

역전파 알고리즘을 이용한 최적의 교통안전 평가 모형개발

김중효* · 권성대** · 홍정표*** · 하태준****

Kim, Joong-Hyo*, Kwon, Sung-Dae**, Hong, Jeong-Pyo***, Ha, Tae-Jun****

Development of Optimum Traffic Safety Evaluation Model Using the Back-Propagation Algorithm

ABSTRACT

The need to remove the cause of traffic accidents by improving the engineering system for a vehicle and the road in order to minimize the accident hazard. This is likely to cause traffic accident continue to take a large and significant social cost and time to improve the reliability and efficiency of this generally poor road, thereby generating a lot of damage to the national traffic accident caused by improper environmental factors. In order to minimize damage from traffic accidents, the cause of accidents must be eliminated through technological improvements of vehicles and road systems. Generally, it is highly probable that traffic accident occurs more often on roads that lack safety measures, and can only be improved with tremendous time and costs. In particular, traffic accidents at intersections are on the rise due to inappropriate environmental factors, and are causing great losses for the nation as a whole. This study aims to present safety countermeasures against the cause of accidents by developing an intersection Traffic safety evaluation model. It will also diagnose vulnerable traffic points through BPA (Back-propagation algorithm) among artificial neural networks recently investigated in the area of artificial intelligence. Furthermore, it aims to pursue a more efficient traffic safety improvement project in terms of operating signalized intersections and establishing traffic safety policies. As a result of conducting this study, the mean square error approximate between the predicted values and actual measured values of traffic accidents derived from the BPA is estimated to be 3.89. It appeared that the BPA appeared to have excellent traffic safety evaluating abilities compared to the multiple regression model. In other words, The BPA can be effectively utilized in diagnosing and practical establishing transportation policy in the safety of actual signalized intersections.

Key words : Signalized intersection, Traffic conflict, Traffic accident, Multiple regression model, Back-propagation algorithm

초록

교통사고 피해를 최소화하기 위해서는 차량과 도로 체계에 대한 공학적인 개선을 통하여 교통사고 원인을 제거해야 한다. 일반적으로 안정성과 효율성이 부족한 도로는 교통사고가 지속적으로 발생할 가능성이 크고 이를 개선하는데 막대한 사회적 비용과 시간이 소요되며, 부적절한 환경 요인으로 발생한 교통사고는 국가적으로 큰 피해를 발생시키게 된다. 따라서 본 연구는 최근 인공지능 분야 중 활발히 연구 중인 역전파 알고리즘(Back-Propagation Algorithm : BPA)을 이용하여 신호교차로를 대상으로 최적의 교통안전 평가기법을 제시하고자 하였다. 본 연구는 광주광역시내 교통혼잡과 교통사고가 빈번하게 발생하고 있는 신호교차로 지점을 대상으로, BPA를 이용하여 보다 신뢰성 높은 교통안전 평가 모형을 개발하고자 다음과 같은 일련의 방법으로 연구를 수행하였다. 첫째, 신호교차로 교통사고와 교통상충간의 순위상관분석을 실시하여 교통사고 순위와 교통상충 순위가 통계적으로 유의함을 확인하였다. 이는 교통상충이 신호교차로 교통안전 평가 변수로 사용될 수 있음에 따라 설명변

* 정회원 · 도로교통공단 교통과학연구원 선임연구원 (KOROAD · ccacca-1@hanmail.net)

** 정회원 · 교신저자 · 전남대학교 공업기술연구소 선임연구원 (Corresponding Author · Cheonnam National University · mjs0124@jnu.ac.kr)

*** 한국도로공사 진안지사 차장 (Korea Expressway Corporation · chamjuin@naver.com)

**** 정회원 · 전남대학교 토목공학과 교수 (Cheonnam National University · tjha@jnu.ac.kr)

Received February 28, 2015/ revised March 23, 2015/ accepted April 20, 2015

수로 입력되고 교통사고가 종속변수인 선형회귀모형을 개발하는데 이용하였다. 둘째, 신호교차로의 교통량과 진입·진출 차로수 차이 등을 교통사고의 설명변수로 간주하여 다중회귀분석을 통해 교통사고 예측모형을 개발하였다. 셋째, 교통량과 도로 기하구조 요소를 모형의 설명변수로 설정하고 교통상충을 종속변수로 하여 BPA를 이용한 최적의 교통안전 평가 모형을 개발하였다. 마지막으로, 교통사고 실측값, 다중회귀모형, BPA에 의한 교통사고 예측값을 평균제곱근오차 방법으로 모형의 적합도 비교·분석을 하였다. 본 연구의 결과, BPA에 의해 도출된 교통사고 예측값과 교통사고 실측값 사이의 평균제곱오차는 3.89로 계산되어 BPA가 다중회귀 모형보다 상대적으로 교통안전 평가능력이 우수한 것으로 나타나 실제 신호교차로 교통안전도를 평가하는데 효과적으로 활용될 수 있을 것으로 판단되고 추후, 교통안전정책 수립시 실질적인 도움이 될 것으로 기대된다.

검색어 : 신호교차로, 교통상충, 교통사고, 다중회귀모형, 역전과 알고리즘

1. 서론

1.1 연구의 배경 및 목적

2012년 우리나라 국내총생산(GDP)은 1,270조를 넘어서 37조 원에 불과하던 1980년에 비해 외형적으로 거의 34배 이상 증가하였고 단기간의 경제성장에 힘입어 자동차 수요역시 폭발적으로 증가한 것으로 나타났다.

이처럼 경제성장과 국민소득의 증가로 인하여 삶의 질이 향상됨에 따라 편리하고 안전한 교통수단으로서 자동차의 보유가 보편화되기 오래되었다.

하지만 우리나라는 세계 10위권의 경제대국이자 세계 5위권의 자동차대국이라는 괄목할 경제성장을 이루었으나 교통사고로 매년 5천명 이상의 사망자와 13조원에 이르는 사회적 비용이 발생하고 있어 여전히 교통후진국을 면치 못하고 있는 실정이다. 이와 같이 교통사고를 선진국 수준으로 줄이기 위해서는 사고 감소를 위해 다각적이고 현실적인 노력을 기울여야하며 각 자치단체별 지역실정에 맞는 교통안전에 대한 관심과 역할이 더욱 중요한 시점임을 인식하고, 현실적인 방안을 모색해야 한다.

지난 2013년도 한 해 동안 우리나라에서 발생한 교통사고 사망자는 2012년 대비 300명(5.6%)이 감소한 5,092명으로 역대 최대치를 기록하였으며 사망자 감소를 또한 최근 10년간 평균 감소율(2.9%)을 크게 상회하였다. 작년 한해 동안 교통사고는 총 215,354건이 발생하였으며 그 중 대상지점 교차로에서 발생한 교통사고는 97,002건으로, 전체 교통사고의 약 43.7%를 차지하고 있어 상대적으로 높은 비중을 나타내고 있다. 특히 광주광역시 교차로 교통사고는 광주광역시 전체 교통사고 8,007건의 약 47.4%(교차로 사고 3,954건)를 차지하는 등 교차로에 대한 근본적인 교통사고 원인 진단 및 대책 마련이 시급하다.

이에 따라 교통사고 피해를 최소화하기 위해서 차량과 도로 체계에 대한 공학적인 개선을 통하여 교통사고 원인을 제거해야 한다. 안정성과 효율성이 떨어지는 도로는 교통사고가 지속적으로 발생할 가능성이 크고 이를 개선하는데 막대한 사회적 비용과 시간이 소요되며, 부적절한 환경 요인으로 발생한 교통사고는 국가

적으로 큰 피해를 발생시키게 된다. 따라서 본 연구에서는 최근 인공지능 분야 중 활발히 연구중인 역전과 알고리즘을 도입하여 신호교차로 대상으로 교통안전 평가모형을 개발하고, 이를 통해 실제 교통안전개선 분야에 적용성 및 효용성을 확인하고자 하였다.

1.2 연구의 내용 및 방법

본 연구는 광주광역시내 교통혼잡과 교통사고가 빈번하게 발생하고 있는 신호교차로 지점을 대상으로 인공지능 분야의 BPA를 이용하여 신뢰성 높은 교통안전 평가기법을 개발하고자 하였다.

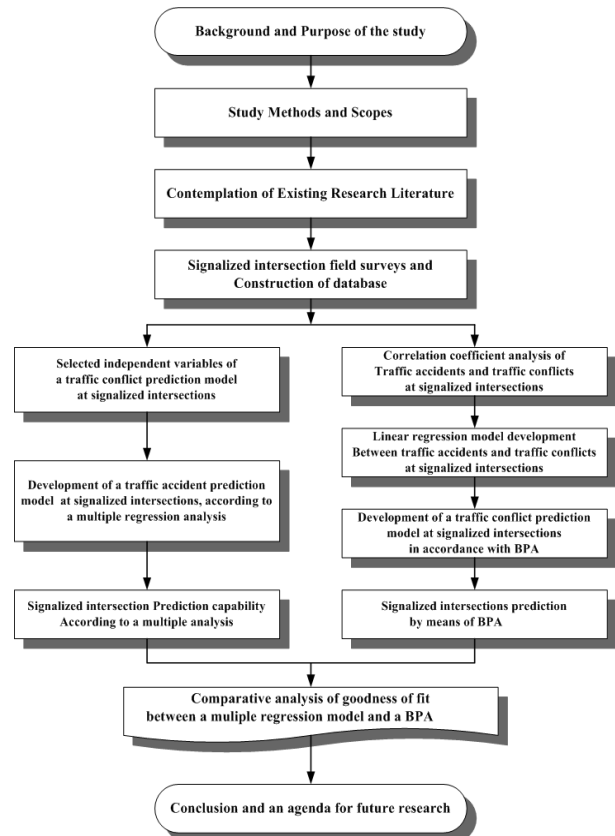


Fig. 1. Flowchart of Research

첫째, 신호교차로를 대상지점으로 접근로별 교통량과 도로의 주요 기하구조 요소를 파악하였으며 잠재적 위험으로 간주되는 교차로의 교통상충을 진행방향별 교통상충 특성을 중심으로 현장 조사를 실시하였다.

둘째, 선정된 교차로의 교통사고와 교통상충간의 상관분석을 실시한 후 신호교차로 교통사고와 교통상충의 선형회귀모형을 구축하였고, 교통사고에 영향을 미치는 교통량 및 도로 기하구조 요소들을 이용하여 다중회귀분석에 의한 교통사고 예측모형을 개발하였다.

셋째, 교통량과 도로 기하구조 요소를 모형의 독립변수로 선정하고 교통상충을 종속변수로 하여, 신호교차로 교통안전도를 평가할 수 있는 역전파 알고리즘의 모형을 구축하였다. 이때 예측된 교통상충수는 사고와 상충간의 선형회귀모형에 적용하여 교통안전도를 평가하였다.

마지막으로, 최적의 신경망이론에 모형을 이용하여 신호교차로 교통안전도를 평가한 결과와 다중회귀분석에 의한 신호교차로 교통안전도를 평가한 결과와의 비교분석을 실시하여 모형에 대한 신뢰성과 적용성을 확인하였다. 본 연구의 수행과정은 Fig. 1과 같다.

2. 기존 연구문헌 고찰

2.1 교차로 교통사고에 관한 연구

2.1.1 한국건설기술연구원

도로 기하구조와 교통사고의 관계에 있어서 중단경사가 교통사고에 미치는 영향은 불명확하고, 곡선반경과 사고율간의 상관관계는 급경사보다 평지부에서 더욱 크며, 중단경사와 사고율간의 상관관계는 곡선반경의 대소에 관계없이 불명확하다고 제시하였다.

2.1.2 Barbaresso

교차로 사고 비교분석, 신호형태에 의한 분류, 교차로 형태, 교차도로의 기능적으로 구분하여 연구한 결과, 두 간선도로의 4지 교차로의 점멸신호를 제거하고 신호화한 결과 야간시간대의 직각 사고 빈도의 감소와 직각사고의 빈도와 비율은 다른 교차로 형태에 비해 비점두시간에 점멸신호를 가진 간선도로 4지 교차로에서 매우 높게 나타난다는 것을 분석하였다.

2.1.3 Hauer et al.

사고 자료는 사고형태에 의해서 분류, 구분(보행자, 단독차량, 다차량)하여 연구하였다. 빈도분석과 변수간의 도표를 통계적으로 설명하였고, 충돌빈도는 교차로에 진입하는 교통류 전체보다는 방향별 교통량과 관련 깊은 것으로 나타났다.

2.2 교차로의 교통사고 예측모형에 관한 연구

2.2.1 Nam

교차로에 진입하는 전체교통량 보다는 교통류에 따른 사고유형으로써 교차로의 안전성을 판단해야 한다는 Hauer의 주장을 수용하여 교통사고 자료와 교통량 자료를 이용하여 교통사고 예측모형을 개발하였다.

2.2.2 Sung

신호교차로에서의 도로조건, 교통조건, 교통운영상의 조건 등을 분석하여 교차로의 설계 안전성을 저해시키는 요인들을 찾아내고, 그 요인들과 사고와의 상관관계를 이용하여 지방부 4지 신호교차로의 안전성 판단을 위한 사고 예측모형을 개발하였다.

2.3 인공지능을 이용한 교통사고 모형에 관한 연구

2.3.1 Lee

신경망 이론을 이용하여 교통사고와 요인과의 관계를 분석하였으며 분석된 모형의 민감도 분석을 통해 교통사고와 개별요인과의 관계를 살펴보았다. 이를 통해 교통사고의 결정적 요인을 추출하였다.

2.3.2 Kim

환경적 요인으로 퍼지적 성격을 가진 교통량, 차로폭, 시거를 입력변수로 비신호 교차로에서의 사고건수 예측을 위한 ANFIS (Adaptive Neuro-Fuzzy Inference System) 모형을 통하여 교통사고 예측모형을 개발하였다.

2.3.3 Kong

도로기하구조특성이 교통사고에 미치는 영향을 분석하기 위해 유전감독-인공신경망이론을 이용하여 교통사고모형을 개발하여 교통사고에 대한 도로기하구조 변수의 특성을 규명하였다.

2.3.4 Kang

퍼지추론 이론과 신경망이론을 이용한 모형을 구축하였고, 마지막으로 구축된 퍼지추론이론 모형 및 신경망이론 모형과 기존 회귀모형인 포아송 회귀모형간의 통계적인 검증과 실제 Data를 이용한 모형의 적정성을 검토하였다.

2.4 역전파 알고리즘 이론

BPA는 다층 신경망을 학습시키기 위한 알고리즘이다. 다층 신경망 학습은 입력층의 입력값과 기울기에 의한 탐색과정에서 기대되는 출력층 노드에서 원하는 목표값(Target)을 가지고 오차함수를 최소화시켜 연결강도를 최적화하는 방법이며, 최적 매개변수로의 탐색방향을 결정하는 강하방법이다(Fig. 2).

이에 본 연구의 인공지능 모형은 다층 신경망 형태로 구성하였고, 입력층, 은닉층, 출력층 3단계로 개체화하였다. 모형의 학습 알고리즘은 모멘텀 상수와 적응식 학습률이 적용된 최급강하법을 이용한 역전파 학습알고리즘이고, 보정을 위한 훈련과정의 효율성을 높이기 위해 조기 정지 학습기법을 도입하여 구현하였다. 역전파 알고리즘에 의한 모형 학습시 발생하는 지역 최소값 문제를 극복하고 학습속도를 향상시키기 위한 방법으로 모멘트-적응식 학습률 방법을 적용하였다.

여기서 모멘트는 오차표면에서 역전파의 민감도를 감소시키는 역할을 하며 네트워크가 낮은 오차의 해를 탐색하는 과정에서 얕은 최소점을 회피하는데 도움을 준다. 또한, 적응식 학습율은 모멘트법과 병행하여 학습계수를 조절하여 학습속도를 향상시키는 역할을 수행한다.

신경망이론과 관련한 선행 연구는 교차로 기하구조, 교통량,

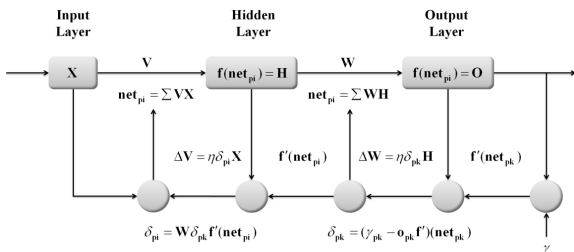


Fig. 2. Learning Process of BPA

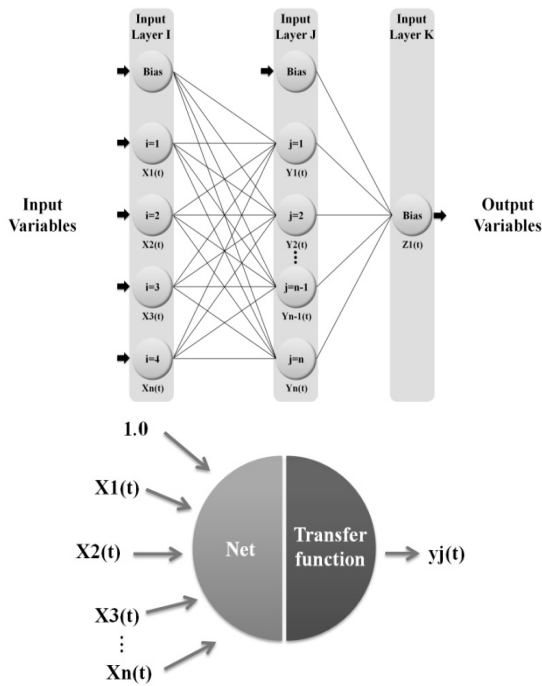


Fig. 3. Structure of BPA and Processing Processor

신호운영방법, 사고심각도 등의 요소를 변수로 교통사고예측모형, 대기행렬예측모형, 교통사고심각도예측 등에 관한 연구가 대부분이며, BPA를 이용한 교통사고와 교통상충예측에 관련된 연구는 전무한 실정이라 교통사고를 예측할 수 있는 교통안전 평가 모형을 개발하고자 BPA를 활용하였다.

3. 신호교차로 현장조사 및 자료수집

본 연구는 광주광역시내 교통사고 잦은 지점 중 최근까지 도로의 기하구조가 크게 변화되지 않고 지난 2009년부터 2010년까지 2년간 자료의 신뢰성이 높으며 교통사고 자료를 용이하게 확보할 수 있는 8개 신호교차로를 대상으로 선정하였다. 대상지점은 Table 1과 같다.

교통안전평가를 위해 대상지점의 교통조건인 교통량 자료는 광주광역시청과 도로교통공단에서 기 조사된 자료를 중심으로 이용하였으며 교통상충은 침두시간대에 현장조사를 통해 파악하였다. 또한 교통사고에 영향을 미치는 도로 기하구조 요소는 해당교차로의 현장조사를 실시하여 모형의 변수로 정리하였으며 이때 사고요인과 상관성이 높은 도로 기하구조 변수는 평가모형에 적용하였다. 대상도로의 교통조건 및 도로 기하구조 요소 자료 선정은 기존 관련 연구문헌을 통하여 교차로 교통사고에 영향을 미치는 교통량을 포함한 중단경사, 진입·진출 차로수 변화차이, 교차로 면적, 좌회전 길이, 횡단보도 길이, 횡단보도 폭 등 7개의 변수를 기본적으로 선정하여 자료를 분석하였다.

4. 교통사고와 교통상충간의 상관분석

4.1 교통사고와 교통상충간의 순위상관분석

교통상충이란 둘 또는 그 이상의 도로이용자 사이의 상호작용으로써 도로이용자들이 사고에 임박한 상황을 피하기 위한 회피행동을 발생시키는 사건이라고 정의한다. 이러한 사고발생 잠재지점과 사고발생지점에 대하여 안전도를 진단하는 하나의 기법을 교통상충기법이라고 한다. 교통상충기법은 사고자료에 의한 분석방법에 비해 신속하고 정확한 분석을 할 수 있다는 장점이 있다. 따라서 본 연구의 교통상충 조사는 진행방향별 교통상충특성을 중심으로

Table 1. 8 Target Locations of Signalized Intersections at Gwang-ju City

Spots	Intersection's Name	Spots	Intersection's Name
1	Chonnamdae	5	Wallsan
2	Sinan	6	Heukseok
3	Seobang	7	Usan
4	Unam	8	Yudong

Table 2. Explanatory Variable of Linear Regression Model

Spots	Traffic Volume (1,000,000/Year)	Longitudinal Slope(%)	The Change Difference of The Number of Entry and Exit Lanes (absolute value)	Intersection Area (m ²)	Left-Turn Length (m)	Crosswalk Length (m)	Crosswalk Width (m)
1	39.49	0	0	450	20	16.38	7
2	53.75	0	3	798	32.28	19.85	7.93
3	43.91	0	2	748	25	23.95	8
4	45.66	0	2	789	35.75	20.48	7.78
5	54.51	0	1	837	2.5	3.38	1.05
6	31.90	0	0	370	20	18.55	7.7
7	41.49	0	2	761	27.5	24.65	8
8	53.08	3	0	534	10	15.38	6

Table 3. Result of Spearman Rank-Order Correlation Analysis

Spearman Rank Correlation Coefficient for A Correlation Analysis		Traffic Conflict	Traffic Accident
The Rho of Spearman	Traffic Conflict	Correlation Coefficient	1.000
		Significance Probability (Both Sides)	.000
		N	8

** The correlation of significance level is 0.01. (Both Sides)

조사하였으며 일반적인 교차로 교통상충의 형태는 크게 다음과 같이 6가지 형태로 분류하여 정리하였다.

- 동일방향 교통상충(Same direction)
- 대향 좌회전 교통상충(Opposing left turn)
- 교차 교통상충(Cross traffic)
- 적신호 우회전 교통상충(Right-turn-on-red)
- 보행자 교통상충(Pedestrian)
- 2차 교통상충(Secondary)

따라서 본 연구는 교통상충이 교통사고의 위험을 대신할 수 있는 척도로 사용이 가능한지 여부를 판단하였다. 이를 위하여 스피어만 순위상관분석을 통한 적합도 검증을 실시하였다. 이는 스피어만 순위상관분석을 통한 적합도 검증을 위한 관측치들이 기수척도가 아닌 서열척도로 되어있는 이변량 모집단으로부터 N개의 짝으로 된 표본들이 선정된 경우라면 두 변수간의 상관관계, 즉 모수적 상관관계는 측정할 수 없기 때문에 판단에 따른 것이다.

이때 순위를 사용한 두 변수 간의 관련도 측정이 스피어만 순위상관계수이며 이는 r_s 로 분석된다. 여기서, 상관분석을 위한 이변량 확률표본의 크기가 n 인 경우에, 이 중 하나의 변수를 X 로 표시하

고 다른 하나의 변수를 Y 로 표시한다. 이때 변수 X 의 i 번째 관측치의 순위를 x_i 로, 변수 Y 의 i 번째 관측치의 순위를 y_i 로 표시한다. 스피어만 순위상관계수는 짝으로 구성된 두 관측치들 순위의 차, 즉 $d_i = x_i - y_i$ 에 중점을 둔다. 또한 분석은 변수의 정규성 분포가정이 필요로 하지 않으며 상관분석의 스피어만의 순위 상관관계수는 다음과 같이 정의된다.

$$\therefore r_s = 1 - \frac{6 \sum_i d_i^2}{n(n^2 - 1)}$$

여기서, $d_i = x_i - y_i$

위 공식에 있어서 x_i 와 y_i 사이의 차, 즉, d_i 가 크면 클수록 $\sum d_i^2$ 도 크게 된다. 또한 모든 짝에서의 차가 0인 경우에는 $\sum d_i^2$ 은 0이 되고, r_s 는 1이 되어 두 변수간의 관계는 완전한 +의 상관관계가 성립된다.

모든 짝에서 x 와 y 의 순위가 반대로 되어 있는 경우에 $\sum d_i^2$ 은 최대가 되어짐과 동시에 r_s 는 -1이 된다. 즉, 위 식 (4.1)에서 $-1 \leq r_s \leq 1$ 의 관계를 알 수 있다.

그리고 $r_s = 0$ 인 경우에는 순위로 표시된 두 변수 간에는 상관관계가 없다. 이는 관측치들이 기수척도가 아닌 서열척도로 되어있는 이변량 모집단으로부터 N개의 짝으로 된 표본들이 선정된 경우라면 두 변수간의 상관관계, 즉 모수적 상관관계는 측정할 수 없기 때문에 판단에 따른 것이다. 따라서 스피어만 순위상관분석을 실시하기 위하여 본 연구의 대상지점인 신호교차로 8개 지점을 선정하였다.

이 때, 분석 자료는 2009년과 2010년의 2년 동안 조사된 각각의 수를 합산하여 평균으로 산출한 교통사고 발생건수와 교통상충수를 이용하였다. 교통사고와 교통상충간의 상관분석을 통한 스피어만 상관관계수는 0.703의 결과가 나타났으며, 이는 교통사고 순위와

Table 4. Data on Analysis of Traffic Accidents at Signalized Intersections and Traffic Conflicts (Current Status of Accidents and Conflicts at Major Intersections in Gwang-ju City in 2009)

Spots	Spot Name	Data on Analysis of Traffic Accidents at Signalized Intersections and Traffic Conflicts in 2009					Data on Analysis of Traffic Accidents at Signalized Intersections and Traffic Conflicts in 2010				
		The Number of Traffic Accidents				Traffic Volume	The Number of Traffic Accidents				Traffic Volume
		Total	Death (Person)	Injury (Person)	Wound (Person)		Total	Death (Person)	Injury (Person)	Wound (Person)	
1	Chonnam Univ.	18	0	3	18	2,512	12	0	4	14	2,838
2	Sinan	18	0	4	23	4,329	21	1	6	24	3,275
3	Seobang	10	0	2	17	1,594	18	0	5	22	2,657.5
4	Unam	24	0	4	42	2,388	32	0	8	49	3,896
5	Wolsan	4	0	1	5	2,591	8	0	2	11	2,504.5
6	Heukseok	2	0	2	0	3,403	7	0	1	8	3,901
7	Usan	7	0	2	10	2,473	7	0	4	8	4,137.5
8	Yudong	7	0	0	11	4,907	4	0	0	23	4,259.5

교통상충 순위 간에 강한 양(+)의 상관관계가 있음을 보여준다. 교통사고와 교통상충의 순위 상관계수가 $\alpha = 0.01$ (양측)에서 유의하여 귀무가설을 기각하고 대립 가설을 채택하게 됨으로써 교통사고와 교통상충의 순위는 상관관계가 있는 것으로 검증되었다. 한편 교통사고와 교통상충간의 상관관계 검증을 위한 귀무가설과 대립 가설은 다음과 같다.

- $H_0 : \rho_s = 0$
(교통사고 순위와 교통상충 순위는 상관관계가 없다.)
- $H_1: \rho_s > 0$
(교통사고 순위와 교통상충 순위는 상관관계가 있다.)
- $N = 8, \alpha = 0.01$

4.2 신호교차로 교통사고와 교통상충간의 선형회귀 모형 개발

교통사고와 교통상충의 단순선형 회귀분석 결과는 다음 Table 5와 같다. 모형식의 설명력인 결정계수 값이 0.703으로 나타났고, 교통상충의 결과가 증가함에 따라 교통사고의 결과가 증가됨을 알 수 있다. 모형의 적합도를 나타내는 Durbin-Watson 값은 1.590으로 나타났다.

Table 5. Summary of Linear Regression Analysis Model (Standard Error of Estimated Value)

Model	R	R ²	Modified R ²	Standard Error of Estimate	Durbin-Watson
1	.839(a)	.703	.699	4.228800	1.590

a - Prediction Value : (Constant), Traffic Value

b - A Dependent Variable : The Number of Traffic Accident

회귀분석결과, 모형의 변수인 교통상충은 사고예측모형에 대한 설명력이 높은 것으로 나타났고, 모형의 변수인 교통상충의 유의확률은 0.000으로 나타나 유의수준 0.01에서 회귀모형이 적합한 것으로 확인되었다. 최종적으로 교통사고와 교통상충 관계에 의한 교통사고 예측모형은 Eq. (1)과 같다.

$$\therefore Y = - 11.932 + 0.017X \tag{1}$$

여기서, X : 교통상충수, Y : 교통사고 발생건수

4.3 신호교차로 교통사고와 교통상충간의 선형회귀 모형 검증

본 연구는 교통사고와 교통상충의 선형회귀모형에 의해 예측된 교통사고 건수가 실측치를 얼마만큼 대표할 수 있는지를 검증하기 위하여 대응표본 T 검증을 실시하였다.

여기서, 짝별로는 어떤 동질성이 유지된다고 가정한다. 예를 들어 표본 1과 2는 두 가지의 약의 성능을 실험하는 경우인데 짝별로는 동일한 환자인 경우이다. 이러한 구조의 표본을 대응표본

Table 6. Coefficient of Linear Regression Analysis (Standardization Coefficient)

Model	Non-standardization Coefficient	Standardization Coefficient		t	Significance Probability	Significance Probability
		B	37,616	β		
1	Constant	-11.932	4.454	-	-2.831	.011
	Traffic Conflict	.017	.002	.838	6.986	.000

a - A Dependent Variable: The Number of Traffic Accident

이라고 부른다. 이때 각 쌍의 차이를

$$D_i = x_{1i} - x_{2i}, \quad i = 1, 2, \dots, n$$

이라고 할 때 D_i 는 $N(\mu_D, \sigma_D^2)$ 로부터의 확률표본으로 가정한다. 다음과 같은 가설 $H_0: \mu_D = 0, H_0: \mu_D \neq 0$ 에 대한 검정통계량은

$$T = \frac{\bar{D}}{\frac{s_d}{\sqrt{n}}}$$

이며, 이때 귀무가설이 참일 때 검정통계량 T 는 자유도 $(n-1)$ 인 t-분포를 따른다. 여기서,

$$s_d^2 = \frac{\sum(D_i - \bar{D})^2}{n-1}$$

으로 σ_D^2 의 불편추정량이다. 이 경우, μ_D 의 $100(1-\alpha)\%$ 신뢰구간은 다음과 같다.

$$\bar{D} \pm t(n-1; \frac{\alpha}{2}) \cdot \frac{s_d}{\sqrt{n}}$$

한편 교통사고와 교통상충에 대해 대응표본 t-검증을 실시하기 위하여 2010년 교통사고 자료와 교통상충 자료를 이용하였다. 교통사고와 교통상충의 선형회귀모형에 의해 예측된 교통사고 건

수가 실측값을 얼마만큼 대표할 수 있는지를 검증하기 위한 귀무가설과 대립가설은 다음과 같다.

H_0 : (교통사고실측값과 교통사고예측값은 차이가 없다.)

H_1 : (교통사고실측값과 교통사고예측값은 차이가 있다.)

$n = 8, \alpha = 0.05$

Table 8은 교통사고 실측값과 교통사고 예측값은 차이가 없다는 귀무가설에 대한 검증 결과이다. 검증결과, t 값은 -1.884이고 이에 대한 유의확률은 0.074로 유의수준 0.05에서 귀무가설을 기각하지 못한다. 즉, 교통사고 실측값과 교통사고 예측값은 차이가 없는 것으로 나타났다. 따라서 교통사고와 교통상충의 선형회귀모형에 의해 예측된 교통사고 건수가 실측값을 대표한다고 할 수 있다.

5. 교통안전 평가모형 개발

5.1 교통상충 예측을 위한 BPA 적용

BPA 구성단계는 훈련, 보정 및 검증의 3단계로 분류하여 구축하였다. 2009년부터 2010년까지 기간의 교통조건인 교통량, 도로 기하구조 요소인 차로수 변화차이와 교통사고의 잠재적 위험을 대신할 수 있는 교통상충을 변수로 사용하였다. 해당 학습 단계는 훈련 및 보정을 위해 2009년부터 2010년까지의 자료를 분류하여 모의하였고, 훈련에 의해 구해진 매개변수인 가중치

Table 7. Paired Sample T-Test Method

Pari number	Sample 1	Sample 2	Difference (sample1-sample2)
1			
2			
⋮	⋮	⋮	⋮
n	x_{1n}	x_{2n}	D_n
mean	\bar{x}_1	\bar{x}_2	\bar{D}

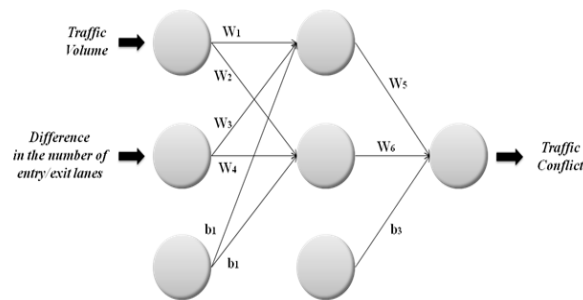


Fig. 4. Structure of Safety Evaluation Model for Signalized Intersections and Processing Processors

Table 8. Paired Sample T-Test Result for the Linear Regression Model

-	-	Correspondence Difference					t	Degrees of Freedom	P-value (both)
		Mean	Standard Residuals	Mean Standard error	95% Confidence Interval of the Difference				
					Lower	Maximum			
CorreSpon Dence 1	Observed Value-Prediction Value	-1.92238	4.67643	1.02048	-4.05107	.20630	-1.884	20	.074

Table 9. List of Training Parameters

Tool	Parameters	Variable	Value
ANN	Learning Rate	lr	0.9
	Momentum Factor	mc	0.7
	Iteration No	epochs	10,000
	Terminate Goal	goal	0.00001

및 임계값은 2010년 또 다른 기간의 자료에 적용하여 검증을 실시하였다.

신호교차로 대상지점의 교통안전 평가 모형은 시간대별로 구성하였으며, 통계적인 검증을 실시하여 모형의 실제 현장 적용성을 확인하였다. 모형훈련을 실시 하였을 때 각 모형에 대한 입력층의 노드 수는 은닉층의 노드 수가 조절되도록 하였다. 즉, 입력층의 노드 수를 n개라고 할 때 은닉층의 노드 수는 n~3n개 까지 변화하면서 그 결과 값을 나타낼 수 있도록 설계하였다. 이러한 일련의 과정을 거친 모형들 중 통계적인 결과가 우수한 모형을 분석하여 교통안전 평가결과를 제시하였다.

본 연구에서는 BPA 성능을 개선하기 위하여 은닉층(Hidden Layer)에 로그-시그모이드(Log-Sigmoid) 함수를 사용하였고, 출력층에 선형전달함수를 가진 네트워크를 구성하였다. 또한 은닉층의 연결강도와 편치를 얻기 위하여 구조층의 입력공간을 교차하였고 각 Neuron의 활동영역을 균등하게 분할하고자 모멘트와 적응 학습률(Adaptive Learning Rate)를 이용하였다. 다음 Table 9는 모형의 학습률과 모멘텀 계수 등의 훈련 매개변수를 나타낸 것이다.

5.2 교통상충 예측을 위한 BPA 검증

신호교차로의 교통조건, 도로의 기하구조 요소, 교차로의 교통상충의 결과를 입·출력패턴으로 간주하여 대상지점의 교통안전도를 평가하고, 선별된 모형에 대한 실측치에 대한 예측력을 파악하기 위하여 모형의 결과 값에 대한 통계적 오차분석을 Table 10과

같이 제시하였다.

학습결과에 나타난 구성별 모형 중 BPA (2,5,1)이 대체로 결정계수가 높은 평가도를 보여주고 있고, 모형의 모의 이후 각 모델에서 발생된 전반적인 결과는 은닉층의 처리소자의 개수 증가에 따라 상관성은 대다수 유사한 결과를 나타내었다. 모형의 교통안전도 평가 결과, 훈련과정의 결정계수는 0.974~0.978의 범위로 전반적으로 우수한 결과를 보여주었고, 모형의 검증과정 역시 결정계수는 0.952~0.959의 범위로 대체로 만족할만한 결과를 나타내었다. 이는 훈련과정에서 구해진 매개변수가 다른 기간에도 잘 적용됨을 보여 주는 것을 의미한다. 한편, 검증결과를 확인하고자 2009년부터 2010년 까지 2년간의 훈련 및 보정을 통해 얻게 된 매개변수인 가중치와 바이어스를 가지고 2010년 다른 기간에 발생된 입력변수와 종속변수의 자료를 설정하여 모의시킴으로써 학습오차 및 모형의 평가능력 정도를 파악하였다. 이와 같이 모형의 교통안전도 평가 결과, 검증과정 역시, 결정계수는 0.931~0.934의 범위로 대부분 우수한 평가능력을 보여주었다.

5.3 교통사고 예측을 위한 다중회귀모형 적용

교통사고에 영향을 미치는 설명변수는 도로교통요인을 기초로 하여 설정하였다. 그 중에서 도로조건에 포함된 요소로는 전체 차로수, 좌회전 차로수, 차로폭, 보·차도 구분, 회전반경, 횡단보도 유·무, 교통상충 지점수, 교차로 면적, 길어깨 폭, 교통섬 유·무, 좌·우회전 차로형태, 방향별 차로수 변화의 차이 유·무, 시거, 중앙 분리시설 유·무, 유턴 유·무, 유도선 유·무 등이 있다. 그러나 현장조사 결과, 95%이상의 교차로가 보도와 차도가 완전 분리되어 있고 교차로의 길어깨는 폭이 0.5m 이하로 범위가 극히 작으며 중앙 분리시설은 대상지점에는 존재하지 않았기 때문에 분석대상에서 제외하였다. 이에 따라 다음 Table 11에 제시된 교통사고의 설명변수는 요인분석, 주성분분석을 통하여 교통사고예측모형의 독립변수를 선정하였다.

Tables 11~14에 제시된 공통성은 추출된 요인들에 의해서 각

Table 10. Result of Statistical Analysis of Signalized Intersections at Sinan Intersection of Gwang-ju City

Model	Configuration of Model network			Error					
	Input Layer	Hidden Layer	Output Layer	R ²			RMSE		
				Tra	Cal	Ver	Tra	Cal	Ver
No.1	2(N)	2(N)	1	0.974	0.952	0.931	7.883	10.23	11.48
No.2		3(N+1)		0.976	0.959	0.932	7.712	10.11	11.34
No.3		4(N+2)		0.976	0.958	0.933	7.601	10.09	11.29
No.4		5(N+3)		0.978	0.957	0.934	7.517	10.00	11.12
No.5		6(N+4)		0.975	0.959	0.931	6.725	9.449	11.07

Source) Tra : Training, Cal : Calibration, Ver : Verification

Table 11. Signalized Intersection Traffic Accident Prediction Model Explanatory Variable

Explanatory Variable	Sign	Explanatory Variable	Sign
Traffic Volume (1,000,000/Year)	Q	Left-Turn Length (m)	L_L
Longitudinal Slope(%)	G	Crosswalk Length (m)	L_{CL}
The Change Difference of The Number of Entry and Exit Lanes (absolute value)	N	Crosswalk Width (m)	L_{CW}
Intersection Area (m ²)	A		

Table 12. Multiple Regression Analysis Base Data

Spots	2009 Year Traffic Accident (Number/Year)	2010 Year Traffic Accident (Number/Year)	Mean (Number/Year)
1	13	8	10.5
2	23	35	29
3	6	14	10
4	7	9	8
5	43	31	37
6	37	27	32
7	24	11	17.5
8	23	16	19.5

Table 13. Multiple Regression Analysis Explanatory Variable

Spots	Traffic Volume (1,000,000/Year)	Longitudinal Slope (%)	The Change Difference of The Number of Entry and Exit Lanes (absolute value)
1	21.53	2.5	0
2	36.5	3.75	0
3	18.21	0	0
4	19.39	0	0
5	54.75	0	3
6	25.92	0	2
7	34.66	0	2
8	31.46	0	3

Spots	Intersection Area (m ²)	Left-Turn Length (m)	Crosswalk Length (m)	Crosswalk Width (m)
1	1,313	15	22.75	7
2	1,593	15	14.75	3.75
3	642	17.5	15.5	6.25
4	450	20	16.38	7
5	798	32.28	19.85	7.93
6	748	25	23.95	8
7	789	35.75	20.48	7.78
8	789	15	21.15	8

Table 14. Multiple Regression Analysis Explanatory Variable Commonality Analysis

Commonality Analysis	Beginning	extraction
Traffic Volume	1.000	.887
Longitudinal Slope	1.000	.210
The Change Difference of The Number of Entry and Exit Lane	1.000	.620
Intersection Area	1.000	.134
Left-Turn Length	1.000	.321
Crosswalk Length	1.000	.109
Crosswalk Width	1.000	.407

Extraction Method : Principal Component Analysis

Table 15. Summary of Multiple Regression Analysis Model (Standard Error of Estimated Value)

Model	R	R ²	Modified R ²	Standard Error of Estimate	Durbin-Watson
1	.717(a)	.514	.488	5.57019	-
2	.783(b)	.612	.569	5.10953	2.003

a - Prediction Value : (Constant), Traffic Value

b - Prediction Value : (Constant), Traffic Value, Entry Difference

c - A Dependent Variable : The Number of Traffic Accidents

변수가 얼마나 설명되는지를 나타낸다. 한 변수의 공통성은 그 변수의 분산이 추출된 요인들에 의해 설명되는 정도를 가리키며 0과 1사이의 값을 갖는다. 해당 추출된 요인들에 의해서 교통량 변수는 88.7%, 진입진출차로수 변수는 62%가 설명됨을 알 수 있다. 따라서 7개 변수의 공통성 중 교통량과 진입진출차로수 만이 0.5 이상의 값을 가지므로 독립변수로 사용하였다.

본 연구는 다중회귀모형을 이용하여 교차로 교통사고 예측모형을 개발하였다. 모형의 기본구조는 Eq. (2)와 같다.

$$\therefore Y = a + b_1 (X_1) + b_2(X_2) + b_3(X_3) + \dots + b_n(X_n) \quad (2)$$

Y : 종속변수 (교통사고 발생건수)

X : 독립변수 (교통량, 교차로 기하구조 요소 등)

a, b : 상수

Table 15는 종속변수인 교통사고와 독립변수들 간의 다중회귀 분석 결과이다. 모형 2의 경우, 모형식의 결정계수 값이 0.783으로 나타났다. 모형의 적합도를 나타내는 Durbin-Watson은 2.003으로 나타나 모형의 적합도가 높음을 알 수 있다.

Table 16. Result of Paired T-Test on Multiple Regression Model

-	-	Pair-Difference					t	The Degree of Freedom	Significance Probability (Both Sides)
		Average	Standard Deviation	Standard Error of The Mean	95% Confidence Interval				
					Min	Max			
Pair	Actual Measurement Value - Predictive Value	1.807	4.70148	1.02595	-0.332	3.947	1.762	20	.093

일반적으로 Durbin-Watson은 2에 가까우면 오차항간에 독립성이 존재하는 것이고, 이 값이 0에 가까우면 인접 오차항간에 양의 상관관계를 나타낸다. 만약 이 값이 4에 가까우면 인접 오차항간에 음의 상관관계가 존재하는 것을 의미한다.

다중회귀분석 결과, 모형의 독립변수인 교통량과 진입-진출차로 차이는 사고예측모형의 설명력을 높이는 것으로 나타났고, 교통량의 유의확률은 0.002, 진입-진출 차로 차이의 유의확률은 0.046으로 나타나 유의수준 0.05에서 회귀모형이 적합한 것으로 도출되었다. 최종적으로 교통사고와 교통량, 진입-진출차로 차이의 관계에 의한 사고예측모형은 Eq. (3)과 같다.

$$\therefore Y = 2.823 + 0.455 X1 + 2.375 X2 \quad (3)$$

Y : 교통사고 발생건수

X1 : 교통량, X2 : 진입-진출차로 차이수

다중회귀모형에 의해 예측된 교통사고건수가 실측치를 얼마만큼 대표할 수 있는지를 검증하기 위해 대응표본 t-검증을 실시하였다. 아래 Table 16은 대응표본 t-검증을 실시하기 위하여 2010년 교통사고 자료와 교통량 및 진입-진출 차로차이 값을 이용하였다.

교통사고와 독립변수간의 다중회귀모형에 의해 예측된 교통사

고건수가 실측값을 얼마만큼 대표할 수 있는지를 검증하기 위한 귀무가설과 대립가설은 다음과 같다.

H0 : 교통사고 실측값과 교통사고 예측값은 차이가 없다.

H1 : 교통사고 실측값과 교통사고 예측값은 차이가 있다.

N = 8, $\alpha = 0.05$

Table 16은 교통사고 실측값과 교통사고 예측값은 차이가 없다는 귀무가설에 대한 검증 결과이다. 검증결과, t 값은 1.762이고 이에 대한 유의확률은 0.093으로 나타나 유의수준 0.05에서 귀무가설을 기각하지 못한다. 즉, 교통사고 실측값과 교통사고 예측값은 차이가 없는 것으로 나타났다. 따라서 다중회귀모형에 의해 예측된 교통사고건수가 실측값을 대표한다고 할 수 있다.

5.4 BPA와 다중회귀모형간의 적합도 비교·분석

본 연구는 다중회귀분석에 의한 모형과 BPA 간의 적합도를 비교·분석하였다. 각각의 신호교차로 교통안전도 평가 모형의 적합도를 분석하기 위하여 실제 실측값과 모형으로부터 도출되는 예측값의 차이를 이용하였다. 실측값과 예측값 사이의 잔차의 제곱합을 산술평균한 값에 대한 제곱근으로 실측값들 사이의 상호간 편차를 나타내는 평균제곱근오차를 이용하여 신호교차로 교통안전도 평가 모형의 적합도를 분석하였다. 일반적으로 평균제곱근오차는 0에

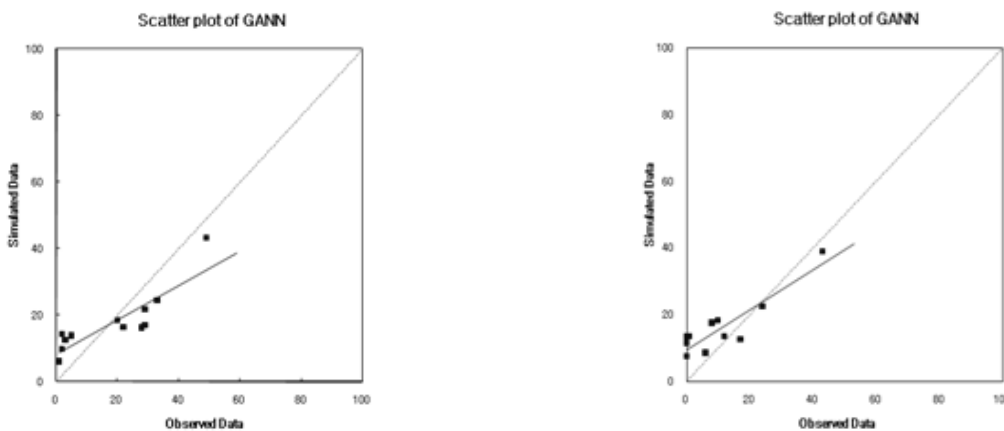


Fig. 5. Dispersion of Measured Values and Simulation Values on Training and Test of Model for Predicting Traffic Conflicts

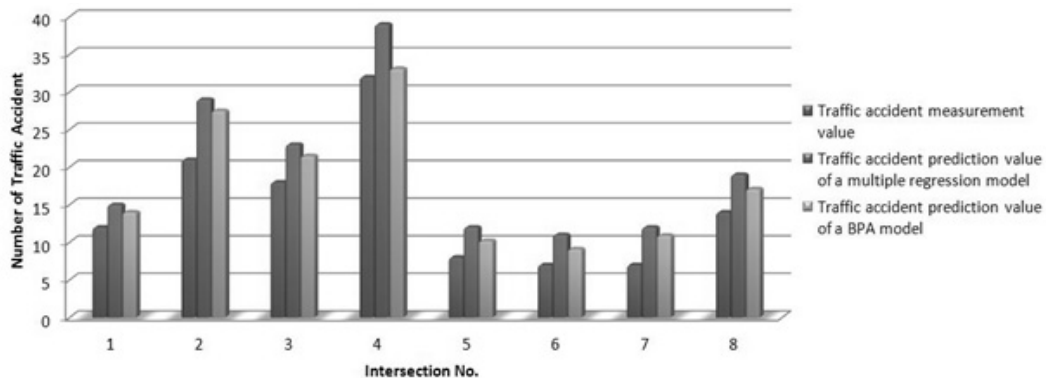


Fig. 6. Comparison of Measurement Values of Traffic Accidents, Prediction Value of Multiple Regression Model and Prediction Value of BPA (Target Locations at Major Intersections)

가까울수록 모형의 적합도가 높다는 것을 의미한다. 따라서 다중회귀모형에서 도출된 교통사고 예측값과 교통사고 실측값 사이의 평균제곱오차는 4.84로 나타났고, BPA에 의해 도출된 교통사고 예측값과 교통사고 실측값 사이의 평균제곱오차는 3.89로 계산되었다. 다중회귀모형과 BPA 적합도를 분석한 결과, BPA 적합도가 상대적으로 우수한 것으로 나타났다.

6. 결론 및 향후 연구과제

6.1 결론

본 연구는 인공지능 분야에서 활발히 연구되고 있는 인공 신경망 이론(역전파 알고리즘)을 도입하여 신호 교차로 대상지점으로 교통 안전 평가모형을 구축하였으며 그 결과는 다음과 같다.

- (1) 신호교차로 교통사고와 교통상충간의 순위상관관계 분석을 실시하여 교통사고와 교통상충이 통계적으로 유의함을 확인하였다. 이는 교통상충이 잠재적 위험으로 판단하고 신호교차로 안전평가 모형의 변수로 사용될 수 있음에 따라 교통상충이 설명변수, 교통사고가 종속변수인 선형회귀모형을 개발하였으며, 이를 통해 교통사고의 결정적인 요인을 추출할 수 있었다.
- (2) 2010년 다른 기간에 발생한 입력변수와 종속변수의 자료를 설정하여 교통상충예측 모형의 분석결과, 검증과정 역시 결정계수는 0.931-0.934의 범위로 대부분 우수한 평가능력을 보여 주었다.
- (3) 다중회귀모형에서 도출된 교통사고 예측값과 교통사고 실측값 사이의 평균제곱오차는 4.84로 나타났고, BPA에 의해 도출된 교통사고 예측값과 교통사고 실측값 사이의 평균제곱오차는 3.89로 계산되었다.

다중회귀모형과 BPA 적합도를 분석한 결과, BPA 적합도가 상대적으로 우수한 것으로 나타났다.

6.2 향후 연구과제

본 연구 수행과정에 있어서의 시·공간적으로 한계점과 향후 연구과제는 다음과 같다.

- (1) 향후 신호교차로 대상지점에서의 다년간의 교통사고, 교통여건과 도로여건 등의 자료를 확보하여 모형의 객관적인 평가능력을 재확인할 필요가 있다.
- (2) 다양한 연구 수행과 정확한 사고자료 수집을 토대로 보다 체계적인 교통사고 DB 구축방안이 마련된다면 정확하고 실질적인 요인들을 반영한 신호교차로 안전 평가모형의 구축이 가능할 것으로 판단된다.
- (3) 일반적으로 개선사업 대상지점 선정은 당해 연도 교통사고 자료조사에 의한 지점별 사고자료를 토대로 대상후보지점을 선정 후 이를 기초로 하여 사업대상지점 선정 우선 순위에 따라 사업대상지점을 선정하지만 비용효율적인 측면에서 최적의 방법이 무엇인지 기존방법과의 비교분석을 위한 연구를 수행할 필요가 있다.

References

Adler, B. and Alfaro, L. (2007). "A content-driven reputation system for Wikipedia." *Proc. of the 16th international conference on World wide web*, NewYork, N.Y., pp. 261-270.

Agichtein, E. (2005). "Web information extraction and user modeling: Towards Closing the Gap." *IEEE Data Engineering Bulletin*, Vol. 28, No. 4.

Agrawal, S., Chaudhuri, S. and Das, G. (2002). "DB explorer : A System for Keyword Search over Relational Databases." *In Proc.*

- of ICDE-02.
- Bang, K. L., Lindberg, G. and Harahap, G. (1997). "Development of guidelines based on life-cycle cost to replace level-of-service conception capacity analysis." *Transportation Research Record : Journal of the Transportation Research Board*, No. 1572, pp. 10-17.
- Chol, J. (1998). "Development of a traffic safety index for urban arterials." *Journal of Korean Society of Transportation*, Vol. 16, No. 3, pp. 47-56.
- Greibe, P. (2003). "Accident prediction models for urban roads." *Accident Analysis Prevention*, Vol. 35, No. 2, pp. 273-285.
- IYM, Y. T. (2011). "Development of traffic accident frequency prediction model in urban signalized intersections with fuzzy reasoning and neural network theories." *Journal of Korean Society of Road Engineers*, Vol. 13, No. 1, pp. 69-77.
- Jacobson, G., Booch, and Rumbaugh, J. (1999). *The unified software development process*, Addison-Wesley Publishing Company.
- Jang, J., Lee, Y., Kim, H. and Jang, M. (2000). "The method of deciding design level of service with optimal user cost." *Journal of Korean Society of Transportation*, Vol. 18, No. 2, pp. 17-26.
- Kang, Y. K. (2011). "Development of traffic accident frequency prediction model in urban signalized intersections with fuzzy reasoning and neural network theories." *Journal of Korean Society of Road Engineers*, Vol. 13, No. 1, pp. 69-77.
- Kim, E. C. (2008). "Development of traffic accident frequency model for evaluating safety at rural signalized intersections." *Journal of Korean Society of Road Engineers*, Vol. 10, No. 4, pp. 53-63.
- Kim, K. W. (2012). "Building a traffic accident frequency prediction model at unsignalized intersections in urban areas by using adaptive neuro-fuzzy inference system." *Journal of Korean Society of Civil Engineers*, Vol. 32, No. 2D, pp. 137-145.
- Kong, S. H. (2003). "Traffic accident frequency model using the genetic-supervised artificial neural network." *Journal of Korean Society of Civil Engineers*, Vol. 23, No. 4, pp. 479-487.
- Lee, C. N. (1998). "Identification of variables influencing traffic accidents on urban arterials using neural network." *Journal of Korean Society of Civil Engineers*, Vol. 18, No. 3, pp. 803-810.
- Lee, J. M. (2008). "Developing the traffic accident prediction model using classification and regression tree analysis." *Journal of Korean Society of Civil Engineers*, Vol. 10, No. 4, pp. 31-39.
- Lee, S. B. (2003). "Development of a traffic accident prediction model by level of roads and traffic characteristics." *Journal of Korean Society of Civil Engineers*, Vol. 23, No. 4D, pp. 495-504.
- Nam, G. H. (2001). "Development and application of traffic accident forecasting model for signalized intersections (Four-Legged Signalized Intersections in Kwang-Ju)." *Journal of Korean Society of Transportation*, Vol. 19, No. 6, pp. 207-218.
- Shim, K. B. (1994). "Analysis of traffic accident by the relation parameters in city arterial roads." *Journal of Traffic Safety Research*, Vol. 13, pp. 125-136.
- Son, E. Y. (1992). *A study of estimating traffic congestion costs*, The Korea Transport Institute (in Korean).
- Sung, N. M. (2008). "Development of severity model for rural unsignalized intersection crashes." *Journal of Korean Society of Road Engineers*, Vol. 10, No. 3, pp. 47-56.
- Trujillo, J., Palomar, M., Gomez, J. and Song, Y. (2001). "Designing data warehouses with conceptual models." *Computer IEEE*, pp. 66-75.