

■ 論 文 ■

# 실시간 교통자료 기반 고속도로 교통사고 발생 가능성 추정 모형

Estimation of Freeway Accident Likelihood using Real-time Traffic Data

**박 준 형**

(한양대학교 교통시스템공학과 석사 과정)

**오 철**

(한양대학교 교통시스템공학과 교수)

**남 궁 성**

(한국도로공사 도로교통기술원 수석연구원)

## 목 차

- |  |  |
|--|--|
| <p>I. 서론</p> <p>1. 연구의 배경 및 목적</p> <p>2. 연구의 범위 및 방법</p> <p>II. 자료구축</p> <p>1. 교통사고자료</p> <p>2. 검지기 자료처리</p> | <p>III. 방법론 및 모형개발</p> <p>1. 이항로지스틱 회귀분석</p> <p>2. 모형개발 결과</p> <p>IV. 모형의 활용방안 및 기술적 이슈</p> <p>V. 결론</p> <p>참고문헌</p> |
|--|--|

Key Words : 교통안전, 이항로지스틱회귀분석, 경고정보, 루프 검지기, 사고발생 가능성

Traffic safety, Binary Logistic Regression, Warning information, Loop Detector, Accident Likelihood

## 요 약

본 연구에서는 실시간으로 수집되는 고속도로의 검지기 자료를 이용하여 교통사고 발생 가능성을 확률적 관측값으로 나타낼 수 있는 모형을 개발하였다. 사고발생 지점을 기준으로 상류부 및 하류부에서 수집된 사고발생 이전의 교통자료를 모형의 독립변수로 설정하였다. 이항 로지스틱 회귀분석 기법을 적용하여 교통사고 발생을 유발할 잠재력이 높은 교통상황을 교통사고와는 무관한 교통상황으로부터 추출하는 분류문제(classification problem)로 설정하고 모형을 개발하였다. 최근 3년간 서해안 고속도로에서 발생한 사고자료와 검지기 자료를 맵핑하였으며, 유효한 검지기 자료를 모형에 적용하기 위하여 이상치 제거 및 결측치 보정을 위한 자료처리 과정을 별도로 수행하였다. 본 연구에서 개발한 모형에서 산출되는 계량화된 교통사고 발생가능성은 고속도로상에서 실시간 경고정보 제공 및 다양한 교통운영관리 전략의 교통안전 측면에서의 효과를 평가하는데 유용하게 적용될 수 있을 것으로 기대된다.

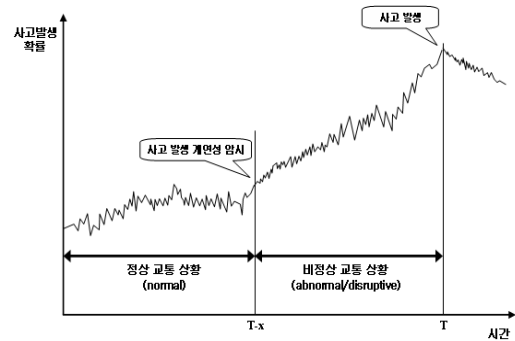
This study proposed a model to estimate traffic accident likelihood using real-time traffic data obtained from freeway traffic surveillance systems. Traffic variables representing spatio-temporal variations of traffic conditions were utilized as independent variables in the proposed models. Binary logistics regression modelings were conducted to correlate traffic variables and accident data that were collected from the Seohaean freeway during recent three years, from 2004 to 2006. To apply more reliable traffic variables, outlier filtering and data imputation were also performed. The outcomes of the model that are actually probabilistic measures of accident occurrence would be effectively utilized not only in designing warning information systems but also in evaluating the effectiveness of various traffic operations strategies in terms of traffic safety.

## 1. 서론

### 1. 연구의 배경 및 목적

인간의 생명과 직결되는 교통사고 예방 및 심각도 감소를 위한 교통안전 관련 연구는 교통공학의 여러 가지 연구 분야 중에서도 가장 중요한 분야라고 할 수 있다. 교통사고 발생 및 원인분석을 위한 대부분의 기존연구에서는 연평균일교통량과 같은 장기간 집계된 교통변수와 도로의 기하구조 등의 주변 환경요인과의 관계를 통계모형으로 설명하려는 노력을 기울여 왔다. 고속도로 인터체인지 연결로에서의 교통사고를 예측하는 모형을 개발한 박효신 등(2007)의 최근 연구가 가장 대표적인 사례라고 할 수 있다. 아울러 도로기하구조 설계요소가 교통사고에 미치는 영향을 분석하기 위한 모형(Zegeer et al., 1988; Vogot and Bared, 1998; Council, 1998; Anderson et al., 1999)에서도 집계된 교통량변수가 독립변수로 널리 사용되어왔다. 한편, 교통상황이 사고에 미치는 영향을 분석한 기존 연구(Gwynn, 1967; Oh et al., 1999)에서는 교통상황을 나타내기 위한 변수로 연평균 시간교통량을 적용하기도 하였다. 이러한 기존연구에서는 AADT 또는 hourly volume 등 장기간에 걸쳐 수집·저장된 교통자료를 활용하였다.

고속도로교통관리시스템(FTMS)와 같은 첨단교통체계 사업의 적극적인 구축 및 운영은 교통연구자에게 실시간 교통자료의 취득을 용이하게 하였으며, 1990년대 후반 이후 이를 활용한 교통안전분야의 새로운 연구를 가능하게 하였다. 실시간 차량검지기 자료를 이용하여 교통사고자료와의 관계분석을 통해 교통사고 발생 개연성(accident likelihood)을 추정하는 모형을 개발하고 이를 활용하는 분야가 바로 그것이다. 이 분야의 연구는 교통사고 발생 이전에 교통사고 발생을 암시하는 비정상적인 교통상황이 전개될 것이고 이러한 비정상적 교통상황이 검지기를 통해 수집되는 교통변수의 분석을 통해 추출 가능할 것이라는 가정을 갖고 출발한다. (<그림 1> 참조). 실시간 교통자료를 교통사고 모형에 응용하여 체계적인 분석을 수행한 연구로 Oh et al. (2001)은 사고발생 이전 5분간의 속도자료의 표준편차가 교통사고 발생을 유발할 수 있는 비정상 교통상황(disruptive traffic conditions leading traffic accident)임을 캘리포니아 I-880 고속도로 상의 교통사고 및 루프검지기 자료의 분



〈그림 1〉 교통 상황 관측을 통한 교통사고 예측 개념도

석을 통해 밝혀낸 바 있다. 교통사고 발생지점에서 가장 인접한 루프검지기에서 수집되는 자료를 이용하였다. Lee et al. (2002, 2003)은 캐나다 Ontario의 도시고속도로를 대상으로 13개월 동안 발생한 교통사고 및 루프검지기 자료를 이용하여 교통사고발생가능성을 예측하는 모형을 개발하였으며, 사고발생지점 상류부 및 하류부 검지기에서 수집되는 속도의 편차를 이용하였다. Golob et al. (2001)은 Southern California의 사고발생시간 이전의 30-sec 루프검지기 자료와 날씨 및 lighting condition이 교통사고 발생에 미치는 영향을 분석하였다. Abdel-Aty et al. (2004)은 플로리다 Orlando에서 수집한 자료를 분석하여 교통사고 발생지점 상류부에서 수집된 사고발생 이전 5~10분 사이의 5분간의 평균 점유율과 하류부 검지기의 5분 동안의 속도의 변동계수가 교통사고 발생을 예측할 수 있음을 보고하였다. 또한 Adel-Aty and Abdalla (2004)는 실시간 교통자료와 도로 기하구조, 교통사고와의 관계를 1999년 발생한 337건의 주건 교통사고를 이용하여 분석하고, 15분 동안의 속도의 변동이 큰 경우 해당 검지기 0.5마일 하류부에서 교통사고 발생가능성이 높게 나타나는 것으로 보고하였다. 위에서 고찰한 본 연구와 관련된 최근의 연구성과를 종합해 보면 실시간 교통변수의 시공간적 변동(Spatio-temporal Variation of real-time traffic parameters)의 효과적인 관측 및 분석을 통해 교통사고를 유발할 수 있는 잠재력이 높은 교통상황을 사전에 어느 정도 검지할 수 있다는 것이다.

본 연구는 기존의 돌발상황관리시스템과 달리 교통사고 발생 이전의 교통상황 분석을 통해 사고 발생가능성을 예측하여 사고위험성이 높은 교통상황을 판단하는 모형을 개발한 측면에서 큰 의의가 있다.

## 2. 연구의 범위 및 방법

본 연구에서는 서해안 고속도로 서울-목포 구간 339.51km에서 2004~2006년 최근 3년 동안 발생한 교통사고와 현장에서 수집된 검지기 자료를 이용하여 교통사고 발생 잠재력을 계량화할 수 있는 모형을 개발하였다. 이를 위해 연구문제(research problem)를 교통사고와 개연성이 없는 정상 교통상황(normal traffic conditions)과 교통사고를 유발할 수 있는 비정상 교통상황(abnormal traffic conditions)을 구분하는 분류문제(classification problem)로 설정하고 이항로지스틱 회귀분석(binary logistic regression analysis) 기법을 적용하였다. 3년간 발생한 전체 교통사고에 대한 분석뿐만 아니라 다양한 분석시나리오를 구상하여 모형 구축 작업을 수행하였다.

본 연구의 구성은 다음과 같다. 2장에서는 모형 개발용 dataset 구축을 위한 자료처리과정을 기술하였으며, 3장에서는 모형구축을 위해 적용된 방법론인 이항로지스틱 회귀분석에 대한 소개와 모형개발 결과를 정리하였다. 4장에서는 모형의 활용방안 및 향후 현장적용을 위한 기술적 이슈들과 함께 모형의 신뢰도를 높이기 위한

방안을 논의하였다. 마지막 장에서는 본 연구의 공헌 및 한계점, 향후 연구를 위한 방향을 제시하였다.

## II. 자료구축

본 연구에서는 교통사고와 무관한(Accident-free traffic conditions) 정상적인 교통상황(Normal)과 교통사고 발생을 유도할 수 있는 비정상(abnormal) 교통상황으로 구분하였다. 교통사고 발생가능성을 측정하는 모형을 위의 두 가지 교통상황을 분류하는 문제(Classification Problem)로 설정하였다. 이를 위해 교통사고 자료와 이에 상응하는 검지기 자료를 이용하여 모형개발을 위한 data set을 구축하였다.

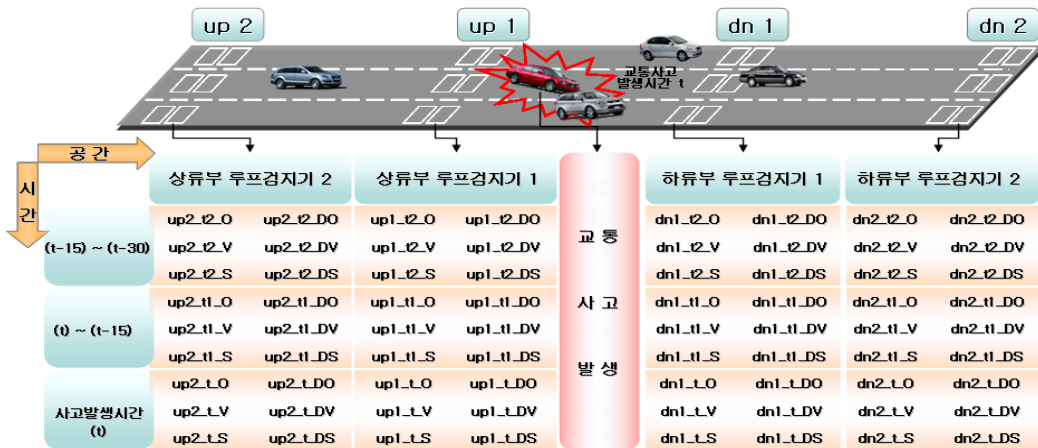
### 1. 교통사고 자료

2004년에서 2006년까지의 최근 3년간 서해안 고속도로에서 발생한 교통사고자료에서 사고일자, 사고시간, 사고지점의 이점, 방향, 사고유형 등의 정보를 추출하였다. <표 1>에 본 연구에서 활용한 구간의 교통사고 자료를 제시하였다.

<표 1> 사고 자료 현황

(단위 : 건)

연도	교통사고 건수	사고유형		주·야간		서울방향		목포방향	
		추돌 및 충돌	기타	주간	야간	주간	야간	주간	야간
2004	348	69	279	224	124	122	72	102	52
2005	311	44	267	200	111	113	61	87	50
2006	280	47	233	199	81	106	37	93	44
합계	939	160	779	623	316	341	170	282	146



<그림 2> 교통사고 관련 루프검지기 자료 처리 개념도

## 2. 검지기 자료 처리

2004년, 2005년, 2006년도의 3년간 서해안고속도로 루프검지기 자료와 교통사고자료를 통합하여 분석을 수행하였다. <그림 2>에서와 같이 사고 발생 지점 인근의 상류부 및 하류부의 2개 검지기로부터 수집된 사고발생시간(t)으로부터 15분전(t-15)까지의 교통상황과 15분전(t-15)으로부터 30분전(t-30)까지의 교통상황을 대표하는 교통량, 평균 속도 및 점유율을 본 연구에서는 이용하였다. 이는 교통사고 발생에 기여할 것으로 판단되는 교통상황을 추출하기 위한 것이다.

<그림 2>와 같이 사고발생지점을 기준으로 “up”은 upstream을 뜻하며 “dn”은 downstream을 뜻한다. t1, t2는 사고 발생시각의 15분전, 30분전을 의미하며, “O”, “V”, “S”는 각각 Occupancy, Volume과 Speed를 의미한다. O, V, S 앞에 붙어있는 “D”는 하루 동안 수집된 검지기 자료의 평균값과 각 시간별 이정별 검지기 자료의 차이값이며 Difference값으로 정의하였다. Difference값도 하루 평균값과의 편차 정도를 변수화하여 교통사고와의 관계를 분석하였다. <그림 3>에 자료처리 절차를 도식화 하였다.

### [1 단계] 노선별 방향별 검지기 자료 추출

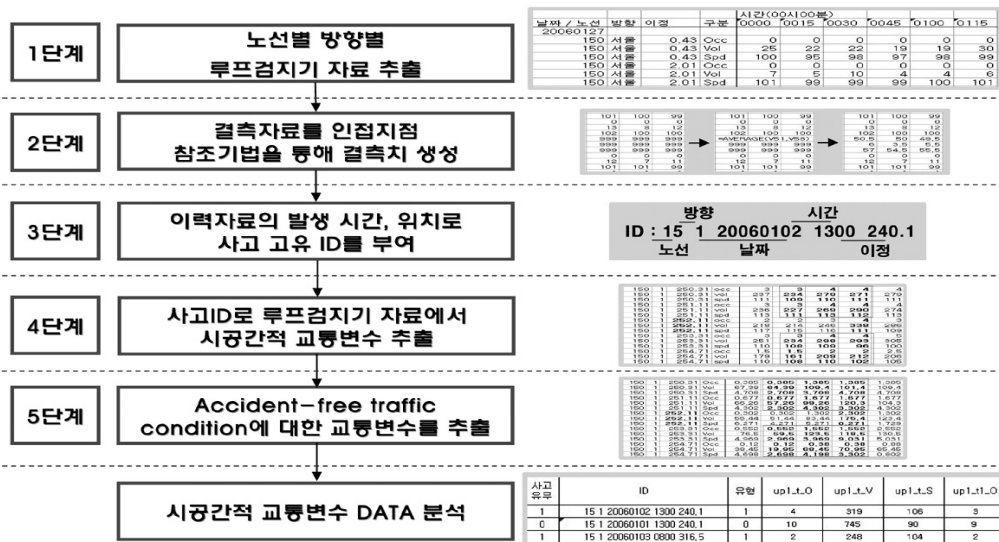
검지기에서 생성되는 원본데이터에서 노선별, 방향별로 자료를 추출하여 검지기의 이정 순으로 나열하고 각

검지기 마다 시간 순으로 교통량, 점유율, 속도 자료를 배열하였다.

### [2 단계] 루프검지기 자료 결측치 보정

사고와 관련된 변수로 추출하는 범위 내에 “-999”로 기록되는 결측자료에 대해 인접지점 참조 기법을 이용한 결측치 생성을 수행하였다. 결측자료를 보정하지 않으면 Difference값이 오류값을 가지게 되고 모형의 정확도가 떨어진다. 인접참조기법의 방법에 대해서는 <그림 3>에 도식화 되어있다. 결측자료가 있을 경우 결측자료를 기준으로 상류부와 하류부의 검지기 자료를 평균하여 입력하고 상류부와 하류부에도 결측자료가 있을 경우 15분 전 15분 후의 검지기 자료에서 평균을 산출하여 입력하였다. 검지기 자체의 고장으로 예상되어 나타난 많은 결측자료가 발생한 부분에 대해서는 보다 정확한 average 값과 difference값을 구하기 위해 계산 시 결측자료를 제외하였다. 그리고 추출 자료가 하루 중 측정시작시간(00시 00분)이거나 측정종료시간(24시00분)일 경우, 시종점에 위치한 검지기 자료일 경우, 인접지점의 자료가 없으므로 결측자료를 제외하고 검지기 자료의 일일평균값과 Difference값을 계산하였다.

결측자료 보정 시 occ와 vol, speed는 각 data의 단위와 성질이 다르므로 각각 같은 종류의 Data끼리만 참조하였다. 결측자료는 주로 하나의 시간대나 하나의 검지기 전체에 분포된 경우가 많은데 결측자료의 분포에



<그림 3> 자료 정리 절차

따라 추출하는 분석자료의 신뢰도가 달라진다. 결국자료에 따른 분석자료의 신뢰도를 평가하기 위해 보정시 “-999”값을 보정한 정도에 따라 사고를 0, 1, 2로 나누어 코딩하였다. 세 가지 Case로 구분한 기준은 다음과 같다.

- 유형 “0” : 자료의 추출 영역에 결측 자료가 없는 경우
- 유형 “1” : 자료의 추출영역의 결측 자료를 인접 참조기법을 통해 보정을 한 경우
- 유형 “2” : 자료의 추출영역에 결측 자료가 많아 모두 보정시키지 못한 경우

**【3 단계】 개별 사고ID 부여**

교통사고 자료를 이용하여 발생한 각 교통사고에 노선, 방향, 날짜, 시간, 이점 등의 자료로 구성된 개별 ID를 부여하였다.

ID의 방향 정보를 나타내는 것으로 서울방향은 “1”로 하며 반대방향은 “0”으로 코딩하였다. 시간은 오전과 오후의 혼동을 없애기 위해 일 24시간 기준의 시간을 적용하였다.

**【4 단계】 사고자료 기반의 교통변수 추출**

교통사고 자료의 사고 ID를 확인하고 서해안고속도로 루프검지기 자료로부터 시공간적 교통변수를 추출하였다. 사고발생지점의 가장 인접한 루프검지기에서 상류부 방향의 2개 검지기, 하류부 방향의 2개 검지기에서 수집된 사고발생시점부터 15분전까지의 자료와 사고발생 15분전부터 30분전까지의 교통량, 속도, 점유율 자료를 추출하였다. 추출된 총 교통변수는 72개로서 각각의 검지기에서 수집된 실제값과 일일평균에 대한 각 교통변수의 difference 값으로 구성되었다. difference값은 각각의 루프검지기의 24시간 자료의 평균값을 각 시간별 자료에서 빼준 값의 절대값이다.

**【5 단계】 정상교통상황(Normal traffic Condition) Data 구축**

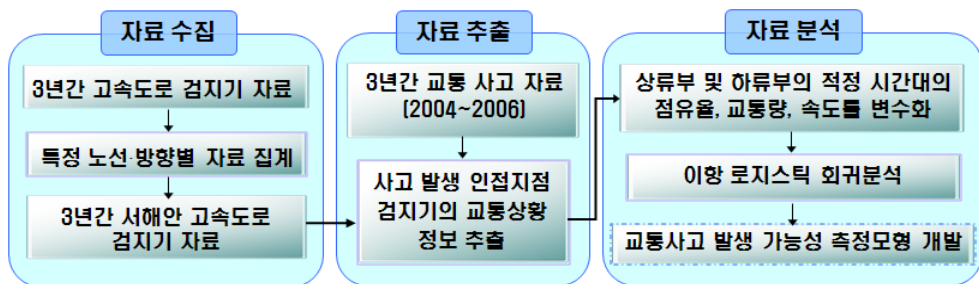
accident-free traffic condition에 대한 교통변수를 추출하였다. 교통사고 발생 유무를 비교하기 위해 교통사고 발생지점에서 발생시간 하루 전날의 동일시간대를 대상으로 72개의 교통변수를 추출하였다. 정상교통상황 선정 시 사고 전날에도 사고가 발생했을 경우 사고일로부터 2일전의 자료에서 추출하였다. 사고발생 자료를 “1”로 표기하고 사고발생 전날의 자료를 “0”으로 표기하였다.

**III. 방법론 및 모형개발**

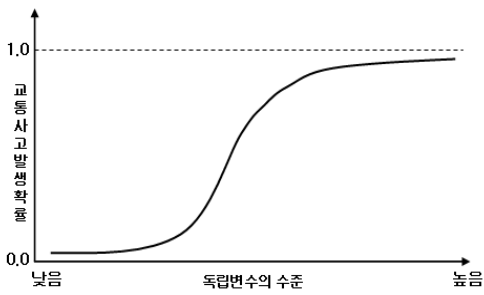
검지기에서 생성된 자료로부터 분석을 통해 사고위험도를 산출하는 과정까지의 자료가공과정이 <그림 4>에 도출되어 있다. MATLAB을 이용하여 전국 고속도로 검지기에서 생성되는 2004~2006년의 3년간 검지기 자료 중 서해안 고속도로 검지기에서 생성되는 교통량, 점유율, 속도 방향별 자료를 추출하는 프로그램을 작성하고 교통사고자료와 비교해 사고와 연관성이 있는 교통상황 자료를 추출하였다. 사고 지점에서 인접한 검지기의 자료를 추출하여 변수화 하고 각 교통사고들의 사고정보를 기준으로 모형 구축을 위한 자료를 그룹핑하였다. 기준으로는 주·야간, 차량진행방향(서울방향, 목포방향), 검지기 자료 내 결측자료에 의한 손실정도(0,1,2), 사고 유형(추돌 및 충돌, 기타)으로 하였다.

**1. 이항로지스틱 회귀분석**

로지스틱 회귀분석은 종속변수가 범주형으로 관측된 데이터를 적합 시킬 때 유용하게 사용되는 기법으로, 모형의 종속변수로 교통사고를 유발할 수 있는 교통상황을 1로,



<그림 4> 검지기 자료의 추출 및 분석 과정



〈그림 5〉 독립변수와 교통사고발생의 로지스틱 관계

교통사고와 무관한 교통상황을 0으로 설정하는 binary classification 문제에 효과적으로 응용 가능하다.

본 과업에서는 binary 로지스틱 회귀분석 기법을 루프 검지기 자료기반의 교통사고발생확률모형에 적용하였는데, 모형의 산출물이 0~1 사이의 probabilistic measure로서 교통사고 발생 가능성을 예측하였다. 종속변수 1은 교통사고 발생 확률이 1을 의미한다. 교통사고확률 값은 0과 1의 범위에서 다양한 값들을 가질 수 있으나 0과 1의 범위를 벗어날 수는 없다. 로지스틱 회귀분석은 독립변수와 0과 1사이의 종속변수의 관계가 〈그림 5〉에서와 같이 S자 형태를 나타내는 것으로 가정한다. 사고 발생가능성 추정 함수는 식 (1)이다.

$$\Pr(ACC_i = 1 | X_i) = \frac{\exp[f(X_i, \beta)]}{1 + \exp[f(X_i, \beta)]} \quad (1)$$

여기서,

$\Pr(ACC_i)$  : 주어진 교통변수  $i$ 가 교통사고를 발생시킬 확률

$X_i$  : 교통사고 발생에 영향을 미치는 교통

변수

$f(X_i, \beta)$  :  $X_i$ 와 파라미터  $\beta$ 로 구성된 함수

## 2. 모형개발 결과

본 연구에서는 〈표 2〉에서와 같이 11개 Case에 대한 모형구축 작업을 수행하였다. 각 Case들은 자료를 필터링한 기준에 따라 나뉘며 주·야간, 사고유형, 방향과 data의 결측 자료에 의한 손실정도에 따라 나뉘었다. R-Square값은 회귀분석을 실행한 후 산출되는 회귀식이 전체 사고 발생 유무와 변수들 간의 관계를 얼마나 설명하는가를 나타내는 것으로 값이 클수록 더 큰 설명력을 가진다 할 수 있다. 이항로지스틱 회귀분석기법을 사용하여 변수들의 유의성을 판단하여 유의하지 못한 변수들을 제거하고 반복하여 회귀분석을 실시하는 과정을 통해 사고가 발생하는 데 연관성이 있는 변수들을 산출하였다. 자료를 집계한 기준에 따라 11개의 Case를 이항로지스틱 회귀분석한 결과 〈표 2〉에서와 같이 회귀식의 R-square값은 case8(야간서울방향)의 R-square이 0.340으로 가장 높고 case4(추돌 및 충돌)이 0.301로 두 번째로 높았다. 이것은 사고 발생에 case4와 case8에 관련된 변수들이 다른 case들의 사고요인에 비해 영향력이 높은 것을 알 수 있다. case4는 추돌 및 충돌시의 교통량과 점유율, 속도자료이다. 추돌 및 충돌 시와 야간의 서울방향의 교통 환경이 교통사고와 연관성이 다른 case에 비해 크다. 이와는 대조적으로 Case3의 R-square값(0.000)을 통해 야간의 교통 환경은 교통사고와 유의한 관계를 보이지 않았으며 Case2와 Case3의 비교결과 R-square값이 0.049와 0.000으로 주야간 구분의 사고 연관성은 매우 떨어진다. 그러나 주·야

〈표 2〉 이항 로지스틱 회귀분석 R-Square

CASE	-2 Log 우도	Cox와 Snell의 R-제곱	Nagelkerke R-제곱	교통사고건수
case 1 : 사고 교통상황 및 사고전날 교통상황	2304.647	0.017	0.022	939 건
case 2 : 주간에 발생한 사고와 사고전날 교통상황	1505.103	0.037	0.049	623 건
case 3 : 야간에 발생한 사고와 사고전날 교통상황	786.013	0.000	0.000	316 건
case 4 : 추돌 및 충돌 사고와 사고전날 교통상황	264.329	0.226	0.301	142 건
case 5 : 기타사고유형의 사고와 사고전날 교통상황	1904.125	0.018	0.024	779 건
case 6 : 서울방향으로 주간의 사고·비사고 교통상황	792.243	0.053	0.070	341 건
case 7 : 목포방향으로 주간의 사고·비사고 교통상황	697.217	0.046	0.062	282 건
case 8 : 서울방향으로 야간의 사고·비사고 교통상황	153.471	0.254	0.340	113 건
case 9 : 목포방향으로 야간의 사고·비사고 교통상황	325.187	0.173	0.231	146 건
case10 : 결측자료가 없거나 보정이 된 경우	1692.932	0.026	0.035	666 건
case11 : 결측자료가 많아 모두 보정하지 못한 경우	587.525	0.035	0.047	273 건

〈표 3〉 이항 로지스틱 회귀분석 분류정확도

CASE	관측		예측값								
			사고여부		분류정확도 (%)	사고여부		분류정확도 (%)	사고여부		분류정확도 (%)
			0	1	0.5	0	1	0.7	0	1	0.8
case4	사고여부	0	89	25	78.070	108	6	94.737	112	2	98.246
		1	37	83	69.167	77	43	35.833	82	38	31.667
	전체 %				73.504			64.530			64.103
case8	사고여부	0	62	17	78.481	93	9	91.176	98	4	96.078
		1	25	38	60.317	66	40	37.736	82	24	22.642
	전체 %				70.423			63.942			58.654

〈표 4〉 이항로지스틱 회귀분석 결과 방정식에 포함된 변수

CASE	변수	B	S.E.	Wald	유의확률	Exp(B)
case 4 추돌 및 충돌	up1_t1_O	0.63607	0.29560	4.63014	0.03142	1.88905
	up1_t2_V	-0.00701	0.00336	4.35420	0.03692	0.99302
	up2_t1_O	-0.54779	0.24958	4.81708	0.02818	0.57823
	up2_t2_V	0.00671	0.00286	5.49887	0.01903	1.00673
	up1_t1_DO	-1.23492	0.48501	6.48305	0.01089	0.29086
	up1_t2_DO	1.11977	0.48792	5.26702	0.02173	3.06416
	up2_t1_DS	0.06987	0.03676	3.61161	0.05738	1.07236
	dn2_t1_DV	0.00816	0.00279	8.54909	0.00346	1.00820
	dn2_t1_DS	0.20489	0.05862	12.21552	0.00047	1.22739
	dn2_t2_DS	-0.12514	0.05715	4.79472	0.02855	0.88238
상수	-1.13773	0.29770	14.60587	0.00013	0.32054	
case 8 야간 서울방향	up1_t1_O	1.89637	0.83917	5.10678	0.02383	6.66169
	up1_t2_O	-3.14948	0.94933	11.00636	0.00091	0.04287
	up2_t2_O	1.33663	0.60218	4.92682	0.02644	3.80620
	dn1_t2_O	-1.41153	0.53297	7.01414	0.00809	0.24377
	dn1_t2_V	0.02321	0.00860	7.28865	0.00694	1.02349
	dn1_t2_S	-0.14279	0.05124	7.76556	0.00533	0.86694
	up1_t1_DO	-2.72418	1.28526	4.49253	0.03404	0.06560
	up1_t1_DV	-0.05901	0.02548	5.36282	0.02057	0.94270
	up1_t2_DO	2.73222	1.28032	4.55401	0.03284	15.36690
	up1_t2_DV	0.06750	0.02528	7.13085	0.00758	1.06983
	up2_t1_DV	0.09030	0.02822	10.24061	0.00137	1.09450
	up2_t2_DV	-0.08813	0.02751	10.26567	0.00136	0.91564
	dn1_t2_DV	-0.02210	0.01122	3.87991	0.04887	0.97814
	dn2_t1_DO	-1.57296	0.65590	5.75126	0.01648	0.20743
	dn2_t2_DV	0.02428	0.01337	3.30142	0.06922	1.02458
	상수	13.90168	5.13050	7.34200	0.00674	1089986

간 구분과 방향 구분을 같이 적용한 경우, 주간에 발생한 사고인 Case6과 Case7의 R-square값이 각 0.07, 0.06이고 야간에 발생한 사고인 Case8과 Case9의 R-square값은 각 0.34, 0.23의 값으로 야간시의 R-square값이 서울방향과 목포방향 모두 주간보다 더 높은 사고연관성을 보였다.

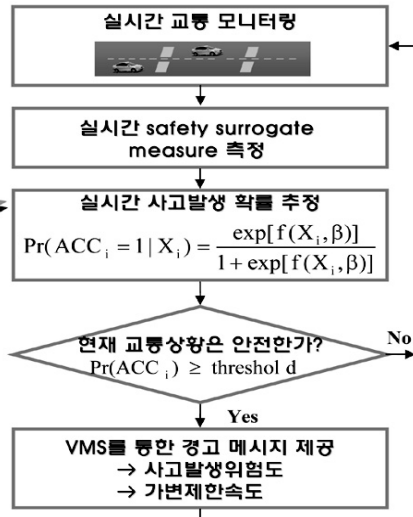
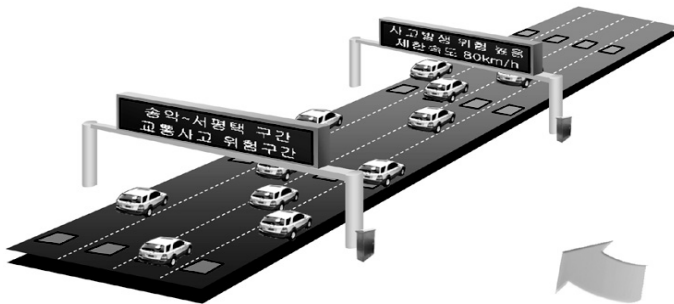
분류표에서는 회귀분석 시 분류분리점을 달리하여 분석하고 유의한 결과를 보인 case를 〈표 3〉에 정리하였다.

회귀분석 시 분류분리점을 0.5, 0.7, 0.8로 달리 설정하여 수행한 결과 〈표 3〉에서와 같이 분류분리점 값을 높게 설정할수록 분류정확도의 값이 낮아졌다. Case4의 경우 분류분리점을 0.5로 설정했을 때 73.5%의 분류정확도를 보였으나 분류분리점을 0.8로 설정했을 때는 64.1%의 분류정확도를 보였다. 2004~2006년의 최근 3년간의 교통사고지점의 검지기 자료를 이항로지스틱 회귀분석한 결과 〈표 4〉에 11개의 Case에 대한 유의한 변수들

공간	상류부 루프검지기 2 (up2)				상류부 루프검지기 1 (up1)				하류부 루프검지기 1 (dn1)				하류부 루프검지기 2 (dn2)			
(t2) (t-15)~(t-30)	up2_t2_O	up2_t2_DO	up1_t2_O	up1_t2_DO	dn1_t2_O	dn1_t2_DO	dn2_t2_O	dn2_t2_DO								
	up2_t2_V	up2_t2_DV	up1_t2_V	up1_t2_DV	dn1_t2_V	dn1_t2_DV	dn2_t2_V	dn2_t2_DV								
	up2_t2_S	up2_t2_DS	up1_t2_S	up1_t2_DS	dn1_t2_S	dn1_t2_DS	dn2_t2_S	dn2_t2_DS								
(t1) (t)~(t-15) (t1)	up2_t1_O	up2_t1_DO	up1_t1_O	up1_t1_DO	dn1_t1_O	dn1_t1_DO	dn2_t1_O	dn2_t1_DO								
	up2_t1_V	up2_t1_DV	up1_t1_V	up1_t1_DV	dn1_t1_V	dn1_t1_DV	dn2_t1_V	dn2_t1_DV								
	up2_t1_S	up2_t1_DS	up1_t1_S	up1_t1_DS	dn1_t1_S	dn1_t1_DS	dn2_t1_S	dn2_t1_DS								
사고발생 시간 (t)	up2_t_O	up2_t_DO	up1_t_O	up1_t_DO	dn1_t_O	dn1_t_DO	dn2_t_O	dn2_t_DO								
	up2_t_V	up2_t_DV	up1_t_V	up1_t_DV	dn1_t_V	dn1_t_DV	dn2_t_V	dn2_t_DV								
	up2_t_S	up2_t_DS	up1_t_S	up1_t_DS	dn1_t_S	dn1_t_DS	dn2_t_S	dn2_t_DS								

〈그림 6〉 추돌 및 충돌사고(Case4) 발생가능성 예측 변수

- 사고발생 개연성(Accident Likelihood)이 높은 교통상황을 모니터링 하여 교통정보로 가공하여 제공하는 Warning Information System
  - 가변제한 속도 및 사고발생 위험도



〈그림 7〉 실시간 교통안전 관리 시스템 개념도

의 산출 결과가 표기되어 있다. Case4(충돌 및 추돌)의 경우 Difference 관련변수 6개와 up1\_t1\_O, up1\_t2\_V, up2\_t1\_O, up2\_t2\_V의 변수들이 통계적으로 유의한 것으로 나타났다. 추돌 및 충돌 교통사고 발생 가능성을 나타낼 수 있는 변수들을 〈그림 6〉에 도식화하였다. Case8의 경우 15개의 변수들이 산출되었으며 Difference 관련변수를 포함한 up1\_t1\_O, up1\_t2\_O, up2\_t2\_O, dn2\_t2\_O, dn1\_t2\_V, dn1\_t2\_S 변수들이 선정되었다. Case4와 Case8을 회귀분석한 결과 중에서, up1\_t1\_O, up1\_t1\_DO, up1\_t2\_DO, up2\_t1\_DV로 4개의 변수가 공통적으로 산출되었다.

#### IV. 모형의 활용방안 및 기술적 이슈

본 연구에서 개발한 모형은 운전자에게 사고 발생 개

연성이 높은 교통상황에 대한 경고정보 제공을 위해 활용될 수 있을 것으로 판단된다. 아울러 가변제한속도와 같은 교통류 관리전략에도 응용될 수 있을 것이다. 예를 들어 〈그림 7〉에서와 같이 실시간으로 수집되는 교통상황을 교통정보센터에서 지속적으로 모니터링하면서 교통사고 발생 가능성을 산출하고 일정 수준 이상의 교통사고 발생 가능성이 관측되면 경고 정보와 함께 가변제한속도를 표출하는 가칭 실시간 교통안전 관리시스템을 고려해 볼 수 있다.

개념도로서 제시한 〈그림 7〉과 같은 시스템의 실제 현장 적용을 위해서는 여러 가지 기술적 이슈가 해결되어야 할 것이다. 본 연구에서 모형개발을 위해 사용한 15분 집계된 검지기 자료를 이용하여 추출된 교통상황 변수들은 급변하는 실시간 교통상황에 민감하게 반응하지 못할 수 있다. 교통환경의 변화에 더 민감하게 반응하기 위해 자료를 수집하는 time slice를 더 짧게 하고



time slice가 짧아짐에 따라 Difference값을 산출하는데 사용되는 평균값의 범위도 다양하게 적용하여 볼 수 있을 것으로 판단된다.

분류분리점의 조정으로 사고와 비사고의 발생 예측 확률이 달라지는 것을 <표 3>에서 볼 수 있는데, Case4의 경우 분류분리점을 0.5에서 0.8로 변화시킬 때 분류정확도가 73%에서 64%로 감소한다. 이것은 사고 위험도 예측 모형이 교통사고 발생 가능성이 있는 교통상황을 검지할 확률이 낮아지는 것을 의미한다. 분류분리점의 조정은 많은 교통사고 위험상황들 중에서 교통사고 발생을 실제로 유도하는 교통상황만을 검지하고 경고메세지를 표출하는 횟수와 내용을 조정하는 Parameter가 될 수 있다. 따라서 이러한 Parameter 및 제한속도의 동적인 조정에 대한 별도의 연구가 수행되어야 할 것이다.

## V. 결론

서해안 고속도로에 설치된 검지기로부터 생성되는 2004년에서 2006년까지 최근 3년간의 검지기 자료와 교통사고자료를 가공한 사고리스트 및 비사고 리스트와 비교하여 사고지점의 검지기를 기준으로 사고 영향권의 교통상황 자료를 추출하였다. 이 자료를 이항 로지스틱 회귀분석 기법에 적용하여 교통사고 발생 가능성을 확률적으로 계량화한 결과, 추돌 및 충돌사고의 경우와 야간 서울방향의 사고가 다른 case에 비해 교통사고와 유의한 연관성이 있는 것으로 나타났다. 추돌 및 충돌의 경우 0.301의 R-square를 보였으며 73.5%의 분류정확도를 보였다. 야간 서울방향의 경우 0.340의 R-square와 70.4%의 분류정확도를 보였다.

교통 상황이 변화함에 따라 운전자에게 도로의 사고 위험도 정보를 제공함으로써 운전자가 현재 교통상황의 위험성을 인식하고 감속 또는 방어운전에 집중하게 함으로써 교통사고의 발생 가능성을 줄일 수 있을 것으로 기대된다. 교통 상황의 사고위험도가 기준보다 클 경우 경고메세지를 표출하면서 도로의 제한속도를 변화시킨다면 사고위험도를 더 낮출 수 있는 방안이 될 수 있다. 그러나 가변제한속도를 적용시킬 때 일정하게 제한속도를 감소시키기 보다는 교통환경에 적합하게 감소시키는 시스템에 대한 연구가 필요하며, 가변제한속도 적용과 관련하여 감소시간, 감속폭 등과 관련된 연구가 추가적으로 수행되어야 할 것이다.

본 연구에서 구축한 모형이 실제 시스템으로 개발되어 현장에서 활용되기 위해서는 몇 가지 기술적 이슈들이 해결되어야 하며 이를 위해 향후 다음과 같은 연구가 추진되어야 할 것이다. 첫째로, 모형이 사고위험상황인지 아닌지를 구분하는데 기준이 되는 분류분리점이 기술적 이슈가 될 수 있다. 분류분리점의 설정에 따라 사고위험상황을 구분하는 분류정확도 및 사고위험상황이라 판단되는 시간간격과 빈도들이 달라지므로 향후 다양한 교통조건에서 분류분리점 설정에 따른 모형의 분류정확도에 대한 연구가 필요하다.

둘째로, 모형이 교통상황의 변화에 민감하게 반응해야 한다. 교통사고는 매우 짧은 순간에 여러 직·간접적인 요인에 의해 발생한다. 이렇게 순간적이고 동적인 요인들의 변화를 보다 정밀하게 관측하기 위해서는 본 연구의 15분 집계간격 보다는 짧은 간격으로 수집된 원시 자료를 이용하는 것이 바람직 할 것이다. 향후 교통정보 센터의 이력자료 저장 및 관리에 시사점을 제공하는 내용이라고 할 수 있다.

이항로지스틱 회귀분석에 의해 추돌 및 충돌, 야간 서울방향 등의 유의한 결과는 얻었으나 여러 유형의 사고 발생가능성에 대한 경고메세지 제공은 불가능하다. 이는 더 많은 검지기 자료의 분석을 통해 다른 case의 경우에 대해서도 사고와의 어떤 연관성이 있는지에 대한 연구가 추진되어야 한다. 또한 본 연구에서 제시한 시스템이 현장에 구현이 될 경우 실제 사고 감소 효과가 발생하는가에 대한 검증작업도 필요할 것이다.

본 연구에서 제안한 실시간 사고 위험도 추정 모형은 운전자에게 방어운전과 사고에 대한 환기를 유도함으로써 운전자의 자발적인 교통사고 감소효과를 불러일으킬 수 있다고 판단된다. 그리고 운전자가 정보를 제공받았을 때 운전자의 방어본능이 차량의 주행에서 명확히 나타나도록 보조지원 시스템이 필요하다고 판단된다.

## 참고문헌

1. 박효신·손봉수·김형진(2007), “고속도로 인터체인지 연결로에서의 교통사고 예측모형 개발”, 대한교통학회지, 제25권 제3호, 대한교통학회, pp.123~135.
2. 오 철·장재남·장명순(1999), “고속도로 시설물 구간의 교통혼잡도와 사고율의 관계 분석 (신갈-안산 고속도로를 중심으로)”, 대한교통학회지, 제17

- 권 제2호, pp.21~27.
3. Abdel-Aty, Nizam Uddin, Anurag Pande, M. Fathy Abdalla, and Liang Hsia(2004), Predicting Freeway Crash from Loop Detector Data by Matched Case-Control Logistic Regression. In Transportation Research Record: Journal of Transportation Research Board, No.1897, TRB, National Research Council, Washington D.C, pp.88~95.
  4. Abdel-Atry, M., and Pande, A.(2004), Classification of real-time traffic speed patterns to predict crashes on freeways. Presented at the 83rd annual Meeting of the Transportation Research Board(TRB), Washington D.C.
  5. Anderson, I. B., Bauer, K. M., Harwood, D. W., and Fitzpatrick, K.(1999), Relationship to Safety of Geometric Design Consistency Measures for Rural Two-lane Highways. Transportation Research Record: Journal of Transportation Research Board, No.1658, TRB, National Research Council, Washington D.C., pp.43~51.
  6. Council, F. M.(1998), Safety Benefits of Spiral Transitions on Horizontal Curves on Two-lane Rural Roads. Transportation Research Record: Journal of Transportation Research Board, No.1635, TRB, National Research Council, Washington D.C., pp.10~17.
  7. Golob, T., and W. Recker.(2001), Relationship Among Urban Freeway Accidents, Traffic Flow, Weather, and Lighting Conditions. Institute of Transportation Studies, University of California, Berkeley.
  8. Gwynn, D. W.(1967), Relationship of Accident Rate and Accident Involvements with Hourly Volumes. Traffic Q., July 407-418.
  9. Lee, C., F. Saccomanno, and B. Hellinga.(2002), Analysis of Crash Precursors on Instrumented Freeways. In Transportation Research Record: Journal of Transportation Research Board, No.1784, TRB, National Research Council, Washington D.C., pp.1~8.
  10. Lee, C., B. Hellinga, and F. Saccomanno.(2003), Real-time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic. In Transportation Research Record: Journal of Transportation Research Board, No.1840, TRB, National Research Council, Washington D.C., pp.67~77.
  11. Oh, C., J.-S. Oh, S. G. Ritchie, and M.-S. Chang.(2001), Real-Time Estimation of Freeway Accident Likelihood.. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C.
  12. Vogot, A., and Bared, J.(1998), Accident Models for Two-lane Rural Segments and Intersections. Transportation Research Record: Journal of Transportation Research Board, No.1635, TRB, National Research Council, Washington D.C., pp.18~29.
  13. Zegeer, C.V., Reinfurt, D. W., Hummer, J., Herf, L., and Hunter, W.(1988), Safety Effects of Cross-section Design for Two-lane Roads. Transportation Research Record: Journal of Transportation Research Board, No.1195, TRB, National Research Council, Washington D.C., pp.20~32.

✉ 주 작 성 자 : 박준형

✉ 교 신 저 자 : 오 철

✉ 논문투고일 : 2007. 11. 10

✉ 논문심사일 : 2008. 3. 10 (1차)

2008. 3. 18 (2차)

✉ 심사판정일 : 2008. 3. 18

✉ 반론접수기한 : 2008. 8. 31

✉ 3인 익명 심사필

✉ 1인 abstract 교정필