

GIS를 이용한 교통사고의 분석 기법 개발

A GIS Based Technique for Analyzing Traffic Accidents

최 기 주 *	박 인 철 **	오 세 창 ***
Choi, Kee Choo	Park, In Chol	Oh, Sei-Chang

要 旨

본 논문은 도시의 가로망에서 발생하는 교통사고의 효과적인 관리와 분석을 수행하기 위한 방안을 제시한 글이다. 즉 지리정보시스템을 교통사고 관련 데이터의 구축과 분석의 기본적인 틀로서 이용하여 교통사고의 분석을 보다 과학적으로 수행할 수 있는 기법을 제시하고자하였다. 교통사고는 네트워크 상에서 발생하는 공간적 특성을 가지고 있다. 공간적 특성 및 관련 데이터와 사고데이터를 지리정보시스템을 이용하여 데이터베이스를 구축하고, 이러한 기반 위에서 점, 선, 면적인 분석을 수행할 수 있는 기법을 제시하였다. 이는 단순히 문자위주의 교통사고 관리시스템보다 진일보한 시스템으로 판단되고, 효과적인 원인 분석 등에 활용 될 수 있다는 측면에서 본 연구에서는 수원시의 특정부분에 있어서의 사고데이터를 지리정보 데이터베이스로 구축하였다. 이를 기반으로 대상지역에 대한 점, 선, 면적인 제반 분석을 수행하였다. 한편, 구축된 데이터베이스를 바탕으로 특정공간 상에서 사고가 얼마나 많이 발생하는지, 또는 어떤 유형의 사고가 많이 발생하는지를 판단할 수 있는 계량적인 판단을 할 수 있는 지수들을 소개하여 교통사고에 대한 각 소구역별 세부분석을 실시하였다. 끝으로, 지형분석자료를 바탕으로 도로의 연장과, 도로의 교차점 수를 설명변수로 하는 교통사고 예측모형에 대한 교통사고빈도 회귀식을 도출하였다.

ABSTRACT

This article aims at presenting a new framework for traffic accident analysis by proposing a new methodology for the management of the accident data and by establishing the relationship between accidents and roadway characteristics within it. For the first issue, authors introduced geographic information system (GIS) into the analysis framework of the accident data since it is believed that analysis based on GIS seems to provide more effective information in reducing accidents. Point-based, line-based, and polygon (grid)-based approaches were set up along with concrete examples. Especially, the location-based scores such as localization, specialization coefficients, and Tress score have been added to identify the intensity of certain accident types within study area or grids. The second issue addressed the equation formulation of accident and fatality numbers with roadway characteristics like number of intersections and road length in a grid with a sense that (1) accidents on roadways are the function of the roadway physical characteristics rather than the socio-economical secondary data, (2) the equation can be applied to the any "suggested" area, not just region or nation, and (3) the accident forecasting model should emphasize the accident location itself more than any other factors. Some equations based on those assumption have been derived along with some future research agenda.

1. 서 론

수년 전부터 우리 사회는 자동차 시대로 불릴 만큼 급속하게 차량들이 증가하였고, 이로 인한 여러 가지 문제들이 발생하고 있다. 1990년대 초까지 지속적인 증가를 보이던 교통사고는 최근 들어 감소 추세를 보

이고는 있으나 정신적, 물질적 피해는 아직도 큰 비중을 차지하고 있다. 1994년 현재 전국의 교통사고 발생 건수는 26만 건에 이르고 있으며, 사망 1만명, 부상 35만명, 재산 피해액 3천5백억 원으로 평균 하루 10억원의 직접적인 손실이 발생하고 있고, 이와 더불어 교통혼잡비용까지 포함하면 교통사고는 생활상의 매우 큰 비중을 차지하고 있다

* 아주대학교 교통공학전공 부교수

** 아주대학교 교통연구센터 연구원

*** 아주대학교 교통공학전공 조교수

(도로교통안전협회, 1995). 날로 심해지는 교통사고는 인적, 물적 손실과 직결되어 있는 것으로 교통사고의 분석과 각종 교통안전시설물의 체계적인 관리를 요구하고 있다. 그러나 현재 우리나라의 각종 교통관련 정보는 일부 전산화되어 있기는 하지만, 종합적인 교통관련정보의 수집과 분석이 요구되는 교통사고분석시스템을 구축·활용하기에는 매우 어려운 상황이다. 특히 교통사고는 공간적, 지형적 발생 요인을 하나의 중요한 사고의 원인으로 볼 수 있는데, 지금과 같은 문자기반의 교통사고에 대한 전산화는 더 이상 효율적인 데이터의 관리 체계로서는 그 기능을 다할 수가 없는 상황이다.

이에 본 연구 논문에서는 최근 여러 부문에서 관심을 갖는 지리정보시스템(GIS : Geographic Information Systems)과 지역분석적 기법을 이용하여 교통사고를 보다 체계적으로 관리·분석 할 수 있는 방안과 그 기법에 대하여 살펴보기로 한다. 또한 지형 분석자료를 바탕으로 도로의 물리적 환경에 기반을 둔 교통사고 예측 모형에 대해 고찰함으로써 도로구조와 교통사고의 상관관계에 대해 살펴보기로 한다.

2. 교통사고 현황과 분석 기법

2.1 교통사고 발생 현황

최근 5년간 우리나라 전체에서 발생한 교통사고 발생건수는 1991년에 26만 여건이었고, 1995년에는 25만 여건으로 75년 공식적인 통계조사 이후 급속한 증가 추세를 보이던 것이 잠시 주춤한 상황이다 (경찰청/도로교통안전협회, 1996). 교통사고발생으로 인한 사망자의 수는 1991년 이후 지속적인 감소 경향을 보이고 있는데 이는 1992년 이후 전개되어 온 교통사고 줄이기 운동의 결과로 인한 영향이라고 보여진다. 관련된 사항을 정리하면 표 2.1과 같다.

표 2.1 교통사고 발생 건수 ('91-'95)

구분	1991년	1992년	1993년	1994년	1995년
발생건수	265,964	257,194	260,951	266,107	248,865
증감율(%)	4.2	-3.3	1.4	2.0	-6.5
사망자	13,429	11,640	10,402	10,087	10,323
증감율(%)	9.0	-13.3	-10.6	-3.0	2.3
부상자	331,610	325,943	337,679	350,892	331,747
증감율(%)	2.3	-1.7	3.6	3.9	-5.5

표 2.2 95년도 서울시 인구, 차량, 교통사고 현황

(단위 : 인, 대, M, 개소)

구 이름	95년 인구	사고건수	사망자	부상자	차량수	승용차	승합차	화물차	도로연장	교차점
송파구	10,229,262	42,228	877	2,019	2,026,962	1,614,646	98,287	312,010	1,959,645	5,270
송파구	636,485	2,427	47	46	148,326	123,989	3,291	21,000	136,269	316
노원구	569,471	1,452	29	19	105,857	88,288	6,710	10,840	97,261	250
관악구	548,420	1,695	29	22	80,755	65,076	2,413	13,244	60,351	138
강남구	534,861	2,464	59	208	162,717	141,593	1,624	19,292	128,349	280
강서구	502,979	1,841	38	74	98,281	78,627	5,200	14,380	97,033	261
성북구	490,389	2,114	22	29	72,692	56,378	6,162	10,123	79,827	201
강동구	486,971	1,646	33	39	99,349	80,719	4,092	14,499	91,915	195
은평구	485,322	1,357	32	35	85,562	68,228	5,877	11,422	58,969	133
양천구	460,593	1,209	38	83	96,290	78,192	4,290	13,725	78,380	234
중랑구	439,945	1,653	42	42	75,767	56,526	6,219	12,980	56,877	108
동작구	422,564	1,182	26	30	69,084	58,513	1,594	8,947	59,519	206
동대문구	415,318	2,506	58	93	66,655	48,865	4,777	12,920	63,247	180
영등포구	413,936	2,096	51	172	93,159	69,156	1,837	21,994	147,179	480
강북구	389,995	1,106	11	27	55,778	42,707	4,700	8,344	40,617	89
마포구	385,917	1,950	26	51	71,931	57,900	2,367	11,613	92,354	266
서초구	384,546	1,915	44	371	119,821	100,492	1,401	17,557	128,310	337
광진구	382,118	909	12	31	71,664	55,891	5,228	10,514	53,567	130
구로구	364,928	1,724	37	81	68,315	50,883	3,415	13,936	67,134	242
서대문구	359,808	1,456	34	21	63,094	50,028	4,295	8,750	50,047	126
도봉구	355,966	1,364	23	40	70,886	57,850	4,307	8,689	56,142	140
성동구	334,320	1,942	38	73	55,308	40,389	4,004	10,842	72,037	198
금천구	284,955	1,157	30	19	48,383	35,737	1,801	10,826	48,643	146
용산구	253,765	1,673	53	58	51,624	39,593	3,594	8,379	73,737	184
종로구	190,243	1,980	30	66	46,552	35,329	4,699	6,458	65,611	211
중구	135,447	1,410	35	289	49,112	33,697	4,390	10,736	56,270	219

위 표의 자료는 인구부문은 1996년 통계청, 교통사고는 1996년 도로교통안전협회의 시군구별 주야별 교통사고 내부자료, 차량등록은 1999년 서울시 교통국 주차계획과 내부자료이다. 한편, 우리나라 인구의 4분의 1이 거주하는 서울시의 경우를 살펴보면, 25개의 행정구로 구분되어 있다. 서울지역의 인구와 자동차 등록 현황 및 서울지역에서 발생한 교통사고의 발생 건수에 대해 살펴보면 표 2.2와 같다. 95년 말 현재 서울시의 인구는 10,229천명이고, 95년 한해 동안 서울에서 발생한 전체 교통사고는 4만2천 여건, 사망자 수는 877명에 이르고 있다. 차량등록현황에 있어서는 96년 2월 현재 자가용 승용차가 161만 여대, 승합차가 10만 여대, 화물차가 31만 여대, 특수차가 0.2만 여대로 총 등록대수는 202만 여대이다. 각 구별로 교통사고의 현황을 살펴보면 동대문구와 송파구가 가장 많은 사고가 발생하고 있고, 광진구가 가장 낮은 사고발생건수를 나타내고 있다. 교통사고 현황에 관한 자료는 교통사고가 발생하는 관할 경찰서 교통사고 조사반에서 개별사고에 대한 조사보고서를 작성하고 이를 토대로 교통사고 통계연표를 작성, 이를 토대로 통계수치화 한다. 수집된 각종 통계자료는 도로교통안전협회에서 일괄적으로 처리하여, 매년 “지역별 교통사고 통계”와 “교통사고 통계분석”이라는 2권의 책자로 정리하여 발간된다.

2.2 교통사고의 전통적 분석 기법

교통사고의 시간적 변화, 공간적 차이 또는 사고요인에 따른 사고율을 비교분석하고, ‘사고 많은 지점’을 판별하며, 사고원인을 규명하는 분석을 모두 합해서 교통사고 분석이라 통칭한다. 교통사고분석의 궁극적인 목적은 사고를 감소시키고, 사고의 심각성을 감소 시킴으로써 사람의 생명과 재산을 보호하기 위한 대책을 수립하는 것이다. (도철웅, 1995; Simon et al., 1991)

교통사고를 분석하기 위한 전통적인 방법으로 다음과 같은 4가지가 있다.

- ① 기본적인 사고통계 비교분석 : 국가, 지역 내, 지역 간, 도로종류별 사고통계. 사고발생주체별 사고통

- 계, 사고발생구간 또는 지점별 사고통계
- ② 사고요인 분석 : 도로, 교통, 차량, 교통안전시설, 교통운영방법과 사고율과의 관계
- ③ 위험도 분석 : 사고 많은 구간 또는 지점을 판별
- ④ 사고원인 분석 : 사고 많은 지점 또는 특정한 사고에 대해서 그 원인을 분석하거나 규명하는 미시적인 분석

그러나 우리나라의 경우 교통사고에 대한 분석은 크게 2가지의 형태로 이루어져 왔다. 하나는 전술한 바와 같이 사고발생에 대한 각종 자료를 통계화하여 각 지역별로 교통사고자료를 집계, 분석하는 것이고¹⁾, 또 다른 하나는 교통사고가 빈번히 발생하는 교차로 및 지점에 대한 개선의 목적으로 특정지점에서 발생하는 교통사고를 사고유형, 발생빈도 등의 자료를 이용하는 것으로서 주로 ①, ④의 분석방법을 이용한다고 볼 수 있다.

즉, 전자는 일반적인 기술 통계적 기법을 이용하여 자료를 분석하는 것이고, 후자는 특정지점에서 발생한 모든 사고자료를 취합하여 충돌도(Collision Diagram)를 작성하고 이를 토대로 사고의 유형과 사고발생지점에 대한 대사고 예방차원의 미시적인 분석을 수행함으로써, 사고의 방지가 주된 목적이 된다.

2.3 지리정보시스템과 교통사고의 분석

교통사고는 어떤 특정구간의 몇 가지 단순한 요인에서 발생하는 것이 아니라, 가로망체계와 교통시설물의 유무, 도로망의 물리적 조건, 운전자의 행태, 기상 등의 특성이 상호 복합적인 관계 하에서 유기적으로 변화되는 과정 중에 발생한다. 그러므로 사고의 분석은 사고발생지점의 지형적인 현황과 사고발생이력은 물론 총괄적인 지리정보가 갖고 있는 속성들을 포함하여 수행되어야 한다.

지리정보시스템이란 공간상에 존재하는 모든 객체를 상대적 공간성을 바탕으로 컴퓨터 상에 입력하고, 그 객체와 관련된 모든 속성자료를 데이터베이스의

1) 매년 도로교통안전협회에서 ‘지역별 교통사고 통계’와 ‘교통사고 통계분석’이라는 연구의 성과로 보고된다

형태로 구축하여 동시에 관리함으로써 데이터 관리의 시너지(Synergy)효과를 창출하기 위한 정보체계이다. 지리정보시스템은 공간적인 정보와 속성정보를 동시에 관리할 수 있는 능력을 가짐으로써, 공간적으로 분포, 발생하는 모든 객체²⁾에 대한 데이터베이스화가 가능하다.

교통사고의 관리 및 분석에 있어서는 지리정보시스템의 기본적인 입력, 표현, 검색 기능과 함께 중첩, 추출, 공간검색 등의 공간분석기능을 활용할 수도 있다. 최근 지리정보시스템의 본격적인 공급과 더불어 이를 교통분야에서 활용하는 사례가 늘고 있는 실정이다. 그러나 교통사고의 관리 및 분석 분야에서는 그 이용도가 미미하고, 아직 체계적인 연구가 없으므로 본 고에서는 이를 중점적으로 검토해 보고자 한다.

교통사고의 분석에 이용될 수 있는 지리정보시스템의 주요한 기능을 정리하면 다음과 같다.

- ① 공간적 중첩
지형정보로서 도로망, 행정구역, 수계, 등고선, 주요 지형지물 등의 기본적인 단위(Layer)를 서로 중첩하여 화면상에 표현함으로써 다양한 정보를 하나의 화면상에 표시하는 기능
- ② 공간적 검색
사용자가 지정하는 임의의 공간을 속성에 기초하여 검색하는 기능
- ③ 공간적 추출
사용자가 공간적으로 검색한 내용을 새로운 데이터 집합(Data Set)으로 저장하거나 재구성하는 기능
- ④ 공간적 측정
사용자가 지형 정보를 입력시 특정하게 사용되는 측지좌계를 이용하여 입력할 경우 실제지형과 동일한 측정 값(예, 거리, 면적, 둘레)을 산출할 수 있는 기능
- ⑤ 공간적 연산
기존의 공간 정보를 부울(Boolean)대수를 이용하여 공간적인 포함관계를 연산하거나, 검색 또는 추출된 공간정보를 이용하여 각종 통계치를 계산하는 기능

2) 본고에서의 사고도 이러한 객체에 속한다.

3. GIS 기반의 교통사고 분석

3.1 분석방법론의 정립

교통사고분석의 목적은 개별적인 교통사고의 발생 현황 및 원인분석 뿐만 아니라, 특정지점에 대한 사고의 유형 및 형태 분석, 또는 도로의 일정구간에 대한 사고의 발생 현황 및 요인 분석, 나아가서는 도시 전체에 대한 사고의 현황 및 분석을 통해 보다 안전한 도시 환경을 구축하는데 있다.

그러나 앞에서도 언급했듯이 현재의 교통사고 분석 방법은 단순히 개별사고의 보고 및 분석에 초점이 맞추어져 있거나 연간 사고 발생건수를 이용한 기술 통계에 대해서만 분석을 하고 있는 것이 주된 현황이므로 실질적인 교통 사고에 대한 대책 및 위험도의 평가에 대해서는 별다른 방법론을 제공해주고 있지 못하다. 따라서 본 고에서는 교통사고를 보다 효과적으로 관리, 분석할 수 있는 체계로서 지리정보시스템을 이용한 사고분석의 새로운 방법론 몇 가지를 제안하고자 한다.

컴퓨터의 발달과 더불어 다양한 정보시스템들이 발달을 거듭하여 왔는데, 그 중 지리정보시스템은 지리 정보와 속성정보를 하나의 단일한 체계 내에서 운영함으로써 속성정보와 관련된 그래픽정보를 동시에 이용하여 미시적인 분석과 거시적인 분석을 동시에 수행할 수 있는 기반을 제공하고 있다. 지리정보시스템은 대상체(Object)를 표현함에 있어서 점(Point-Zero dimension), 선(Line-One dimension), 면(Polygon-Two dimension)으로 나타낸다(Bernhardsen, 1992). 따라서 지형 및 관련속성정보의 구축과 분석에 있어서도 위의 세 가지 표현 방법을 기초로 한다. 본 연구에 있어서도 위의 3가지 기본 체계를 교통사고를 관리, 분석하는 근본적인 틀로서 활용하기로 하며 분석의 대상은 수원시로 하고 이 지역의 교통사고자료를 활용하였다.

3.2 점적 분석

점(Point)적 분석은 개별 교통사고의 점적인 분석으로서 대개의 교통사고 분석에 흔히 활용되는 분석 방

법이다. 기존의 교통사고 분석을 위한 주된 방법으로 활용되었던 이 분석 방법은 교통사고가 빈번히 발생하는 지점에서 충돌도(Collision Diagram)를 그리고, 그 지점에서 발생하는 사고의 유형과 패턴에 대해 분석함으로써 어떠한 사고가 많이 발생하는가를 규명하는 일반적인 방식이다.

교통사고분석에 기본적으로 활용되는 충돌도는 화살과 기호로 사고에 관련된 차량이나, 보행자의 경로, 사고의 유형 및 정도를 도식화한 것을 일컫는다. 그림 3.1은 충돌도의 한 예를 시한 것이다. 충돌도는 대개 정해진 축척이 없으며, 분석가가 임의의 축척으로 그리는 것이 보통이다. 이로 인해 사고가 많이 발생하는 지점의 경우 표기되는 화살표나 기호가 중복되거나 뒤섞이기 때문에 정확한 사고의 형태나, 유형을 판단하기가 어려운 경우도 있고, 사고차량이나 운전자의 특성과 같은 속성적인 정보는 나타낼 수 없는 한계가 있다(김경환, 1995 ; ITE, 1965).

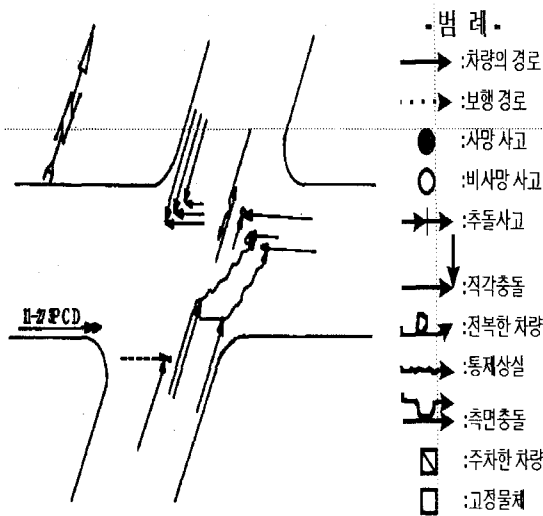


그림 3.1 교통사고 충돌도

반면 지리정보시스템을 이용한 점적 분석은 대상지역을 상대로 사고발생과 관련속성정보를 동시에 구축할 수 있고, 실제의 지형조건을 그대로 반영할 수 있을 뿐만 아니라, 다양한 특성정보를 데이터베이스화

여 동시에 나타낼 수 있으므로 사고의 발생형태나 유형을 손쉽게 알아볼 수 있는 장점이 있다.

점적 분석에서는 특정지역의 부분적인 교통사고 발생 현황뿐만 아니라, 도시 전체의 사고발생 정도를 알 수 있고, 지리정보시스템이 가지고 있는 속성별 분류에 의한 교통사고 발생 유형을 표현할 수 있으므로, 사고의 유형별 분류나, 정확한 사고의 기록 및 관리가 가능해진다. 즉 사고의 발생정도가 도시의 어떤 부분에서 밀집해 있는지, 또는 어떤 사고가 도시의 어느 부분에서 많이 발생하는지, 그리고 그 지역의 개략적인 기하구조와 어떠한 관계성을 갖는지를 파악할 수 있다.

지리정보시스템을 이용하여 사고를 분석하는데 있어 가장 이로운 점은 사고위치의 확인 및 위치 입력 작업이다. 교통사고 발생시 운전자나, 목격자들이 사고 발생지점을 스프레이나 표시도구를 이용하여 발생지점을 표시하고 있으나, 교통사고의 조사에 있어서는 사고조사원부에 사고지점의 대강의 위치³⁾나 그림으로만 표현되어 있을 뿐, 정확한 위치를 표현하고 있지는 못하다. 그러나 최근 측지측량 기술이 발달하면서, 선진 외국의 경우 범세계측위시스템(GPS: Global Positioning System)이라고 하는 장비를 이용하여 사고발생지점의 정확한 위치를 확인하는 작업이 동시에 이루어지고 있다⁴⁾. 따라서 앞으로는 정확한 위치자료를 토대로 한 사고의 분석이 이루어져야 하며 첨단장비와 기술을 이용한 분석방법 등이 지속적으로 요구된다고 볼 수 있다.

점적 분석의 활용은 주로 관리적인 측면이 강조된다. 이러한 관리적인 면에 있어서 지리정보시스템은 발생한 사고를 그 내용과 발생위치에 따라 컴퓨터 상에 입력하고 이를 데이터베이스화하여 관리하는 것이다. 그림 3.2는 점적분석의 예로 교통사고관리를 위한 지리정보시스템의 관리화면을 나타낸 것이다.

- 3) 예를 들면, 조사원부의 위치 기록란에는 “강남역 사거리 앞, 테헤란로 시작지점 부근...”이라고 위치를 현하거나 약도를 그리는 것이 대부분이다.
- 4) 최근 도로교통안전협회에서도 GPS를 이용하여 교통사고 및 교통시설물관리를 위한 장기연구가 수행 중에 있다.

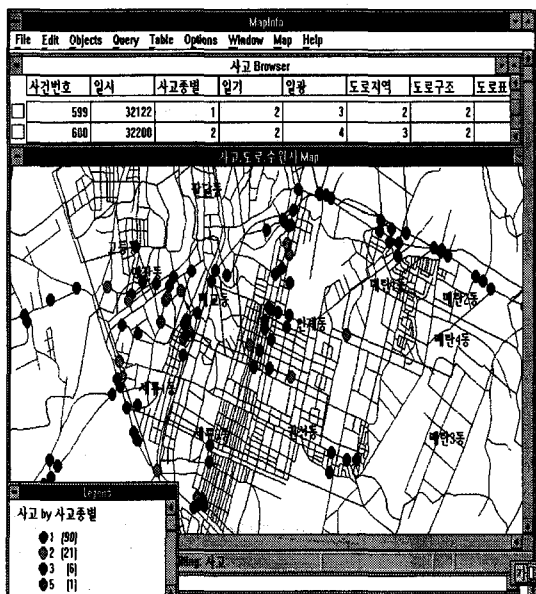


그림 3.2 교통사고의 점적 분석과 사고유형의 표시

- 1: 차와 차 사이의 사고 2: 차와 사람간의 사고
- 3: 차와 기타 물체와의 사고 4: 전도
- 5: 도로상의 전복 6: 도로의 전복 7: 추락

3.3 선적인 분석

GIS를 이용한 교통사고 분석의 두 번째 기법은 사고데이터의 선적 정보화 및 분석이다. 교통사고는 “도로”라는 선적인 객체 상에서 발생하는 것이 대부분이므로 이를 지도상에 표현하면, 도로의 형태를 따라 선형으로 분포하게 된다. 교통사고는 도시의 동맥이라고 할 수 있는 간선교통망을 중심으로 발생하는 사고가 대부분을 이루고 있으므로, 교통사고의 분석은 단순한 지점에 머무는 것이 아니라 여러 지점을 이어서 하나의 선상에 놓고 공간적인 상호관련성을 분석할 필요가 발생하며 바로 이런 이유에서 선적인 분석의 필요가 등장하게 된다.

선(Line)적인 분석을 위한 기법으로서 지리정보시스템에서 많이 활용되고 있는 것이 버퍼링(Buffering) 또는 버퍼분석(Buffer Analysis)이다. 버퍼분석은 점이나 선, 면(Polygon)을 중심으로 일정한거리의 거리 안에 드는 영역을 새로운 구역으로 설정하고, 이 구역에 존재하는 각종 대상체(Object)를 분석하는 기법이다. 사

용자는 분석의 목적에 따라서 도로의 네트워크를 기준으로 임의의 거리를 지정하여 새로운 구역을 생성하고 이 구역과 기존의 교통사고 발생데이터를 중첩(Overlay)시킴으로써 새로운 정보를 도출해 낼 수 있다. (물론, 이러한 buffering의 단점도 있을 수 있음. 즉, 간선 및 이면도로의 사고가 함께 추출되어 희석되어질 수 있음. 그러나 본고에서는 주간선을 따라 그것도 약 30미터 내외로 buffering을 수행하였음.)

예를 들어 선적인 분석은 사용자가 1번 국도와 42번 국도의 수도권 구간에서 발생한 교통사고를 확인하고자 할 경우 지리정보시스템은 공간적인 질의와 버퍼링 기법을 이용하여 2개의 국도 상에서 발생한 교통사고를 공간적으로 검색하여 2개 국도상의 개별 사고에 대한 정보를 사용자의 요구에 따라 화면상에 도출하거나 출력할 수 있고, 2개의 국도 상에서 발생한 총 발생건 수 등 사용자의 조건에 부합하는 사고의 표출도 가능하게 해준다.

따라서 선적인 분석은 연구대상지역의 도로 축을 대상으로 어떤 도로 축에서 교통사고가 많이 발생하는가를 규명해 볼 수 있는 기능을 제공한다. 그림 3.3은 대상지역의 주요 간선도로를 따라 선적인 분석을 수행한 결과이다. 즉, 주요간선 축에서 발생한 교통사고를 일정한 기준에 따라 화면장에서 표시하고 이를 출력한 예이다.

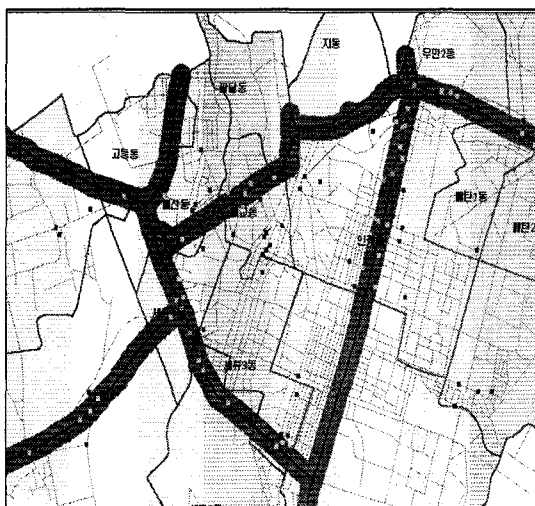


그림 3.3 교통사고의 선적인 분석

3.4 면적인 분석

면(Polygon)적 분석은 두 가지로 구분할 수 있다. 행정구역이나 교통분석 존과같이 부정형의 면을 분석의 기본단위로 설정하는 것과 일정격자의 격자(Grid) 상에서 정형화된 면을 분석의 기본단위로 설정하는 것이 있다. 전자의 분석 방법은 행정구역 단위의 각종 통계자료와 사고발생자료를 연계하여 분석할 수 있는 환경을 제공하고, 후자는 전체 대상지역을 하나의 동일한 분석단위로 설정하여 분석을 수행할 수 있는 환경을 제공한다. 면(Polygon)적 분석은 사고가 많이 발생하는 지역, 또는 특정사고가 편중되어 있는 지역에 대한 분석과 같이 지역(또는 격자 별) 특성분석을 수행할 수 있는 장점이 있다. 그림 3.4는 면적인 분석의 예로서 그리드로 구분한 각 존의 사고발생빈도를 지도상에 표출한 것이다.

면적 분석은 앞에서 제시한 점적 분석이나, 선적 분석과는 달리 교통사고의 발생빈도에 따른 공간적 위험도의 평가를 가능하게 해준다. 특히 면적인 분석 기법은 교통사고의 지역적인 발생현황과 각 지역에 대한 위험도를 평가함으로써 의사결정자가 개별지역에 대한 위험도를 평가하고 이를 토대로 교통안전시설의 지역적인 확충에 대한 판단을 수행하는데 기본적인 의사결정의 자료로 활용할 수 있다는 장점을 가지고 있다.

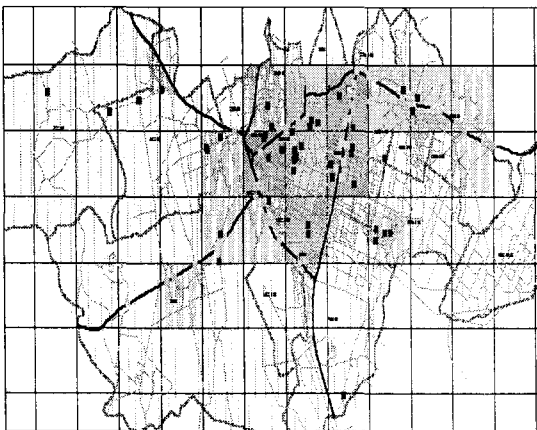


그림 3.4 교통사고의 면적인 분석 예(그리드별 발생 빈도)

면으로 구성된 대상체의 통계적 분석기법으로는 많은 방법들이 있으나, 본 연구에서는 계량지리학에서 주로 이용되는 공간분석법들을 모태로 하여 세 가지 분석기법을 제안하고자 한다.

4. GIS-GRID기반의 교통사고 지표분석

4.1 위험도 평가 기법 1: 집중계수

대개의 경우 사고가 어디에 집중적으로 발생하는가를 분석하는 데에는 사고의 발생 빈도 수를 우선적으로 고려한다. 그러나 사고의 빈도 수는 지역의 규모나, 적용하는 기준에 따라 많은 차이가 발생할 수 있다. 따라서 단순히 빈도 수를 이용하여 지역의 위험도를 평가하는 것은 외적인 영향을 고려하지 못할 수도 있다. 따라서 연구지역전체에 대한 해당지역의 사고의 빈도 수를 지표화 함으로써 사고의 발생정도를 규명하는 것이 보다 합리적이라고 할 수 있을 것이다. 따라서 본 연구에서는 사고의 공간적 집중도를 파악하기 위하여 집중계수와 특화계수 등을 위험도를 평가하기 위한 기법으로 제시하였다.

첫 번째 위험도 평가기법은 사고의 공간적 분포와 집중의 정도를 알아보기 위한 것으로 흔히 '집중계수(Coefficient of Localization)'라고 하는 계수 추정법을 이용한 것이다. 집중계수는 원래 특정산업의 지역분포, 즉 산업입지의 공간적인 분산정도를 일정시점 또는 시점간에 있어 비교하는 수단으로서 뿐만 아니라 특정산업의 입지분포 변동을 알 수 있게 하는 척도로서 이용되어 왔다(국토개발연구원, 1981).

본 고에서는 이를 이용하여 사고의 유형별로 공간적 분포 특성이 어떠한가를 살펴보고자 한다. 즉 교통사고를 사용자의 분석목적에 따라 유형별로 분류한 뒤, 어떤 특정한 유형이 대상지역 전체에 걸쳐 어느 정도 집중되어 나타나는가를 판단하기 위한 분석 방법이다.

일반적으로 집중계수는 0과 1사이에 존재한다. 구해진 값이 0에 가까우면 특정사고가 전 지역에 걸쳐 고르게 발생하고 있다는 것을 의미하며, 구해진 값이 1에 가까우면 특정사고가 한 지역에 집중하여 발생하

고 있음을 의미한다. 이는 점적 분석에 있어서도 시각적으로 확인할 수 있으나, 집중계수를 이용할 경우 계량화 할 수 있는 장점이 있다 (Choi and Park, 1996).

본 연구에서 이용한 집중계수의 공식은 다음과 같다.

$$C^k = 1/2 \sum_i \left| \frac{e_i^k}{e^k} - \frac{e_i}{e} \right| \quad (4.1)$$

(단, $0 \leq C^k \leq 1$)

여기서,

C^k = 집중계수

e_i^k = 일정시점의 존 i 의 사고유형 k 의 발생회수

e^k = 일정시점의 전체지역의 사고유형 k 의 발생회수

e_i = 일정시점의 존 i 의 총사고 발생회수

e = 일정시점의 전체지역의 총사고 발생회수

표 4.3 집중계수의 결과 값

사고의 유형	집중계수
차 대 차 사고	0.11619
차 대 보행자 사고	0.47911
차 대 기타사고 (이륜차, 자전거 등)	0.77778
차의 전도(단독)	-
도로 위 전복(단독)	0.97619
도로 밖 전복(단독)	-

연구대상지인 수원시 팔달구 일부지역의 교통사고를 기초로 하여 분석한 결과 각 사고의 유형별로 얻어진 집중계수는 표 4.3와 같다.

이상의 결과를 보면 ‘차 대 차 사고’는 도로의 전지역에 고루 분포하는 반면, ‘차 대 기타사고’와 ‘도로 위 전복사고’는 대상지역 중 특정지역에 집중적으로 분포한다는 것을 알 수 있다.

4.2 위험도 평가 기법 II: 특화계수

앞에서 살펴 본 집중계수는 사고의 집중의 정도를

계량적으로 산출하는 것으로 교통사고의 공간적 분포를 보여주지만 각각의 면적 구분(Zone)내에서 특화된 교통사고의 유형이 무엇인가에 대해서는 밝혀주지 못한다. 반면에 특정지역에 있어 어떤 교통사고가 많이 발생하는가를 판별하기 위해서는 새로운 계량의 척도가 필요하다. 이처럼 개별교통사고의 발생확률을 판단하는 것이 특화계수(Coefficient of Specialization)이다.

표 4.4 특화계수의 산출의 예

존	차대차	차대 사람	차대 기타	차의 전도	도로위 전복	도로밖 전복	Ci
18	0.206	0.151	0.048	0.000	0.008	0.000	0.206
19	0.794	0.151	0.952	0.000	0.008	0.000	0.952
20	0.294	0.151	0.452	0.000	0.008	0.000	0.452
21	0.294	0.349	0.048	0.000	0.008	0.000	0.349
23	0.194	0.249	0.048	0.000	0.008	0.000	0.249
24	0.206	0.151	0.048	0.000	0.008	0.000	0.206
25	0.040	0.016	0.048	0.000	0.008	0.000	0.056
26	0.206	0.151	0.048	0.000	0.008	0.000	0.206
27	0.206	0.151	0.048	0.000	0.008	0.000	0.206
28	0.206	0.151	0.048	0.000	0.008	0.000	0.206
37	0.206	0.151	0.048	0.000	0.008	0.000	0.206
38	0.206	0.151	0.048	0.000	0.008	0.000	0.206
38	0.238	0.294	0.048	0.000	0.008	0.000	0.294
39	0.238	0.294	0.048	0.000	0.008	0.000	0.294
40	0.101	0.003	0.106	0.000	0.008	0.000	0.109
41	0.127	0.183	0.048	0.000	0.008	0.000	0.183
43	0.206	0.151	0.048	0.000	0.008	0.000	0.206
44	0.206	0.151	0.048	0.000	0.008	0.000	0.206
52	0.206	0.151	0.048	0.000	0.008	0.000	0.206
53	0.194	0.049	0.152	0.000	0.008	0.000	0.202
54	0.079	0.008	0.095	0.000	0.008	0.000	0.095

$$C_i = 1/2 \sum_k \left| \frac{e_i^k}{e_i} - \frac{e^k}{e} \right| \quad (4.2)$$

(단, $0 \leq C_i \leq 1$)

여기서,

C_i = 특화계수 이고 나머지는 식 (4.1)에서의 그것들과 동일함.

특화계수도 집중계수와 같이 0과 1사이의 값으로 존재한다. 임의의 Zone에서 특정계수를 구했을 경우 0에 가까우면 특정 Zone *i*에서 발생한 사고유형의 빈도수가 비슷하다는 것을 나타내고, 특정계수가 1에 가까우면 사고유형 *k*가 빈번하게 발생함을 나타낸다.

표 4.4는 대상지역의 특화계수 산출결과이다. 대상지역의 특화계수를 검토해 보면 19번 Zone의 특화계수는 '차대 기타' 사고에 대해 0.95로 19번 Zone에서는 차대 기타사고가 매우 빈번하게 발생함을 알 수 있고, 21번 Zone은 특정 유형의 사고가 지배적으로 발생하지는 않으나 '차대 사람'의 사고가 다른 사고에 비해 상대적으로 빈번히 발생함을 계량적으로 알 수 있다. 54번 Zone의 경우는 0.095로 특별히 어떤 사고가 빈번하게 발생한다고는 보기 어려운 상황이다. 산출된 특화계수를 지리정보시스템을 이용하여 표현한 것이 그림 4.1 이다.

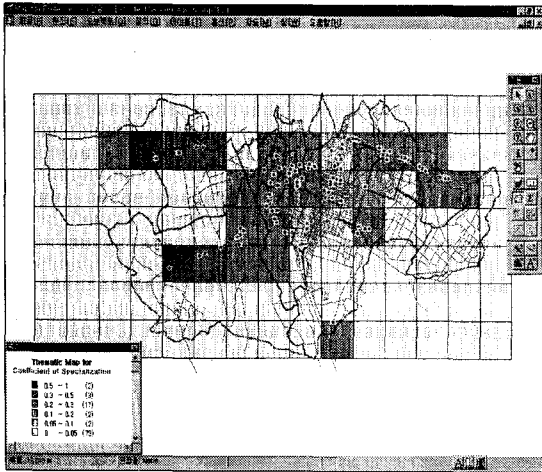


그림 4.1 특화계수 분포도

4.3 위험도 평가 기법III: 트레스 점수

트레스 점수(Tress Score)는 앞에서 제시한 집중계수나, 특화계수가 대상지역의 전체와 개별 Zone의 관계 속에서 발생빈도의 분포를 파악하는 것과 달리, 대상지역 전체와 비교하지 않고 각각의 Zone 내에서 교통사고의 균형적인 발생 패턴을 분석하기 위한 것이다.

이는 단순히 각 교통사고 유형의 구성비(%)를 낸 후 구성비가 큰 사고유형에서부터 차례로 누계를 낸 다음 다시 그 누계들의 합계를 구하는 것이다. 예컨대 어떤 Zone에 있어 6개의 사고유형 중 1개 사고유형밖에 없을 경우 구성비는 100%, 0%, 0%, 0%, 0%, 0%로 나타나고, 그 누계는 100%, 100%, 100%, 100%, 100%, 100%로 기록되며 이 경우 트레스 점수는 600% 이다. 트레스 점수가 낮은 Zone 일수록 그 지역은 모든 사고의 유형이 고루 발생한 지역이라고 말할 수 있다.

트레스 점수를 이용하여 각 Zone의 사고발생 정도를 정리하면 다음의 표 4.5와 같다.

표 4.5 대상지역의 트레스 점수

존	차대차	차대 사람	차대 기타	차의 전도	도로위 전복	도로밖 전복	TRESS
18	100%	100%	100%	100%	100%	100%	600%
20	50%	50%	100%	100%	100%	100%	500%
21	50%	100%	100%	100%	100%	100%	550%
23	60%	100%	100%	100%	100%	100%	560%
24	100%	100%	100%	100%	100%	100%	600%
25	83%	100%	100%	100%	100%	100%	583%
26	100%	100%	100%	100%	100%	100%	600%
27	100%	100%	100%	100%	100%	100%	600%
28	100%	100%	100%	100%	100%	100%	600%
38	100%	100%	100%	100%	100%	100%	600%
38	56%	100%	100%	100%	100%	100%	556%
39	56%	100%	100%	100%	100%	100%	556%
40	69%	86%	100%	100%	100%	100%	554%
41	67%	100%	100%	100%	100%	100%	567%
43	100%	100%	100%	100%	100%	100%	600%
44	100%	100%	100%	100%	100%	100%	600%
52	100%	100%	100%	100%	100%	100%	600%
53	60%	80%	100%	100%	100%	100%	540%
54	71%	86%	100%	100%	100%	100%	557%
56	100%	100%	100%	100%	100%	100%	600%
65	100%	100%	100%	100%	100%	100%	600%
66	67%	67%	67%	67%	100%	100%	467%
67	100%	100%	100%	100%	100%	100%	600%
68	100%	100%	100%	100%	100%	100%	600%
100	100%	100%	100%	100%	100%	100%	600%

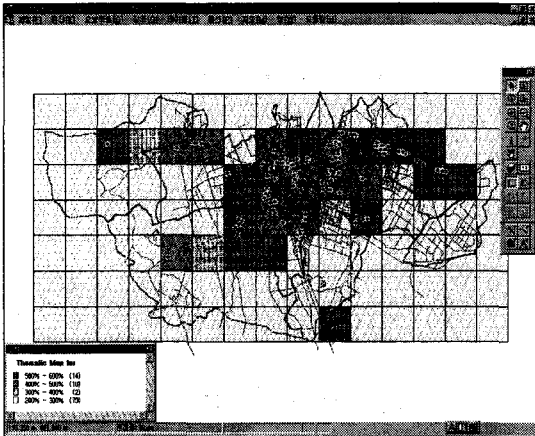


그림 42 트레스 점수 분포도

표 4.5를 살펴보면 24번 Zone은 트레스 점수가 600점으로 나타났다. 이러한 경우는 24번 Zone에서 6개의 사고유형 중 1개의 사고유형에 집중적으로 사고가 발생했음을 나타내는 것이다. 반면 66번 Zone은 트레스 점수가 467점으로 다른 Zone에 비해 다양한 유형의 사고가 발생하고 있음을 알 수 있다. 이를 지리정보시스템을 이용하여 나타내면 그림 4.2와 같다.

5. 교통사고 예측모형의 개발

5.1 기존 모형의 검토

5.1.1 스미스(Smith) 모형

국가간의 비교나 국가간 공통으로 적용될 교통사고 사망자수 설명모형을 구축할 경우는 대개 스미스모형이나 이로부터 변형된 스미스모형을 기초로 한다. 자동차 보유대수와 인구수를 독립변수로 하는 스미스모형은 1970년대 초부터 선진국을 중심으로 자동차대수가 증가함에도 불구하고 교통사고 사망자수가 감소하는 현상으로 인해 1980년 이후 스미스모형에 대한 논박이 지속되고 있다.

1949년에 스미스는 교통사고사망자 예측모형을 다음과 같이 일반화하여 제시하였다.

$$D = AN^aP^b \quad (5.1)$$

여기서,

D = 주어진 연도에 한나라의 교통사고사망자수

N = 등록된 자동차대수

P = 인구수

A, a, b = 상수

스미스는 1938년도의 16개 유럽국가와 미국, 캐나다, 호주, 뉴질랜드 등의 인구수와 자동차대수를 표본점으로 하여 교통사고사망자수를 추정하였는데 그 결과 식은 다음과 같다.

$$D = 0.0000993 N^{0.3377} P^{0.7323} \quad (4.4)$$

1970년 초기부터 유럽국가를 중심으로 사망자수가 감소하는 현상을 스미스모형으로는 설명하기 어려워 이에 대한 해석에 대해 많은 토론이 제기되고 있으나, 개발도상국 등은 여전히 시계열자료를 이용하여 원래의 모형식이나 유사한 모형식으로 구축된 예측모형이 응용되고 있다.

5.1.2 파티카모형 (Partyka, 1991)

미국의 교통사고사망자수 예측모형으로서 파티카모형은 1960~1982년까지의 미국의 교통사고 사망자수의 추이를 설명하기 위해 3의 경제활동인구변수(취업자수, 비경제활동인구, 실업자수)와 1974년에 실시한 최고속도 제한의 하향으로 인한 장기효과와 1974년에 석유수출금지로 인한 단기효과를 설명하기 위해 가변수(dummy variables)를 부가한 모형으로서, 가변수에 의해 모형의 적합성이 실질적으로 개선된다는 것이다. 그러나 이 모형은 최근의 실측치와 추정치를 비교한 결과 상당한 불일치를 보여 파티카는 불일치를 최소화하기 위해 원래 모형을 변형하여 보다 단순화된 새 모형을 구축하였다. 파티카의 교통사고 사망자수 추정을 위한 경제모형 회귀식은 다음과 같고, 이 모형에서 실업률변수와 2개의 가변수는 교통사고사망자와 부의 관계를 보이며, 독립변수조합이 종속변수에 대한 설명력은 매우 높음을 알 수 있다.

$$F_t = -96781U_t + 0.4971E_t + 0.9616N_t - 3995D_{1t} - 4824D_{2t} \quad (5.2) \quad (R^2 = 0.98)$$

여기서,

- F_t = t 년도의 도로교통사고사망자수
- U_t = t 년도의 실업자수(1,000명)
- E_t = t 년도의 취업자수(10,000명)
- N_t = t 년도의 비경제활동인구수(1,000명)
- D_{1t} = 74년 중동산석유수출금지 영향을 나타낸 가변수
- D_{2t} = 74년 50mph이하 속도제한 영향을 나타낸 가변수

1991년 파티카는 80년대 후반에 들면서 점점 커지는 추정치와 실제치의 불일치는 새로운 변수의 추가로 인한 것이며 그 추가요인은 안전벨트 착용율의 증가와 음주운전의 감소에서 발생하기 시작했다고 주장했다.

파티카는 회귀식의 설명력이 점차 떨어짐으로 해서 시계열을 89년까지 연장하고 수정된 인구변수로 당초의 모형식에 적용한 결과 그 결정계수는 0.64로서 당초의 0.98보다 매우 낮았다. 이에 그녀는 다음의 간단한 모형식 즉, 교통사고사망자수 추정을 실업자수와 비경제활동인구인 두 개의 독립변수로 하는 새로운 모형을 제시하였다.

$$F_t = -100631 - 1.3686 U_t + 1.2934 N_t \quad (5.3) \quad (R^2 = 0.60)$$

5.1.3 채남희 & 심민수 모형

채남희 & 심민수(1996)는 모형의 구축에 있어서 기본적으로 스미스와 파티카모형에 의한 이론과 방법을 바탕으로 하는 새로운 모형을 정립하였다. 이들의 모형에서 교통사고사망자 예측은 사회·경제변수에 의해 구축되는 회귀식에 의해 추정하였고 예측치와의 비교를 위해 ARIMA기법을 사용하였다.

즉, 교통사고사망자와 관련이 깊을 것으로 판단되는 사회·경제변수로서 각 국 공통으로 시계열자료가 확보 가능한 다음의 13개 변수들이 선정되었다. 변수 13개가 선정된 사유는 자동차화 초창기 교통사고사망 현상을 설명할 수 있는 시계열의 확보가 가능하고, 또한 교통사고 사망자의 변동추이 특징을 나타낼 수 있는 즉, 교통사고사망자의 증대요인인 경제활동증대, 사회적 난폭성 등과 감소요인인 사고에 대한 비용 및

심리적 부담감, 경제정책, 교통약자의 수 등을 고려하였다.

- 증가요인 : 자동차 대수, GNP, 취업자수, 경제활동인구, 총인구, 14세 이하 인구, 15~64세 인구, 자살자, 타살자
- 감소요인 : 실업자수, 비경제활동인구, 65세 이상 인구, 교통사고비용지수⁵⁾

분석대상기간은 국가별로 자동차화 초기의 시계열 자료를 포함한 다음의 기간으로 하였다.

- 미국 : 1923 ~ 1990년
- 영국 : 1946 ~ 1990년
- 일본 : 1955 ~ 1990년
- 한국 : 1975 ~ 1990년

모형의 형태 : 교통사고사망자수는 여러 요인으로 부터의 영향을 받으면서 일정한 변동양태를 보이기 때문에 자동차화 초기단계에서 현재까지 비교대상국가의 사망자수 증가와 감소추세를 설명할 수 있는 두 개이상의 변수조합에 의한 중회귀모형을 이용하였다.

$$Y = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_k x_{ki} + \epsilon_i \quad (i = 1, 2, 3, \dots, n) \quad (5.4)$$

여기서,

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ 는 모집단회귀계수(parameter regression coefficient)로서 i 번째 주어진 독립변수 $x_{1i}, x_{2i}, \dots, x_{ki}$ 의 주어진 상수

ϵ_i 는 i 번째 측정된 Y_i 의 오차항으로서, $\epsilon_i = N(0, \sigma^2)$, $Cov(\epsilon_i, \epsilon_j) = 0$ (단, $i \neq j$)인 정규 분포확률.

종속변수를 설명하는 독립변수조합의 선택은 13개 사회·경제변수들의 다중상관분석에 의한 상관계수에

5) 교통사고비용은 교통사고사망과 관련된 제비용 즉 자동차안전운행과 관리에 소요되는 비용과 교통사고사망으로부터 입게되는 경제적·심리적 손실비용이 포함된다. 비용과 관련된 재화와 서비스의 종류와 구조가 너무도 다양하기 때문에 표준화된 지수 시계열이나 유사 시계열을 4국가 모두의 통계에서 발견할 수 없었다. 따라서 교통사고비용지수를 대신하여 각국별 소비자물가지수를 대용변수(proxy variables)로 이용한다.

의해 선택하고 다중상관계수는 회귀방정식의 설명력을 의미하므로 잠정회귀모형의 결정계수가 높고, 국가간의 비교에서 공통성을 얻기가 용이한 R^2 선택법 (R^2 selection)을 이용하였으며, 최적의 변수선택 기준은 가장 좋은 축소모형, 각 국에 공통으로 적용될 변수조합, 잠정회귀계수의 부호(±)가 각 국에 공통이어야 한다는 기준을 책정하였다.

각 국별로 변수선택 기준에 적합한 변수조합은 133개 조합 중 유일하게 자동차보유대수(V)와 교통사고비용지수(C)에 자연대수를 취한 변수조합이었고 공통적으로 결정계수가 높았다. 변수조합의 선정에 따라 교통사고 사망자수를 종속변수로 하고 자동차 보유대수와 교통사고 비용지수를 독립변수로 하는 일반 예측모형식은 다음과 같다.

$$\text{Ln}F = a + a\text{Ln}V - \beta\text{Ln}C \quad (5.5)$$

$$\text{또는 } F = a' + V^a C^{-\beta} \quad (5.6)$$

여기서,

V = 자동차보유대수

C = 교통사고비용지수

a, a' = 절편상수

α, β = 회귀모수

모형식에서 4개국 공통으로 자동차보유대수는 양(+), 관계, 사고비용지수는 음(-)의 관계로 작용하여 교통사고사망자수의 변동추이를 설명할 수 있는 논거를 제시하였다. 여기서 사고비용지수는 안전운행과 교통사고사망자에 관련한 제비용으로서 관련되는 모든 재화와 서비스의 가중가격을 총칭하며, 이러한 지수는 실질적으로 자동차화 초창기부터 통계화되지 않기 때문에 모든 재화와 서비스를 망라해 조사되는 소비자물가지수가 대용변수로 사용되었다.

인구 10만 명당 사망자수를 종속변수로 하여 똑같은 독립변수의 시계열에 의해 회귀모수를 추정할 결과도 결정계수가 높게 나왔고 두 변수의 부호도 각 국에 공통으로 자동차보유대수는 +, 사고비용지수는 -로 나타났다.

예측모형식 (5.7)을 구성하는 독립변수의 시계열이 시간에 의해 각각의 함수로 표현되며 사망자 예측함수는 다음의 식 (5.10)과 같이 변형할 수 있다.

$$\text{Ln}F = k + \alpha\text{Ln}V - \beta\text{Ln}C \quad (5.7)$$

$$\text{Ln}V = c_1 t \quad (5.8)$$

$$\text{Ln}C = c_2 t \quad (5.9)$$

위의 식(4.2), (4.3)을 식(4.1)에 대입하면 다음과 같다.

$$\text{Ln}F = k + (\alpha \times c_1 - \beta \times c_2) \times t \quad (5.10)$$

따라서 $\alpha \times c_1 > \beta \times c_2$ 이면 사망자수 증가

$\alpha \times c_1 = \beta \times c_2$ 이면 사망자수 불변

$\alpha \times c_1 < \beta \times c_2$ 이면 사망자수 감소

5.2 기존 모형의 평가

기존의 모형은 우선 시간적으로 일정기간동안 발생한 사고에 대한 분석에서 출발하고 있다. 또한 공간적으로는 일국(一國)을 기본 단위로 하여 일국 또는 다국(多國)의 교통사고를 예측하기 위한 모형식을 정립한 것이 대부분의 형태이다. 변수의 사용에 있어서는 주로 인구, 자동차 보유대수, 실업률 등의 사회경제적인 지표가 주된 설명변수로 이용되었다.

그러나 이러한 모형식의 정립은 다음과 같은 문제점이 있다.

① 광범위한 대상지역

대상지역이 일국에 대한 분석 위주이므로 너무 광범위한 문제점이 있다. 따라서 이들 모형을 이용하여 시·도 또는 그 이하의 규모를 분석할 경우 모형의 설명력이 적절하지 않은 경우가 발생한다.

② 사회경제지표의 설명 변수화

이들 모형은 대부분이 사회경제지표라 할 수 있는 인구, 자동차 보유대수, 물가지수 등의 지표를 설명변수로 활용하고 있다. 이러한 사회경제지표는 여타의 교통사고 요인 변수 중에서 쉽게 구할 수 있고, 확실하고, 지속적인 통계자료를 구할 수 있다는 장점은 있으나, 이들 지표들이 교통사고를 직접적으로 설명하기보다는 사회 전반적인 발전추세에 의한 간접적인 지표일 경우가 많다.

그러나 교통사고는 인구나 차량이 많다고 해서 발생한다기보다는 교통량의 규모와 일인당 운행거리의 규모가 사고와 더 많은 관련성이 있을 것이라고 판단된다. 즉 교통운행 여건 속(즉, 사고가 발생할 개연성이 있는 공간)에 얼마만큼 많이 노출되어 있는가가 교통사고 발생 확률의 정도를 결정하는 요인이 된다는 것이며, 이는 바로 인구나 차량보다는 도로의 물리적 여건에 대한 자료를 바탕으로 예측을 하는 편이 더 바람직할 것이란 대안을 제시케 한다.

③ 시계열모형

앞에서 제기된 모형은 대부분 일정기간동안 발생한 교통사고에 대한 시계열모형의 형태를 띠고 있으므로 새로운 개발지나 신도시와 같은 지역의 교통사고예측에 대해서는 매우 설명하기 어렵게 된다.

그러므로 기존의 모형이 나름대로 유의성을 가지고 있다고는 판단되나 위에서 제시한 세 가지 문제점을 해결하지 못할 경우 적절한 설명력을 갖추지 못한 것으로 판단된다.

5.3 GIS와 통계적 기법을 이용한 교통사고 예측모형 개발

5.3.1 예측모형 개발을 위한 전제 조건

본 연구에서는 교통사고를 분석함에 있어서 사회경제적인 변수를 지양하고 도로교통사고와 실제적인 관련성이 깊을 것이라고 판단되는 도로의 연장, 도로의 교차점 수와 같이 설명력이 있는 변수를 사용함으로써 기존의 예측모형에 비해 보다 신빙성 있는 모형을 개발하고, 다양한 규모의 지역에 활용될 수 있는 모형을 제시하는데 그 목적이 있다.

도로교통사고 현상을 설명하기 위해 많은 변수를 이용하기보다는 적은 변수로 많은 설명을 가능케 하는 모형 구축이 더욱 현명하다고 할 것이다. 이는 적은 변수로 현상에 대한 원인과 결과의 체계를 명확히 구축하고 이를 토대로 하여 예측력이 높다면 적은 독립변수를 선택하는 것이 더욱 유용하고 타당성이 있기 때문이다. 즉, 모형들은 그 유용성의 크기에 따라 가치를 인정받게 된다.

일반적으로 교통사고는 도로이용자, 자동차 운행대수, 교통안전체계, 사회심리적 요인 등을 복합적으로 가지는 함수식으로 표현될 수 있다. 이를 수식으로 표현하면 다음과 같다.

$$F = g(X_j) \tag{5.11}$$

여기서,

F = 교통사고 사망자수

X_j = 변수(도로이용자, 교통환경, 자동차, 교통안전체계)

본 연구에서는 다양한 교통변수 중에서 도로의 구조와 관련된 변수를 이용하여 교통사고를 예측하려고 하였다. 따라서 다음과 같은 전제조건을 이용하여 교통사고예측을 위한 모형을 개발하였다.

① 교통사고는 도로의 물리적 제반여건의 함수이다.

$$T = F(\text{도로의 물리적여건})$$

교통사고는 도로상에서 발생하는 것이므로 도로와 가장 깊은 관련성을 가지고 있다. 즉 도로의 연장이 길수록, 그리고 도로상의 교차점 또는 속성변화점이 많을수록 사고가 발생할 개연성은 많아진다. 이는 운전자 또는 도로의 이용자가 위험환경에 노출되는 시간 또는 공간적인 구조가 교통사고에 가장 민감하게 영향을 미칠 것이라는 판단에 기인한다.

② 교통사고 예측모형은 지역의 규모에 관계없이 적용될 수 있어야 한다.

기존의 교통사고 예측모형은 국가 또는 폐쇄된 지역을 분석 단위로 하는 매우 거시적인 관점에서의 예측모형이었다. 그러나, 교통사고에 대한 예측은 규모의 크기에 관계없이 예측하고자 하는 대상지역에 적절하게 적용될 수 있어야 한다. 즉 1개 구를 대상지역으로 할 수도 있고, 5개동을 대상지역으로 할 수도 있어야 한다.

③ 교통사고 예측모형은 교통사고 발생지점에 대한 분석이어야 한다.

기존의 예측모형의 경우 행정구역의 인구나, 등록된 차량대수 또는 기타의 사회경제적인 지표를 설명변수로 활용함으로써, 교통사고 발생지점이라는 공간적인 기준을 무시하였다. 예를 들어 경기도에 살고 있

는 사람이 서울에서 교통사고를 유발했을 경우 사고 통계로는 서울시에서 발생한 것으로 집계되나, 기존 모형에서는 분석을 위한 변수를 경기도의 것으로 이용하게 되는 불합리한 경우가 있었다.

5.3.2 예측모형의 구조와 형태

분석대상 자료는 95년도의 통계청 인구센서스 조사의 잠정보고서를 기준으로 한 서울특별시 및 부산광역시, 대구광역시의 구별인구와, 교통사고 발생건수, 교통사고 사망자수, 도로의 연장, 도로의 교차점수 이었다. 이는 이들 지역이 인구 300만 이상의 대도시들로 교통사고의 발생 유형이 서로 비슷하고, 도로의 연장이나, 교차점의 수, 사고발생 건수 등의 정확한 자료를 얻을 수 있었다는데 기인한다. 이들 도시 중 몇몇 구는 행정구역의 변화 등으로 인해 자료의 정리 및 구분이 불충분하여 제외되었다. 인구와 사고지표는 통계청과 경찰청의 자료를 이용하였고, 도로의 연장과 도로의 교차점 수는 지리정보시스템에 구축된 서울과 부산, 대구의 행정구역을 토대로 어떤도로 이상의 모든 도로(폭 6m 이상)를 1/25,000 축척의 지도를 이용하여 직접 디지털화하여 수치지도를 제작하고, 이를 토대로 지리정보시스템의 공간적 연산기능을 이용하여 각 구별로 공간추출 기법을 이용하여 구의 도로연장과 교차점 수를 산출하였다.

우선 교통사고발생건수에 대한 각 변수의 피어슨 상관관계를 살펴보면, 다음의 표 5.1과 같다. 도로의 연장은 교통사고 사망자 수와 0.74, 도로의 교차점 수는 0.64, 구별인구는 0.42로 기존에 교통사고 예측 모형의 가장 중요한 변수 중의 하나로 인식되었던 인구는 교통사고를 설명할 수 있는 변수로 사용하기에는 도로의 연장이나, 도로의 교차점 수보다는 비중이 떨어진다는 것을 알 수 있다. 사고의 발생건수에 대해서는 도로연장이 0.66, 도로교차점의 수가 0.64로 상관관계가 있음을 알 수 있다. 기존의 모형에서 인구를 중요한 설명변수로 사용하였던 것은 구하기가 쉽다는 측면도 있으나, 이는 기존의 모형이 시계열모형으로서 실지로는 인구수 자체를 설명변수로 썼다기보다는 인구의 증가치를 설명변수로 썼다고 보아야 할 것이다.

표 5.1 각 변수간의 상관관계

	인구수	사고발생	사망자	교차로수	도로연장
인구수	1	0.55317	0.42149	0.49539	0.5403
t	0	0.0002	0.0068	0.0012	0.0003
사고발생	0.55317	1	0.71795	0.63686	0.66184
t	0.0002	0	0.0001	0.0001	0.0001
사망자	0.42149	0.71795	1	0.67032	0.74357
t	0.0068	0.0001	0	0.0001	0.0001
교차로수	0.49539	0.63686	0.67032	1	0.96075
t	0.0012	0.0001	0.0001	0	0.0001
도로연장	0.5403	0.66184	0.74357	0.96075	1
t	0.0003	0.0001	0.0001	0.0001	0

5.3.3 예측모형의 개발

예측모형의 개발을 위한 통계적 도구로는 분석도구로 널리 이용되고 있는 SAS (김충련, 1994) 를 이용하여 각종 자료의 처리와 분석을 수행하였다. 도로의 연장과 도로의 교차점 수를 이용한 stepwise방식으로 적합한 모형을 도출한 바, 예측 모형식은 단순 선형회귀 방정식으로 도출되었고 그 회귀식과 결정계수는 표 5.2와 같다.

표 5.2 예측모형의 회귀식

모형구분	독립변수	식의 형태	회귀식	결정계수 (r ²)
사고건수	도로연장	선형	A = 847.355569 + 4.072552(L)	0.4380
	도로교차점수	선형	A = 809.006731 + 0.796817(I)	0.4056
사망자수	도로연장	선형	F = 9.906904 + 0.025886(L)	0.5529
	도로교차점수	선형	F = 9.511079 + 0.141226(I)	0.4493

본 연구에서는 교통사고를 규명하기 위한 모형식으로 통시적 데이터를 이용한 (longitudinal approach)가 아니라 특정시점의 데이터를 이용한 (latitudinal approach)-데이터의 제약으로-를 이용하여 교통사고 발생건 수와 교통사고 사망자 수를 종속변수로 한 2개의 모형을 도출하였다. 하나는 도로의 연장을 설명변수로 한 회귀모형과 또 다른 하나는 도로의 교차점 수를 설명변수로 한 회귀모형이다.

5.3.3.1 교통사고 발생 건수 모형

교통사고의 발생건수를 예측하기 위한 모형은 도로의 연장을 독립변수로 한 모형과 도로의 교차점 수를 독립변수로 한 모형의 2가지가 있다.

먼저 도로의 연장을 독립변수로 한 모형을 살펴보면, 40개의 Grid를 기반으로 각 Grid의 특성이 설명변수로 이용되었고, 99%의 신뢰구간에서 F값 26.619로 회귀식의 적합도가 검증되었다. 회귀계수의 경우에도 99%의 신뢰구간에서 T값이 각각 5.594, 5.442로 회귀계수 역시 유용성이 있다는 것을 알 수 있다. 결정계수는 0.4380으로 다소 낮은 듯하지만 회귀모형으로의 유용성은 충분하다고 판단되며, 한편 구해진 회귀식은 다음과 같다.

$$A = 847.355569 + 4.072552*(L) \quad (5.12)$$

여기서, A : 교통사고 발생건수
L : 도로연장

도로의 교차점 수를 독립변수로 한 예측모형에 있어서는 40개의 설명변수가 이용되었고, 99%의 신뢰구간에서 F값 25.930으로 회귀식의 적합도가 검증되었다. 회귀계수의 경우에도 99%의 신뢰구간에서 T값이 각각 4.822, 5.092로 회귀계수 역시 유용성이 있다는 것을 알 수 있다. 결정계수는 0.4056으로 나타났다 식(5.13).

$$A = 809.006731 + 0.796817*(I) \quad (5.13)$$

여기서, A : 교통사고 발생건수
I : 도로교차점의 수

위에서 제시한 두 식은 교통사고 발생건 수를 종속변수로 설정한 것이다. 이 두 식을 이용하여 적합성을 검증하기 위해 교통사고 발생건수를 예측해보면 표 5.3과 같다.

표 5.3 교통사고 발생건수의 실측치와 예측치

(단위 : 건, km, 개)

	실측치	예 측 치			실측치	예 측 치	
부산시		도로연장	교차점수	서울시		도로연장	교차점수
중구	695	1,077	1,204	중랑구	1,653	1,464	1,391
서구	739	1,183	1,254	강북구	1,106	1,188	1,117
동구	970	1,170	1,184	강동구	1,646	1,668	1,600
영도구	714	1,220	1,180	강남구	2,464	2,167	2,117
부산진구	1,645	1,630	1,747	강서구	1,841	1,860	1,886
동래구	3,373	1,955	2,032	금천구	1,157	1,371	1,438
남구	3,387	1,545	1,571	구로구	1,724	1,442	1,544
북구	754	1,372	1,352	관악구	1,695	2,333	2,501
해운대구	1,511	1,701	1,686	광진구	909	1,460	1,421
사하구	1,265	1,655	1,492	마포구	1,950	1,663	1,729
금정구	1,498	1,825	1,759	노원구	1,452	1,522	1,423
대구시				서대문구	1,456	1,428	1,501
북구	2,373	2,460	2,272	서초구	1,915	2,019	2,054
달서구	2,530	2,290	2,393	송파구	2,427	1,921	1,823
동구	1,835	2,324	2,038	성북구	2,114	1,389	1,294
중구	1,288	1,268	1,355	성동구	1,942	1,381	1,289
남구	1,348	1,539	1,728	동대문구	2,506	1,355	1,382
서구	2,234	1,621	1,652	도봉구	1,364	1,352	1,352
수성구	2,040	2,174	2,224	동작구	1,182	1,459	1,525
서울시				양천구	1,209	1,629	1,728
은평구	1,357	1,565	1,655	영등포구	2,096	2,032	2,146
종로구	1,980	1,313	1,276	용산구	1,673	1,399	1,355
중구	1,410	1,278	1,306				

5.3.3.2 교통사고 사망자 모형

교통사고 사망자의 경우 역시 도로의 연장과 도로의 교차점 수라는 2개의 독립변수를 이용하여 예측모형을 개발하였다. 교통사고 사망자의 경우 교통사고 발생건 수 보다 더 정확한 실측치를 얻을 수 있으나, 도시부의 경우 도로환경이나, 속도의 제한으로 인해 지방부 도로보다는 상대적으로 적은 실측치를 보이고 있다.

우선 도로의 연장을 이용한 교통사고 사망자 모형식을 살펴보면, 40개의 데이터 세트가 이용되었고, 99%의 신뢰구간에서 F값 46.991로 회귀식의 적합도가 검증되었다. 회귀계수의 경우 도로연장에 대한 회귀계수는 99%의 신뢰구간에서 T값이 6.855였으나, 절편의 경우 95% 신뢰구간에서 2.281로 약간의 차이가 있으나, 유용성은 있다고 판단된다. 결정계수는 0.5529로 높은 값을 나타내고 있다. 이는 교통사고 발생 건수 보다 훨씬 변수의 설명력이 있음을 의미한다. 교통사고 사망자의 경우 정확한 실측이 가능하기 때문에 모형식에서도 보다 정밀도 있는 모형식이 산출된 것으로 판단된다. 도출된 회귀식은 다음과 같다.

$$F = 9.906904 + 0.025886(L) \quad (5.14)$$

표 5.4 교통사고 사망자 수의 실측치와 예측치
(단위 : 건, km, 개)

여기서, F : 교통사고 사망자 수
L : 도로연장

도로의 교차점 수를 독립변수로 하여 교통사고 사망자 수를 예측하는 모형식을 살펴보면, 99%의 신뢰구간에서 F값이 31.006으로 회귀식의 적합도가 검증되었다. 회귀계수의 경우 도로교차 점의 수에 대한 회귀계수는 99%의 신뢰구간에서 T값이 5.568이었으나, 절편의 경우 90% 신뢰구간에서 1.988로 약간의 차이가 있으나, 유용성은 있다고 판단된다. 결정계수는 0.4493으로 도로의 연장을 독립변수로 사용한 것에 비해 다소 설명력이 부족한 것으로 나타났다.

$$F = 9.511079 + 0.141226(I) \quad (5.15)$$

여기서, F : 교통사고 사망자 수
L : 교차점의 수

위에서 제시한 두 식은 교통사고 사망자 수를 독립변수로 설정하고, 도로의 연장과 도로의 교차점 수를 설명변수로 하여 추정된 모형식이다. 이 두 모형식을 이용하여 교통사고 사망자 수를 실측치, 예측치 관점에서 살펴보면 표 5.4와 같다.

	실측치			예측치			
	도로연장	교차점수	사망자수	도로연장	교차점수	사망자수	
서울시							
중구	4	17	23	중랑구	42	21	29
서구	14	21	24	강북구	11	31	20
동구	18	21	22	강동구	33	38	36
영도구	13	22	22	강남구	59	55	52
부산진구	30	37	40	강서구	38	45	45
동래구	68	48	50	금천구	30	28	30
남구	70	34	35	구로구	37	30	34
북구	41	28	28	관악구	29	61	65
해운대구	57	39	38	광진구	12	31	30
시하구	44	38	32	마포구	26	38	40
금정구	35	43	41	노원구	29	33	30
대구시				서대문구	34	30	32
북구	65	65	57	서초구	44	50	50
달서구	58	60	61	송파구	47	47	43
동구	66	61	50	성북구	22	28	26
중구	17	24	28	성동구	38	28	26
남구	10	34	40	동대문구	58	27	29
서구	48	36	37	도봉구	23	27	28
수성구	52	56	56	동작구	26	31	33
서울시				양천구	38	37	40
은평구	32	35	37	영등포구	51	51	53
종로구	30	26	25	용산구	53	29	28
중구	35	24	26				

사고의 효과적인 관리와 분석을 위한 패러다임이고, 다른 하나는 교통사고에 대한 예측을 하기 위한 모델식의 개발이 그것이다.

6. 결론

현재 우리나라는 날로 증가하는 차량으로 인해 교통환경이 매우 악화되어 있는 상황이다. 그러나 이러한 교통환경을 효과적으로 유지 관리하기 위해서는 교통안전부분에 있어 많은 부분에 대한 연구와 투자가 요구된다. 특히 교통사고는 도로교통환경에서 인적, 물적 피해와 직결되는 것이므로 교통사고에 대한 관리와 분석은 교통운영 및 적절한 교통환경을 조성하는데 매우 필수적인 것이다.

본 고에서는 두 가지 측면에서 교통사고에 대한 새로운 패러다임(Paradigm)을 제시하였다. 하나는 교통

교통사고의 관리와 분석에서는 최근에 많이 활용되고 있는 지리정보시스템을 바탕으로 기존의 단순 통계적 방법보다 합리적인 분석방법으로서 점적 분석기법과, 선적 분석기법, 면적 분석기법 등을 제시하였고, 특히 면적인 분석 방법에서는 지역의 위험도 평가를 수행할 수 있는 집중계수, 특화계수, 트레스 점수 기법 등을 제시하였다.

기존의 교통사고 예측모형은 거시적인 모형으로, 한 국가 또는 폐쇄적인 지역단위에서만 분석이 가능한 모형이었다. 그러나 본 연구에서는 지역의 규모와 상관없이 교통사고의 예측이 가능한 사고예측 모형을 작성하는데 중점을 두었다.

교통사고 예측 모델의 개발에서는 기존의 예측모형

이 사회경제적인 변수를 이용하여 사고를 예측하여왔던 것과는 달리 도로의 물리적 여건에 중점을 두어 도로의 연장과 도로의 교차점 수를 설명변수로 한 회귀모형식을 개발하였다. 도로의 기하구조에 있어서 도로의 굴곡도도 교통사고 발생의 중요한 요인 중의 하나이나, 자료의 취득이 어려워 2가지 변수만을 설명변수로 활용하였다. 향후에는 도로의 굴곡도 및 선형에 대한 연구를 지속적으로 진행함으로써 본 연구를 더욱 발전시킬 수 있도록 하여야 될 것으로 판단된다.

참 고 문 헌

1. 경찰청·도로교통안전협회, 1995년 교통사고 발생 현황, pp.5-23 1996.
2. 국토개발연구원, 지역분석을 위한 계량적 접근 방법, pp.251-255, 1981.
3. 김경환, 교통안전공학, 태림문화사, 1991.
4. 김충런, SAS라는통계상자, 데이터리서치, pp.122-143, 1994.
5. 도로교통안전협회, 지역별 교통사고통계, pp.7-26, 1995.
6. 도로교통안전협회, 교통사고 통계분석, pp.23-25, 1995.
7. 도철웅, 교통공학원론-하권, 청문각, pp.722-732, 1995.
8. 통계청, 1995년 인구주택총조사 잠정보고서, pp.40-77, 1996.
9. 채남희, 심민수, "교통사고 사망자수 예측모델 구축", 교통개발연구, pp 104-114, 교통개발연구원, 1996.
10. Keechoo Choi and Incheol Park, "A Development of Traffic Safety Investigation Tool Coupled with GIS and Statistical Analysis", Geographic Information Systems for Transportation Symposium, pp.352-362, 1996.
11. Hakim Simon et al, "A Critical Review of Macro Models for Road Accidents", Accident Analysis and Prevention, Vol.23, No.5, pp.379-400, 1991.
12. ITE, Transportation and Traffic Engineering Handbook, 1965.
13. S.C. Partyka, "Simple Model of Fatality Trends Revisited Seven Years Later.", Accident Analysis and Prevention, vol.16. p422, 1991.