

퍼지 및 신경망이론을 이용한 도시부 신호교차로 교통사고예측모형 개발

Development of Traffic Accident Frequency Prediction Model in Urban Signalized Intersections with Fuzzy Reasoning and Neural Network Theories

<p>강 영 균 Kang, Young Kyun 김 장 옥 Kim, Jang Wook 이 수 일 Lee, Soo Il 이 수 범 Lee, Soo Beom</p>	<p>현대건설 개발사업본부 인프라팀 과장 · 주저자 (Email : ykkang@hdec.co.kr) 정회원 · 현대해상교통기후환경연구소 연구위원 · 교신저자 (Email : newaxlrose@hotmail.com) 정회원 · 현대해상교통기후환경연구소 연구위원 (Email : sooil99@hanmail.net) 정회원 · 서울시립대학교 교통공학과 교수 (Email : mendota@uos.ac.kr)</p>
--	---

ABSTRACT

This study is to suggest a methodology to overcome the uncertainty and lack of reliability of data. The fuzzy reasoning model and the neural network model were developed in order to overcome the potential lack of reliability which may occur during the process of data collection. According to the result of comparison with the Poisson regression model, the suggested models showed better performance in the accuracy of the accident frequency prediction. It means that the more accurate accident frequency prediction model can be developed by the process of the uncertainty of raw data and the adjustment of errors in data by learning. Among the suggested models, the performance of the neural network model was better than that of the fuzzy reasoning model. The suggested models can evaluate the safety of signalized intersections in operation and/or planning, and ultimately contribute the reduction of accidents.

KEYWORDS

traffic accident frequency model, poisson model, fuzzy reasoning model, neural network theory model, signalized intersections

요지

본 연구는 신호교차로 교통사고예측모형 구축 과정 중 일반적으로 제한된 변수의 선정 및 모형의 구축에만 주로 초점이 맞추어진 기존 방법론의 문제점을 개선하고, 자료조사 및 수집 과정에서 발생하는 자료의 불확실한 상태를 인정하면서 자료의 불확실성을 최소화하여 이용할 수 있는 방법론을 개발하는데 연구의 주안점을 두었다. 퍼지추론이론과 신경망이론을 이용한 모형을 구축하였고, 마지막으로 구축된 퍼지추론이론 모형 및 신경망이론 모형과 기존 회귀모형인 포아송 회귀모형간의 통계적인 검증과 실제 Data를 이용한 모형의 적정성을 검토하였다. 모형의 통계적인 검증시 기존모형에 비해 퍼지추론모형과 신경망이론모형이 더 설명력이 높은 것으로 나타났고, 검증에서도 퍼지추론이론과 신경망이론이 적절한 것으로 나타났으며 기존모형보다 사고건수를 예측하는 설명력이 높은 것으로 입증되었다. 본 연구에서 개발된 모형은 계획 및 운영단계에서 신호교차로의 안전성을 측정하는데 활용될 수 있으며, 궁극적으로는 신호교차로에서 교통사고를 줄이는데 기여할 수 있을 것으로 판단된다.

핵심용어

사고빈도모형, 포아송 모형, 퍼지추론모형, 신경망이론, 신호교차로

1. 서론

1.1. 연구의 배경 및 목적

도로교통체계란 사람, 도로, 및 차량이 복합적으로 어우러

져서 운영된다고 볼 수 있다. 교통사고란 위의 3가지 요소 중
에 한 가지 또는 여러 가지가 불완전할 경우 발생한다고 볼
수 있다. 이제까지 전 세계적으로 교통사고예측 모형은 많이

개발되었다. 그러나 이제까지의 모형은 주로 도로기하구조적 특성과 교통특성 등을 설명 변수로 이용하여 모형을 개발하였으며, 데이터 수집과정에서의 오류에 대한 보정에 관해서는 크게 신경을 쓰지 못한 것이 사실이다. GIGO(Garbage In Garbage Out) 원리에 의하면 수집자료의 신뢰성이 떨어지면 그 자료를 이용한 모형의 결과에 대한 신뢰성도 떨어지게 된다. 즉, 수집자료의 신뢰성 문제를 해결하지 못하면 교통사고의 원인을 정확히 규명하는 일과, 정확한 사고 예측에 한계가 있다는 의미이다. 교통사고 관련 자료 수집 시 조사원의 숙련도 및 조사원간의 차이로 인하여 객관적이며 정확한 자료 수집이 한계가 있는 것이 현실이다. 이러한 경우에 교통사고의 주원인을 명확히 찾아내기가 어렵기 때문에 데이터의 불확실한 상태를 인정하면서 데이터의 불확실성을 최소화하여 이용할 수 있는 방법의 개발이 절실하다. 기존의 방법들은 데이터 수집과정 속에 내포되어 있는 불확실성과 비선형성 및 시공간적 다양성에 의해 발생될 수 있는 오차를 규명하는데 한계가 있다. 이에 본 연구에서는 데이터의 불확실성 상태를 합리적으로 처리할 수 있는 퍼지추론이론(Fuzzy Reasoning Theory) 및 인간의 신경계를 수학적으로 모형화하여 학습에 의한 예측이 뛰어난 것으로 알려져 있는 신경망이론(Neural Network Theory)을 적용한 도시부 신호교차로 교통사고예측 모형을 구축하여 주요 영향요인과 예측력 등을 비교·검토하였다.

1.2. 연구의 방법

기존의 교차로 교통사고 예측모형은 적용되고 있는 변수와 모형의 구조에 있어서 두 가지 문제점을 제기할 수 있다. 첫째로 모형 구축시 조사자료의 중요성 인식 부재에 따른 신뢰성 저하이고, 두 번째 문제점은 변수 선정 및 모형의 형태만을 중시하는 경향에 따른 조사자료의 수집 및 가공과정에 대한 신뢰성 부족이다. 이러한 문제점을 해결하기 위하여 자료의 불확실성 상태를 합리적으로 처리할 수 있는 퍼지추론이론과 자료의 학습과 반복을 통한 안정화 상태를 유지할 수 있는 신경망이론을 이용하여 모형을 구축하는 대안을 설정하였다. 대안별 모형을 구축하기 위해 필요한 교통사고자료와 신호교차로의 교통, 도로 및 주변 환경 조건에 대한 영향변수 자료를 수집하였다. 수집한 교통사고자료와 영향변수자료간의 상관분석을 통해 교통사고와의 연관성을 검증하고, 다중공선성을 고려하여 모형구축을 위한 영향변수를 선정하였다. 선정된 영향변수를 이용하여 지능형 시스템이론인 퍼지추론과 신경망이론을 적용하여 모형을 구축하였다. 본 연구에서 구축된 퍼지추론모형과 신경망이론모형은 두 단계에 걸쳐 검증이 이루어졌다. 1단계에서는 기존방법론인 선형 및 비선형 회귀분석 모형과의 비교를 통하여 적합성을 검증하였다. 2단

계에서는 본 연구에서 모형정산(Calibration)을 위하여 사용된 사고자료 Database 이외에 2차 검증을 위한 별도의 자료를 본 연구에서 개발된 모형과 타 기존 모형을 적용하여 비교함으로써 보다 신뢰성 있는 모형의 적합도 검증을 수행하였다. 본 연구의 방법은 기존문제분석, 대안설정, 조사자료 수집, 영향변수 선정, 모형개발, 모형검증의 총 6단계로 이루어지며 이를 그림 1에 도식화 하여 나타내었다.

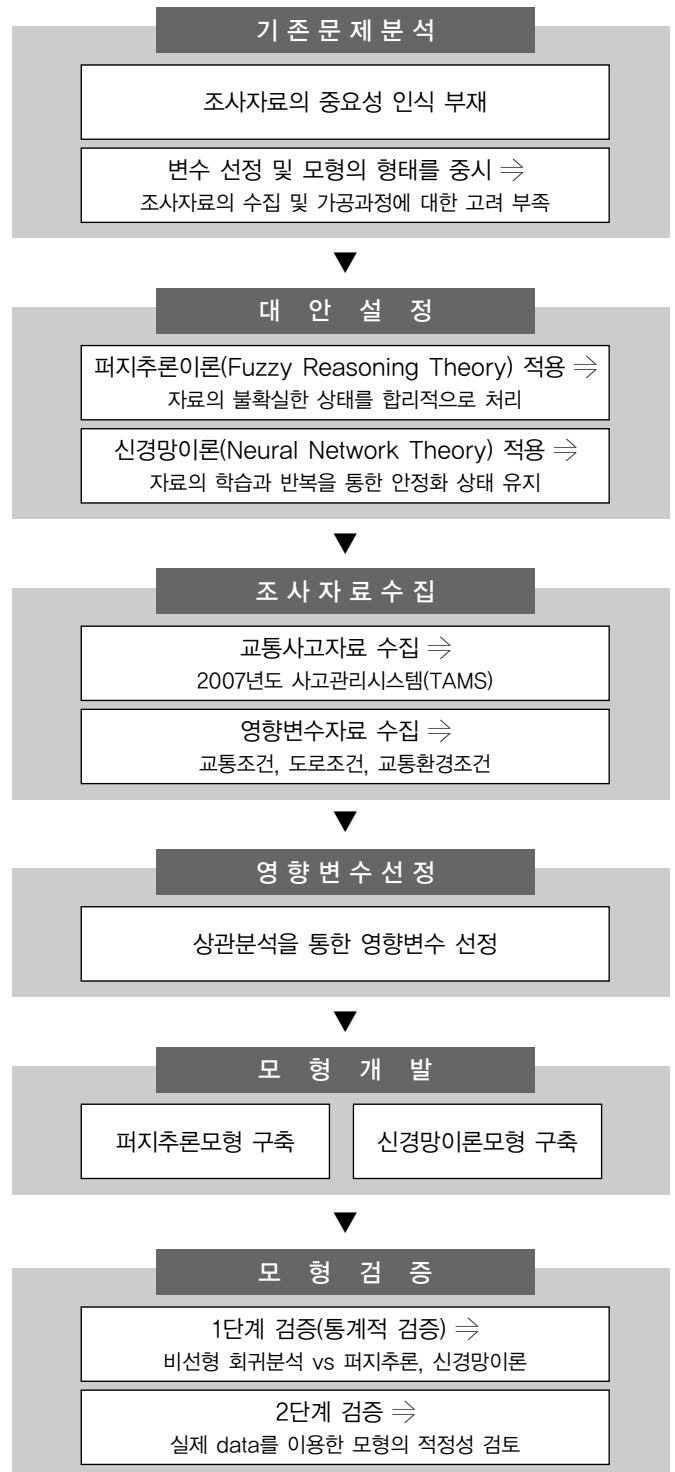


그림 1. 연구의 방법 절차

2. 기존연구 고찰

2.1. 신호교차로 교통사고 영향변수

신호교차로에서 발생하는 교통사고요인은 크게 교통류의 특성과 관계된 교통조건과 교차로의 기하구조 및 환경과 관련된 도로환경조건으로 나누어질 수 있다.

2.1.1. 교통조건

James A. Bonneson(1993)의 연구에서는 125개 교차로의 교통량과 사고 데이터를 분석한 결과 양방향 정지 교차로의 교통량과 사고와의 관계를 음지수 분포로 분석하여 주도로와 부도로 교통량과 사고와의 관계를 정의함으로써 사고예측모형을 개발하였다. 주도로, 부도로 교통량이 많아질수록 사고는 지수관계의 형태로 증가하게 되며, 사고율은 주도로의 교통량 보다 부도로 교통량에 더 민감하게 영향을 받는 것으로 나타나고 있다. 김효종(1997)은 교차로 및 횡단보도 교통사고 예측모형에 관한 연구를 통해 교차로 교통량의 증가가 교통사고와 연관성이 있음을 보여주었다.

Sheffer와 Janson(1999)의 연구에서는 신호현시 방법 중 선행 좌회전 현시와 후행 좌회전 현시방법에 따라 교차로 사고율의 영향도를 판정하였다.

2.1.2 도로환경조건

James A. Bonneson(1993) 연구에서는 접속도로의 중앙분리대 설치로 양방향 좌회전 차선으로 분리되었을 때와 분리되지 않았을 때를 구분하여 각각의 조건에 따른 사고예측모형을 개발하였다. 이 연구에서는 같은 교통량을 가지고 있다는 전제 하에서 중앙이 분리되지 않은 교차로의 사고건수가 가장 많았고, 그 다음으로 두 방향 좌회전 차선으로 분리했을 경우, 중앙분리시설로 분리했을 경우의 교차로 순으로 사고건수가 적어지는 것을 발견하였다.

임운택(1993)은 서울시 교통사고 많은 지점 147개소를 대상으로 교통량, 횡단보도수, 교차각 등의 도로 특성이 교통사고에 미치는 영향을 분석한 결과 전체적으로 교통량이 많고, 교차로의 횡단보도 수가 많은 지점에서 사고가 많이 발생하는 것으로 나타났다.

Hoong Chor Chin과 Mohammed Abdul Quddus(2001)의 연구에서는 싱가포르의 52개 4지 신호교차로의 교통사고건수, 도로기하구조, 교통류 통제 특성 관계에 대해 연구하였는데, 조사된 32개의 설명변수 중 교차로 교통사고에 영향을 미치는 변수는 버스정류장, 단속카메라, 접근교통량, 회전교통량, 좌회전 차로수 등으로 분석되었다.

2.2. 기존교통사고 예측모형

교통사고 예측모형 개발과 관련된 국내외 기존 연구는 비

선형회귀분석을 이용한 연구가 주를 이루었다. 사고에 영향을 미치는 다양한 변수를 고려하여 연구가 진행되어 왔으나, 자료의 수집 및 가공과 관련한 자료의 수준 특성을 고려하지 않고 단순·다중회귀분석을 이용한 연구들이 주를 이루고 있다. 또한 변수선정에 있어서 교통량, 중차량비율, 교통섬과 같은 변수에 국한된 연구들이 주를 이루고 있다.

표 1. 국내 신호교차로 관련 교통사고 예측모형

연구자	분석방법	변 수
김효종 (1997년)	주성분분석	• 교통량, 중차량비율 • 교통섬 수
하태준 (2001년)	교차분석 다중회귀분석	• 방향별 교통량
김원철 (2001년)	다중회귀분석 수량화이론	• 교통량, 주행속도 • 종단경사, 교통섬

표 2. 국외 신호교차로 관련 교통사고 예측모형

연구자	분석방법	변 수
A. Bonneson (1993년)	비선형회귀모형	• 주도로/부도로 교통량
Seffer, Janson (1999년)	범주형회귀모형	• 신호현시
Hoong chor chin(2001년)	비선형회귀모형	• 교통량, 좌회전 차로수 • 버스정류장, 단속카메라

따라서 본 연구에서는 기존의 교통량 위주의 분석결과 한계에서 벗어나도록 자료의 불확실성과 가공과정의 오류를 최소화하는 방법인 퍼지추론 및 신경망 이론을 적용하였다.

3. 자료의 수집

퍼지 및 신경망이론을 이용한 도시부 신호교차로 교통사고 예측모형 구축에 사용될 기초자료를 수집한다. 자료의 수집은 문헌조사 결과를 근거로 조사항목을 결정하고 현장조사를 통해 수집한 후, 이를 Database화 한다.

3.1. 조사지점의 선정

본 연구에서 조사지점 선정을 위한 표본추출방법은 비확률 표본추출방법인 판단표본추출방법(Judgement Sampling)을 이용하여 표본을 선정하였고 조사지점의 선정기준은 표 3과 같다.

표 3. 조사지점(신호교차로)의 선정기준

<ul style="list-style-type: none"> • 기 준 1 : 신호교차로 여부(교통사고 수와 상관 없음) • 기 준 2 : 주·부도로 구분 여부 및 주도로 4차로 이상 여부 • 기 준 3 : 지역적으로 도시부 특성의 존재 여부

조사 표본수의 결정은 분산 값을 추정하기가 곤란하므로 모비율에 대한 사전정보가 없는 경우를 적용하여 표 4와 같이 최대허용오차(10%)와 신뢰수준(95%)을 고려하여 표본수 96개를 선정하였다.

표 4. 표본수 선정

신뢰수준	최대허용오차(%)	20	15	10	5	2.5
90%일 때	필요한 표본수(개)	16	30	67	270	1,081
95%일 때		24	42	96	384	1,534
98%일 때		33	60	135	542	2,166

3.2. 조사방법

본 연구에서 조사대상 교차로는 위치 및 교통특성이 도시부의 전형적인 모습을 나타낸다고 볼 수 있는 서울 및 수도권과 부산권의 교차로 중 4지 신호교차로를 대상으로 하였다. 조사대상 지점수는 판단표본추출방법에 따라 결정된 96개 지점과 모형의 검증에 위해 사용할 10개 지점을 선정하였다. 그리고 기존문헌 고찰을 통하여 교차로 교통사고에 영향을 미치는 것으로 규정되어 있는 다양한 요소를 검토하여, 교통조건과 도로환경조건을 포함한 54개의 항목을 선정하였다. 교통사고자료는 교통사고관리시스템(TAMS)*에 있는 2007년도 교통사고 자료와, 도로교통공단의 사고충돌도 자료를 활용하며, 공단에서 자체 조사한 『2007년 전자교통신호체계 운영 및 교통량자료』를 이용하여 신호운영 및 교통조건을 파악하였다.

4. 변수의 선정

4.1. 교통사고와 교통사고요인과의 상관분석

신호교차로 사고예측모형을 개발하기 위해 교통사고빈도를 종속변수로, 신호교차로에 대한 교통조건, 도로조건, 교통환경조건으로 구축된 교통사고요인 자료를 독립변수로 하여 상관분석을 수행하였다. 이때 변수의 유의성을 검증하기 위한 신뢰수준은 95%($\alpha=0.05$)로 하였다. 그 결과 신뢰수준내에서 총 18개의 교통사고와 상관이 있는 요인을 찾아내었다. 교통사고와 상관이 있는 요인으로는 주·부도로 교통량, 차로수, 부도로 우회전 시거 등이다. 교통사고와의 상관성이 높은 변수를 적용한 모형이 교통사고예측모형의 신뢰성을 보다 제고할 수 있다는 가정하에 변수를 선정하였다. 표 5에서 제시된 18개 변수 중 상관계수가 0.3 이상인 변수를 살펴보면 부도로 교통량, 주도로 차로수, 주도로 교통량, 교차로 넓이, 부도로 딜레마구간 길이를 포함하여 총 10개의 영향변수가 이

에 해당된다. 선정된 10개의 변수를 이용하여 모형을 구축할 경우, 상관관계가 높은 독립변수들간에 발생하는 다중공선성(Multi-Collinearity)의 문제가 생기게 된다. 따라서 상관관계가 높은 변수가 존재할 경우에는 이 중 하나의 변수만을 이용하든지 두 변수를 합하여 다른 새로운 변수를 생성한다면 가 하는 변수변환의 작업을 거쳐 처리를 해주어야 합리적인 모형이 구축된다.

표 5. 상관계수가 높은 독립변수

영향변수	상관계수	유의확률	영향변수	상관계수	유의확률
부도로교통량	0.52	0.00	주기길이	0.31	0.00
주도로차로수	0.48	0.00	주도로지형	0.30	0.00
주도로교통량	0.44	0.00	부도로차로교통량	0.24	0.02
교차로넓이	0.41	0.00	부도로제한속도	0.23	0.03
딜레마구간길이(부)	0.41	0.00	부도로좌회전신호	0.22	0.03
부도로차로수	0.39	0.00	주도로유출입구	0.21	0.04
주도로횡단보도	0.38	0.00	부도로우회전시거	-0.25	0.02
부도로횡단보도	0.38	0.00	부도로좌회전시거	-0.27	0.01
딜레마구간길이(주)	0.33	0.00	주도로좌회전시거	-0.30	0.00

본 연구에서는 첫째, 우선 영향변수와 교통사고와의 상관성이 높은 변수를 선택하고 둘째, 변수들 간의 상관분석을 수행하여 변수간의 다중공선성이 발생하지 않는 변수를 가지고 모형을 구축한 후, 마지막으로 적용된 변수의 유의성을 검토하여 변수를 선정하고자 한다.

변수간의 상관분석을 수행한 결과는 표 6과 같다. 교통사고 모형에 적용될 영향변수의 선정을 위해 교통사고와 상관계수가 큰 순서대로 영향변수간의 상관관계를 살펴보면 다음과 같다.

- (a) 부도로 교통량의 경우는 주도로 교통량, 부도로 차로수, 주도로 횡단보도, 부도로 횡단보도, 주도로 딜레마구간 길이와 상관성이 높은 것으로 나타났고,
- (b) 주도로 교통량의 경우는 부도로 교통량, 주도로 차로수, 주기길이, 부도로 딜레마구간과의 상관성이 높으나, 그 중 부도로 교통량이 교통사고와의 연관성이 더 높아 주도로 교통량은 변수선정 시 제외하였다.
- (c) 주도로 차로수의 경우는 부도로 차로수, 주기길이, 주도로 딜레마구간 길이와 상관성이 높은 것으로 나타났으며,
- (d) 교차로 넓이의 경우도 주도로 교통량, 부도로 차로수, 주도로 딜레마구간 길이와 상관성이 높았다.
- (e) 부도로 딜레마구간의 길이도 주도로 교통량과 주도로 딜레마구간의 길이와 상관성이 높은 것으로 나타났다.

이러한 과정을 거쳐서 부도로 교통량, 주도로 차로수, 교차로 넓이, 부도로 딜레마구간 길이 등 총 4개의 영향변수를 선정하였다.

* 교통사고관리시스템(TAMS : Traffic Accident Management System)

표 6. 1차 선정된 설명변수간 상관분석 결과

구 분		주도로 교통량	부도로 교통량	주도로 차로수	부도로 차로수	주도로 횡단보도	부도로 횡단보도	주기길이	교차로 넓 이	주도로 딜레마	부도로 딜레마
주도로 교통량	상관계수	1	0.564	0.876	0.195	0.136	0.142	0.330	0.322	0.282	0.375
	유의수준		0.000	0.000	0.057	0.187	0.167	0.001	0.001	0.005	0.000
부도로 교통량	상관계수		1	0.207	0.627	0.376	0.410	0.136	0.346	0.424	0.326
	유의수준			0.033	0.000	0.000	0.000	0.186	0.000	0.000	0.001
주도로 차로수	상관계수			1	0.554	0.220	0.162	0.425	0.340	0.482	0.317
	유의수준				0.000	0.031	0.116	0.000	0.000	0.000	0.000
부도로 차로수	상관계수				1	0.121	0.199	0.227	0.827	0.642	0.450
	유의수준					0.241	0.052	0.026	0.000	0.000	0.000
주도로 횡단보도	상관계수					1	0.326	-0.006	0.122	0.048	0.145
	유의수준						0.000	0.951	0.237	0.644	0.160
부도로 횡단보도	상관계수						1	-0.015	0.184	0.085	0.140
	유의수준							0.886	0.073	0.412	0.175
주기길이	상관계수							1	0.352	0.198	0.159
	유의수준								0.000	0.053	0.121
교차로 넓 이	상관계수								1	0.666	0.398
	유의수준									0.000	0.000
주도로 딜레마	상관계수									1	0.704
	유의수준										0.000
부도로 딜레마	상관계수										1
	유의수준										

4.2. 영향변수와 신호교차로 교통사고

선정된 영향변수와 신호교차로 교통사고와의 관계를 검증하기 위해 신호교차로에서 발생하는 교통사고 유형과 변수와의 연관성을 검토해 보았다. 영향변수의 선정과정에서 제시된 54개의 변수 중에는 교통사고의 유형보다는 사고의 심각도에 더 많은 영향을 미치는 변수가 존재하였다. 본 연구에서는 사고의 빈도예측을 위한 모형을 구축하는 것이 목적이므로 주로 사고의 심각도에 영향을 미치는 변수는 배제하고자 한다.

표 7. 변수와 교통사고와의 연계성

	직각 충돌	좌-정면	추 돌	우-추돌	보행자	측 면
부도로 교통량	○		▲	○		○
주도로 차로수		▲		○	○	
교차로 넓이	○	○			▲	○
부도로 딜레마			○	▲		

※ ○ : 연관성이 높음 / ▲ : 연관성이 있음

신호교차로에서 발생하는 교통사고의 유형은 크게 직각 충돌, 좌회전정면 충돌, 추돌, 우회전 추돌, 보행자, 측면 사고의 5가지로 분류할 수 있는데, 이 모든 유형을 반영할 수 있

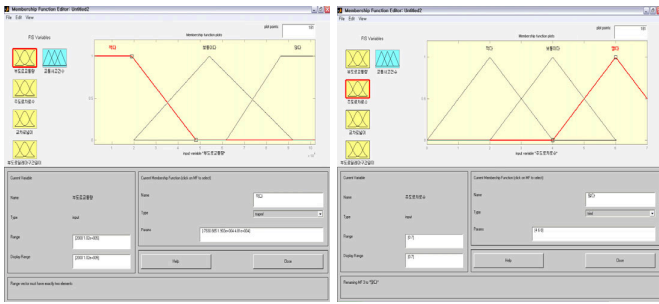
는 모형이어야 신호교차로의 교통사고예측에 적절할 것이다. 표 7에 나타난 바와 같이 선정된 변수와 사고유형간의 관계를 검토한 결과 제시된 영향변수가 신호교차로의 교통사고 유형을 잘 반영하는 것으로 나타났다.

5. 모형 구축

5.1 퍼지추론을 이용한 교통사고 예측모형

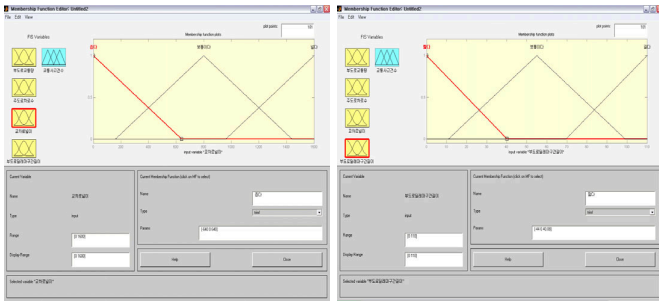
5.1.1. 퍼지추론 적용을 위한 멤버쉽 함수 작성

퍼지추론을 적용하기 위해서 입력변수로는 부도로 교통량, 주도로 차로수, 교차로면적, 부도로 딜레마구간길이를 사용하였다. 출력변수로는 교통사고건수를 이용하여 멤버쉽 함수를 작성하였다. 부도로 교통량은 연평균 일교통량을 이용하여 {적다, 보통이다, 많다}인 3개의 멤버쉽 함수로, 주도로 차로수도 {적다, 보통이다, 많다}로, 교차로 면적은 {넓다, 보통이다, 좁다}로, 부도로 딜레마구간 길이는 {길다, 보통이다, 짧다}로 멤버쉽 함수를 작성했으며, 교통사고건수는 {매우적다, 적다, 보통이다, 많다, 매우많다} 다섯 가지로 멤버쉽 함수를 작성하였다.



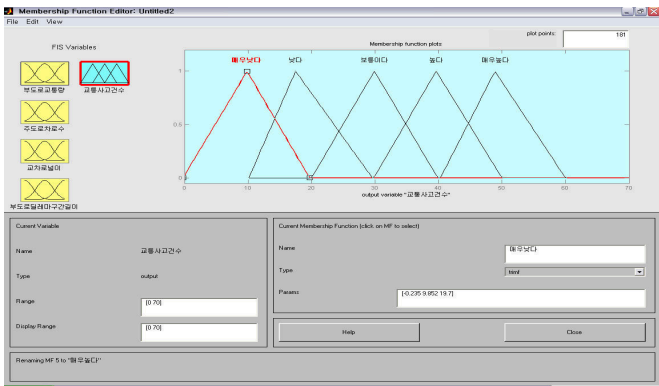
부도로 교통량

주도로 차로수



교차로 넓이

부도로 딜레마구간 길이



교통사고건수

그림 2. 퍼지추론 적용을 위한 멤버십 함수

5.1.2. 연산방법

퍼지추론의 연산을 위해 작성된 81가지의 퍼지추론 규칙의 예를 나타내면, “부도로 교통량은 많고 주도로 차로수가 많으며, 교차로 면적이 넓고, 부도로 딜레마구간 길이가 길면 교통사고건수는 매우 높아진다.”와 같은 퍼지추론 규칙을 작성하였다.

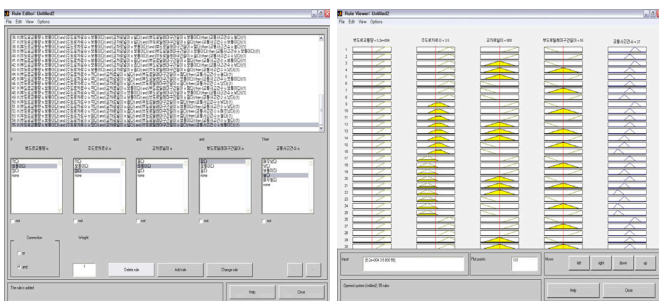


그림 3. 교통사고건수 예측 규칙

퍼지추론 방법으로는 Mamdani의 Min-Max중심법 (Centroid)을 적용하여 계산하였다. 또한, 각 변수에 대한 멤버십 함수의 범위와 중복도는 전문가의 경험을 이용하여 각 데이터의 평균값과 50% 중복도를 사용하였다. 이와 같은 방법에 의하여 퍼지추론에 의한 교통사고 예측모형을 추론하였다. 퍼지추론모형의 평균값은 적합성을 나타내는 것으로, 식 (1)과 같다.

$$\Pi(average) = \frac{\sum_i^n |R_i - P_i|}{n} \quad (1)$$

여기서, n : 자료의 수, R_i : 실측값, P_i : 예측값

표 8을 보면 모형 설명력을 나타내는 P_i 값은 2.64이었고, MSE 는 0.60로 분석되었다. 그리고 퍼지추론을 이용한 모형의 결정계수(R^2)값은 0.741로 포아송회귀모형을 이용하였을 때보다 높게 나타났다.

표 8. 퍼지추론을 이용한 모형 구축결과

설명변수	
부도로 교통량(적다, 보통이다, 적다)	
주도로 차로수(적다, 보통이다, 적다)	
교차로 면적(적다, 보통이다, 적다)	
부도로 딜레마구간 길이(적다, 보통이다, 적다)	
P_i (평균)	2.64
MSE	0.60
적중률(%)	62.00
R^2	0.74

5.2. 신경망을 이용한 교통사고 예측모형

본 연구에서 신경망 모형을 학습시키기 위한 학습자료 (Learning Data)로서 부도로 교통량, 주도로 차로수, 교차로 면적, 부도로 딜레마구간 길이를 사용하였다. 모든 학습자료는 전처리과정(Pre-process)인 정규화(Normalization)를 한다. 즉, 평균과 표준편차가 [0,1]인 정규화된 자료로 변환하였다. 이는 후에 후처리 과정(Post-process)을 거쳐 실제 값으로 다시 변환되어진다. 모형을 학습시키기 위한 방법으로 일반적으로 널리 응용되고 있는 오차 역전파 알고리즘을 적용하여 모형에 따른 특성을 분석하였다. 예측치와 실측치를 비교하여 적중률을 알아보았다. 신경망 모형은 각각의 변수들을 적용하여 구성하였다. 입력층의 노드(Node)의 수를 n 이라 하면, 은닉층의 노드 수는 n 부터 $6n$ 까지의 변화를 시키면서 최적의 신경망모형을 선별하였다. 또한 은닉층 처리소자의 개수에 따른 학습효과를 분석하기 위해 동일한 반복횟수 10,000회에 대해 각 모형별로 모의하였다. 이는 10,000회 이상의 기간으로 설정할 경우, 최대 반복횟수까지

훈련이 반복되어 모형의 일반화 및 효율성이 떨어지는 것으로 나타났으나 10,000회 정도의 기간에서 가장 좋은 결과를 얻을 수 있었다. 따라서 본 연구에서는 모형에 대한 선정기간을 10,000회의 기간으로 설정하였다. 신경망의 보정을 위한 훈련에 사용된 학습방법은 알고리즘 적용시 오차가 지역 최소값(Local Minima)에 수렴하는 것을 방지하고 학습의 효율을 높이기 위해 모멘텀-적응식 학습를 방법을 사용하였다. 모멘텀 상수와 초기 학습률은 민감도 분석을 통해 얻었고 모든 모형에서 각각 0.1를 일괄적으로 사용하였다. 일련의 과정을 거쳐 얻어진 모형들의 평가기준은 수치적 기준을 적용하여 살펴보았다. 수치적기준인 MSE(평균제곱오차)는 모형 적합도의 전체적인 우수성을 보여주기 위한 것으로 식 (2)와 같다. MSE값이 0이라면 관측치와 예측치가 완전히 일치하는 경우라고 할 수 있다.

$$MSE = \frac{\sum_{i=1}^n (obs_i - exp_i)^2}{n} \quad (2)$$

여기서, n : 자료의 수, obs_i : 관측값, exp_i : 출력값

표 9를 보면 모형 설명력을 나타내는 R^2 값은 1.91이었고, MSE는 0.30으로 분석되었다. 그리고 신경망을 이용한 모형의 결정계수(R^2)값은 0.83으로 앞의 2가지 모형을 이용하였을 때보다 높게 나타났다.

표 9. 신경망을 이용한 모형 구축결과

설명변수	
부도로 교통량, 주도로 차로수 교차로 면적, 부도로 딜레마구간 길이	
R^2 (평균)	1.91
MSE	0.30
적중률(%)	72.28
R^2	0.83

6. 모형 검증

6.1. 기존 방법론에 의한 모형 구축

기존 신호교차로 교통사고 예측모형 구축에 쓰이는 방법 중 교차로 교통사고 예측모형에 가장 적합하다고 판단되는 모형을 검토하였다. 일반적으로 다중선형회귀분석 모형과 비선형회귀분석 모형 중 포아송 회귀모형이 가장 일반적으로 많이 쓰이고 있으나, 적합도 면에서는 포아송 회귀모형이 더 좋은 것으로 판단되어 비선형 회귀분석 중 포아송 회귀모형을 적용하였다. 4개의 영향변수(부도로 교통량, 주도로 차로수, 교차로 넓이, 부도로 딜레마구간 길이)를 가지고 구축된 모형은 식(3)으로 표현된다.

$$A_{main} = \exp \left\{ \begin{aligned} & -1.9882 + (0.3734 \times X_1) + (0.0827 \times X_2) \\ & + (0.0002 \times X_3) + (0.0087 \times X_4) \end{aligned} \right\} \quad (3)$$

여기서, X_1 = 부도로 교통량(대/일, Natural Logarithm 대입)

X_2 = 주도로 차로수(lane)

X_3 = 교차로 넓이(m²)

X_4 = 부도로 딜레마구간 길이(m)

표 10. 포아송 모형의 설명 자료

변 수	개발모형		
	Coeff.	Std.Err.	P-value
상 수	-1.99	0.44	0.00
부도로 교통량	0.37	0.04	0.00
주도로 차로수	0.08	0.01	0.00
교차로 넓이	0.00	0.00	0.02
부도로 딜레마구간 길이	0.01	0.00	0.00
LL($\hat{\xi}$)	-446.7546		
LL(0)	-591.4069		
ρ^2	0.2445		
MPB/year	-6.8431		
MAD/year	10.6328		

6.2. 모형간 결과값 비교

선정된 변수로 구축된 모형간의 비교를 위하여 다항로짓모형의 유효성 지표인 적합도와 퍼지추론모형의 유효성 지표인 R^2 값 그리고 신경망 이론 모형의 유효성 지표인 MSE(평균 제곱오차: Mean Square Error)를 이용하여 모형의 적합성을 비교하였다. 표 11의 부도로 교통량, 주도로 차로수, 교차로 넓이, 부도로 딜레마구간 길이를 변수로 하여 구축한 교통사고건수모형을 보면 퍼지추론에 의한 모형의 적중률은 기존의 포아송 회귀모형에 비하여 높은 적중률을 보이고 있으며, 퍼지모형의 평가지표인 평균 R^2 값도 포아송 회귀모형에 비하

표 11. 추정된 모형의 비교

변 수	포아송회귀모형		퍼지추론모형	신경망모형
	Coeff	P-값	변 수	
상수	-1.99	0.00	① 부도로 교통량 ② 주도로 차로수 ③ 교차로 넓이 ④ 부도로 딜레마구간	
부도로 교통량	0.37	0.00		
주도로 차로수	0.08	0.00		
교차로 넓이	0.00	0.02		
부도로딜레마구간	0.01	0.00		
R^2 (평균)	3.30		2.64	1.91
적중률(%)	30.00		62.00	72.28
MSE	2.00		0.60	0.30
R^2	0.68		0.74	0.83

여 낮게 나타나고 있다. 또한 신경망이론에 의한 모형의 적중률은 72.28%로 나타나 포와송 회귀모형과 퍼지추론모형에 비하여 아주 높은 적중률을 보이고 있으며, 신경망모형의 평가지표 MSE(평균제곱오차)가 포와송 회귀모형과 퍼지추론모형과 비교하면 아주 낮게 나타나고 있다. 그러므로 신호교차로에서의 도로기하구조와 교통운영특성을 변수로 교통사고건수를 예측하는 모형으로써 지능형 이론을 사용한 모형이 높은 적중률을 나타내고 있다고 볼 수 있다.

6.3. 2차 검증

6.3.1. 모형별 사고빈도 예측결과

교통사고예측모형의 적정성을 2차적으로 검증하기 위하여 모형 구축시 조사하였던 96개 지점과는 별개로 추가적으로 10개의 조사지점을 선정하여 실제 교통사고건수와 각 예측모형별 추론결과를 표 12와 그림 4에 나타내었다. 실제 교통사고건수와 3가지 이론에 의해 구축된 예측모형의 구현결과를 통해서, 퍼지추론 및 신경망에 의해 구축된 사고건수예측모형이 사고의 재현성면에서 기존에 사용되어온 선형·비선형 회귀모형으로 만들어진 모형보다 그 적용성이 유용하다는 것으로 판단할 수 있다. 모형별 교통사고 예측결과 신경망 모형이 가장 높았으며, 실제사고건수 대비 40%정도 정확히 예측하였다.

표 12. 모형별 교통사고 예측 결과

검증지점	실제사고건수	포와송 모형	퍼지추론 모형	신경망 모형	
1	충랑구청사거리	11	17	12	11
2	퇴계로5가사거리	17	20	18	16
3	서울병원사거리	11	9	10	10
4	신양초교사거리	18	16	18	18
5	방이동사거리	8	9	8	7
6	일산병원사거리	16	21	17	16
7	용화사사거리	21	25	23	24
8	호계2동사거리	16	21	18	17
9	셋별사거리	27	26	28	26
10	모라사거리	15	19	16	15

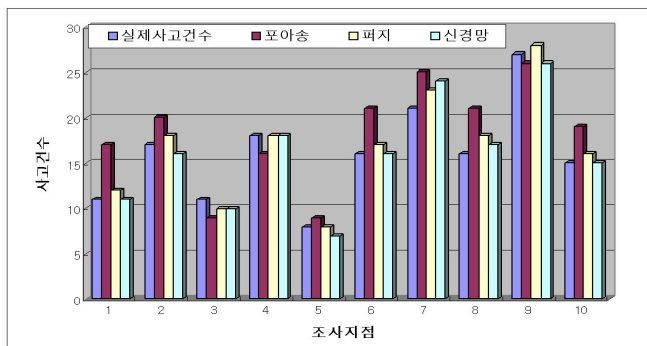


그림 4. 실제사고건수와 모형예측결과 비교

6.3.2 예측결과에 대한 통계적 검증

앞에서 제시된 실제사고건수와 각 모형에 따라 산정된 사고건수에 대한 동질성 검증을 위하여 카이스퀘어 검증(X^2-Test)을 수행하였다. 본 검증의 가설은 실제사고건수와 예측사고건수가 유의수준 안에서 동일하다고 판단할 수 있는지 여부를 알기 위함이다.

$$H_0 : \text{실제사고건수} = \text{예측사고건수}$$

$$H_1 : \text{실제사고건수} \neq \text{예측사고건수}$$

$$x^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \geq x^2((r-1)(c-1), \alpha) \quad (4)$$

$$\hat{o}_{ij} = \text{관측도수}, \hat{E}_{ij} = \text{기대도수}$$

표 13. 실제사고건수와 각 모형간 카이스퀘어 결과

변 수	포와송모형	퍼지모형	신경망모형
검증방법	X^2-Test		
자유도	9	9	9
유의수준	$\alpha = 0.05$		
X_2	2.56	1.78	0.81

X^2 검증 결과 신경망 모형의 X^2 값이 가장 낮아 기무가설이 채택될 확률이 제일 높은 것으로 나타났다.

7. 결론

본 연구에서는 신호교차로 교통사고 예측모형 구축과정 중 일반적으로 제한된 변수의 선정 및 모형의 구축에만 주로 초점이 맞추어진 기존 방법론의 문제점을 개선하고, 자료조사 및 수집과정에서 발생하는 자료의 불확실한 상태를 인정하면서 자료의 불확실성으로 인한 문제를 최소화할 수 있는 방법론을 개발하고자 하였다. 이를 위하여 첫째로 변수의 선정의 합리성을 제고하기 위하여 도로조건, 교통조건, 주변환경조건 등을 고려한 총 54개의 다양한 영향요인을 고려하여 분석하였다. 교통사고와 변수간의 상관관계를 분석하여 18개의 영향요인을 선정하였다. 그리고 신호교차로 교통사고와의 상관정도 및 변수간의 상관관계를 분석하여 부도로 교통량, 주도로 차로수, 교차로 넓이, 부도로 딜레마구간 길이 등 4개의 영향변수를 선정하였다. 두 번째로는 자료 수집과정에서 발생하는 불확실성을 극복할 수 있는 방법론으로서 퍼지추론모형과 신경망이론모형을 적용하였다. 그리고 기존에 가장 많이 적용되고 있는 포와송 회귀모형과의 비교를 통하여 이 모형의 적합성을 검증하였다. 그 결과 포와송 회귀모형보다는 퍼지추론모형 및 신경망이론모형이 보다 적합한 것으로 나타

났다. 이는 조사자료의 불확실성을 합리적으로 처리하고, 자료의 학습을 통한 조사자료의 가공을 통해 조사자료의 오류를 보정하여 모형을 구축할 경우 보다 사고빈도예측의 정확성이 뛰어난 모형을 얻을 수 있다는 것을 의미한다. 그리고 신경망이론모형이 퍼지추론모형보다 설명력이 높은 모형으로 나타났다. 본 연구에는 교통조건, 도로환경조건 등만을 고려하여 변수가 선정되었으나, 향후에는 이외에도 운전자에 대한 인적요소와 같이 본 연구에서 고려하지 못한 주요 영향 변수에 대한 연구도 필요할 것이다.

참고 문헌

김원철, 외(2001), “교차로 안전진단을 위한 교통사고건수예측모형화 수법에 관한 연구”, *대한토목학회지* 제21권 제4호.

김장욱 외 3인(2006), “퍼지 및 신경망 이론을 이용한 교통사고예측모형 개발에 관한 연구”, *대한교통학회지* 제24권 제7호

김효종(1997), “교통사고 유형에 미치는 영향요인 분석에 관한 연구”, 전남대학교 대학원

임윤택, “도로특성이 교통사고에 미치는 영향분석”, 연세대학교 대학원 건축공학과, 1993

하오근(2005), “국도변 신호교차로 안전성 평가를 위한 사고예측모형 개발과 심각도 분석”, 한양대 대학원

하태준 외 3인(2001), “신호교차로 교통사고예측모형의 개발 및 적용(광주광역시 4지신호교차로를 중심으로)”, *대한교통학회지*, 제19권 제6호 대한교통학회

홍정열(2002), “신호교차로에서의 사고예측모형개발 및 위험수준 결정 연구”, 한양대학교 대학원

Chin, H. C., and Quddus, M.A. (2002). , “Applying the Random Effect Negative Binomial Model to Examine Traffic Accident Occurrence at Signalized Intersections” , *Accident Analysis and Prevention 34*

Daniel, J., Tsai, C., and Chien, S. (2002). , “Factors Influencing Truck Crashes on Roadways with Intersections” , *Paper submitted to the Transportation Research Board, Washington, D.C.*

Hoong Chor Chin, Mohammed Abdul Quddus(2001), “Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersection” , *November, Accident Analysis & Prevention*

James A. Bonneson(1993), “Estimation of safety at two-way stop-controlled intersections on rural highway” , *TRR 1401*

Maher, M. J. and Summersgill, I. (1996). A comprehensive Methodology for the Fitting of Predictive Accident Models. *Accident Analysis & Prevention 28(3)*, pp 281-296

Miaou, S. P. Hu, Wright, T., Davis, S.C., and Rathi, A.K. (1993), “Development of Relationship between Truck Accidents and Geometric Design” , *FHWA-RD-91-124*, Federal Highway Administration, Washington, D.C.

Retting, R., Williams, A. F., Prusser, D. F., and Weinstein, H. B. (1995), “Classifying Urban Crashes for Countermeasure Development” , *Accident Analysis & Prevention 27(3)*, pp 283-294

Sheffer and Janson(1999), “Accident and capacity comparisons of leading and lagging left-turn signal phasings” , January, *Transportation Research Board*, National Research Council, Washington, DC

접 수 일 : 2010. 8. 11
 심 사 일 : 2010. 8. 12
 심사완료일 : 2011. 2. 21