

GPS 프로브 차량 속도자료를 이용한 고속도로 사고 위험구간 추출기법

Extraction of Hazardous Freeway Sections Using GPS-Based Probe Vehicle Speed Data

박재홍*
(Jac-Hong Park)

오철**
(Cheol Oh)

김태형***
(Tae-Hyung Kim)

주신혜****
(Shin-Hye Joo)

요약

본 연구에서는 고속도로에서 GPS(Global Positioning System)수신기를 장착한 프로브차량을 이용하여 수집한 속도자료를 이용하여 사고 위험구간을 추출하는 방법론을 제시하였다. 위험구간 추출을 사고발생 유·무를 판단하는 분류문제(Classification)로 정형화하고 베이지안 신경망을 적용하였다. 개별차량의 속도자료를 이용하여 다양한 잠재적 독립변수를 설정하고 이항 로지스틱 회귀분석을 이용하여 통계적으로 유의미한 변수만을 추출하여 베이지안 신경망의 입력자료로 사용하였다. 제안된 방법론의 성능 평가를 위해 사고 발생 경험이 있는 위험구간을 정확히 추출하는 분류정확도를 효과적으로 활용하였다. 본 연구에서 제안한 방법론의 타당성을 60%의 분류정확도를 통해 확인할 수 있었다. 고속도로 신설노선의 교통안전성을 평가하고 사고예방을 위한 대응책 개발 및 적용에 본 연구의 결과가 효과적으로 활용될 것으로 기대된다.

Abstract

This study presents a novel method to identify hazardous segments of freeway using global positioning system(GPS) based probe vehicle data. A variety of candidate contributing factors leading to higher potential of accident occurrence were extracted from the probe vehicle dataset. The research problem was defined as a classification problem, then a well-known classifier, bayesian neural network was adopted to solve the problem. A binary logistic regression technique was also used for selecting salient input variables. Test results showed that the proposed method is promising in extracting hazardous freeway sections. The outcome of this study will be effectively used for evaluating the safety of freeway sections and deriving countermeasures to prevent accidents.

Key words: Traffic safety, speed, BLR(Binary Logistic Regression), probabilistic neural network, probe vehicle, CCR(Correct Classification Rate)

* 주저자 : 한양대학교 교통공학과 석사과정
** 공저자 및 교신저자 : 한양대학교 교통공학과 교수
*** 공저자 : 한국교통연구원 첨단교통연구실 책임연구원
**** 공저자 : 한양대학교 교통공학과 석사과정
† 논문접수일 : 2010년 3월 30일
† 논문심사일 : 2010년 5월 1일
† 게재확정일 : 2010년 5월 3일

I. 서 론

고속도로에서 발생하는 교통사고의 원인을 분석하고 교통사고를 예방하기 위한 노력은 국내·외에서 지속적으로 진행되고 있다. 교통사고는 속도, 기하구조등과 관련성이 높으며, 특히 속도는 규정속도 이상의 속도로 주행하여 발생하는 과속사고뿐만 아니라 안전거리(정지거리)와 관련된 안전거리미확보등의 사고의 원인이다. 또한, 사고심각도와 관련이 높은 것으로 기존의 연구에서 분석되므로, 속도는 교통사고와 밀접한 관계를 가지는 것으로 나타난다.

따라서 본 연구에서는 GPS(Global Positioning System)수신기를 장착한 프로브차량을 이용하여 속도자료를 수집하고, 수집한 속도자료를 베이지안 신경망(Probabilistic Neural Network)에 적용하여 사고 위험구간을 추출하는 방법론을 제시하였다.

고속도로에서 발생한 사고와 속도와의 관계를 조사하기 위해 제한속도가 110km/h인 전국의 고속도로 중에서 4개 노선(서해안 고속도로, 중부내륙 고속도로, 대구포항 고속도로, 청원상주 고속도로)을 선택하였다. 본 연구에서는 수집되는 속도자료들의 교통류에 대한 동질성을 확보하기 위해 제한속도가 110km/h인 고속도로만을 대상구간으로 선정하였다.

본 연구에서는 단위분석구간을 1km로 설정하였다. 또한, 분석구간의 개별차량 속도자료를 이용하여 평균속도, 85th Percentile 속도, 표준편차, 최고속도와 최저속도의 차이, 최고속도와 평균속도의 차이를 산출하고 분석구간별 변수값들의 차이를 통해 변수를 도출하였다. 사고자료는 4개노선의 2~3년간의 사고 자료를 이용했으며, 제안한 방법론의 성능평가를 위해 비교적 신규 개통노선을 선정하였다.

베이지안 신경망의 입력변수 설정을 위해 이항 로지스틱 회귀분석을 사용하였으며, 베이지안 신경망을 사용하여 신설되는 고속도로의 사고위험구간을 예측할 수 있는 방법론을 제시하였다.

고속도로에서 발생한 사고에 영향을 미치는 속도와 관련된 변수를 도출하기 위하여 이항 로지스틱 회귀분석을 사용하였다. 이항 로지스틱 회귀분석은 확률적모형을 통해 독립변수와 종속변수의 관계를

규명하는데 유용한 분석방법이다. 일반적으로 독립변수의 개수가 적을수록 모형의 이식성이 높게 나타나므로, 이항 로지스틱 회귀분석을 사용하여 속도와 관련된 변수 중에서 통계적으로 유의한 변수를 도출하였다.

이항 로지스틱 회귀분석을 통해 도출된 6개의 변수를 베이지안 신경망에 적용하여 분류정확도(CCR:Correct Classification Rate)를 산출하였다. 분류정확도는 대상노선의 ‘사고발생’ 및 ‘사고미발생’ 구간의 개수중에서 베이지안 신경망에 의해 정확히 분류된 ‘사고발생’ 및 ‘사고미발생’ 구간의 수에 대한 백분율을 의미한다.

본 연구의 2장에서는 속도와 사고의 관계에 관한 국내·외의 연구사례를 제시했으며, 3장에서는 프로브 차량을 통해 수집된 속도자료들과 대상노선에서 발생한 사고자료의 변수설정방법을 제시하였다. 4장에서는 본 연구에서 사용되는 이항 로지스틱 분석과 베이지안 신경망의 이론적 배경을 제시했으며, 5장에서는 분석방법에 따른 분석결과를 나타냈다. 마지막으로 6장에서는 본 연구에서 도출된 결론과 향후 과제를 제시하였다.

II. 기존문헌 고찰

1. 제한속도와 사고위험의 관계

이남수(2007)는 제한속도가 설계속도보다 높고 통행속도가 높아 사고가 발생하는 문제를 해결하기 위해 제한속도 하향조정 후 교통사고변화를 교통관리 시스템자료를 이용하여 분석하였다. 분석결과 통행속도 감소로 교통사고감소효과가 있다는 결과를 제시하였다[1]. Ezra Hauer(2006)는 전체 사고 수가 변화 없다고 가정하고 제한속도를 증가시킬 경우 부상사고의 변화를 연구하였다. 제한속도가 높아지면, 심각도가 높은 부상사고가 높아지는 것으로 나타났다[2]. Nilsson(1982, 2004)은 제한속도 변화에 따른 안전성효과를 평가하기 위하여 운동에너지를 적용한 경상(Injury)/중상(Severe)/치명(Fatal)사고율을 예측하는 Power function 모형을 정립하였다. 연구에서 개발

된 모형은 지방부도로와 도시부도로에서 제한속도 변화에 따른 결과를 예측하는데 매우 우수한 모형으로 분석되었다[3, 4]. Fildes et al.(1991)은 제한속도가 60km/h인 도시부도로와 100km/h인 지방부도로를 대상으로 개별차량의 속도자료를 수집하고 저속주행운전자(V15 이하)와 고속주행운전자(V85 이상)를 선출하여 5년간의 사고이력을 조사하였다. 이 자료를 바탕으로 제한속도와 개별차량속도 차이와 사고위험과의 관계를 설명하는 지수함수모형을 정립하였으며 제한속도와 개별차량속도 차이가 크거나 제한속도가 높은 도로일수록 사고위험이 높다는 결과를 제시하였다[5]. Kloeden et al.(1997, 2002)은 평균속도와 개별차량속도를 이용하여 제한속도가 다른 도시부도로와 지방부도로의 주행속도와 부상사고위험과의 관계를 설명하는 지수함수모형을 정립하였다. 연구결과 도시부도로가 지방부도로보다 속도변화에 민감한 것으로 나타났다[6]. Hauer and Bonneson(1997)은 평균속도변화와 안전성변화의 관계를 제시하기 위해 평균속도를 적용한 사고수에 관한 Power function 모형을 제시하였고 모형의 전이성을 위해 AMF(Accident modification function)을 도출하였다[7]. Baruya(1998a, 1998b)는 평균속도와 사고빈도의 관계를 도출하기 위해 도로유형별 제한속도위반비율, 교통량, 구간길이, 교차점, 도로폭, 제한속도 등의 변수를 사용하여 사고빈도를 산출함으로써 평균속도뿐만 아니라 도로 및 교통특성을 반영한 사고빈도 모형을 제시하였다 [8, 9].

2. 속도분산과 사고위험의 관계

하태준, 이석(2002)은 속도분산차, 차량 진행방향 가속도차, 곡선반경방향 가속도차를 이용하여 새로운 안전성 평가방법을 제안하였다. 기존의 안전성평가방법과 비교한 결과 제안된 방법이 보다 적절하며, 속도분산이 도로의 위험도(EPDO)와 통계적으로 유의한 상관관계를 나타내었다고 제시하였다[10]. Garber and Gadiraju(1989)는 도로유형별 속도편차와 사고율의 관계를 나타내기 위해 주간선도로, 고속도로, 집산도로에서 수집된 지점속도와 교통량, 사고자

료를 이용하였다. 분석결과 속도편차가 증가함에 따라 사고율이 증가하며, 제한속도가 설계속도보다 5~10mph 정도 낮을 경우 속도편차가 작아 사고율이 낮은 것으로 나타났다. 속도분산의 증가에 따라서 사고율은 증가한다는 결과를 도출함으로써 속도자체가 교통사고 증감요인이 아니라 개별차량의 속도편차가 사고발생과 밀접한 관계가 있음을 제시하였다[11]. Taylor et al.(2000)은 도로유형을 도시 내 혼잡도로, 시내도로, 교외도로, 외곽고속도로로 구분했으며, 도로유형에 상관없이 평균속도가 커질수록 사고빈도가 증가한다는 결과를 제시하였다. 또한, 혼잡도로에서 사고빈도수와 사고증가율이 높게 나타났으며, 표준편차와 평균속도를 고려한 사고빈도 산출모형을 정립하였다[12].

속도와 사고와의 관련성을 분석하기 위한 기존연구의 교통자료 수집은 설문조사, 루프검지기 또는 영상검지기등을 이용하였다. 하지만 본 연구에서는 개별차량의 순간속도 및 속도변화량을 분석하기 쉽고, 속도측정이 정확한 개별차량을 이용하여 속도를 수집하였다.

기존연구에서는 평균속도, 제한속도 등과의 차이를 통계적 분석기법을 이용하여 속도와 사고와의 관계를 분석하였다. 그런데 본 연구에서는 개별차량의 속도를 구간의 대표속도, 구간내 속도의 변화, 구간간 속도의 변화로 구분하고 사고자료와 매칭시켜 사고에 영향을 미치는 요인을 찾고 위험구간을 추출하는 새로운 방법론을 제시하는데 차별성을 부여하고 연구를 진행하였다.

3. 기존 연구와의 차별성

속도와 사고와의 관련성을 분석하기 위한 기존연구의 교통자료 수집은 설문조사, 루프검지기 또는 영상검지기등을 이용하였다. 하지만 본 연구에서는 개별차량의 순간속도 및 속도변화량을 분석하기 쉽고, 속도측정이 정확한 개별차량을 이용하여 속도를 수집하였다.

기존연구에서는 평균속도, 제한속도 등과의 차이를 통계적 분석기법을 이용하여 속도와 사고와의 관

<표 1> 분석구간의 사고건수, 연장, AADT, 차로수, 이점 및 수집시기
 <Table 1> # Of accident, length, AADT, lane station and date

대상구간	사고수 (건)	연장 (km)	AADT			차로수	이점	수집시기
			구간 수	최소	최대			
중부내륙고속도로	184	299	16	35,841	55,783	42,310	양방향 4차로 김천방향(Sta.263~115) 여주방향(Sta.114~263)	2009년 7월 6일 월요일
서해안고속도로	197	342	12	24,705	117,570	55,876	양방향 4차로 서울방향(Sta.146~316) 목포방향(Sta.315~145)	2010년 2월 4일 목요일
대구포항고속도로	70	154	4	19,040	35,656	24,481	양방향 4차로 대구방향(Sta.77~0) 포항방향(Sta.0~77)	2009년 7월 24일 금요일
청원상주고속도로	13	154	7	17,890	24,741	21,358	양방향 4차로 청원방향(Sta.77~0) 상주방향(Sta.0~77)	2009년 7월 23일 목요일

계를 분석하였다. 그런데 본 연구에서는 개별차량의 속도를 구간의 대표속도, 구간내 속도의 변화, 구간간 속도의 변화로 구분하고 사고자료와 매칭시켜 사고에 영향을 미치는 요인을 찾고 위험구간을 추출하는 새로운 방법론을 제시하는데 차별성을 부여하고 연구를 진행하였다.

III. 자료수집 및 변수설정

1. 자료수집

본 연구의 조사대상구간으로 선정한 중부내륙고속도로, 서해안고속도로, 대구포항고속도로, 청원상주고속도로의 분석구간을 <그림 1>에 제시하였다. 또한, 분석구간의 사고건수, 연장, AADT, 차로수, 이

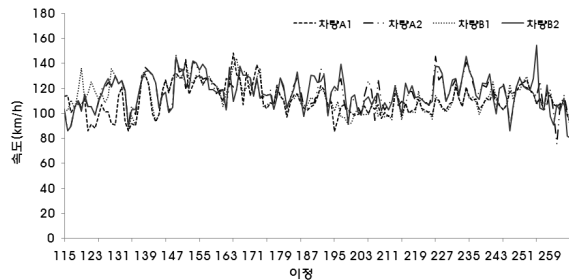
점 및 수집시기를 <표 1>에 제시하였다. AADT의 경우, 본 연구의 분석구간(km)과 AADT를 제시하는 ‘고속도로 통계연보’의 구간(IC)연장의 차이로 인하여 최소, 최대 및 평균 AADT를 제시하였다. 본 연구



<그림 1> 분석대상구간
 <Fig. 1> Freeways for data collection



(a) GPS 수집장치
 (a) GPS-equipped probe vehicle



(b) 중부내륙고속도로의 속도 변화(여주방향-상행)
 (b) Example of speed profile (toward Yeosu)

<그림 2> 조사장비 및 수집자료
 <Fig. 2> Investigation equipment and Collecting data

에서는 프로브차량을 이용하여 4개 노선(중부내륙고속도로, 서해안고속도로, 대구포항고속도로, 청원상주고속도로)의 속도자료를 GPS수신기를 장착한 프로브차량을 이용하여 수집하였다.

프로브 차량을 이용하여 속도자료를 수집하는 경우, 개별차량의 순간속도를 측정할 수 있는 장점을 가지고 있다. 4개 노선의 속도자료는 정체가 없는 교통상황에서(오전 9시~오후 5시) 속도자료를 수집하여 정체에 의한 속도변화의 영향을 제거하였다. <그림 2-a>에는 GPS 수집장치와 GPS 수집장치에서 수집되는 자료화면을 제시했으며, <그림 2-b>에는 프로브 차량이 중부내륙고속도로를 주행하여 수집한 이점별 속도변화를 제시하였다. 또한, 본 연구에서는 오차범위가 1Hz에서 10m~100m로 나타나는 일반적인 상용GPS 수집장치를 이용하여 자료를 수집하였다.

GPS수신기를 장착한 프로브차량을 통해 수집된 속도(km/h)는 위도-경도에 따라 수집되므로 Map-Matching작업을 통해 4개 노선의 위도-경도를 이점으로 변환하였다.

2. 변수설정

일반적으로 평균속도, 가속도의 변화 및 가속소음을 통해 교통류의 상태를 판단할 수 있다. 특히, 가

속소음의 변화가 큰 것은 순간적인 차량의 급감속 또는 급가속을 나타내므로 정상적인 주행상황이 아닌 것을 의미한다. 따라서, 개별차량의 속도변화등을 측정함으로써 교통안전상태의 분석이 가능하다.

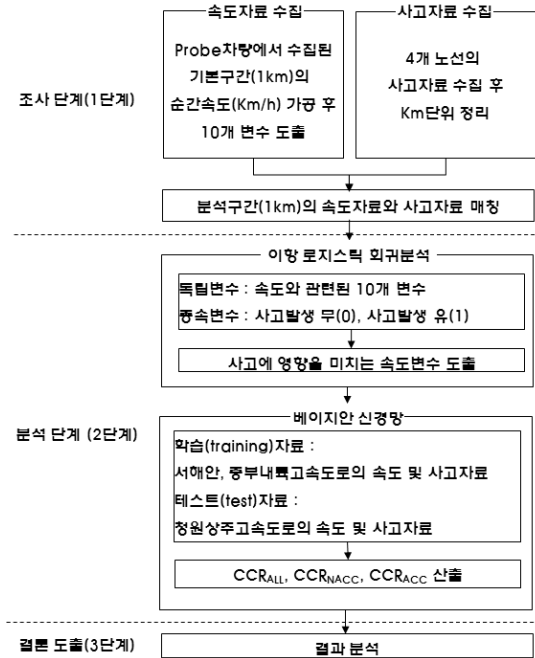
본 연구에서는 GPS수신기를 장착한 프로브차량을 통해 수집된 속도를 가공하여 10개의 변수를 설정했으며, 변수 및 변수에 대한 설명을 <표 2>에 제시하였다. 또한, 본 연구에서 설정한 10개의 속도관련변수는 변수들의 특성에 따라 속도의 변화, 변동 및 대표성을 나타내는 3개의 그룹으로 구분하였다. 구간(i)의 대표속도를 나타내는 $\bar{V}(i)$, $V(i)^{85}$ 와 구간(i) 내 속도의 변화를 의미하는 $V(i)^{SD}$, $V(i)^{MM}$, $V(i)^{MA}$ 로 그룹이 형성된다. 또한, 구간간 속도의 변화를 나타내는 $\bar{V}^d(i)$, $V^d(i)^{85}$, $V^d(i)^{SD}$, $V^d(i)^{MM}$, $V^d(i)^{MA}$ 이 하나의 그룹을 형성한다. $\bar{V}(i)$ 는 구간 i를 주행한 4대의 프로브차량에서 수집된 평균속도를 의미하며, $\bar{V}^d(i)$ 는 $\bar{V}(i)$ 와 $\bar{V}(i+1)$ 의 차이의 절대값을 의미한다. $V(i)^{85}$ 는 구간(i)를 주행한 4대의 프로브차량에서 수집된 순간속도들의 85th percentile속도를 의미하며, $V^d(i)^{85}$ 는 $V(i)^{85}$ 와 $V(i+1)^{85}$ 의 절대값의 차이를 나타낸다. $V(i)^{SD}$ 는 구간(i)을 주행한 4대 차량에서 수집된 순간속도의 표준편차를 나타내며, $V^d(i)^{SD}$ 는 $V(i)^{SD}$ 와 $V(i+1)^{SD}$ 의 절대값의 차이를 나타낸다. $V(i)^{MM}$ 는 구간(i)에서 주행한 4대 차량에서 수집된

변수	설명	구분
$\bar{V}(i)$	구간 i에서 프로브 차량의 평균속도	구간의 대표속도
$\bar{V}^d(i)$	$ \bar{V}(i) - \bar{V}(i+1) $	구간간 속도의 변화
$V(i)^{85}$	구간 i에서 프로브차량의 85th percentile 속도	구간의 대표속도
$V^d(i)^{85}$	$ V(i)^{85} - V(i+1)^{85} $	구간간 속도의 변화
$V(i)^{SD}$	구간 i에서 프로브차량 속도의 표준편차	구간내 속도의 변화
$V^d(i)^{SD}$	$ V(i)^{SD} - V(i+1)^{SD} $	구간간 속도의 변화
$V(i)^{MM}$	구간 i에서 프로브차량 속도의 최대값과 최소값의 차이	구간내 속도의 변화
$V^d(i)^{MM}$	$ V(i)^{MM} - V(i+1)^{MM} $	구간간 속도의 변화
$V(i)^{MA}$	구간 i에서 프로브차량의 최대속도와 평균속도의 차이	구간내 속도의 변화
$V^d(i)^{MA}$	$ V(i)^{MA} - V(i+1)^{MA} $	구간간 속도의 변화

d = 구간 i와 i+1구간의 차이 i = 분석구간(km) SD = 표준편차(Standard Deviation)

<표 2> 변수설명 및 변수구분

<Table 2> Variable descriptions



<그림 3> 연구 흐름도
<Fig. 3> Research procedure

순간속도 중에서 최고속도 및 최저속도의 차를 의미하며, $V^{(i)MM}$ 는 $V(i)^{MM}$ 와 $V(i+1)^{MM}$ 의 절대값 차이를 나타낸다. 마지막으로 $V(i)^{MA}$ 는 분석구간(i)을 수행한 4대의 프로브차량에서 수집된 순간속도의 최고속도와 분석구간(i) 평균속도의 차이를 나타내며, $V(i)^{MA}$ 와 $V(i+1)^{MA}$ 의 차이를 나타내는 $V^{(i)MA}$ 를 변수로 사용하였다.

베이지안 신경망을 이용하여 사고위험구간을 추출하기 위해서는 속도자료와 사고자료 설정이 필요하다. 속도자료는 앞에서 언급한 10개의 속도변수를 사용했으며, 사고자료는 속도자료와 동일하게 1km를 분석구간으로 설정하고 사고발생유·무를 구분한 자료를 사용하였다. 완공년도가 3년 이상된 3개노선(중부내륙고속도로, 서해안고속도로, 대구포항고속도로)은 3개년의 사고자료를 이용하였으며, 3년간 1건 이상의 교통사고가 발생한 구간을 1, 교통사고가 발생하지 않은 구간을 0으로 설정하였다. 단, 청원상주고속도로의 경우 완공년도가 3년 이하이므로 최근 2년간의 사고자료를 사용하였으며, 2년간 1건 이상의

교통사고가 발생한 구간을 1, 교통사고가 발생하지 않은 구간을 0으로 설정하였다.

IV. 방법론

1. 분석절차

본 연구에서는 고속도로 위험구간을 추출하기 위한 방법론을 도출하기 위하여 <그림 3>에 제시한 3단계의 분석절차(조사단계(1단계), 분석단계(2단계), 결론도출(3단계))로 진행하였다.

조사단계(1단계)에서는 GPS수신기를 장착한 프로브 차량을 통해 4개 노선의 속도자료를 수집하고 사고자료의 정리를 수행하였다. 수집된 속도자료를 이용하여 10개의 변수를 도출하고 4개 노선의 사고자료를 이용하여 분석구간의 사고유·무를 결정하고 속도자료와 사고자료를 매칭하였다.

분석단계에서는 베이지안 신경망의 성능을 향상시키고 통계적으로 유의한 변수를 설정하기위해 이항 로지스틱 회귀분석을 우선 시행하였다. 속도와 관련된 10개의 변수중에서 고속도로 사고발생에 영향을 미치는 변수를 도출하였으며, 종속변수는 사고발생유·무에 따라 사고무 0, 사고유 1로 적용하였다.

또한, 이항 로지스틱 회귀분석 결과에서 유의하다고 도출된 변수를 베이지안 신경망에 적용하였다. 베이지안 신경망에서는 서해안고속도로와 중부내륙고속도로의 속도와 사고자료를 학습(training)후, 완공년도가 3년 이하인 청원상주고속도로의 사고자료를 테스트(test)하여 Smoothing Parameter에 따른 분류정확도를 산출하였다.

결론도출에서는 속도관련변수와 사고발생유·무를 베이지안 신경망에 적용하여 고속도로 위험구간을 추출하는 방법론에 대한 결론을 제시하였다.

2. BLR(Binary Logistic Regression)

이항 로지스틱 회귀분석은 종속변수가 범주형인 자료를 분석하는데 적합한 분석방법으로써, 확률에 대한 로지스틱 변환을 고려한 분석방법이다. 따라서,

본 연구에서는 4개 노선의 사고발생여부를 범주형 종속변수를 사고무 0, 사고유 1로 설정하고, 속도와 관련된 10개의 변수를 독립변수로 설정하여 이항 로지스틱 회귀분석을 사용하였다. 식(1)에는 이항 로지스틱 회귀모형식을 제시하였다.

$$P(Y_i = 1/X_i) = \frac{\exp[f(X_i, \beta)]}{1 + \exp[f(X_i, \beta)]} \quad \text{식(1)}$$

Y_i : 독립변수 대하여 사고가 발생할 확률
 X_i : 사고가 발생할 확률에 영향을 미치는 독립변수
 $f(X_i, \beta)$: X_i 와 β 로 구성된 함수

3. 베이지안 신경망

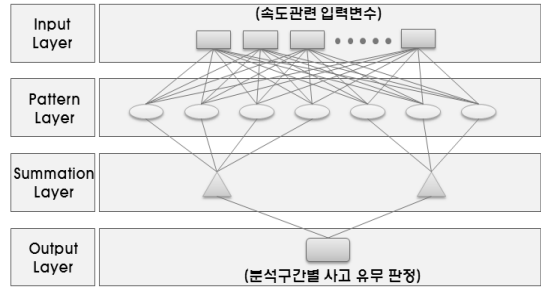
베이지안 신경망(Probabilistic Neural Network)은 인공지능(AI : Artificial Intelligent)기법 중의 하나이며, 베이지안 신경망의 주요 구성요소 중의 하나는 그룹별 확률밀도 함수이다. 가장 일반적으로 사용되는 가우시안 함수를 적용한 식을 식(2)에 제시하였다(오철 등, 2006)[13].

$$g(X) = \frac{1}{(2\pi)^{\frac{p}{2}} n \sigma^2} \sum_{i=1}^n \exp\left(-\frac{\|X - X_i\|^2}{2\sigma^2}\right) \quad \text{식(2)}$$

X = 입력변수
 X_i = i 번째 학습(training) 벡터
 σ = 입력변수의 표준편차(Smoothing parameter)
 p = 입력변수의 차원(dimensinality)

또한, 베이지안 신경망은 연결강도의 조정과정이 필요하지 않으므로 학습이 간단하며, 훈련(training)에 사용되는 자료를 1회만 사용하므로 학습속도가 빠르다는 장점을 가진 분석방법이다.

베이지안 신경망은 a 개의 입력 유닛을 포함하는 input layer, n 개의 패턴 유닛을 포함하는 pattern layer, b 개의 부류유닛을 포함하는 summation layer와 결과 유닛을 포함하는 out layer로 구성된다. 입력유닛은 패턴층의 모든 패턴유닛에 연결되며, 패턴유닛은 b 개의 부류유닛 중에서 하나의 부류유닛에 연결된다.



<그림 4> 베이지안 신경망 구조
 <Fig. 4> Architecture of bayesian neural network

<그림 4>에는 베이지안 신경망 신경망구조에 대한 그림을 제시하였다. Pattern layer에서는 input layer에 존재하는 입력유닛을 가중치 벡터 W_i 와 연결하여 Z_i ($Z_i=XW_i$)를 생성한다. 또한, output layer에서는 식(2)에 제시된 확률밀도함수를 이용하여 입력 유닛들이 각각의 그룹에 소속된 확률을 산출한다. <그림 5-a>에는 베이지안 신경망 훈련 알고리즘, <그림 5-b>에는 베이지안 신경망 분류 알고리즘을 제시하였다. <그림 5-a>에서 j 는 패턴유닛 수, i 는 부류유닛 수, x_{jk} 는 j 번째 패턴의 요소, w_{jk} 는 j 번째 패턴 유닛의 가중치를 의미한다. 베이지안 신경망훈련은 각 패턴의 정규화 후, 가중치가 설정되며 부류유닛에 도달하는 반복작업이 이루어진다. <그림 5-b>의 net_k 는 각 패턴 유닛의 내적 값, $e^{(net_k-1)/\sigma^2}$ 는 비선형 함수 나타내며, σ 는 유효 가우스 윈도우의 폭(Smoothing Parameter), $g_i(x)$ 는 판별함수를 의미한다(Richard O. Duda et al., 2006)[14].

V. 적용

1. BLR을 이용한 분석결과

본 연구에서는 상용적인 통계 프로그램 중 하나인 SPSS 17.0을 사용하여 이항 로지스틱 회귀분석을 실시하였다. 이항 로지스틱 회귀분석은 입력선택과 전진/후진단계(wald)선택법등을 통해 모형에 유의한 변수를 도출 할 수 있다. 본 연구에서는 wald 통계량을 이용하여 사고발생 확률과 관련된 유의한 변수를 도

출하는 전진단계선택(wald)방법을 사용하였다.

<표 3>에는 이항 로지스틱 회귀모형의 Initial Model과 Final Model을 제시하였다. 또한, Final Model에서는 속도와 관련된 10개의 독립변수 중에서

<표 3> 초기모형과 최종모형
<Table 3> BLR analysis results

Variable	초기모형			최종모형		
	B	p-value	Exp(B)	B	p-value	Exp(B)
$\bar{V}(i)$	0.345	0.557	-	-	-	-
$\bar{V}^d(i)$	5.780	0.016	-	-	-	-
$V(i)^{S5}$	0.289	0.591	-	-	-	-
$V^d(i)^{S5}$	0.001	0.976	-	-0.014	0.030	0.986
$V(i)^{SD}$	1.149	0.284	-	-0.029	0.042	0.971
$V^d(i)^{SD}$	-0.081	0.000	0.923	-0.072	0.000	0.931
$V(i)^{MM}$	21.883	0.000	-	0.031	0.000	1.031
$V^d(i)^{MM}$	4.980	0.026	-	-	-	-
$V(i)^{MA}$	15.320	0.000	-	-0.033	0.000	0.968
$V^d(i)^{MA}$	20.004	0.000	-	0.030	0.022	1.030
상수항	0.398	0.000	1.489	1.697	0.026	5.456

고속도로 사고발생에 영향을 미치며, 통계적으로 유의한 6개 변수($V^d(i)^{S5}$, $V(i)^{SD}$, $V^d(i)^{SD}$, $V(i)^{MM}$, $V(i)^{MA}$, $V^d(i)^{MA}$)를 제시하였다. 본 연구에서는 모형의 적합도 보다는 종속변수에 영향을 주는 속도요인을 찾고 고속도로의 위험구간을 추출하기 위한 방법론에 관한 연구이므로 독립변수의 유의확률에 초점을 맞추었다.

2. 베이지안 신경망을 이용한 분석결과

본 연구에서는 고속도로 위험구간을 추출하기 위한 방법론의 성능평가를 위해 중부내륙고속도로와 서해안고속도로의 속도 및 사고자료를 학습 자료(training dataset)로 설정하였다. 또한, 비교적 최근에 개통한 청원상주고속도로의 속도 및 사고자료를 테스트 자료(test dataset)로 사용하여 베이지안 신경망 분석을 시행하였다.

베이지안 신경망분석에서는 분석 초기에 설정한 10개의 속도관련변수를 독립변수로 사용하지 않고, 이항 로지스틱 회귀분석 결과를 통해 통계적으로 유의하다고 도출된 6개의 변수($V^d(i)^{S5}$, $V(i)^{SD}$, $V^d(i)^{SD}$,

$$CCR_{ALL}(\%) = \frac{\text{베이지안신경망에 의해 정확히 분류된 '사고발생' 및 '사고미발생' 구간의 개수}}{\text{전체 분석구간개수}} \times 100 \dots \text{식(3)}$$

$$CCR_{NACC}(\%) = \frac{\text{베이지안신경망에 의해 정확히 분류된 '사고미발생' 구간의 개수}}{\text{'사고미발생' 구간의 개수}} \times 100 \dots \text{식(4)}$$

$$CCR_{ACC}(\%) = \frac{\text{베이지안신경망에 의해 정확히 분류된 '사고발생' 구간의 개수}}{\text{'사고발생' 구간의 개수}} \times 100 \dots \text{식(5)}$$

<pre> Begin initialize j ← 0, n, a_{ji} ← 0 for j=1, ..., n; i=1, ..., c do j ← j+1 x_{jk} ← x_{jk} / (∑ a_i x_{ij}²)^{1/2} (normalize) w_{jk} ← x_{jk} (train) if x ∈ w_i then a_{ji} ← 1 until j=n end </pre>	<pre> Begin initialize k ← 0, x ← test patten do k ← k+1 net_k ← w_kx if a_{ki} = 1 then g_i ← g_i + exp[(net_k - 1) / σ²] until k = n return class ← arg max g_i(x) end </pre>
--	--

(a) 베이지안 신경망 훈련 알고리즘
(a) Training algorithm

(b) 베이지안 신경망 분류 알고리즘
(b) Classification algorithm

<그림 5> 베이지안 신경망 알고리즘

<Fig. 5> Bayesian network algorithm

$V(i)^{MM}$, $V(i)^{MA}$, $V^{d(i)^{MA}}$ 를 이용하여 사고발생에 대한 분류정확도(CCR:Correct Classification Rate)를 산출하였다. 특히, 분류정확도는 신뢰성 있는 모형개발 및 평가를 위해 CCR_{ALL} , CCR_{NACC} , CCR_{ACC} 로 구분하고, 식(3), 식(4), 식(5)에 제시하였다. CCR_{ALL} 은 고속도로의 전체 분석구간 개수중에서 베이지안 신경망을 통해 정확히 분류된 ‘사고발생’ 및 ‘사고미발생’ 구간의 개수를 의미한다. CCR_{NACC} , CCR_{ACC} 은 ‘사고미발생’과 ‘사고발생’ 구간의 각각의 개수중에서 베이지안 신경망을 통해 정확히 분류된 ‘사고미발생’ 또는 ‘사고발생’구간의 개수를 의미한다.

중부내륙고속도로와 서해안고속도로의 속도 및 사고자료 641개의 2/3에 해당하는 426개를 학습(training)에 사용하였고, 1/3에 해당하는 215개를 테스트(test)에 사용하여 분류정확도를 산출하였다. 학습(training)과 테스트(test)에 사용된 자료는 신뢰성 확보 및 표본의 다양성을 위해 랜덤함수를 이용하여 30회 반복 추출하고 베이지안 신경망모형에 적용하였다. 또한, 분석에 사용된 Smoothing Parameter는 0.0001~0.4000으로 설정하였으며, 0.0001씩 증가시켰다. <그림 6>에는 베이지안 신경망의 30회 반복학습의 시행 결과중에서 Smoothing Parameter에 따른 CCR_{ALL} , CCR_{NACC} , CCR_{ACC} 변화 예를 제시하였다.

본 연구에서는 중부내륙고속도로와 서해안고속도로를 학습(train)자료로 사용하고 청원상주고속도로를 테스트(test)자료로 사용하여 CCR_{ALL} , CCR_{NACC} , CCR_{ACC} 을 산출하였다. Smoothing Parameter가 0.0272인 경우에는 CCR_{ALL} 은 63.87%, CCR_{NACC} 64.08%,

CCR_{ACC} 61.54%의 분류정확도를 얻을 수 있었다.

본 연구에서 제시한 방법론의 성능평가를 위해 효과적으로 적용한 분류정확도가 60%이상 산출되므로, 이항 로지스틱 회귀분석을 이용하여 통계적으로 유의미한 속도관련변수를 베이지안 신경망에 적용하여 사고위험구간을 추출하는 방법론은 타당성이 있는 것으로 분석되었다.

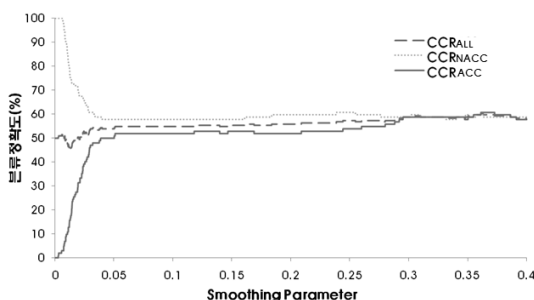
VI. 결론

고속도로에서 발생하는 사고발생여부와 속도와의 관계를 분석하고 신설 고속도로의 위험구간을 사전에 도출하기 위한 연구를 시행하기 위해 서해안고속도로, 중부내륙고속도로, 청원상주고속도로, 대구포항고속도로의 속도 및 사고자료를 사용하였다. 각 노선의 사고자료는 대상노선에 발생한 2년 또는 3년간의 사고자료를 이용하였으며, 프로브 차량에 GPS수신기를 장착하고 대상노선의 속도자료를 측정하였다.

본 연구에서는 GPS수신기를 장착한 프로브 차량을 이용하여 수집된 속도를 이용하여 10가지 독립변수로 설정하고, 사고유무를 종속변수로 설정하여 이항 로지스틱 회귀분석을 실시하였다. 분석결과, 10가지 변수중에서 6개 변수($V^{d(i)^{SS}}$, $V(i)^{SD}$, $V^{d(i)^{SD}}$, $V(i)^{MM}$, $V(i)^{MA}$, $V^{d(i)^{MA}}$)가 사고에 영향을 주는 유의미한 변수로 도출되었다. 이항 로지스틱 회귀분석을 통해 도출된 6개의 변수를 베이지안 신경망에 적용하여 분류정확도(CCR)를 산출하였다. 중부내륙고속도로와 서해안고속도로의 속도 및 사고자료를 랜덤함수를 이용하여 임의로 배정한 후, 학습(training)자료로 30회 사용했으며, 청원상주고속도로의 속도 및 사고자료를 테스트(test)자료로 사용하여 분류정확도를 산출하였다.

본 연구에서 제안한 방법론의 효과적으로 활용한 분류정확도가 60%이상으로 타당성이 확보되었다. 따라서, 본 연구에서 제안한 방법론은 고속도로 신설노선의 교통안전성을 평가하고 사고예방을 위한 대응책 개발 및 적용에 활용될 수 있다.

본 연구에서 도출된 분류정확도가 60%내외로 나



<그림 6> Smoothing Parameter에 따른 분류정확도 변화
<Fig. 6> CCRs by smoothing parameter

타나므로 방법론에 대한 적용 타당성 및 가능성을 제시하였다. 그러나 보다 통계적으로 유의하며 높은 수준의 분류정확도를 확보하기위해 아래에 제시한 추가적인 연구가 필요하다.

첫째, 본 연구에서 사용된 프로브차량의 속도자료는 2명의 운전자가 대상 고속도로를 2회 주행한 속도자료를 이용하여 분석을 실시하였다. 또한, 오전·오후에 각각 1회 운행함으로써 대상구간의 개별차량 속도를 수집하여 속도와 관련된 다양한 독립변수를 도출하여 분석하였다. 하지만, 추가연구에서는 다양한 운행패턴을 가진 운전자가 운행하는 프로브차량을 주행시킴으로써 다양한 개별차량의 속도자료를 수집하고 분석을 실시해야 한다.

둘째, 본 연구에서 단순히 사고발생유·무만을 종속변수로 설정하였다. 하지만 종속변수를 사고발생에 따른 사망건수비교, 사망, 부상과 같은 사고심각도에 대한 분류등으로 그루핑(grouping)하여 추가적인 분석을 실시해야 한다.

셋째, GPS수집장치를 이용하여 속도자료를 수집하는 경우, 교량이나 산악지형등에서 결측자료가 발생할 수 있다. 본 연구의 대상구간에서 수집된 위도, 경도에 따른 속도자료를 분석한 결과 결측자료는 발생하지 않았지만, 본 연구에서 제안한 방법론을 타 고속도로에 적용하는 경우 결측자료가 발생할 수 있으므로, 결측자료 보정방법에 대한 추가적인 연구가 진행되어야 한다.

넷째, 본 연구에서는 속도자료 및 사고자료의 분석을 km단위로 구분하여 분석하였다. 하지만, 통행특성의 동질성이 고려된 구간으로 구분하거나, km단위를 세분화하여 분석하는 등의 구간변화에 따른 추가적인 연구가 필요하다.

참고문헌

[1] 이남수, “제한속도 변경에 따른 교통사고 효과분석”, 시립대학교 석사학위논문, 2007.
 [2] E Hauer, “The frequency-severity indeterminacy,” *Accident Analysis and Prevention*, vol.38, no.1, pp.78-83, Jan. 2006.

[3] G. Nilsson, “The effects of speed limits on traffic crashes in sweden,” Organisation for Economy, In: *Proceedings of the international symposium on the effects of speed limits on traffic crashes and fuel consumption*, Dublin. Co-operation, and Development (OECD), 1982.
 [4] G. Nilsson, “Traffic safety dimensions and the power model to describe the effect of speed on safety,” Lund Institute of Technology, Bulletin 221, 2004.
 [5] B. N. Fildes, G. Rumbold, and A. Leening, “Speed behavior and drivers’ attitude to speeding,” *VIC Roads*, General report, no.16, June 1991.
 [6] C. N. Kloden, G. Ponte, A. J. McLean, and G. Glonek, “Reanalysis of travelling speed and the rate of crash involvement in adelaide south australia,” *Australian Transport Safety Bureau*, no. CR 207, 2002.
 [7] E. Hauer, and J. Bonneson, “An empirical examination of the relationship between speed and road accident based on data by elvik,” Chistensen and Amundesen, Report ptrpared for project NCHRP pp.17-25, Sep. 2006.
 [8] B. Baruya, “Speed-accident relationships on european roads,” *In:Proceedings of the conference ‘Road safety in Europe’, Bergisch Gladbach, Germany, VTI Konferens*, no. 10A, Part 10, pp. 1-17, Sep. 1998.
 [9] B. Baruya, “MASTER: Speed-accident relationship on European roads,” Technical Research Centre of Finland VTT, Working Paper R 1.1.3, Deliverable D7, 1998b.
 [10] 하태준, 이석, “지방부 2차로 안전성 평가에 관한 연구”, *대학교통학회지*, 제20권, 제1호, pp. 121-130, 2002. 4.
 [11] N. J. Garber, and R. Gadiraju, “Factors affecting speed variance and its influence on accidents,” *Transportation Research Record*, no.1213, pp. 64-71, 1989.

- [12] M. C. Taylor, D. A. Lynam, and A. Baruya, "The effects of drivers' speed on the frequency of road accidents," *Transport Research Laboratory*, no.421, March 2000.
- [13] 오철, 강연수, 김범일, "베이지안 신경망을 이용한 보행자 사망확률모형 개발," *대한교통학회지*, 제24권, 제2호, pp.139-145, 2006. 4.
- [14] R. O. Duda, P. E. Hart, and D. G. Stock, "*Pattern classification, 2nd edition*" John Wiley&Sons, Inc. pp.196-199, 2001.

저자소개



박 재 홍 (Park, Jae-Hong)

2003년 3월 ~ 2009년 2월 : 한양대학교 교통시스템공학과 공학사
2009년 3월 ~ 현재 : 한양대학교 대학원 교통공학과 석사과정



오 철 (Oh, Cheol)

1989년 3월 ~ 1993년 2월 : 한양대학교 교통공학과 공학사
1993년 3월 ~ 1997년 8월 : 한양대학교 대학원 교통공학과 석사
1999년 9월 ~ 2003년 12월 : University of California, Irvine, 토목환경공학과 공학박사
(교통시스템전공)
1998년 9월 ~ 1999년 8월 : 한국건설기술연구원 도로연구부 연구원
2004년 1월 ~ 2004년 3월 : Post-Doctorate Researcher, Institute of Transportation Studies,
University of California, Irvine, CA, USA
2004년 4월 ~ 2006년 2월 : 한국교통연구원 첨단교통기술연구실 책임연구원
2006년 3월 ~ 현재 : 한양대학교 교통공학과 교수



김 태 형 (Kim, Tae-Hyung)

1989년 3월~1993년 2월 : 한양대학교 교통공학과 학사
1993년 3월~1995년 2월 : 한양대학교 대학원 교통공학과 석사
1994년 4월~1998년 7월 : 서울시정개발연구원 도시교통연구부 연구원
1998년 8월~2005년 8월 : University of Maryland, College Park, 토목환경공학과 공학박사
(교통공학 전공)
2005년 8월~현재 : 한국교통연구원 첨단교통연구실 책임연구원



주 신 혜 (Joo, Shin-Hye)

2006년 3월 ~ 2010년 2월 : 한양대학교 교통시스템공학과 공학사
2010년 3월 ~ 현재 : 한양대학교 대학원 교통공학과 석사과정