적

2\_연구의 주요내용 및 방법

# 01 | 연구의 개요

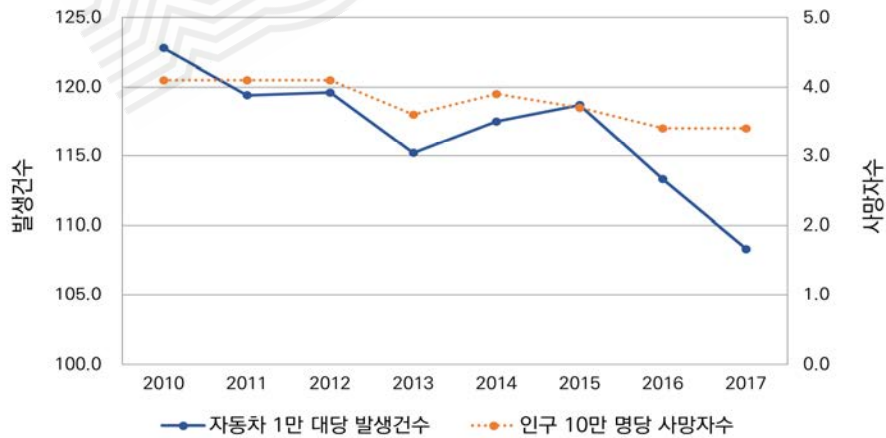
## 1\_연구의 배경 및 목적

중앙정부와 서울시 차원에서 교통사고 사망자 줄이기에 노력한 결과, 서울시의 인구 10만 명당 사망자수는 2010년 4.1명에서 2017년 3.4명으로 연평균 2.6%의 감소세를 보여 지속적으로 감소한 것으로 나타났다. 자동차 1만 대당 사고건수 역시 2010년 122.8건에서 2017년 108.3건으로 연평균 1.8%의 감소세를 보이고 있다.

[표 1-1] 서울시 교통사고 발생건수 및 사망자수 추이

(단위: 건, 명)

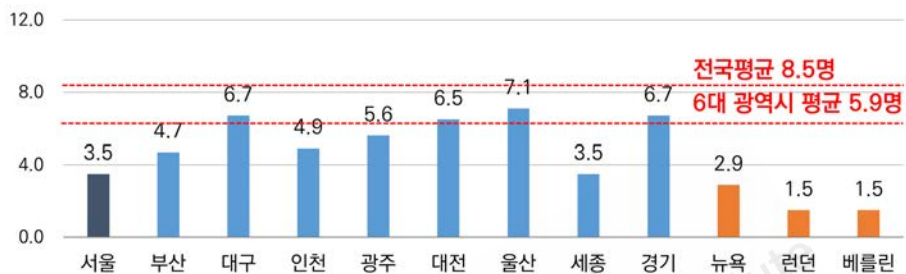
기간	2010	2011	2012	2013	2014	2015	2016	2017	연평균 감소율
사고건수	41,662	40,451	40,829	39,439	40,792	41,665	40,039	38,625	-1.1%
1만 대당 사고건수	122.8	119.4	119.6	115.2	117.5	118.7	113.3	108.3	-1.8%
사망자수	429	435	424	378	400	376	348	343	-3.1%
10만 명당 사망자수	4.1	4.1	4.1	3.6	3.9	3.7	3.4	3.4	-2.6%



[그림 1-1] 서울시 교통사고 발생건수 및 사망자수 추이

자료: 교통사고분석시스템 (도로교통공단, 2018)

인구 10만 명당 교통사고 사망자수는 국내 특별광역시도 중 가장 낮은 수준으로 서울시의 교통안전개선 노력이 성과를 거두고 있는 것으로 볼 수 있으나, 국외 주요 선진도시들에 비해서는 여전히 높은 사망자수 비율을 나타내고 있다(2016년 기준 인구 10만 명당 교통사고 사망자수: 서울 3.5명, 뉴욕 2.9명, 런던 1.5명, 베를린 1.5명 등). 더욱이 최근 들어 서울시의 교통사고 사망자수 감소 추세가 이전에 비해 둔화되고 있어 지금까지와는 다른 차별화된 접근 방법의 검토가 필요한 시점이다.



[그림 1-2] 주요도시 인구 10만 명당 사망자수

자료: 2018년 서울시 교통안전사업 추진계획 (서울시, 2018)

지금까지 교통안전 개선사업은 교통사고가 잦은 곳을 중심으로 교통안전시설 설치 및 신호운영체계 조정 등의 개선방안을 제시하는 방향으로 추진되고 있다. 1980년대 후반에 도입된 ‘교통사고 잦은 곳 개선사업’은 국내 교통사고를 획기적으로 줄이는 데에 지대한 공헌을 해왔으나, 이미 발생한 교통사고 통계를 기초로 하는 특성상 결과적, 사후적 (retroactive) 대응에 그칠 수밖에 없는 근본적인 한계를 지닌다.

교통사고는 교통체계뿐 아니라 주변 토지이용과 사람들의 행태 간의 복합적인 관계에 영향을 받으므로 잠재적인 교통사고 위험도를 판단하기 위해서는 교통사고를 사회환경의 맥락 안에서 이해하려는 노력이 필수적이다. 실제 교통사고가 많이 발생하지 않았더라도 교통여건과 사회환경을 고려할 때 교통사고가 발생할 개연성이 높은 지역에 대해서는 예방적(proactive)으로 대응함으로써 현재의 교통안전 개선체계를 한 단계 발전시키는 계기로 삼아야 할 것이다.

본 연구에서는 교통공학적 관점의 전통적인 교통사고 분석에서 벗어나 교통체계-사회환

경-인구활동을 연계한 교통사고 위험도 평가방법을 제시하고자 한다. 이를 위해 현재 서울시에서 수집 가능한 자료를 바탕으로 교통사고 위험도 평가 및 개선방향을 도출하는 일련의 과정을 제시하고, 향후 교통사고와 관련하여 추가로 수집해야 할 기초자료와 이를 활용한 평가방법 개선방향을 검토하였다.

## 2\_연구의 주요내용 및 방법

### 1) 주요내용

2장에서는 서울시 교통사고 특성을 분석하고 유형별 교통사고 발생추이와 기존 교통안전 개선대책 수립 절차를 살펴본다. 3장에서는 서울시 교통사고 위험도 평가기반을 분석하기 위하여 교통사고 위험도 평가지표 검토, 교통사고 영향요인 검토를 시행하고 교통사고와 관련된 서울시 기초자료 및 기구축 DB를 조사하여 연구 추진방향을 설정한다. 4장에서는 교통사고 위험도 평가모형 수립을 위하여 기존 문헌에서 제시한 유형을 검토하고, 교통사고 위험도 평가모형 후보군을 선정한다. 이후 기초자료 DB를 구축하여 모형별 사례분석 및 특징분석을 수행하고, 시사점을 비교 검토하여 제시한다. 마지막으로, 결론 및 정책제언에서는 서울시 교통사고 위험도 분석 고도화를 위한 발전전략을 제시하였다.

### 2) 연구방법

도로위험도에 영향을 미치는 요인을 분석하고 이에 대한 기존 평가기법을 검토하기 위해 관련 분야의 연구보고서와 논문을 검토하였고, 이를 토대로 빅데이터를 활용한 변수 간 복합적 효과를 분석하기 위한 모형을 선정하였다. 또한 서울시의 교통사고 발생현황 및 원인 분석을 위하여 구득가능한 최신의 교통체계, 사회환경, 인구활동 등의 관련 자료를 수집하고 분석하였다. 마지막으로 기초자료의 적합성, 모형의 합리성, 분석결과의 적절성 등을 검토하기 위하여 단계별로 외부 전문가 자문을 통해 의견을 수렴하여 보완하였다.



# 02

---

## 서울시 교통사고 현황 분석

- 1\_교통사고 특성 분석
- 2\_기존 교통안전 개선사업 검토

## 02 | 서울시 교통사고 현황 분석

### 1\_교통사고 특성 분석

앞서 살펴본 바와 같이 서울시 교통사고 발생건수 및 사망자수는 2010년부터 2017년까지 지속적인 감소세를 보이고 있으나 최근 들어 감소율이 둔화되고 있다. 세부적인 교통사고 특성을 살펴보기 위하여 교통사고분석시스템(TAAS, 도로교통공단<sup>1)</sup>)에서 제공하는 서울시 상세 교통사고 자료를 분석하였다.

2014년 기준으로 서울시 총 교통사고 발생건수는 40,792건으로 나타났다. 시간대별로는 주간과 야간의 비율이 거의 동일하게 나타났고, 요일별로는 금요일(15.7%)과 토요일(15.0%)의 비율이 다소 높게 나타났고, 월요일~목요일까지는 13.6~14.9% 사이의 값을 보여 대체로 유사하게 나타났다.

지역별로는 활동인구가 많을 것으로 추측되는 강남구(8.9%), 송파구(6.4%), 서초구(5.4%), 영등포구(5.4%)의 순으로 사고건수가 높게 나타나 토지이용과 교통사고의 관계 검토가 필요함을 시사하고 있다.

사고유형으로는 차대차 사고가 전체의 69.1%를 차지하였고, 승용차 교통량 비율(71.3%, 2017년 전국 평균 기준, 교통량정보제공시스템)이 압도적으로 높다는 사실에서 유추할 수 있듯이 1당사자의 66.6%, 2당사자의 41.8%가 승용차 관련 사고로 나타났다. 이에 비해 경찰에 신고되지 않은 단순 접촉사고가 포함되지 않았음에도 불구하고 차대사람 사고의 비율이 27.3%로 비교적 높게 나타나 교통안전 개선에 있어 보행사고의 중요성을 간접적으로 보여주고 있다.

---

<sup>1)</sup> 교통사고분석시스템은 교통안전법에 의거하여 경찰, 보험사, 공제조합 등의 교통사고 자료를 수집하고 제공하는 시스템으로 도로교통공단이 경찰청으로부터 위탁받아 운영하고 있다. 그러나 개별사고의 위치정보를 포함한 사고유형 등 세부정보는 경찰에 신고된 자료에 한해 제공되고 있다.

[표 2-1] 서울시 교통사고 기초통계 (2014년 기준)

구분	변수명	건수(건)	비율	구분	변수명	건수(건)	비율	
시간대	주간	20,683	50.7%	1당사자 차종	건설기계	188	0.5%	
	야간	20,109	49.3%		농기계	4	0.0%	
	소계	40,792	100.0%		화물차	3,039	7.4%	
요일	월	5,547	13.6%		승합차	2,976	7.3%	
	화	5,865	14.4%		원동기자전거	1,620	4.0%	
	수	6,086	14.9%		이륜차	3,183	7.8%	
	목	6,032	14.8%		승용차	27,168	66.6%	
	금	6,415	15.7%		자전거	2,065	5.1%	
	토	6,113	15.0%		특수차	54	0.1%	
	일	4,734	11.6%		기타/불명	495	1.2%	
	소계	40,792	100.0%		소계	40,792	100.0%	
계절	봄	10,279	25.2%	1당사자 성별	남자	33,908	83.1%	
	여름	10,872	26.7%		여자	5,569	13.7%	
	가을	10,836	26.6%		기타/불명	1,315	3.2%	
	겨울	8,805	21.6%		소계	40,792	100.0%	
	소계	40,792	100.0%		2당사자 차종	건설기계	61	0.1%
행정구역	강남구	3,624	8.9%	보행자		11,118	27.3%	
	강동구	1,442	3.5%	불명		39	0.1%	
	강북구	1,313	3.2%	승용차		17,055	41.8%	
	강서구	1,806	4.4%	승합차		1,752	4.3%	
	관악구	1,716	4.2%	원동기자전거		1,894	4.6%	
	광진구	1,480	3.6%	이륜차		3,730	9.1%	
	구로구	1,566	3.8%	자전거		2,449	6.0%	
	금천구	1,029	2.5%	특수차		36	0.1%	
	노원구	1,516	3.7%	화물차		1,195	2.9%	
	도봉구	919	2.3%	기타/불명		1,463	3.6%	
	동대문구	1,848	4.5%	소계		40,792	100.0%	
	동작구	1,399	3.4%	2당사자 성별		남자	29,220	71.6%
	마포구	1,759	4.3%			여자	9,998	24.5%
	서대문구	1,122	2.8%			기타/불명	111	0.3%
	서초구	2,217	5.4%			기타(NA)	1,463	3.6%
	성동구	1,189	2.9%	사고유형		차대사람	11,118	27.3%
	성북구	1,664	4.1%		차대차	28,201	69.1%	
	송파구	2,608	6.4%		차량단독	1,473	3.6%	
	송파구	2,608	6.4%		소계	40,792	100.0%	
	양천구	1,466	3.6%	기상상태	눈	302	0.7%	
영등포구	2,193	5.4%	맑음		35,339	86.6%		
용산구	1,350	3.3%	비		2,463	6.0%		
은평구	1,325	3.2%	안개		7	0.0%		
종로구	1,314	3.2%	흐림		2,011	4.9%		
중구	1,305	3.2%	기타/불명		670	1.6%		
중랑구	1,622	4.0%	소계		40,792	100.0%		
소계	40,792	100.0%	도로종류		고속국도	114	0.3%	
고속국도	114	0.3%		특별광역시도	37,833	92.7%		
특별광역시도	37,833	92.7%		기타	2,845	7.0%		
기타	2,845	7.0%		소계	40,792	100.0%		
소계	40,792	100.0%						

자료: 교통사고분석시스템(도로교통공단)

사고유형별 교통사고 발생건수 중 차대차 사고가 69.1%, 차대사람 사고가 27.3%를 차지하는 반면, 사망자수에서는 오히려 차대사람 사고의 경우에 53.5%로 높게 나타나 차대사람 사고의 치사율(사망자수/발생건수)이 4배가량 높은 것으로 분석되었다. 발생건수 대비 중상자수를 살펴보더라도 차대차 사고건수 대비 중상자수가 33.1%인 것에 비해 차대사람 사고의 경우에는 44.9%로 나타나 차대사람 사고에 대한 교통안전 개선의 우선순위가 높은 것으로 판단된다.

차량단독 사고의 경우에는 일반적으로 운전자 부주의, 차량 및 도로시설 결함에 의해 발생하는 경우가 많아 별도의 교통안전 개선사업보다는 캠페인, 운전자 교육과 시민참여형 도로시설물 모니터링 등을 통해 개선해가는 것이 적합할 것이다.

**[표 2-2] 서울시 사고유형별 교통사고 (2014년 기준)**

사고유형	차대사람	차대차	차량단독	합계
발생건수	11,118	28,201	1,473	40,792
구성비(%)	27.3	69.1	3.6	100.0
사망자수	214	150	36	400
구성비(%)	53.5	37.5	9.0	100.0
부상자수	11,621	44,016	1,708	57,345
구성비(%)	20.3	76.8	3.0	100.0
중상자수	4,995	9,342	557	14,894
구성비(%)	33.5	62.7	3.7	100.0
경상자수	5,653	30,076	864	36,593
구성비(%)	15.5	82.2	2.4	100.0
부상신고자수	973	4,598	287	5,858
구성비(%)	16.6	78.5	4.9	100.0

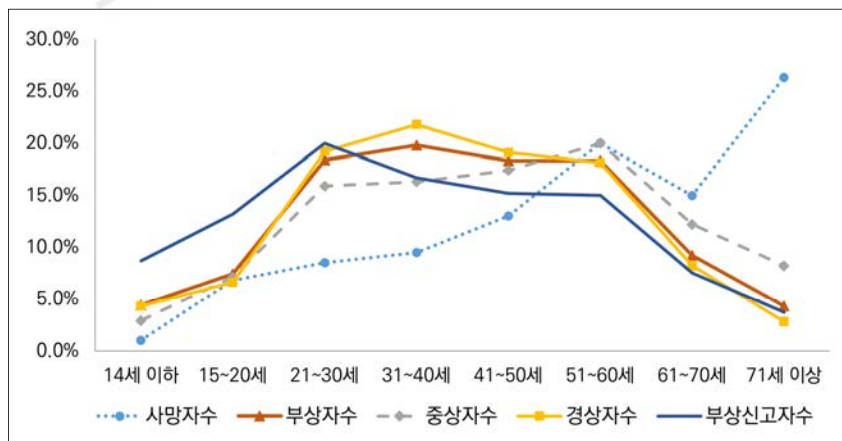
자료: 교통사고분석시스템(도로교통공단)

연령층별 사상자수를 살펴보면 사망자와 중상자수는 연령대가 증가할수록 전반적으로 증가하는 것으로 나타난다. 두 값이 60대에 다소 감소하는 것은 60대 이후에 통행량이 감소함에 따라 사고건수가 감소하기 때문인 것으로 보이고, 70대 이후에는 대상인구가 증가하여 다시 증가세를 나타내는 것으로 판단된다.

반면에 부상자와 경상자수는 20~30대에 가장 높게 나타나 유사한 수준의 교통사고가 발생하더라도 고령자의 상해 정도가 더 심각한 것으로 추정할 수 있고, 빠른 속도로 고령화가 진행되고 있는 우리나라 여건상 향후 고령자 교통사고에 대한 세심한 접근이 필요할 것으로 판단된다.

[표 2-3] 서울시 연령층별 사상자 (2014년 기준)

연령대	~14	15~20	21~30	31~40	41~50	51~60	61~70	61~70		71~
								61~64	65~70	
사망자수	4	27	34	38	52	80	60	29	31	105
구성비(%)	1.0	6.8	8.5	9.5	13.0	20.0	15.0	7.3	7.8	26.3
부상자수	2,515	4,230	10,551	11,372	10,469	10,497	5,248	2,599	2,649	2,463
구성비(%)	4.4	7.4	18.4	19.8	18.3	18.3	9.2	4.5	4.6	4.3
중상자수	436	1,051	2,362	2,429	2,598	2,981	1,821	826	995	1,216
구성비(%)	2.9	7.1	15.9	16.3	17.4	20.0	12.2	5.5	6.7	8.2
경상자수	1,567	2,407	7,018	7,962	6,983	6,637	2,989	1,568	1,421	1,030
구성비(%)	4.3	6.6	19.2	21.8	19.1	18.1	8.2	4.3	3.9	2.8
부상신고자수	512	772	1,171	981	888	879	438	205	233	217
구성비(%)	8.7	13.2	20.0	16.7	15.2	15.0	7.5	3.5	4.0	3.7



[그림 2-1] 서울시 연령층별 사상자 (2014년 기준)

가해운전자의 법규위반사항을 살펴보면 모든 항목에 대하여 다소 포괄적인 ‘안전운전 의무 불이행’이 절반 이상을 차지하고 있어 교통사고 조사자료를 통해 원인을 분석하기에는 미흡한 것으로 나타났다. 이는 교통사고 조사원이 현장에서 사고원인을 파악하고 기록하는 데 현실적인 어려움이 있는 한계도 있지만, 교통사고가 도로환경, 교통시설 여건, 운전자 과실 등의 복합적인 원인에 의해 발생한다는 근본적인 문제도 내재되어 있다.

**[표 2-4] 가해운전자 법규위반별 교통사고 (2014년 기준)**

(단위: 건, 명)

구분	발생건수	사망자수	부상자수	중상자수	경상자수	부상 신고자수
과로	2	0	2	1	1	0
과속	66	23	108	50	48	10
앞지르기 방법위반	5	0	6	0	5	1
앞지르기 금지위반	6	0	7	2	4	1
중앙선 침범	1,536	25	2,338	695	1,359	284
신호위반	5,008	44	7,887	2,425	4,717	745
안전거리 미확보	5,498	9	8,951	1,560	6,467	924
서행 및 일시정지위반	25	0	41	10	18	13
부당한 회전	262	6	338	94	194	50
통행우선 순위위반	1	0	2	0	2	0
진로양보 의무 불이행	6	0	6	1	2	3
안전운전 의무 불이행	22,644	264	30,503	7,925	19,440	3,138
교차로 통행방법 위반	1,495	2	2,055	460	1,335	260
보행자 보호의무 위반	1,429	17	1,527	709	732	86
차로위반(진로변경 위반)	669	1	1,063	209	757	97
직진/우회전차 통행방해	504	2	688	158	448	82
철길건널목 통과방법위반	-	-	-	-	-	-
긴급자동차 피양의무 위반	-	-	-	-	-	-
기타	1,634	7	1,815	591	1,060	164
합계	40,790	400	57,337	14,890	36,589	5,858

## 2\_기존 교통안전 개선사업 검토

정부는 국가 교통안전을 체계적으로 관리하기 위해 교통안전법에 의거하여 교통안전에 관한 중장기 종합계획을 5년 단위로 수립하여 시행하고 있다. 현재 '제8차 국가교통안전 기본계획'(2017~2021)이 수립되어 시행 중이다. 이 계획에 의거하여 매년 지정행정기관의 교통안전 계획을 제출받아 국가교통안전시행계획을 수립하고, 매년 시행결과를 점검하여 교통안전연차보고서를 작성하고 있다.

국가교통안전기본계획이 수립된 이후 이에 의거하여 각 지자체장은 지역교통안전기본계획을 5년 단위로 수립하여 시행하도록 되어 있다. 이에 따라 서울시는 현재 '제3차 서울특별시 교통안전기본계획'(2017~2021)을 수립하여 시행 중이다. 이 계획은 서울시의 교통안전정책 수립 및 추진을 위한 기본방향과 지침을 제시하는 기본계획으로서, 서울시 교통안전정책심의위원회의 심의를 거쳐 교통안전시행계획을 수립하게 된다.

이와 같이 교통안전 개선사업은 주로 중앙정부가 종합계획을 수립하면 지자체에서 상세 실행계획을 수립하여 시행하는 방향으로 이루어지고 있다. 현재 시행되고 있는 주요 교통안전 개선사업은 다음과 같다.

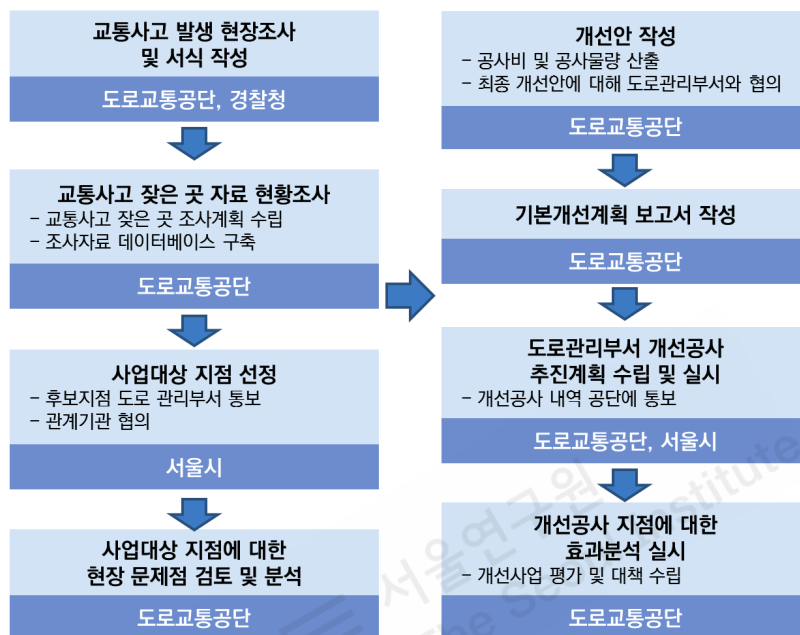
### (1) 교통사고 잦은 곳 개선사업

교통사고 잦은 곳 개선사업은 교통사고가 발생하는 지점 또는 교통안전 취약지점을 대상으로 교통안전시설, 신호운영체계 등을 개선하는 교통안전대책을 시행하는 사업으로, 서울시는 교통사고 잦은 곳 관리지점 약 1,400개소를 관리하고 있다.

교통사고 잦은 곳은 교통사고 발생건수가 동일 지점에서 1년간 5건 이상 발생한 곳을 말하며, 전년도에 발생한 교통사고를 전수 조사하여 선정하게 된다. 공간적으로는 교차로 및 횡단보도는 차량정지선 후방으로 30m 이내, 시가지 단일로는 반경 100m 이내, 기타 단일로와 고속도로는 반경 200m 이내를 기준으로 하고 있다.

교통사고 잦은 곳 개선사업의 추진절차를 살펴보면, 1) 도로교통공단과 경찰청이 교통사고 현장조사자료를 기반으로 기초 통계자료를 구축하고, 2) 서울시는 통계를 기반으로

사업대상 지점을 선정한다. 이후 3) 도로교통공단 주관으로 원인을 조사하고 개선안을 작성하면, 4) 서울시에서는 추진계획을 수립하여 개선사업을 시행하고, 5) 최종적으로 도로교통공단에서 사후평가를 시행하게 된다. 상세 추진절차는 다음 그림과 같다.



[그림 2-2] 교통사고 잦은 곳 개선사업 추진절차

자료: 제3차 서울시 교통안전기본계획 (서울시, 2017)

## (2) 어린이보호구역 정비사업

2003년 자동차교통관리개선특별회계법이 개정됨에 따라 어린이보호구역 정비사업을 추진을 위한 지원 근거가 마련된 이후, 2004년에 사업권한이 경찰청에서 행정안전부로 이관되었다. 이후 2011년에 어린이·노인·장애인보호구역에 대한 지정 권한이 다시 경찰청에서 지자체로 이양되어 운영되고 있다.

행정안전부는 전국의 유치원, 초등학교, 보육시설 및 특수학교 주변의 어린이보호구역을 지정하고 있으며, 지정된 곳에 과속방지시설, 보차분리시설, 횡단보도, 방호울타리 등 도로교통 안전시설을 지속적으로 설치하고 있다. 최근에는 어린이보호구역에서의 과속단속을 위한 다기능 단속카메라 설치사업 추진계획을 수립하였다.



### (3) 노인보호구역 지정 등 노인교통안전대책

2006년 도로교통법이 개정되면서 노인보호구역에 대한 제도가 마련되었으나, 투자예산 확보가 미흡하여 사업추진이 다소 부진한 실정이다. 노인복지시설 등의 주변 지역을 노인보호구역으로 설정할 수 있으며, 설정될 경우 대상 시설의 주출입문을 중심으로 반경 300m 이내의 도로 중 일정 구간을 보호구역으로 지정한다. 또한 주변 교통여건을 고려하여 필요시에는 출입문 중심 반경 500m 이내를 보호구역으로 지정할 수도 있다.

최근 행정안전부에서는 고령화로 인한 노인인구 증가에 대비하여 노인 교통사고 현황분석, 교통안전시설 진단 및 사고위험요인 분석 등을 시행하고, 2019년부터 별도 예산을 투입하여 사고다발지역에 대한 체계적인 정비계획을 수립할 예정이다.

### (4) 안전한 보행환경 조성사업

2009년에 시작된 안전한 보행환경 조성사업은 보도단절구간 연결, 차로수 조정 등 보행환경을 조성하기 위한 사업으로, 2013년에 ‘보행안전 및 편의증진에 관한 법률’이 발효됨에 따라 각 지자체별로 ‘보행환경개선 기본계획’을 수립하여 체계적으로 사업이 추진되고 있다. 특별히 보행환경개선이 요구되는 지역을 보행환경개선지구로 지정하여 종합적인 보행환경 개선계획을 수립하고, 필요한 사업비를 집중적으로 지원함으로써 투자효율성을 극대화할 수 있다.

위와 같이 중앙정부와 서울시에서는 교통안전 개선을 위해 다양한 제도를 운영하고 있다. 그러나 각 사업별로 주관부처, 목표 대상, 개선 시설물이 달라 근본적으로 상호 유기적인 추진절차를 갖기 어렵다. 또한 사업별 개선대책 수립에 있어서도 지점별 여건에 따른 차이가 커 대부분 현장조사 및 전문가의 판단에 의존하는 경향이 커 체계적인 관리에는 한계가 있는 실정이다.

# 03

---

## 교통사고 위험도 평가를 위한 기반 검토

- 1\_교통사고 위험도 평가 기법
- 2\_교통사고 영향요인 검토
- 3\_서울시 기초자료 조사
- 4\_시사점 및 연구방향 설정

## 03 | 교통사고 위험도 평가를 위한 기반 검토

### 1\_교통사고 위험도 평가 기법

#### 1) 교통사고 통계자료 검토

교통사고 통계의 표현법으로는 사고건수법(accident frequency), 사고율법(crash rate), 사고심각도법(Equivalent Property Damage Only, EPDO)이 있다. 사고건수법은 단순히 건수만을 가지고 나타내는 교통사고 통계의 표현법이며, 사고율법은 차량의 주행거리와 통과차량대수를 고려한 교통사고 통계의 상대적 표현법이다. 이러한 방법들은 여러 경우의 사고 통계를 상호 비교하기 위하여 사용된다. 사고심각도법은 모든 피해를 물자의 피해로 환산한 표현법이다.

이와 같은 표현방법과 관련하여 교통량이 많은 구간에서는 교통사고도 더 많이 발생할 개연성이 있으므로 단순 사고건수보다 사고율이 더 적합한 지표라는 관점과 교통량이 많은 구간은 교통안전 개선 우선순위가 더 높기 때문에 노출(exposure)을 고려하지 않은 사고건수를 활용하는 것이 바람직하다는 서로 다른 관점이 제시되고 있다. 본 연구는 공공성 측면에서 한정된 예산으로 교통사고를 효과적으로 개선하기 위한 방법론을 제시하는 것이 목적이므로, 공리적인 관점에서 개선 우선순위가 높은 곳을 대상으로 하는 것이 적합할 것이다. 따라서 본 연구에서는 교통사고 통계자료를 활용할 때 사고건수를 기준으로 분석하는 것이 타당할 것이다.

최근 미국 교통부(Department of Transportation)에서는 교통사고 예측 방법과 교통안전 관리절차를 제시한 도로안전편람(Highway Safety Manual)(US DOT, 2014)을 발간하였다. 이 편람에서는 교통사고의 심각도에 따라 발생원인과 개선대책이 크게 다르게 나타나는 점을 지적하며, 사상사고(Fatal and Injury, FI)와 물피사고(Property Damage Only, PDO)를 구분하여 분석할 것을 제안하고 있다. 따라서 본 연구에서도 교통사고 통계를 사고 심각도에 따라 사상사고와 물피사고로 구분하여 집계하는 것이 바람직할 것으로 판단된다.

## 2) 교통사고 위험도 평가모형 비교

### (1) 분석목적별 분류

교통사고 위험도 평가모형은 분석목적에 따라 크게 교통사고 발생건수를 예측하는 모형과 교통사고 발생원인을 분석하는 모형으로 구분할 수 있다. 교통사고 발생건수를 예측하는 모형은 교통안전 개선사업을 시행했을 때 예상되는 개선 정도나 사회경제적 여건이 바뀌었을 때 교통안전 측면의 변화를 전망하기 위해 사용될 수 있다.

또한 교통사고 발생원인을 분석하는 모형은 교통사고에 영향을 주는 사회경제적 여건, 교통환경 및 인적요소와 관련된 변수와 그 영향의 정도를 밝혀 교통안전 개선을 위한 근거로 활용할 수 있다. 본 연구에서는 각각의 유형이 갖는 정책적 활용성을 고려하여 두 가지 형태 모두를 검토 대상에 포함하는 것이 적합할 것으로 판단된다.

### (2) 분석방법별 분류

분석방법은 전통적인 통계적 기법과 최근 활용도가 높아지고 있는 기계학습 기법으로 구분할 수 있다. 통계적인 기법으로는 회귀분석, 상관분석, 요인분석, 군집분석 등이 널리 활용되고 있으나 모형별로 수리적 가정이 있는 경우가 많은데, 대부분의 교통사고 관련 자료는 비선형인 경우가 많아 활용에 제약이 많다. 이러한 한계로 인하여 최근에는 수리적 모형에 의존하지 않는 기계학습 기법을 적용하는 사례가 증가하고 있다.

기계학습은 여러 데이터를 이용하여 학습한 내용을 기반으로 새로운 데이터에 대한 적절한 작업을 수행할 수 있도록 하는 알고리즘과 기술을 개발하는 분야이다. 기계 학습은 학습 방식에 따라 지도 학습(supervised learning, 감독 학습), 준지도 학습(semi-supervised learning), 비지도 학습(unsupervised learning, 자율 학습), 강화 학습(reinforcement learning)으로 분류된다.

지도 학습은 미리 구축된 학습용 데이터(training data)를 활용하여 모델을 학습하며, 준지도 학습은 학습용 데이터와 정리되지 않은 데이터를 모두 훈련에 사용하는 방법이다. 비지도 학습은 별도의 학습용 데이터를 구축하는 것이 아니라 데이터 자체를 분석하거나 군집(clustering)하면서 학습한다. 강화 학습은 학습 수행 결과에 대해 적절한 보상을

주면서 피드백을 통해 학습한다(TTA정보통신용어사전). 본 연구에서는 향후 다양한 데이터가 수집될 경우의 확장성과 잠재적 활용성을 고려하여 기계학습 기법을 중심으로 검토하되, 계량적인 모형과의 비교 검토를 위하여 전통적인 통계적 기법도 병행하여 검토하도록 한다.

### (3) 분석대상별 분류

교통사고 위험도 평가모형의 분석대상으로는 일반적으로 사고건수, 사고 심각도, 사고 지속시간, 위험도 등이 있다.

가장 대표적인 평가모형은 사고건수를 기반으로 한 모형으로 실제 데이터를 근거로 분석하며 직관적으로 이해하기 쉽다는 장점이 있다. 또한 구체적인 지점, 시점 및 주변 상황에 대한 데이터를 활용하여 신뢰도가 높은 편이다. 그러나 총 통행량에 비해 발생 가능성이 매우 낮은 교통사고의 특성상 관측 데이터수가 극히 제한되어 분석의 신뢰성이 저하될 우려가 있다.

사고 심각도 기반의 모형은 사고유형, 사고형태, 도로 기하구조 등 사고발생 원인에 따라 교통사고의 심각도(사망-중상-경상-대물피해 등)를 판별하기 위해 사용되며, 사고 지속시간 기반의 모형은 사고발생 원인에 따라 교통사고의 지속시간(사고발생-현장접근-사고처리-교통재개 등)을 예측하기 위해 사용된다.

마지막으로 교통사고 위험도를 평가하는 모형은 실제 교통사고 외에 사고가 일어날 뻔했거나(아차사고), 사고 유무와 관계없이 잠재적으로 위험하다고 판단되는 지점을 대상으로 하는 특징이 있다. 이러한 접근은 특히 과거 데이터를 통해 파악하기 어려운 전조증상(precursor)을 고려할 수 있는 장점이 있다. 반면에 평가지표에 대한 신뢰도가 확보되지 않을 경우 위험도를 정의하는 것에 대한 합의가 어려울 수 있다.

교통사고 위험도를 평가하기 위해 실제 발생한 교통사고 자료 대신 잠재적인 위험도를 나타낼 수 있는 교통안전대체지표(Surrogate Safety Measure)를 활용할 수 있다. 대표적인 교통안전대체지표 중 비집계자료로는 개별차량의 주행궤적(DGPS, CCTV 등)을 들 수 있고, 집계자료로는 차량검지기(loop detector, DSRC RSE 등)에서 검지되는 개

별 차량의 주행정보를 들 수 있다. 그러나 현재 교통안전대체지표를 산정하기 위해 서울시 도로 전역에 걸쳐 사용할 수 있는 데이터는 매우 제한적이다. 향후 C-ITS, V2X 등이 보편화되면 첨단 교통정보의 활용이 확대될 것으로 전망된다.

따라서 본 연구에서는 신뢰할 수 있는 통계자료인 사고건수를 기반으로 하여 분석하되, 앞서 도로안전편람(US DOT, 2014)에서 제시한 바와 같이 사고유형을 사상사고(FI)와 물피사고(PDO)로 구분하여 심각도를 일부 반영하는 것이 타당할 것으로 판단된다. 다만, 모형구조를 설정하는 단계에서는 앞으로 수집이 가능할 것으로 전망되는 교통안전대체지표들을 반영할 수 있도록 확장성을 고려하여 향후 발전방향을 제시하도록 한다.

**[표 3-1] 교통안전대체지표의 사례**

구분	지표	세부내용	필요자료
비집계자료	Stopping Distance	주행 중인 선-후행 차량의 정지거리를 비교하여 산출	실시간 속도, 위치
	Time to Collision	주행 중인 선-후행 차량의 정지시간을 비교하여 산출	
	기타	상충, 심각한 돌발상황, 도로이탈 등	주행영상
집계자료	Speed Variation	분석주기 동안 주행한 개별차량 속도의 표준편차	개별차량 검지자료
	Acceleration Noise	분석주기 동안 주행한 개별차량 가속도의 표준편차	

## 2\_교통사고 영향요인 검토

교통은 사람들이 주어진 사회환경 및 교통체계와 유기적으로 영향을 주고 받는 일련의 과정에서 발생하는 이동과 관련된 활동을 의미한다. 교통사고 또한 이러한 맥락에서 요소들 간의 복합적인 상호관계에 의해 발생하는 것으로 이해해야 할 것이다. 따라서 교통사고의 영향요인은 사회환경, 교통여건, 인구활동 등을 포괄적으로 검토하여야 한다.

먼저 사회환경을 구성하는 가장 큰 요소는 인구구조와 토지이용을 들 수 있다. 지역별 거주자수, 종사자수는 지역 내의 활동인구를 결정하는 중요한 요인으로 고려되어야 한다. 특히 고령자수와 어린이수는 통행패턴상의 특수성뿐만 아니라 교통약자로서 교통사고에 취약한 점을 고려하여 별도로 다루어야 할 것으로 판단된다.

토지이용은 지역 내에서 발생하는 인구활동의 유형 및 총량을 결정하는 요소이며, 전통적으로 활동인구의 생산 및 소비활동을 매개로 하여 교통체계와 상호 영향을 주고받는 관계로 알려져있다. 따라서 분석단위 내의 용도별 토지이용 특성은 사회환경의 주요 요인으로 반영되어야 한다.

교통사고에 영향을 주는 교통여건으로는 우선 공급 측면에서 등급별 도로연장 또는 차로 폭을 고려한 도로면적을 고려할 수 있다. 사고율을 고려하는 경우 서울시와 같이 모든 도로구간에서의 교통량을 알기 어려운 때에는 도로연장 또는 도로면적이 간접적인 노출(exposure) 변수로 활용될 수 있을 것이다.

또한 차대차, 차대사람 간 상충이 발생하는 지점인 횡단보도와 교차로의 개수도 교통사고에 영향을 주는 요인으로 고려할 수 있다. 특히 횡단보도 간격이 길면 보행자의 무단횡단 경향이 증가하여 교통사고를 유발한다는 관련 연구를 고려하면 차대사람 사고의 경우에 그 중요성이 더욱 증가할 것으로 예상된다.

마지막으로 도로안전시설물의 경우 현재 시설물 설치 중심으로 시행되는 교통안전 개선사업의 효과를 위해서는 중요한 변수가 될 것이다. 그런데 교통사고가 많이 발생하여 도로안전시설물이 설치된 것인지, 도로안전시설물의 설치로 인해 교통사고가 감소하였는지 인과관계를 파악하기 위해서는 시계열 자료를 결합한 패널분석이 추가로 필요하다. 따라서

본 연구에서는 도로안전시설물 설치에 따른 영향은 분석에서 제외하고 도로연장, 횡단보도수, 교차로수와 같은 일반적인 도로여건에 대해서만 고려하는 것으로 분석의 범위를 한정하였다.

[표 3-2] 분야별 교통사고 영향요인 검토

분류		상세정보	수집주기	수집단위
교통여건		교통량	시간	점
		속도	수시	선
사회환경	인구구조	거주자수(밀도, 비율)	연, 월	면
		종사자수(밀도, 비율)	연, 월	
		고령자수(밀도, 비율)	연, 월	
		어린이수(밀도, 비율)	연, 월	
		차종별 등록대수	연, 월	
	토지이용	구획별 토지이용	연	면
		구획별 대지면적/연면적		
도로시설		등급별 도로연장/면적	연	선
		도로안전시설물	분기	점
		횡단보도	연	점
		중앙버스전용차로	수시	점

### 3\_서울시 기초자료 조사

앞서 살펴본 교통사고 영향요인과 서울시에서 구득할 수 있는 기초자료를 검토한 결과, 잠정적으로 분석에 활용할 수 있는 기초자료 및 자료 제공기관은 다음과 같이 나타났다. 대부분의 인구구조, 토지이용에 관한 사회환경 관련 자료들은 공간적 수집단위가 행정동인 것으로 나타났다. 도로시설에 관한 자료는 GIS로 구축되어 있는 경우가 많아 필요에 따라 공간적으로 집계하거나 재분석하는 것이 비교적 용이한 편이다. 마지막으로 교통여건에 관한 자료는 대중교통을 제외하면 서울시 전역의 교통량, 속도 등의 자료를 확보하는 것이 어려울 것으로 보여 별도의 대리변수를 활용해야 할 것으로 판단된다.



[표 3-3] 서울시 기초자료 검토 결과

분야	자료명	시간범위	제공기관	비고
교통사고	사고유형별 교통사고건수, 사망자수	2007~2016	도로교통공단	속성자료
토지	서울시토지이용현황도	1992~2006	서울시	공간자료
	토지임야대장	1997~2009	서울시	속성자료
	국토교통부지적통계	1970~2013	국토해양부	속성자료
	지적도	1999~2016	서울시	공간자료
	수치표고모델	2009	서울시	공간자료
	수치지형도	1997~2016	서울시	공간자료
지형	국가기초구역별 인구총조사 집계자료	2014~2016	서울시	공간자료
인구	사망원인통계	1999~2015	통계청	속성자료
	통계청 집계구별 통계(인구, 가구)	2000~2015	통계청	속성자료
	통계청 인구총조사	1995~2015	통계청	속성자료
	주민등록인구	1992~2016	통계청	속성자료
기상	기상자료(강수량, 풍향, 풍속 등)	~2016	기상청	속성자료
도시시설	도로기하구조(선형, 포장, 교차로 개수)		국토교통부	속성자료
통행	서울시 교통량	2005~2009	서울시	속성자료
	서울시 유동인구	2009~2015	서울시	속성자료
교통	택시·대중교통 DTG	~2015	서울시	속성자료
	고속도로 통행 데이터		한국도로공사	속성자료
	교통소통이력자료(DSRC)		국토교통부	속성자료
	버스노선운행정보, 운행자료, 이용객	2007~2011	서울시	속성자료
	지하철 수송실적, 교통카드거래내역	2010~2012	서울시	속성자료
	수도권 가구통행실태조사 자료	1996~2016	서울연구원	속성자료
	수도권 차량 통행속도 DB	2001	서울시	공간자료
토지이용	서울시 정기속도조사	2004~2007	서울시	속성자료
	서울시 도로관리 시스템	2010~2015	서울시	공간자료
	도시계획현황도	2001~2006	서울연구원	이미지
경제	도시계획 용도지역지구도	2007~2015	서울시	공간자료
	전국사업체조사자료	1994~2015	서울시	속성자료
	지역별고용조사	2008~2015	통계청	속성자료
	가계동향조사	1982~2015	통계청	속성자료
	건축물대장 집계자료	2015	서울연구원	공간·속성자료
건축	부동산114 아파트 시세	1999~2016	부동산114	공간·속성자료
	통계청 집계구별 통계(주택)	2000~2015	통계청	속성자료
	주택 실거래자료	2007~2016	서울시	속성자료
	재산세 과세자료	2000~2016	서울시	속성자료
	건축물대장	2007~2016	서울시	속성자료
	주거실태조사	2006~2015	국토연구원	속성자료

자료: 기관별 홈페이지, 공공데이터포털(www.data.go.kr)

## 4\_ 시사점 및 연구방향 설정

본 장에서는 교통사고 예방대책 수립을 위한 기반을 다양한 측면에서 검토하였다.

먼저 교통사고 위험도 평가와 관련하여, 교통사고 통계자료는 도로교통공단에서 운영하는 교통사고분석시스템의 자료를 활용하되 공리적인 관점에서 사고건수를 기준으로 적용하는 것이 바람직하다. 다만 교통사고의 심각도에 따라 발생원인과 개선대책이 크게 다르게 나타나므로 사상사고(Fatal and Injury, FI)와 물피사고(Property Damage Only, PDO)를 구분한 사고건수를 활용하도록 한다. 특히, 앞서 검토한 바와 같이 서울시 사망사고의 81%, 사상사고(FI)의 47%가 보행자 관련 사고임을 고려하여 보행사고를 중심으로 접근한다.

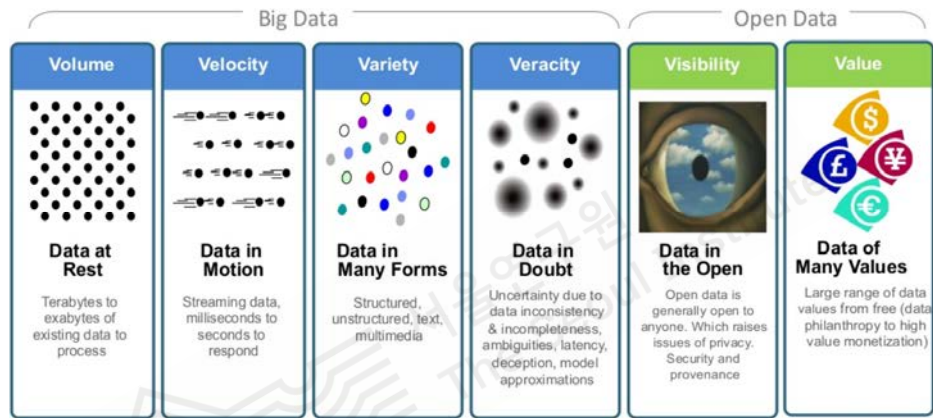
종합적인 정책적 시사점 도출을 위하여 교통사고 위험도 평가모형은 교통사고 발생건수 예측과 교통사고 발생원인 분석 모두를 포함하도록 구성하고, 전통적인 통계적 기법 외에 최근 활용도가 높아지고 있는 기계학습 기법을 중심으로 검토한다. 마지막으로 평가대상은 분석근거가 비교적 명확한 사고건수를 중심으로 하되, 실제 교통사고가 발생하지 않았더라도 잠재적으로 위험지역에 대한 분석이 가능하도록 향후 수집이 가능할 것으로 전망되는 자료들을 통해 평가모형을 확대 적용할 수 있는 방안을 고려하도록 한다.

기존 문헌검토와 전문가 자문을 통해 교통사고에 대한 영향요인을 검토한 결과, 사회환경, 교통체계 및 인간활동에 대한 종합적인 고려가 필요한 것으로 나타났다. 이에 따라 서울시의 기초자료를 조사하여 활용가능한 후보군을 제시한다.

위와 같이 대책 수립을 위한 서울시의 기반을 살펴보았을 때, 최근 다양한 분야에서 언급되고 있는 ‘빅데이터’로서의 교통안전 분야의 데이터는 이제 막 생성되고 있는 단계에 있는 것으로 판단된다. 즉, 데이터의 양(volume), 데이터의 생성 속도(velocity), 데이터 종류의 다양성(variety), 데이터의 정확성(veracity)을 특징으로 하는 빅데이터(IBM, 2014)로서의 데이터 중에 실제 공공분야에서 저장, 관리하는 것은 스마트카드 이용내역 등 일부에 국한되며, 통신사 이용내역, 보험사 정보 등은 한시적일회성으로 가공되어 공공에 제공되고 있는 수준이다.

이에 따라 보다 고도화된 빅데이터 활용을 위해서는 공공분야의 자체적인 데이터 수집, 가공뿐 아니라 민간분야의 참여를 유도할 수 있는 방안을 구축하는 것이 급선무일 것으로 판단된다. 따라서 보다 활용도 높은 기초자료 확보를 위해서 지속적으로 1) 정보수집원의 다양화, 2) 정보분석의 고도화, 3) 민간 정보교류 등을 추진해야 할 것이다.

- 정보수집원의 다양화 : DSRC RSE 설치(대구광역시 사례), 공공 IoT센서망 구축 등
- 정보분석의 고도화 : CCTV 영상분석, DTG 자료 분석 등
- 민간자료의 결합 : 민간 정보교류, 민간자료 구매 등(통신사, 보험사 등)



[그림 3-1] 빅데이터와 공공데이터의 정의

자료: IBM, 2014

# 04

---

## 교통사고 위험도 평가모형 구축

- 1\_평가모형 선정
- 2\_분석자료 구축
- 3\_사례분석 결과
- 4\_평가모형 검토 결과

## 04 | 교통사고 위험도 평가모형 구축

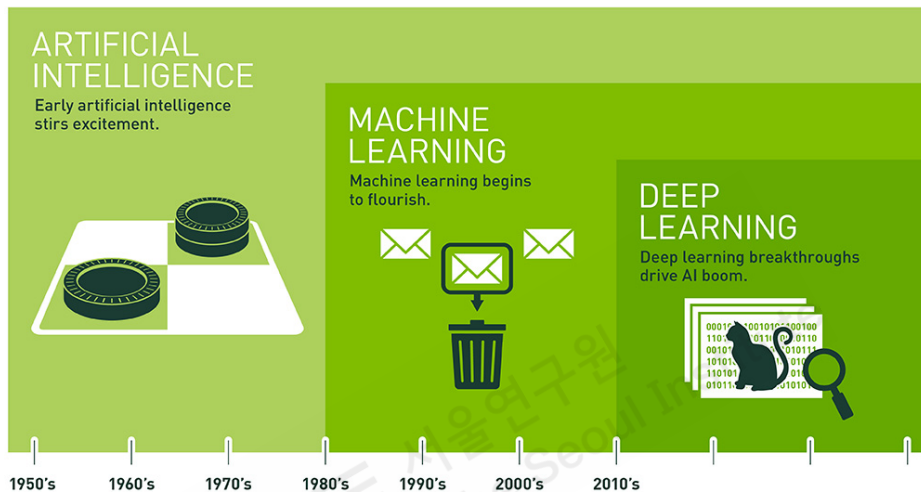
### 1\_평가모형 선정

기존 관련 문헌을 검토한 결과 교통사고 위험도 분석에 적용 가능한 모형은 크게 다음 표와 같다. 연구방향 설정 과정에서 기술한 바와 같이 본 연구에서는 향후 다양한 데이터가 수집될 경우의 확장성과 잠재적 활용성을 고려하여 기계학습 기법을 중심으로 검토하도록 한다.

[표 4-1] 교통사고 위험도 평가모형

대분류	중분류	소분류	알고리즘	적용분야
통계 기법	모수	회귀분석	단순회귀 (Ordinary Least Square, OLSQ)	수치예측
			Geologically Weighted Regression (GWR)	
		요인분석	-	분류예측
		원단위법	-	수치예측
	비모수	검정	Kolmogorov-Smirnov test	
			Cochran's Q	
		비교	Logrank test	
		측정	Cohen's kappa	
기계 학습 기법	지도	회귀	Classification And Regression Tree (CART)	수치예측
			Logistic Regression	분류예측
		분류	K-Nearest Neighbors (K-NN)	분류예측
			Naïve Bayesian	수치예측
	Decision Tree (DT)		분류예측	
	블랙박스	Random Forestm XGBoost	다중용도	
		Support Vector Machine	분류예측	
	비지도	군집모델	K-means Clustering	군집예측
		연관모델	연관규칙분석 (A Priori Algorithm)	연관예측
	강화	인공신경망	ANN, DNN, CNN, RNN	다중용도

기계학습(Machine Learning)이란 1950년대부터 인공지능(Artificial Intelligence) 분야에서 시작된 연구방법으로, 알고리즘 기반으로 자료의 패턴, 규칙, 의미 등을 컴퓨터가 스스로 학습하고 이를 토대로 결과를 예측하는 기법이다. 최근 컴퓨터의 성능이 발전함에 따라 다층적(multi layer) 분석이 용이해져 딥러닝(Deep Learning)을 중심으로 기계학습에 대한 관심이 커지고 있다.



[그림 4-1] 인공지능, 기계학습 및 딥러닝의 관계

자료: NVIDIA 공식 블로그

앞서 표로 제시한 바와 같이 기계학습은 지도학습(Supervised Learning), 비지도학습(Unsupervised Learning), 강화학습(Reinforcement Learning)으로 구분할 수 있다. 지도학습은 학습자료(Training Data)가 정해져 있는 형태이며, 비지도학습은 학습자료가 정해져 있지 않은 형태로 자체적인 자료집합에서 규칙 및 그룹을 지어 학습하는 방식을 의미한다. 마지막으로 강화학습은 선택 가능한 행동들 중 보상을 최대화하는 행동 혹은 행동순서를 선택하도록 학습하는 방식이다.

본 연구의 주요 목적은 1) 교통체계-사회환경-인구활동의 상호관계를 고려한 교통사고 발생원인 분석, 2) 공간적 특성을 반영한 교통사고 영향요인 분석, 3) 장래 여건변화를 고려한 교통사고 위험도 예측 등이다. 이러한 관점에서 선정한 모형은 지리적 이질성을 고려한 회귀분석 기법인 지리가중회귀분석(Geographically Weighted Regression),

분석대상 간에 존재하는 패턴을 예측 가능한 여러 가지 규칙들의 조합으로 분류하는 의사결정나무(Decision Tree) 및 자료 간의 상대거리를 통해 분석대상을 분류하는 K-최근접이웃(K-Nearest Neighbors) 등이다.

각 모형은 각각 지역별 특성을 반영할 수 있다는 장점(지리가중회귀분석), 교통사고 발생 원인 간의 상관관계를 분석할 수 있다는 장점(의사결정나무), 장래 여건변화가 교통사고 위험도에 미치는 영향을 예측할 수 있다는 장점(K-최근접이웃)을 고려하여 선정하였다.

[표 4-2] 선정된 평가모형

구분	내용	지역별 특성반영	변수 간 상관관계	장래 변화예측
지리가중회귀분석 (GWR)	지리적 이질성을 고려한 회귀분석	○	X	○
의사결정나무 (DT)	다차원의 기준으로 분석대상을 분류	X	○	△
K-최근접이웃 (K-NN)	자료 간의 거리를 통해 분석대상을 분류	X	X	○

따라서 전체적인 분석과정은 다음과 같다. 먼저 지리가중회귀분석을 통해 지리적 특성에 따른 교통사고 발생원인별 영향을 살펴보고, 의사결정나무를 통해 변수 간의 조합을 통해 단계적으로 대상을 분류하여 변수 간의 독립성을 가정하는 회귀분석의 한계를 보완하도록 한다. 마지막으로 K-최근접이웃을 통해 개별 사고요인이 장래에 변화할 경우의 사고발생 위험도 변화를 예측하여 제시할 계획이다.

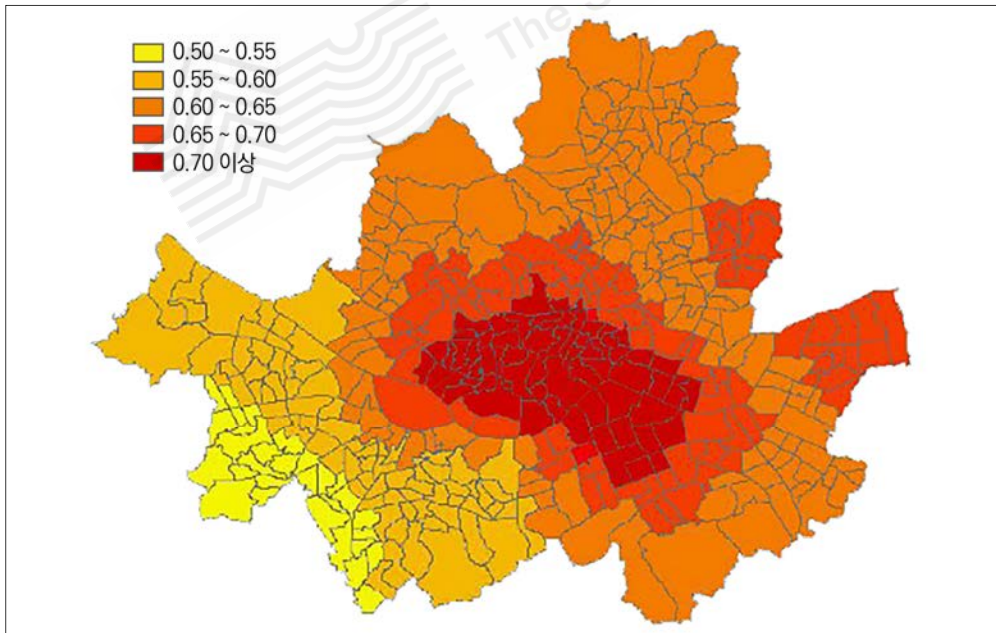
### 1) 지리가중회귀분석(Geographically Weighted Regression)

지리가중회귀분석은 공간회귀분석(Spatial Regression)의 일종으로, 회귀분석에서 자료들이 공간적인 요소에 영향을 받을 때 자료 간의 공간적 상관성을 분석하고 이 상관성에 따라 변수의 유의성이 낮아지는 문제를 해결하기 위해 사용된다. 특히 지리가중회귀분석은 공간 단위로 별도의 변수별 계수값을 제시하는 특성을 가지고 있다. 이를 통해 독립변

수와 종속변수의 관계에서 공간적 이질성을 평가할 수 있는 장점이 있다.

GWR은 모든 위치에서 종속변수와 독립변수 간의 관계가 동일하지 않기 때문에 독립변수의 변화 정도가 동일하여도 종속변수의 변화 정도는 공간적 위치에 따라 달라진다고 가정하는 모델이다. 지리가중회귀분석은 공간행렬에 의해 결정되는데, 이는 지리적 가중함수의 대역폭(Bandwidth)을 고정할 것인가 가변적으로 설정할 것인가가 중요하다. 대역폭 설정 문제에 따라 고정적 커널(fixed kernel)과 적응적 커널(adaptive kernel)로 구분된다. 고정적 커널은 모든 지역에 대해 동일한 대역폭을 적용하고, 적응적 커널은 지역에 따라 가변적으로 커널을 적용하는 방식이다(정효잔이지영, 2015).

선행연구에서도 행정동별로 집계된 교통사고 사고건수 간에 공간적 상관관계가 존재하며, 이를 고려한 지리가중회귀분석이 이반적인 회귀분석에 비해 향상된 설명력을 보이는 것으로 나타났다. 특히 지리가중회귀분석을 통해 특정 변수가 교통사고 발생에 미치는 영향이 행정동마다 다르고, 특정 행정동의 변수 변화가 해당 동뿐만 아니라 인근 행정동의 교통사고 발생에 미칠 수 있음을 제시하였다(홍지연, 2013).



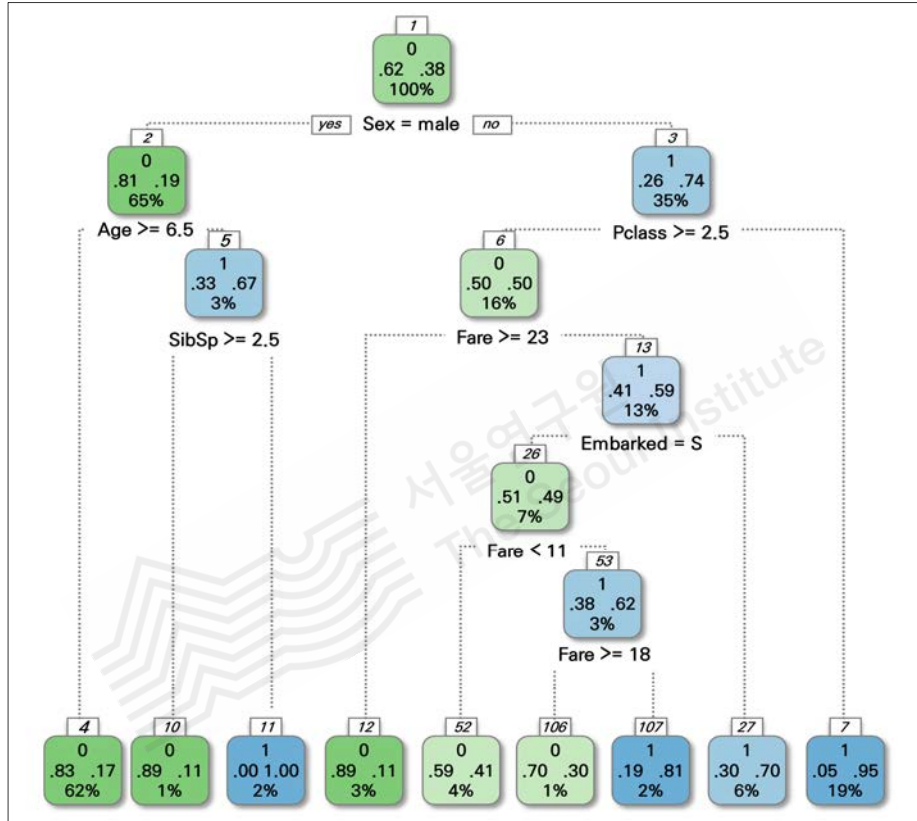
[그림 4-2] 지리가중회귀분석의 지역별 적합도( $r^2$ ) 분석 사례

자료: 공간계량분석을 통한 도시교통사고 예측모형 (홍지연, 2013)



## 2) 의사결정나무(Decision Tree)

의사결정나무는 자료를 분석하여 변수 간의 패턴을 예측 가능한 조건부 규칙들의 조합으로 나타내며, 분석결과가 트리 구조로 나타난다. 의사결정나무는 분류(classification)와 회귀(regression) 분석이 모두 가능하며, 범주형 및 연속형 자료에 적용 가능하다.



[그림 4-3] 의사결정나무 분석결과 예시

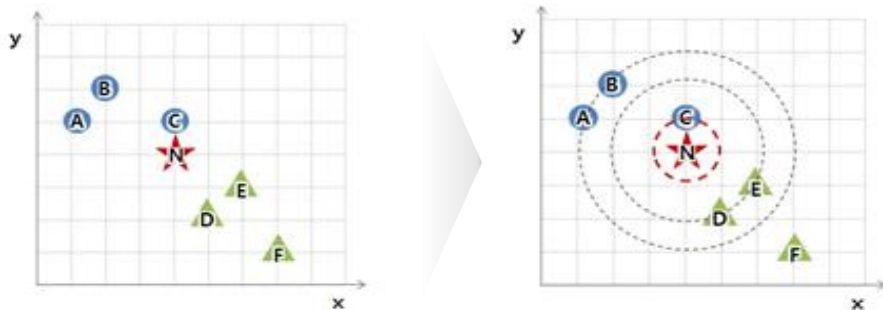
의사결정나무의 가장 큰 장점은 트리 구조의 조건부 규칙들을 제시하여 직관적으로 이해하기 쉽다는 점이다. 또한 수집가능한 자료의 샘플수가 비교적 적은 경우에도 자료들 간의 상관관계와 종속관계에 대한 정보를 추출할 수 있다. 다만 이러한 경우에는 해당 분야 전문가의 판단을 통해 분석결과를 해석하는 과정이 필수적이다. 마지막으로 다른 모형과의 결합이 용이하여 의사결정을 위한 보조수단으로서 활용도가 높은 편이다.

이 모형의 한계는 입력자료의 변화에 따라 의사결정 과정이 쉽게 변경될 수 있다는 점이다. 그러나 이러한 문제는 자료 기반의 모형에서 일반적으로 발생하는 문제이며, 이를 해결하기 위해 구성 가능한 다수의 의사결정나무를 만들고 각 모형별 분석결과를 종합하여 최종 결과를 도출하는 랜덤포레스트(Random Forest) 모형이 활용되고 있다. 다만 랜덤포레스트의 경우에는 보다 신뢰성 있는 예측결과를 나타낼 수 있지만, 의사결정 과정을 명확히 파악하기 어려워 자료에 기반한 정책 방향을 도출하는 데에는 한계가 있다.

### 3) K-최근접이웃(K-Nearest Neighbors)

K-최근접이웃은 기계학습을 활용한 분류(classification) 문제에 쓰이는 대표적인 기법이다. 이 기법은 범주를 알지 못하는 자료의 범주를 분류하고 예측하는 데 사용된다. 즉, 기존의 학습자료(Training Data)를 통해 각 범주 간의 분류기준을 설정한 후에 새로운 자료가 입력되었을 때 기존 범주 중 어떤 것에 가장 가까운지를 판별하는 것이다.

분석과정에서는 새로운 자료가 입력되었을 때 이 값과 가장 가까운 K개의 학습자료를 사용하여 새로운 자료의 범주를 예측한다. 이때 거리를 측정하는 방법으로는 2차원(평면)의 유클리디안(Euclidean) 거리를 이용하는 방법이 일반적으로 사용되지만, 분석목적에 따라 3차원(공간)의 거리 또는 그 이상의 초평면(Hyper Plane)의 거리를 고려할 수 있다. 예를 들어 다음 그림에서  $K=1$ 이고 유클리디안 거리를 사용하는 경우를 가정하면, 새로운 자료 N이 입력되었을 때 거리가 가장 가까운 1개의 자료( $K=1$ )를 고려하여 C와 같은 범주(A, B, C)에 속하는 것으로 분류하는 것이다.



[그림 4-4] K-최근접이웃의 분석 개념도

K-최근접이웃은 분석과정이 단순하고 효율적이며, 훈련 데이터의 양이 많을수록 효과적이다. 이 기법은 사례기반추론(Case-Based Reasoning, CBR)의 일종으로 과거에 적용되었던 사례와 그 결과들을 참조하여 새로운 사례에 대한 결과값을 예측하는 방법이다. 사례기반추론은 상대적으로 적은 데이터를 가진 분야에서도 적용이 가능하며, 복잡하거나 구조화가 부족한 분야에서도 쉽게 적용할 수 있다.

이 기법의 장점은 비교적 간단하며 연산시간이 짧고 대용량의 데이터에서 예측성능이 뛰어나다는 점이다. 또한 본 연구에서 다루는 것과 같은 수치기반 자료의 분류에서 성능이 뛰어난 것으로 알려져있다. 반면에 학습자료가 많아지거나 차원의 크기가 커질수록 분류 속도가 느려지는 단점이 있다. 따라서 적절한 K값을 선정하는 것이 중요하며, 일반적으로는 전체 자료수의 제곱근을 초기 K값으로 하여 오차가 줄어드는 방향으로 조정하는 방식으로 수행한다.



## 2\_분석자료 구축

### 1) 교통사고 발생건수 (종속변수)

교통사고자료는 교통사고분석시스템(TAAS)에서 제공하는 2014년도 서울시 교통사고자료(총 40,437건)를 기준으로 하였다. 교통사고분석시스템에서 제공하는 자료는 다음 표와 같이 총 17가지 변수로 구성된다. 변수는 주로 사고위치, 시간, 사고유형, 당사자, 도로조건 등에 관한 항목으로 구성된다.

[표 4-3] 교통사고분석시스템의 주요 변수

구분	변수	내용
시간	연도	-
	계절	봄 / 여름 / 가을 / 겨울
	주야	주 / 야
공간	위경도	X-Y 좌표
	행정구역	시도명 / 시군구명 / 행정동
기상	-	맑음 / 눈비 / 흐림안개
사고유형	-	차대차 / 차대사람 / 차량단독
사고1당사자	성별	남 / 여
	연령	연령
	사고정도	사망 / 중상 / 경상 / 부상 / 상해없음
	수단	수단
사고2당사자	성별	남 / 여
	연령	연령
	사고정도	사망 / 중상 / 경상 / 부상 / 상해없음
	수단	수단
도로조건	평면	직선 / 곡선 / 기타
	종단	평지 / 오르막 / 내리막 / 기타

주: 음영은 본 연구에 반영한 변수임.

본 연구에서는 최근 교통사고 발생비율이 증가하고 사상사고 비율이 높은 차대사람 사고, 특히 보행자 사고를 분석대상으로 하기 위해 사고유형이 차대사람이고 사고2당사자의 수단이 보행인 경우(10,177건)로 자료를 추출하였다. 또한 사고 심각도에 따라 발생

원인이나 개선대책의 유형이 달라지는 점을 고려하여 사고2당사자의 사고정도를 기준으로 사상사고(FI)와 물피사고(PDO)로 유형을 구분하였다. 또한 교통사고 발생건수를 행정동별로 집계하여 분석하기 위해 행정구역 자료를 활용하였다.

## 2) 교통사고 영향요인 (독립변수)

본 연구의 주요 목적이 정확한 분석결과를 도출하는 것보다 장래 서울시 교통사고 위험도 평가방법을 수립하고 그 적용성을 확인하는 것이므로, 모형의 특성에 따라 최적 변수 조합을 선정하지 않고 심각도별 교통사고 발생건수와 관련이 높은 독립변수 조합을 선정하고 동일한 조합을 각 모형에 적용하는 것이 적합할 것으로 판단된다.

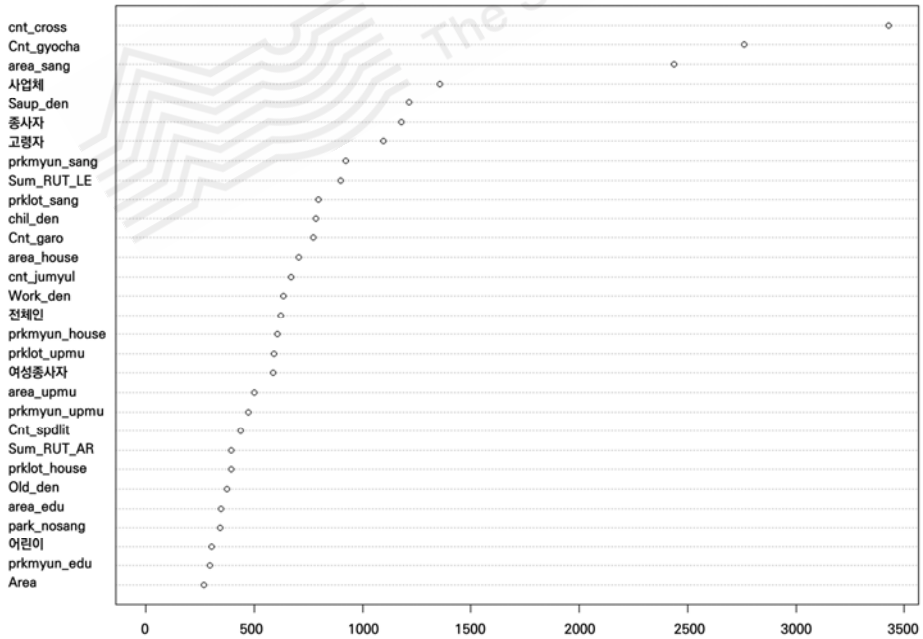
이에 따라 교통사고 영향요인을 앞서 살펴본 바와 같이 크게 사회환경-교통여건-인구활동 분야로 구분하여 30여 개의 변수를 검토하였다. 수집된 자료와 심각도별 교통사고 발생건수를 랜덤포레스트(Random Forest) 기법을 활용하여 변수별 중요도를 평가하였다.

사상사고(FI)의 경우 변수별 중요도는 [그림 4-5]와 같다. 분석결과에 정책적 필요성을 반영한 독립변수로 사회환경(고령자수, 주거지역 연면적, 상업지역 연면적, 업무지역 연면적), 교통여건(도로연장, 교차로수, 횡단보도수, 노상주차면수), 인구활동(종사자수)이 선정되었다.

물피사고(PDO)의 경우 변수별 중요도는 [그림 4-6]과 같다. 분석결과에 정책적 필요성을 반영한 독립변수로 사회환경(어린이수, 주거지역 연면적, 상업지역 연면적, 업무지역 연면적), 교통여건(도로연장, 교차로수, 횡단보도수, 노상주차면수), 인구활동(사업체수, 종사자수)이 선정되었다.

[표 4-4] 사상사고 영향요인 변수 선정결과

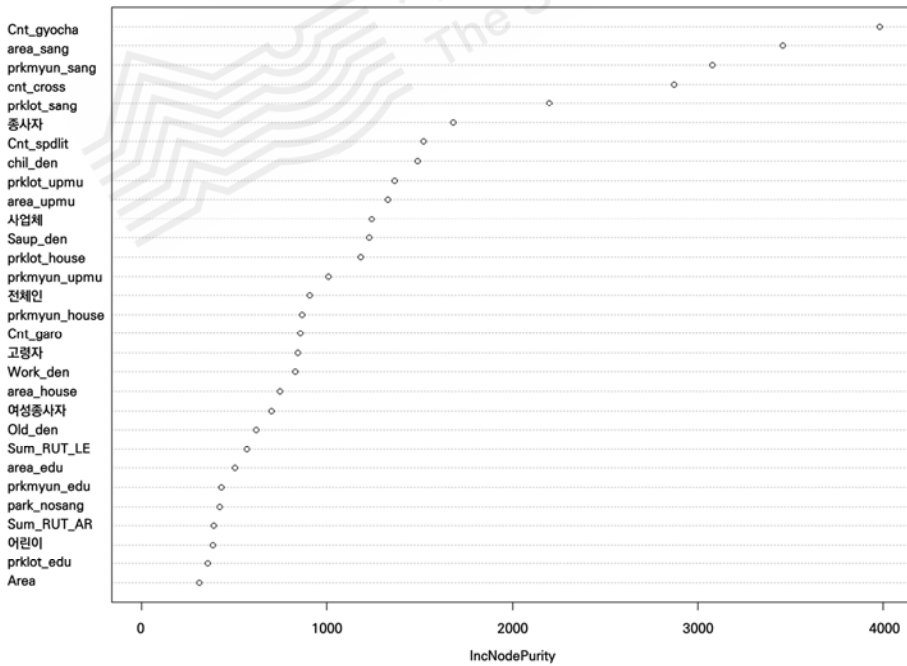
대분류	소분류	변수	자료출처
사회환경	인구구조	행정동별 인구	인구센서스
		고령자수, 비율	
		어린이수, 비율	
	토지이용	행정동별 면적	도시계획도
		주거지역 연면적	
		상업지역 연면적	
업무지역 연면적			
교육시설 연면적			
교통여건	도로시설	연장	도로관리시스템
		교차로수	
		횡단보도수	
	주차환경	노상주차면수	주차정보안내시스템
		노외주차면수	
인구활동	경제활동	사업체수, 밀도	사업체조사
		종사자수, 비율	



[그림 4-5] 사상사고의 변수별 중요도 분석결과

[표 4-5] 물피사고 영향요인 변수 선정결과

대분류	소분류	변수	자료출처
사회환경	인구구조	행정동별 인구	인구센서스
		고령자수, 비율	
		어린이수, 비율	
	토지이용	행정동별 면적	도시계획도
		주거지역 연면적	
		상업지역 연면적	
업무지역 연면적			
교육시설 연면적			
교통여건	도로시설	연장	도로관리시스템
		교차로수	
		횡단보도수	
	주차환경	노상주차면수	주차정보안내시스템
		노외주차면수	
인구활동	경제활동	사업체수, 밀도	사업체조사
		종사자수, 비율	



[그림 4-6] 물피사고의 변수별 중요도 분석결과

### 3\_사례분석 결과

#### 1) 지리가중회귀분석(Geographically Weighted Regression)

지리가중회귀분석을 활용하면 교통사고 발생요인(독립변수)이 행정동별 교통사고 발생건수에 어떤 영향을 미치는지를 계량적으로 분석할 수 있다. 특히 서울시 전역에 대한 분석이 아닌, 행정동별 지리적 특성을 반영할 수 있어 지역별 교통안전 개선대책을 수립하는데 있어 정책적 시사점을 제시할 수 있을 것이다.

앞서 모형 선정단계에서 살펴본 바와 같이 교통사고 심각도에 따라 발생원인 및 개선대책이 크게 달라지는 점을 반영하여 사상사고(Fatal and Injury, FI)와 물피사고(Property Damage Only, PDO)로 구분하여 분석을 수행하였다. 또한 동일한 자료를 일반회귀분석한 결과를 같이 제시하여 두 모형의 차이점을 분석하였다. 또한 독립변수에 대한 사전 검토를 통해 지리적인 상관성이 없는 전역변수(Global Variable)와 지리적인 상관성을 보이는 지역변수(Local Variable)를 구분하여 모형을 구성하였다. 지역변수의 경우에는 추정계수의 범위만을 요약하여 제시하였다.

##### (1) 사상사고(FI)

사상사고의 경우에 일반회귀분석을 수행하면 조정된 결정계수(Adjusted  $R^2$ )가 0.56인 것에 반해 지리가중회귀분석을 수행하면 0.64로 증가하여 전반적인 모형의 설명력이 높아진 것으로 나타났다.

사상사고의 경우에 지리적 상관성을 보이는 변수는 노상주차면수, 상업지역 연면적, 업무지역 연면적, 교차로수 등으로 나타났다. 특히 이 변수들은 일반회귀분석에서 양(+)의 값을 보이는 것에 비해 지리가중회귀분석에서는 행정동에 따라 양(+)의 값과 음(-)의 값을 보이는 경우가 다르게 나타나, 행정동별로 교통안전 개선대책에 대해 반대 방향에서 접근할 수 있음을 시사하고 있다.



[표 4-6] 지리가중회귀분석과 일반회귀분석 결과 비교 (사상사고)

구분	지리가중회귀분석			일반회귀분석		
	추정계수	표준오차	표준화계수	추정계수	표준오차	표준화계수
상수항	9.44~15.34	-	3.10~20.24	11.71	0.26	45.47
고령자수	1.51	0.33	4.51	1.36	0.31	4.40
종사자수	0.70	0.27	2.58	0.49	0.28	1.74
도로연장	0.09	0.34	0.27	0.41	0.33	1.23
노상주차면수	-3.96~6.45	-	-3.50~4.88	0.75	0.30	2.51
주거지 연면적	1.07	0.39	2.72	1.04	0.36	2.90
상업지 연면적	-1.01~6.42	-	-1.12~6.08	1.76	0.37	4.80
업무지 연면적	-13.80~15.10	-	-2.48~5.02	0.54	0.35	1.54
횡단보도수	1.70	0.40	4.28	2.11	0.38	5.53
교차로수	-1.30~4.21	-	-1.33~6.24	1.07	0.31	3.48
통계량	- 잔차제곱합: 8,014 - AICc: 2,597 - Adjusted R <sup>2</sup> : 0.64			- 잔차제곱합: 11,636 - AICc: 2,630 - Adjusted R <sup>2</sup> : 0.56		

특히 노상주차면수와 교차로수와 같이 교통안전 개선대책에 직접적으로 활용할 수 있는 변수의 계수가 양(+) 또는 음(-)의 값을 보임에 따라 행정동의 어떠한 교통체계, 사회환경, 인구활동이 이러한 차이를 나타내는지를 밝히는 것이 향후 정책 방향을 설정하는 데에 중요할 것으로 판단된다. 사례로서 일부 지역을 살펴보면, 대조동, 녹번동, 응암동 등의 지역에서는 노상주차면수의 증가가 사상사고 위험도를 높일 것으로 분석되는 반면, 잠실동, 대치동, 서초동 등의 지역에서는 노상주차면수의 증가가 사상사고 위험도를 낮추는 것으로 나타났다.

[표 4-7] 노상주차면수의 계수값 (일부 지역)

구분	대조동	녹번동	응암동	홍제2동	잠실2동	대치4동	서초2동	도곡1동
노상주차면수	4.07	4.34	4.61	4.88	-2.45	-3.31	-3.45	-3.50

또한 이화동, 창신동, 종로 지역에서는 교차로수가 증가할수록 교통사고 위험도가 증가하였지만, 잠실동, 방이동, 송파동 등의 지역에서는 교차로수가 증가하면 사상사고 위험도가 감소할 것으로 분석되었다.

**[표 4-8] 교차로수의 계수값 (일부 지역)**

구분	이화동	창신1동	종로5,6가동	혜화동	잠실6동	방이2동	송파1동	가락1동
교차로수	6.24	6.01	6.09	5.73	-1.33	-1.22	-1.22	-1.19

(2) 물피사고(PDO)

물피사고의 경우에도 지리가중회귀분석의 조정된 결정계수가 0.52로 다소 증가하여 모형의 설명력이 높아진 것으로 나타났다.

**[표 4-9] 지리가중회귀분석과 일반회귀분석 결과 비교 (물피사고)**

구분	지리가중회귀분석			일반회귀분석		
	추정계수	표준오차	표준화계수	추정계수	표준오차	표준화계수
상수항	11.77~12.64	-	27.47~31.13	12.46	0.33	37.82
사업체수	-0.96~0.52	-	-1.90~0.93	-0.32	0.43	-0.75
종사자수	-0.55~1.61	-	-0.98~3.02	0.59	0.42	1.42
도로연장	-0.19~0.68	-	-0.37~1.17	0.18	0.42	0.43
어린이 인구밀도	-1.66	0.40	-4.13	-1.56	0.40	-3.93
노상주차면수	-1.00~1.02	-	-2.10~2.24	0.21	0.39	0.53
주거지 연면적	-0.07~1.68	-	-0.13~2.56	0.63	0.47	1.34
상업지 연면적	2.12~2.94	-	3.75~4.87	2.69	0.47	5.75
업무지 연면적	-0.28~1.70	-	-0.49~3.45	0.97	0.43	2.25
횡단보도수	2.21~3.94	-	3.11~5.92	2.52	0.50	5.05
교차로수	0.67~1.37	-	1.32~2.94	0.98	0.39	2.48
통계량	- 잔차제곱합: 17,622 - AICc: 2,832 - Adjusted R <sup>2</sup> : 0.52			- 잔차제곱합: 18,995 - AICc: 2,840 - Adjusted R <sup>2</sup> : 0.50		

반면 물피사고의 경우 어린이 인구밀도만 전역변수로 선정되었고, 나머지 변수는 모두 지역변수로 나타나 물피사고의 발생유형이 행정동별로 매우 다양하게 나타난다는 것을 시사하고 있다.

앞서 사상사고에서와 마찬가지로 일반회귀분석이 변수별로 단일한 계수값을 보이는 것에 비해, 지리가중회귀분석에서는 행정동별로 다른 계수값을 보였다. 다만 사업체수, 종사자수, 도로연장, 노상주차면수, 주거지역 연면적, 업무지역 연면적은 행정동별로 양(+) 또는 음(-)의 계수값을 나타내 교통안전 개선대책 수립방향이 달라야 할 것으로 분석되었다. 하지만 상업지역 연면적, 횡단보도수, 교차로수는 값은 다르지만 모두 양(+)의 값을 보여 이들 변수는 공통적으로 물피사고 위험도를 증가시키는 것으로 나타났다.

예를 들어 행정동별 노상주차면수의 계수를 살펴보면 진관동, 갈현동, 불광동 등에서는 노상주차면수의 증가가 물피사고 위험도를 높일 것으로 나타났지만, 옥수동, 청담동, 잠원동 등에서는 오히려 물피사고 위험도를 감소시킬 것으로 분석되었다.

[표 4-10] 노상주차면수의 계수값 (일부 지역)

구분	진관동	갈현1동	불광2동	구산동	옥수동	청담동	잠원동	역삼1동
노상주차면수	2.14	2.24	2.03	2.14	-1.53	-1.72	-1.75	-1.94

또한 횡단보도수의 계수를 살펴보면 모든 행정동에서 횡단보도수의 증가가 물피사고 위험도를 높이는 것으로 나타났지만, 행정동별로 그 영향의 정도는 다르게 나타났다. 도봉동, 쌍문동, 방학동 등에서는 그 영향이 비교적 작은 반면에, 소공동, 회현동, 효창동 등에서는 횡단보도수가 늘어날수록 물피사고 발생건수가 증가할 경향이 강한 것으로 나타났다.

[표 4-11] 횡단보도수의 계수값 (일부 지역)

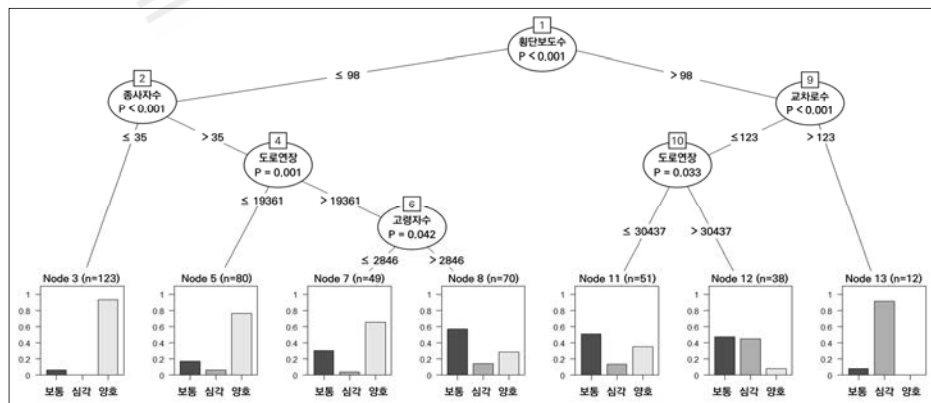
구분	도봉1동	쌍문4동	방학2동	우이동	소공동	회현동	효창동	도화동
횡단보도수	3.11	3.42	3.29	3.39	5.45	5.33	5.83	5.80

2) 의사결정나무(Decision Tree)

앞서 지리가중회귀분석을 통해 교통사고 발생요인이 지역별 특성에 따라 교통사고 발생건수에 미치는 영향이 다를 수 있음을 살펴보았다. 그러나 지리가중회귀분석을 통해서 지역별 영향이 다른 원인에 대한 정보를 알 수 없는 한계가 있었다. 의사결정나무는 변수 간의 종속관계를 고려하여 순차적인 조건부 판별기준을 통해 교통사고 발생건수를 분류하게 된다. 이로써 개별 변수 간의 조합이 최종결과에 미치는 영향을 구조적으로 보여줄 수 있고, 각 교통사고 발생요인이 지역별로 다른 영향을 미치는 원인에 대한 정보를 분석가에게 제시할 수 있다.

(1) 사상사고(FI)

사상사고에 대한 의사결정나무 분석결과는 다음 그림과 같다. 첫 번째 판별기준(=노드1) 으로부터 순차적인 분류를 통하여 최종적인 사상사고 발생건수를 결정하게 된다. 사상사고에 있어서 첫 번째 판별기준은 횡단보도수이다. 횡단보도수가 임계수준(98개)을 넘고, 교차로수가 임계수준(123개)을 넘는 경우에 교통사고 위험도가 가장 높을 것으로 분류되었다. 분석자료 내에서 이러한 범주에 속하는 관측수는 12개 행정동임을 보여주고 있다. 마찬가지로 횡단보도수가 임계수준 이하이고, 종사자수가 임계수준(35명) 이하인 경우에는 교통사고 위험도가 가장 낮을 것으로 나타났다. 인구구조에 관한 변수인 고령자수는 중간 단계의 위험도(노드7, 노드8)를 구분하는 분류기준으로 사용되었다.

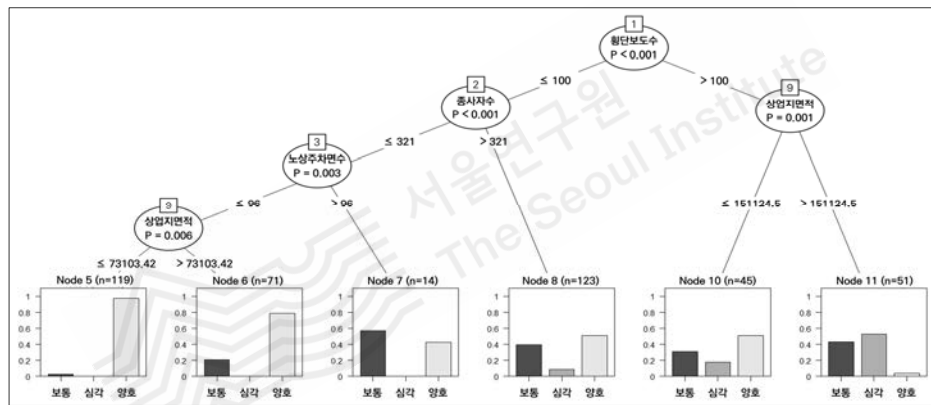


[그림 4-7] 의사결정나무 분석결과 (사상사고)

(2) 물피사고(PDO)

물피사고에 대한 분석결과를 살펴보면, 첫 번째 판별기준은 사상사고와 동일하게 횡단보도수인 것으로 나타났다. 횡단보도수가 임계수준(100개)을 넘고, 상업지역 연면적이 임계수준(151,124m<sup>2</sup>)을 넘는 경우에 교통사고 위험도가 가장 높을 것으로 분류되었다. 반면에 행정동 내의 횡단보도수, 종사자수, 노상주차면수, 상업지역 연면적이 적을 경우에는 교통사고 위험도가 가장 낮을 것으로 나타났다.

또한 노상주차면수는 횡단보도수가 충분한 경우에는 교통사고 위험도에 별다른 영향을 미치지 않지만, 횡단보도수가 적은 곳에서는 사고위험도 판별기준으로 사용되었으며 노상주차면수가 임계수준보다 높을 경우 교통사고 위험도도 증가하는 것으로 분석되었다.



[그림 4-8] 의사결정나무 분석결과 (물피사고)

사상사고와 물피사고의 위험도 판별에 쓰인 변수들을 살펴보면 사상사고(횡단보도수, 종사자수, 도로연장, 교차로수, 고령자수), 물피사고(횡단보도수, 종사자수, 노상주차면수, 상업지역 연면적)가 다소 다른 경향을 보이는 것으로 나타난다. 즉, 횡단보도수와 종사자수는 사상사고와 물피사고의 위험도를 판별하는 데에 공통적으로 사용된 반면, 사상사고는 도로연장, 교차로수와 같은 도로시설 및 고령자수에 의해 영향을 받는 것으로 나타났다. 반면 물피사고는 노상주차면수, 상업지역 연면적과 같이 노변에서 일어나는 활동에 더 영향을 받는 것으로 분석되었다.

### 3) K-최근접이웃(K-Nearest Neighbors)

앞서 교통사고 발생요인이 행정동별로 다른 영향을 미칠 수 있고(지리가중회귀분석), 각 발생요인들이 상호 간에 조건부로 연관되어 교통사고 위험도를 결정한다는 것(의사결정 나무)을 알 수 있었다. 이러한 과정을 거쳐 행정동별로 감소시키거나 분리해야 할 교통사고 발생요인들을 선정한 이후에는 개선사업의 효과분석이 수반되어야 한다. 즉, 특정 행정동의 교통체계, 사회환경, 인구활동이 변화되면 교통사고 위험도가 어떻게 바뀔 것인지에 대한 예측이 필요하다.

K-최근접이웃을 활용하면 신규 자료를 입력하였을 때 기존 자료 중에서 어떤 범주에 속하는지를 판단할 수 있으므로, 개선사업이 시행된 행정동별 특성을 통해 기존에 설정된 교통사고 위험도를 평가할 수 있다. 따라서 본 절에서는 특정 독립변수의 값을 임의로 +10%, +20% 변화시킨 뒤 교통사고 위험도가 어떻게 달라지는지를 제시하여 모형의 적용성을 검토하였다.

#### (1) 사상사고(FI)

사상사고에 대한 분석결과를 살펴보면 도봉동, 창동, 수유동 등에서는 노상주차면수를 추가 공급함에 따라 사상사고 위험도가 감소하는 것으로 전망되었다. 반면에 미아동, 상암동, 대치동 등에서는 오히려 노상주차면수를 공급하는 것이 사상사고 위험도를 증가시키는 것으로 나타났다. 이러한 전망을 활용하여 행정동별로 수립된 교통안전 개선사업의 적절성을 판단하고 개선사업별로 행정동별 최적의 예산 배분을 할 수 있을 것이다.

또한 도봉동, 창동, 수유동 등에서는 상업지역이 증가해도 사상사고 위험도에 큰 변화가 없거나 오히려 감소되는 반면, 미아동, 상암동, 대치동 등은 현재보다 상업지역 연면적이 증가할 경우 사상사고 위험도가 높아질 것으로 전망되었다.

앞선 분석결과들을 통해 유추할 때, 이러한 행정동별 차이는 해당 인구, 토지이용 대비 도로시설이나 교통시설이 충분한지 여부와 관계가 있는 것으로 보인다. 그러나 보다 구체적인 행정동별 특성에 대한 분석을 위해서는 향후 보다 심도있는 연구가 필요할 것이다.

[표 4-12] 장래 여건변화에 따른 사상사고 위험도 사례분석

구분	기존위험도	노상주차면수		상업지역 연면적	
		+10%	+20%	+10%	+20%
도봉1동	심각	양호	양호	양호	양호
창5동	보통	양호	양호	양호	양호
수유3동	심각	보통	보통	보통	보통
서초4동	심각	양호	양호	보통	보통
미아동	보통	심각	심각	심각	심각
상암동	보통	심각	심각	심각	심각
대치4동	보통	심각	심각	심각	보통
빙배동	양호	심각	심각	양호	양호

(2) 물피사고(PDO)

물피사고에 대한 분석결과는 다음 표와 같다. 물피사고는 발생원인 및 개선대책이 사상사고와 다른 것으로 알려져 있다. 사상사고와 비교를 위해 동일한 행정동의 분석결과를 제시한 본 분석에서도 사상사고와는 매우 다른 개선효과를 보이고 있다.

[표 4-13] 장래 여건변화에 따른 물피사고 위험도 사례분석

구분	기존위험도	노상주차면수		상업지역 연면적	
		+10%	+20%	+10%	+20%
도봉1동	양호	양호	양호	양호	양호
창5동	양호	양호	양호	양호	양호
수유3동	심각	양호	양호	양호	보통
서초4동	심각	양호	심각	심각	심각
미아동	심각	심각	심각	심각	심각
상암동	보통	심각	심각	심각	심각
대치4동	심각	심각	심각	심각	심각
빙배동	양호	양호	양호	보통	양호

## 4\_평가모형 검토 결과

본 연구에서는 교통사고 위험도 평가를 위한 모형으로 지역별 특성을 반영할 수 있는 지리가중회귀분석(Geographically Weighted Regression), 교통사고 발생원인 간의 상관관계를 분석할 수 있는 의사결정나무(Decision Tree), 장래 여건변화가 교통사고 위험도에 미치는 영향을 예측할 수 있는 K-최근접이웃(K-Nearest Neighbors)을 중심으로 검토하였다.

각 모형이 갖는 가정과 한계점에 따라 평가결과에서 얻을 수 있는 정책적 시사점이 다르게 나타나므로, 중앙정부와 서울시에서 운영하는 각종 교통안전 개선사업의 목적과 대상에 적합한 복수의 평가모형을 선정하여 상호 보완적으로 활용할 필요가 있다. 특히 평가모형에서와같이 변수 간의 독립성을 가정하는 전통적인 회귀분석과 변수 간의 이론적 관계를 고려하지 않는 기계학습은 분석자료에 대한 기본적인 관점의 차이가 있으므로 각 분석결과를 비교함으로써 보다 다양한 의미를 파악할 수 있다.

따라서 본 연구에서 제시한 바와 같이 1) 교통사고 위험도의 지역적 특성 평가, 2) 교통사고 발생원인 간의 상호연관성 분석, 3) 개선사업 시행 시 교통사고 위험도 전망의 순으로 단계적인 평가체계를 구성하되, 분석목적과 특성에 맞게 평가모형을 추가 또는 제외하는 것이 적절할 것으로 판단된다. 또한 장래 수집자료의 확대와 이를 실제 평가에 적용하는 반복적인 과정을 통해 평가체계를 지속적으로 검증하고 보완하는 작업이 필요할 것이다.



# 05

---

## 교통사고 예방을 위한 정책제언

- 1\_결론 및 정책제언
- 2\_향후 발전방향

## 05 | 교통사고 예방을 위한 정책제언

### 1\_결론 및 정책제언

본 연구에서는 교통공학적 관점의 전통적인 교통사고 분석에서 벗어나 교통체계-사회환경-인구활동을 연계한 교통사고 위험도 평가방법을 제시하고, 이러한 평가에 활용할 수 있는 서울시 기초자료를 검토하였다.

평가모형은 교통사고 발생건수를 기초자료로 하되 도로안전편람(US DOT, 2014)에서 제시한 바와 같이 사고유형을 사상사고(FI)와 물피사고(PDO)로 구분하여 심각도를 일부 반영하였다. 교통사고 영향요인은 교통체계, 사회환경, 인구활동을 반영할 수 있는 변수를 선정하되, 도로안전시설물은 교통사고 발생과의 인과관계가 모호하여 분석에서 제외하였다. 서울시에서 구득 가능한 기초자료를 검토한 결과, 대부분의 인구구조, 토지이용에 관한 사회환경 관련 자료들은 공간적 수직단위가 행정동인 것으로 나타났다. 교통체계에 관한 자료는 GIS로 구축되어 있는 경우가 많아 필요시 공간적으로 집계하거나 재분석하는 것이 비교적 용이한 것으로 조사되었다.

또한 현재 수집 가능한 자료를 바탕으로 교통사고 위험도를 평가하고 개선방향을 도출하는 일련의 과정을 설정하고, 분석결과가 교통안전 개선대책 수립에 시사하는 바를 제시하였다. 교통사고 발생요인이 교통사고 발생에 미치는 영향은 행정동별로 다르게 나타났다. 특히 노상주차면수나 교차로수와 같이 교통안전 개선대책에 직접적으로 활용할 수 있는 변수의 경우에도 행정동에 따라서는 상반된 영향을 미치는 것으로 분석되었다. 이는 유사한 교통사고 발생건수에도 불구하고 행정동별로 적용되어야 하는 개선대책은 전혀 반대 방향에서 접근되어야 할 필요가 있음을 시사하고 있다.

또한 교통체계-사회환경-인구활동에 관한 각 변수는 교통사고 발생에 독립적으로 영향을 미치지 않고 상호 조건부의 관계를 맺고 있는 점도 밝혀졌다. 예를 들어 노상주차면수는 횡단보도수가 충분한 경우에는 교통사고 위험도에 별다른 영향을 미치지 않지만, 횡단보도수가 적은 곳에서는 교통사고 위험도의 판별기준으로 사용되었으며 노상주차면수가 임

계수준보다 높을 경우 교통사고 위험도도 증가하는 것으로 분석되었다. 이는 교통안전 개선대책을 수립할 때 주변 교통체계와 토지이용 등을 복합적으로 고려해야 함을 의미한다.

교통사고 심각도별로 영향을 미치는 요인도 다르게 나타났다. 사상사고(Fatal and Injury, FI)의 경우에는 도로시설의 특징과 고령자수에 의해 주로 영향을 받는 것으로 분석된 반면, 물피사고(Property Damage Only, PDO)의 경우에는 노변에서의 인구활동에 의해 주로 영향을 받는 것으로 나타났다.

또한 교통안전 개선대책이 수립되었을 때 장래 개선효과를 전망하기 위한 분석기법을 제시하였으며, 각 행정동별로 독립변수를 임의로 증가시켰을 때의 변화 양상에 대한 결과를 제시하여 모형의 적용성을 검토하였다.

## 2\_향후 발전방향

본 연구에서는 서울시 교통사고 위험도를 평가할 수 있는 기법을 검토하고 현재 수집가능한 기초자료를 활용하여 각 모형의 적용성과 도출 가능한 정책적 시사점을 검토하였다. 향후 자료에 기반한 교통사고 위험도 평가체계를 고도화하기 위해서는 중장기적으로 기초자료 수집과 평가모형 구축 측면에서의 개선이 필요할 것으로 생각된다.

첫째, 다양하고 구체적인 교통사고 자료의 수집이 필요하다. 본 연구에서 활용한 교통사고 자료는 교통사고분석시스템(TAAS)에서 제공하는 경찰청 데이터베이스 자료이다. 전체 교통사고 중 경찰에 접수된 사고가 약 20%, 자동차 손해보험사가 약 75%, 공제조합이 5%의 비중(2014년 교통사고 통계분석, 도로교통공단, 2015)을 차지하는 것을 고려할 때, 상당수의 소규모 보행자사고, 특히 이면도로에서의 보행자사고는 본 연구에서 활용한 자료에 포함되어 있지 않을 가능성이 크다. 따라서 단기적으로는 각 보험사에서 구축하고 있는 교통사고 데이터베이스를 수집하여 분석에 활용한다면 보다 폭넓은 교통사고 발생유형을 검토하여 신뢰도 있는 분석결과를 도출할 수 있을 것이다.

둘째, 잠재적인 위험도 평가를 위한 첨단 교통정보의 활용이 필요하다. 교통사고 자료에

기반한 교통사고 위험도 분석은 위험했지만 사고는 발생하지 않은 이른바 ‘아차사고’ (Near Accident) 등 잠재적인 위험도를 반영할 수 없는 근본적인 한계를 갖는다. 따라서 장기적으로는 교통안전 분야에서 널리 활용하는 잠재적 위험도 측정방법인 교통안전대체 지표(Surrogate Safety Measure)에 대한 고려가 필요하다. 이러한 지표는 대부분 개별 차량의 위치나 궤적 등을 활용하여 수집하기 어려운 단점이 있지만 통신사카드사 등 민간 자료의 결합, 상용차량 관리시스템의 활용 등을 통해 일부 획득이 가능하며, 장기적으로 C-ITS, V2X 등이 보편화되면 첨단 교통정보의 활용이 확대될 것으로 전망된다. 유경상 (2017)이 제시한 현재 생성되고 있는 교통 관련 빅데이터 원자료는 다음 [표 5-1]과 같다.



[표 5-1] 교통 관련 빅데이터 원자료 주요 내용

구분	자료명	주요 포함 정보	보유기관
1	교통카드 기록	교통카드 태그 일시, 승하차 정류장, 태그 단말기 정보, 적용요금 구분(어린이, 일반 등), 추가 요금 정보	교통카드사
2	포털지도 경로 검색 기록	출-도착지 검색기록, 수단별 경로 정보 등	포털지도 운영업체
3	앱택시 요청 및 연결 자료	승객 위치, 요청 목적지, 요청 시각, 배차차량 정보, 콜 수락 시각, 승하차 시각, 결제 요금	애플리케이션 운영업체
4	대리운전 요청 및 연결 자료	차량 출도착지, 배차 기사, 이용시간, 요금 및 결제 방식	대리운전업체
5	카드매출 기록	카드 사용 일시, 카드 사용자 정보, 가맹점 정보, 거래 금액, 결제 유형(할부 여부 등)	카드사
6	통신사 기지국 접속 기록	이용자 정보, 단말기 정보, 기지국 정보, 기지국 연결 일시, 연결 시간	이동통신사
7	버스운영시스템(BMS) 데이터	버스 및 업체 관련 정보, 정류장교차로별 차량 도착 정보	TOPIS
8	택시 운행 데이터	택시 등록정보, 운전자 정보, 시각별 차량 위치, 속도, 진행방향, 공차 여부, 감·가속 정보 등	한국 스마트카드사
9	상용차 운행 데이터	차량운전자 정보, 운송사업자 등록번호, 시각별 차량 위치, 속도, 진행방향, 브레이크 신호, 가속도 등	교통안전공단
10	내비게이션 애플리케이션 기록	이용자 정보, 출-도착지 검색기록, 시각별 차량 위치, 이동 속도	내비게이션 운영업체
11	공영주차장 출입 기록	주차장명 및 위치, 차량번호, 입·출차 시각, 총 주차시간, 정산요금, 정기권 여부	주차장 운영 주체
12	유료도로 요금소 이용 기록	출발·도착 일시 및 요금소, 차량 정보(번호판), 차종 구분, 요금 정보	한국도로공사 민간운영업체
13	검지기 자료	검지기 위치, 검지기 도달·통과 시점, 검지기 점유시간, 통행속도	경찰청 국토교통부
14	교통정보 CCTV 자료	카메라 설치 위치, 녹화일시	서울시설공단, 경찰청 등
15	교통사고 정보	사고 발생일시 및 장소, 사고내용, 피해자가해자 정보, 상해 정도	도로교통공단
16	불법 주정차 단속 정보	소속 자치단체, 차량번호, 단속 차종, 단속 구역, 단속 위치, 위반 일시, 단속 구분	지방자치단체
17	속도위반 단속 기록	차량번호, 단속 일시, 단속지점, 단속 시 통과속도, 운전자 정보 등	경찰청
18	택배 물품 추적 기록	발신자, 수신자, 물품 정보, 접수·집화·도착시간, 수송차량 및 담당자 정보	택배업체

자료: 빅데이터와 교통정책의 연계 방향 (서울연구원, 2017)

셋째, 자료 기반의 의사결정체계 구축을 위한 지속적인 노력이 요구된다. 서울시에서는 교통안전 개선을 위해 다양한 제도를 운영하며 예산을 투입하고 있다. 각각의 제도가 대상으로 하는 목표나 개선 시설물이 다르고 지점별 여건의 차이가 커 개선대책 수립은 대부분 전문가의 판단에 의존하는 경향이 크다. 따라서 제도별 목적과 대상에 적합한 평가 모형을 지속적으로 발굴하고 현장 데이터에 적용하여 검증하는 노력을 통해 현재 운영되는 개선사업별로 특화된 분석체계를 수립해가는 것이 필요하다. 또한 모형별로 갖는 한계점과 시사점이 상이하므로, 본 연구에서 제시한 바와 같이 전통적인 분석기법 외에도 다양한 분석기법을 교차 적용하여 분석결과를 비교하는 과정이 필요할 것이다.

넷째, 교통사고 분석체계 운영을 위한 데이터 플랫폼 구축이 시급하다. 현재 서울시에서 수집되는 자료 외에 민간에서 수집되는 자료, 또는 자료로 활용되기 위해서 추가 가공이 필요한 자료를 포함한 데이터 플랫폼이 구축되어야 한다. 특히 공공에서 직접 수집하는 정보를 다양화하는 방안, 기존에 가공하지 않고 버려지던 정보를 고도화하는 방안, 민간 분야의 플랫폼 참여를 유도할 수 있는 협력 방안 등이 검토되어야 한다. 이를 통해 확장된 기초자료는 다시 평가모형을 고도화하는 데에 기여할 것이다.

**[표 5-2] 단계별 데이터 플랫폼 구축 방안**

구분	참여대상	관련 데이터	비고
1단계	서울연구원	- 과거 사고자료, 속도 - 기상, 도로기하구조, 토지이용 등	본 연구 수행
2단계	정부/지자체	- 실시간 사고자료, 교통량/속도, 교통카드데이터, - CCTV 영상분석, DTG 자료분석 등	플랫폼 구축
3단계	민간	- 내비게이션, 통신사, 카드사 자료 - 사고영상 블랙박스 등	민간 참여

## 참고문헌

- 강민지. (2018) “DSRC 데이터를 활용한 K-NN 알고리즘 기반 교통사고 위험예측 기법 연구”, 계명대학교대학원 학위논문
- 김도훈, 이동민, 성낙문. (2010) “지방부 3지 교차로에서의 사고예측모형 개발”, 교통연구
- 박나영, 김태양, 박병호. (2017) “국내 교통사고 밀도 모형 개발”, 한국안전학회지
- 박준태, 이수범, 이동민. (2011) “도시부 도로구간 사고예측모형 개발”, 교통연구
- 박준태, 장일준, 손의영, 이수범. (2017) “토지이용 및 교통특성을 반영한 교통사고 예측모형 개발 연구”, 대한교통학회지
- 백승걸, 장현호, 강정규. (2004) “교통량과 통행거리를 고려한 고속도로 교통사고예측모형 개발”, 대한토목학회 학술대회
- 신강원, 심상우, 최기주, 김수희. (2014) “KNN 알고리즘을 활용한 고속도로 통행시간 예측”, 대한토목학회논문집
- 이승봉, 한동희, 이영인. (2015) “사고등급별 고속도로 교통사고 처리시간 예측모형 개발”, 대한교통학회지
- 정효진, 이지영. (2015). “지리가중회귀분석을 이용한 은평뉴타운 지가 분석”, 한국공간정보학회지
- 최재역. (2016) “고속도로 통행정보를 활용한 KNN 기반 교통정체 예측 기법”, 서강대학교 학위논문
- 한상진, 김근정, 오순미. (2008) “전통적 사고예측모형의 한계 및 개선방안 : Hauer 사고예측모형의 소개 및 적용”, 한국도로학회지
- 홍지연. (2013). “공간계량분석을 통한 도시교통사고 예측모형”, 서울시립대학교 박사학위논문
- 홍지연, 이수범, 김정현. (2015). “소규모 지역단위 교통사고예측모형 개발 - 서울시 행정동을 대상으로”, 대한토목학회논문집
- Honglei Ren, You Song, JingXin Liu, Yucheng Hu and Jinzhi Lei. (2017) “A Deep Learning Approach to the Prediction of Short-term Traffic Accident Risk” (arXiv:1710.09543v1)
- 경기연구원. (2017). “경기도 어린이공원 주변 교통안전 증진방안”
- 경기연구원. (2016). “미국의 안전한 교통체계 사례 연구”
- 한국교통연구원. (2017). “보험 교통사고 정보공유 시스템 개발 및 활용방안”

한국교통연구원. (2015). “보행교통사고 감소를 위한 종합적 법체계 정비방안”

한국교통연구원. (2015). “사업용 화물자동차의 선제적 교통사고 예방체계 강화 방안 연구”

<http://bus.go.kr/>(서울대중교통)

<http://data.seoul.go.kr/>(서울 열린데이터 광장)

<http://stat.seoul.go.kr/>(서울통계)

<http://taas.koroad.or.kr/>(도로교통공단 교통사고분석시스템)

<http://www.data.go.kr/>(공공데이터 포털)

<http://www.seoul.go.kr/>(서울특별시)

<https://www.si.re.kr/>(서울연구원 인포그래픽스)





## Abstract

---

### A Study on Risk Assessment Technique for Prevention of Road Traffic Accident

Hong, Sang Yeon · Chung, Jae Hoon

Current traffic accident improvement projects have been conducted in a way that suggest improvement measures, such as traffic safety facilities installation and signal operation adjustment, for the traffic accidents black spots. However, the suggested measures are to be fundamentally retroactive since the inputs are the traffic accident data that has already occurred.

In order to determine the risk for potential traffic accidents, it is essential to understand traffic accidents within the context of social environment. Traffic accidents are not only influenced by the traffic system, but also by the complex relationship between land use and activities of people on them. Even if there is not a lot of actual traffic accidents, it is necessary to prevent them by taking proactive measures against accidents where they are likely to take place, taking into consideration the traffic conditions and social environment as well.

In this research, we present a traffic accident risk assessment model which takes into account the traffic system, social environment and human activity together.

As a result of the research, the effect of the traffic accident factors on traffic accidents was different for each administrative region. It was also found to be influenced by the traffic systems, social environment and human activity.

## Contents

---

### 01 Introduction

- 1\_Background and Purpose
- 2\_Main Contents and Method of Research

### 02 Traffic Accident Analysis for the City of Seoul

- 1\_Analysis of Traffic Accident Characteristics
- 2\_Review of the Existing Traffic Safety Improvement Project

### 03 Examination of the Basis for Risk Assessment

- 1\_Traffic Accident Risk Assessment Models
- 2\_Traffic Accidents Influential Factors
- 3\_Data Analysis for the City of Seoul
- 4\_Implications for the Research

### 04 Risk Assessment Modelling

- 1\_Model Selection
- 2\_Data Collection
- 3\_Case Study
- 4\_Discussion

### 05 Conclusion

---

서울연 2018-BR-08

도로교통사고 예방 위한  
위험도 평가기법 모색

발행인 \_ 서왕진

발행일 \_ 2018년 8월 31일

발행처 \_ 서울연구원

ISBN 979-11-5700-355-6 93330 6,000원

06756 서울특별시 서초구 남부순환로 340길 57

---

본 출판물의 판권은 서울연구원에 속합니다.