

도로 네트워크를 따른 교통사고 핫스팟의 시각화*

A Visualization of Traffic Accidents Hotspot along the Road Network

조나혜** · 전철민*** · 강영옥****
Cho, Nahye · Jun, Chulmin · Kang, Youngok

Abstract

In recent years, the number of traffic accidents caused by car accidents has been decreasing steadily due to traffic accident prevention activities in Korea. However, the number of accidents in Seoul is higher than that of other regions. Various studies have been conducted to prevent traffic accidents, which are human disasters. In particular, previous studies have performed the spatial analysis of traffic accidents by counting the number of traffic accidents by administrative districts or by estimating the density through kernel density method in order to identify the traffic accident cluster areas. However, since traffic accidents take place along the road, it would be more meaningful to investigate them concentrated on the road network. In this study, traffic accidents were assigned to the nearest road network in two ways and analyzed by hotspot analysis using Getis-Ord G_i^* statistics. One of them was investigated with a fixed road link of 10m unit, and the other by computing the average traffic accidents per unit length per road section. As a result by the first method, it was possible to identify the specific road sections where traffic accidents are concentrated. On the other hand, the results by the second method showed that the traffic accident concentrated areas are extensible depending on the characteristic of the road links. The methods proposed here provide different approaches for visualizing the traffic accidents and thus, make it possible to identify those sections clearly that need improvement as for the traffic environment.

Keywords: Traffic Accidents, Road Network, Hotspot Analysis, Visualization, Hotspot zones of traffic accidents

* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015S1A5B8046775).

** 서울시립대학교 공간정보공학과 박사후연구원 (First Author, Post Doctor, Department of Geoinformatics, University of Seoul, cho.nahye@gmail.com)

*** 서울시립대학교 공간정보공학과 교수 (Second Author, Professor, Department of Geoinformatics, University of Seoul, cmjun@uos.ac.kr)

**** 이화여자대학교 사회과교육과 교수 (Corresponding Author, Professor, Department of Social Studies, Ewha Womans University, ykang@ewha.ac.kr)

1. 서론

경제가 발전하고 도시화가 진행되면서 자동차 보유 대수의 증가는 교통사고 또한 증대시키는 현상을 낳고 있다. 최근 국가적 차원에서 교통안전 의식 향상을 위한 다양한 예방활동으로 교통사고 발생건수는 지속적으로 감소하고 있지만(Figure 1), 서울의 경우 다른 지역에 비해 등록된 자동차 1만대 당 교통사고 발생건수는 전국 평균 대비 광주와 함께 가장 높게 나타났다(Figure 2, 도로교통공단, 2016)

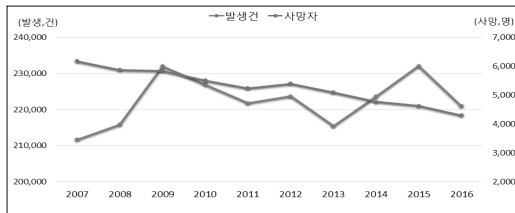


Figure 1. The number of traffic accidents and fatalities' trend.

Source: <http://www.index.go.kr>

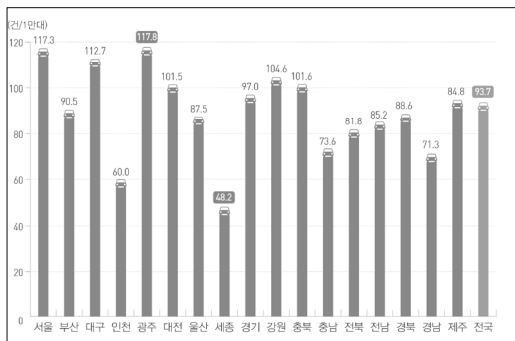


Figure 2. The number of traffic accidents per 10,000 cars each municipality.

Source: <http://www.index.go.kr>

이처럼 인적 재난에 해당하는 교통사고를 저감시키기 위하여 교통사고의 요인을 분석하고, 예방을 위해 다양한 연구들이 진행되어 왔다(강영욱 외, 2017;

Elvik, 2013; Chen, et al., 2016). 많은 연구들이 교통사고의 일반적인 속성을 분석하는데 초점을 두었다면 교통사고의 공간적 특성을 분석하는데 초점을 둔 연구들도 있다.

교통사고의 공간적 특성분석은 주로 2차원의 교통사고 클러스터 지역을 분석하기 위해 커널밀도 추정을 하거나 격자망 형태의 핫스팟 분석을 수행하는 경우가 다수를 이루었다(Anderson, 2009; Prasannakumar, et al., 2011; Xie, and Yan, 2008). 위의 연구들은 교통사고의 공간적 집중 지역을 보는데 효과적이지만, 교통사고 시간의 흐름에 따른 공간적 변화나 패턴 등을 파악하는 데는 한계가 있다. 이에 따라 최근 시간의 흐름에 따른 변화를 보여줄 수 있는 시간차원을 고려한 시각화 방법에 대해 다양한 연구가 진행되고 있다. 교통사고 연구에서도 이러한 방법으로 시공간 패턴과 시간의 흐름에 따른 공간적 변화를 분석한 연구가 있다(강영욱 외, 2017; 손세린·강영욱, 2017; Kang, et al. 2018).

교통사고를 분석함에 있어서 중요한 측면 중 하나는 교통사고 집중 지역 (hotspot, black zone)을 찾는 것이다. 기존의 교통사고를 분석하고 시각화한 연구들은 주로 행정경계나 그리드 단위의 밀집 지역을 추출하였기 때문에 도로를 따라 발생하는 교통사고 집중 구간을 찾기에는 한계가 있었다. 이에 따라 교통사고 집중 구간을 분석하기 위해서 도로 네트워크를 따라 핫스팟 지점을 찾는 것이 필요하며, 교통사고 예방 대책을 수립함에 있어서도 구체적인 사고발생구간을 확인하기 용이하다는 장점이 있다(Yamada and Thill, 2004).

따라서 본 연구는 도로 네트워크를 따른 교통사고 밀집 구간을 분석하기 위하여 도로 링크를 기반으로 교통사고 핫스팟 분석을 통한 교통사고 집중지역을 시각화하고자 한다. 또한 정확한 교통사고 집중 구간을 분석하기 위한 데이터 처리방법을 제안하고, 다른 연구에서 제안한 방법과의 비교분석을 통해 이를 검

증하고자 한다.

2. 선행연구

2.1 교통사고 시각화 연구

교통사고의 공간적 특성을 분석하고 시각화한 연구들은 2차원 상의 점 데이터의 밀도 추정을 통해 교통사고 발생의 집중지역을 분석하려는 연구가 주를 이루었다. 이수진 외(2015)는 2010~2014년까지 서울시에서 발생한 교통사고의 공간 특성과 그에 따른 영향요인을 분석하였다. 교통사고 근집패턴을 파악하기 위한 연구지역은 법정동을 기준으로 하였으며, 법정동 단위의 Getis-Ord G_i^* 를 이용하여 핫스팟 지역을 찾아냈다. 그 결과 교통사고 다발 지역으로 용산구, 중구, 영등포구, 강남구와 같은 회사 밀집 지역에서 사고 발생률이 높으며, 시간대별 사고 발생건수를 보았을 때 오후 11~12시경에 사고가 많이 발생하는 것을 확인하였다. 성병준 외(2015)은 진주시를 대상으로 5년간(2009년~2013년) 진주시 행정동별 교통사고 데이터를 주거지역과 상업지역으로 구분하여 교통사고의 공간분포특성을 분석하였다. 커널 밀도 분석을 통해 교통사고의 밀집지역을 확인하였고 km^2 당 사고 건수를 계산하여 가장 많은 사고가 발생한 동을 중심으로 시각화 및 결과를 분석하였다. 그 결과 차 대 사람의 횡단 중 교통사고는 상업지역이 높았고, 차 대 차의 측면직각 추돌사고는 토지이용 구분 없이 광범위하게 분포되어 발생하는 것으로 나타났다. Erdogan et al.(2008)은 터키의 교통사고를 고속도로에 지오레퍼런싱을 한 후 커널 밀도 분석을 통해 교통사고 핫스팟을 도출하고, 핫스팟에서의 교통사고 원인을 조사하였다. 교차로나 합류지점과 같이 교통사고가 많이 발생하는 곳을 커널 밀도로 강조하여 시각화 할 수 있었다. Anderson(2009)은 런던에서 발생한 교통사고의 공간 분포 패턴을 분석하기 위해 도시 전체를 셀 단위

로 나누고 커널 밀도 분석을 수행하여 교통사고 다발 지역을 셀 형태의 핫스팟으로 시각화 하였다.

강영욱 외(2017), 손세린·강영욱(2017)과 Kang et al.(2018)의 연구는 교통사고와 같이 시공간정보를 포함하는 데이터의 경우 공간적 밀집 분석도 중요하지만 시간의 흐름에 따른 공간적 밀집 지역의 변화, 시간 차원을 고려한 시공간 밀집 지역 분석의 중요성을 강조하면서, 여성 운전자 교통사고, 노인 운전자 및 피해자의 교통사고 데이터를 시공간적으로 분석하여 기존의 연구들에 비해 보다 정확한 교통사고 특성 분석 및 시각화를 수행하였다. 그러나 이러한 연구들은 도로 상에서 발생하는 교통사고 데이터를 그리드 셀(grid cell)단위로 시각화하여 표현함으로써 교통사고가 많이 발생하는 도로 구간을 표현하는데 한계가 있다.

2.2. 도로 네트워크 기반 교통사고 시각화연구

교통사고 데이터의 경우 도로 상에서 발생하기 때문에 단순한 행정구역이나 그리드 형태로 핫스팟 지점을 도출하는 것보다 도로 네트워크상에서 이를 시각화할 필요가 있다(Yamada and Thill, 2004; Xie and Yan, 2008). 국내 연구에서는 홍상기(1998) 연구에서 이러한 시도가 이루어졌는데, 선형 도로 네트워크상에서 교통사고 데이터의 시공간적 패턴을 분석하기 적합한 시각화 시스템을 개발하였다. 시각화 시스템을 통해 교통사고 점 데이터를 3차원으로 시각화한 뒤 특정 공간패턴을 가지는지 여부와 시간 슬라이드 쇼를 통해 시간의 흐름에 따른 점 데이터 분포의 변화를 통해 시간차원의 탐구를 진행하였다.

해외에서는 공간 통계방법에 기초한 밀도 분석을 통해 도로 네트워크에 교통사고 집중 구간을 시각화하는 연구들이 다양하게 시도되고 있다. 크게 커널 밀도 분석(kernel density estimation method), 클러스터 분석(cluster analysis) 등으로 분류할 수 있다. Erdogan et al.(2008)은 커널밀도 분석방법을 적용하

여 뉴질랜드 동부의 도시인 크라이스트처치와 터키의 Afyokarahisar의 교통사고 연구를 수행하였다. 클러스터 분석은 Levine et al.(1995)이 하와이의 호놀룰루에서 발생한 교통사고 분석에 적용되었다. 이들은 교통사고 발생 지점의 반경 내에 포함되는 사고 건수를 기반으로 커널 밀도, 클러스터 분석을 각각 적용하여 밀도를 계산하였고, 이 값을 도로 링크에 할당함으로써 도로 상에서의 교통사고 밀집 지역을 추론하였다. 그러나 이러한 방법들의 일반적 오류는 밀도 추정 에 의한 값을 도로 네트워크 데이터에 할당하였기 때문에 앞서 행정구역이나 지역 단위로 분석한 결과와 같이 도로 네트워크 상에서 발생한 교통사고 밀집 구간을 고려하기 보다는 단순히 하나의 교통사고 발생 지점을 중심으로 밀도가 높은 주변을 찾아 도로의 형태로 시각화하는 것이기 때문에 도로 네트워크를 따른 분석이라고 하기는 어렵다(Romano, and Jiang, 2017).

이후 교통사고와 같은 점 데이터를 선형으로 정제하여 분석 및 시각화하려는 연구가 시도되어왔다. Xie and Yan(2008)은 교통사고의 밀도 추정방법과 관련하여 기존의 커널 밀도 함수를 통한 표현은 선형 도로 공간에서 발생하는 점 데이터를 나타내기에는 적절하지 않다고 주장하며, lixel형태의 네트워크 커널밀도 분석을 제안하였다. 이를 통해 네트워크상에서 도로를 따른 선형 단위의 토폴로지로 교통사고를 시각화하여 어느 도로에 교통사고가 밀집되는지를 2차원의 선 두께로 표현하였다. 이 연구는 네트워크 길이를 2차원의 래스터 격자의 셀과 같은 기본 선형 단위로 나누어 분석하였고 이를 linear와 pixel을 합쳐 lixel로 정의하였다. 네트워크를 따라 일정한 간격을 둔 lixel을 기반으로 커널 분석을 수행함으로써 네트워크 토폴로지를 표현하면서 계산의 효율성을 향상시켰다. Okabe et al.(2006)은 네트워크 데이터를 기반으로 다양한 공간분석이 가능한 기능들을 포함하는 SANET 시스템을 개발하였고, 이를 바탕으로 도로 네트워크에서 발생한 교통사고의 강도에 비례하여 z축의 세로

높이로 시각화한 연구로 발전시켰다(Okabe and Sugihara, 2012). McSwiggan, et al.(2017)은 호주의 동남부 멜버른 서남쪽의 항구도시인 질롱(Geelong)의 주요 도로에서 발생한 3년 동안의 심각한 도로 교통사고의 위치를 보여주지 위하여 선형의 네트워크 커널 밀도 분석에 대한 통계적 접근법을 제안하였다. 수집된 샘플 점 데이터를 네트워크상에 매핑하고 이를 바탕으로 네트워크 기반의 커널 분석을 수행하였고, Xie and Yan(2008)과 Okabe and Sugihara(2012)의 연구방법을 응용하여 이를 시각화하였다(McSwiggan, et al., 2017).

Okabe et al.(2006)와 McSwiggan, et al.(2017)은 도로 네트워크를 따라 교통사고 점 데이터를 집계하여 커널분석을 수행하였고, 이를 다양하게 시각화하였지만, 도로의 길이가 상이하기 때문에 정확히 어떤 위치에서 교통사고가 집중 발생하는지를 나타내기에는 한계가 있었다. Xie and Yan(2008)은 일정한 단위의 lixel을 기반으로 밀집 구간을 확인하고자 하였고, 래스터 형태의 lixel을 구축함으로써 계산의 효율성을 높였다는 장점이 있지만, 서로 다른 도로의 폭을 가지는 도로 네트워크를 일정한 lixel로 정의하였다는 점과 실제 벡터의 선형 데이터를 래스터로 표현함으로써 오류가 발생하였다.

이처럼 도로 네트워크를 따라 발생하는 교통사고를 시각화하기 위해서는 교통사고의 점 데이터를 선형의 네트워크에 할당하는 작업이 이루어지게 되며, 어떤 방법으로 선형의 네트워크 데이터에 점 데이터를 할당하느냐에 따라서 결과가 크게 달라질 수 있다. 도로 네트워크를 따라 교통사고를 시각화하는 초기 연구는 교통사고 점 데이터를 불규칙한 네트워크에 단순히 집계(aggregate)하는 방법이 주를 이루었다(Ceder and Livneh, 1978; Ng and Hauer, 1989). 그러나 도로의 길이가 다른 상태에서 도로 네트워크에 교통사고 점 데이터를 할당하면, 도로 네트워크 길이에 따른 MAUP(Modifiable Area Unit Problem)2)가 발생할

수 있는 문제점이 있다(Openshaw, 1984). 즉, 도로 네트워크 길이가 각 구간별로 상이하기 때문에 길이가 긴 구간일수록 교통사고 건수가 많을 가능성이 높고 정확한 밀도 구간을 추정하기에 한계가 있을 수 있다. 따라서 일정한 구간으로 나누어 교통사고 점 데이터를 집계하여 분석하는 것이 필요해졌다(Erdogan et al., 2008; Yamada and Thill, 2010).

3. 실험

3.1. 데이터 전처리

본 연구는 2013년 서울시에서 발생한 교통사고를 분석 사례로 하였다. 교통사고 데이터는 대한민국 경찰청에서 제공한 자료를 활용하였다. 경찰청이 제공하는 교통사고 데이터에는 발생지점의 위치를 비롯하여 운전자의 성별 및 연령정보, 피해자의 성별과 연령 정보, 교통사고 발생시의 시간, 날짜, 기상상태 등에 대한 정보를 포함하고 있다. 2013년 서울시에서 발생한 교통사고 건수는 총 39,439건이다. 이중 위치 값에 오류가 있는 데이터를 제외한 총 39,010건을 분석에 활용하였다. 2013년 교통사고 발생 지점을 매핑한 것은 Figure 3과 같다.

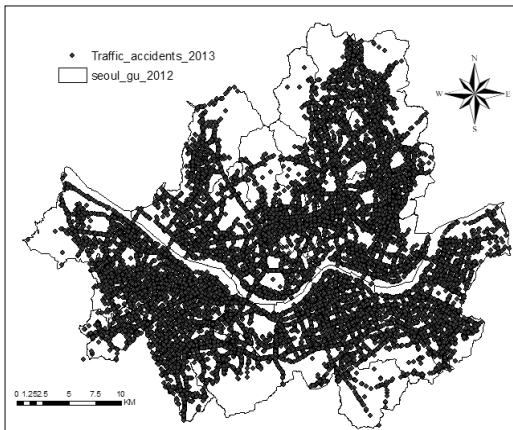


Figure 3. Traffic Accidents of Seoul in 2013.

도로 네트워크 데이터의 경우 2013년 국토교통부의 국가교통정보센터에서 제공하는 도로 링크 데이터를 사용하였다(Figure 4). 2013년 11월에 갱신된 표준 노드 및 링크 데이터는 Link ID, 도로 번호와 이름, 최고 속도, 길이 등에 대한 정보를 포함하고 있으며, 링크 길이는 km로 구축되어 있다.

본 연구에서 도로 링크는 다른 데이터와의 비교를 위해 미터단위로 변경하였다. 도로 링크 데이터의 평균 길이는 약 309.5m, 가장 긴 도로는 약 6,513.3m, 가장 짧은 도로는 약 9.7m이며, 총 22,060개의 도로 링크로 구성되어 있으며(Figure 5), 짧은 도로 9.7m를 제외한 나머지 모든 도로 링크의 길이는 10m 이상인 것으로 확인되었다. 본 연구는 자동차가 통행할 수 있는 1차선 이상의 모든 도로, 즉 이면도로도 포함하여 분석을 수행하였다.

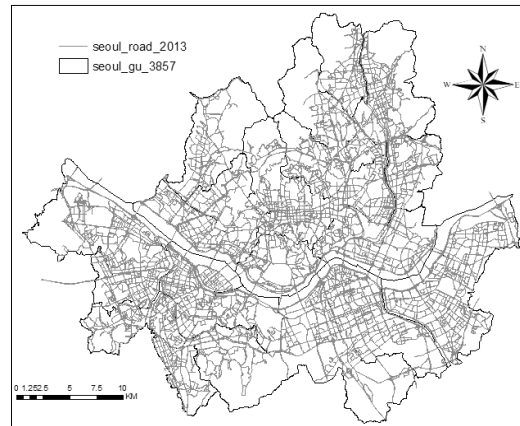


Figure 4. Road network data of Seoul.

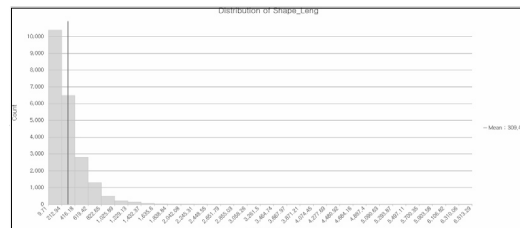


Figure 5. The length of the road link used for the analysis.

3.2. 도로 링크별 교통사고 데이터 할당

본 연구에서 도로 네트워크상에서 교통사고 밀집지역을 분석하기 위해 Figure 6과 같은 과정으로 두 가지 방법에 의해 데이터를 처리하였다. 첫 번째 방법은 불규칙한 길이를 가진 도로 링크를 일정한 간격으로 분할하였다. 도로 링크 길이에 따른 MAUP 발생을 배제하기 위해 도로 링크 길이 간 편차를 가장 작게 줄일 수 있는 10m 단위로 길이를 나누었다. 예를 들어 20m 간격으로 도로 링크를 분할하면 분할 후 도로의 길이가 20m 미만인 경우도 다양하게 나타나므로 길이 간 편차가 커지는 문제가 발생한다. 그러나 10m의 일정한 간격으로 데이터를 분할하면 분석 가능한 범위 내에서 도로 링크 길이의 평균이 약 9.3m로 도로 링크 길이 간 편차가 0.99로 가장 작게 나타난다.

10m 단위로 나눈 도로 링크에 교통사고 점 데이터를 할당하는데, 교통사고 점 데이터에서 가장 가까운 링크를 찾도록 하였다. Snap 기능을 사용하여 가장 가까운 도로 링크 데이터에 교통사고 데이터가 할당될 수 있도록 하였다. 그리고 교통사고 건수가 할당된 도로 링크를 중심으로 Getis-Ord G_i^* 통계에 의한 교통사고 핫스팟 분석을 하고 교통사고 발생 집중 구역을 시각화하였다. 사용된 Getis-Ord G_i^* 통계 공식은 식 1과 같다.

$$G_i^* = \frac{\sum_{j=1}^n w_{ij}x_j - \bar{X}\sum_{j=1}^n w_{ij}}{S\sqrt{\frac{n\sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij}\right)^2}{n-1}}} \quad (1)$$

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n}, \quad S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2},$$

x_j = j 에 대한 속성 값

$w_{i,j}$ = i 와 j 사이의 공간 가중치

n = 피처의 총 수

두 번째 방법은 도로 링크 데이터에 교통사고 데이터를 Snap 기능을 활용하여 할당하였다. 모든 교통사고 데이터가 도로 링크 데이터에 할당되었으며, 공간 결합(Spatial Join)을 통해 도로 링크에 미터 당 교통사고 발생건수를 계산하여 속성 값으로 입력하였다. 도출된 도로 링크별 미터 당 교통사고 발생건수를 이용하여 Getis-Ord G_i^* 에 의한 핫스팟 분석을 하고 시각화하였다.

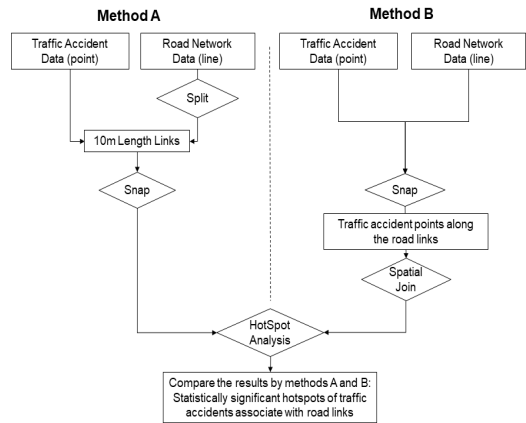


Figure 6. Data Processing

교통사고 점 데이터를 도로 링크 데이터에 할당하는 과정에서 가장 가까운 거리를 산출하여 할당하도록 하였고, 평균적으로 16.6m 내외에 위치하고 있으며, 가장 가까운 경우는 약 1.45m, 가장 멀리 위치하는 경우는 약 1,709.84m로 나타났다. 그러나 이러한 과정에서 도로를 벗어난 곳에서 발생되었다고 추정되는 데이터를 찾아 이를 제거하는 것이 분석의 정확도를 높인다고 판단하였다. 본 연구는 하나의 교통사고 점 데이터로부터 가장 가까운 도로 링크가 약 100m 이상 떨어져 있다는 것은 해당 도로에서 발생한 사고라고 정의하기에 어려움이 있다고 판단하였고(Figure 7), 이러한 사고 데이터 1,736건(전체 데이터의 약 4.45% 차지)을 제외한 37,274건을 대상으로 분석하였

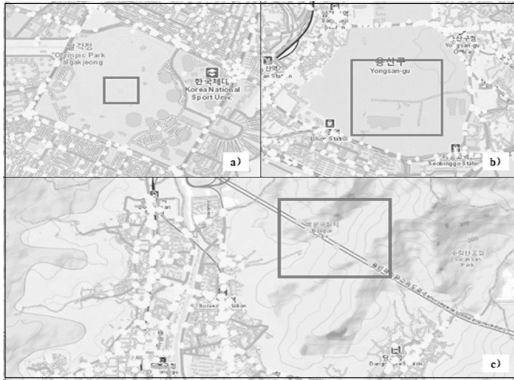


Figure 7. the accident point data they are more than 100 meters away from the nearest road link data. Some typical cases: a) the Olympic park in Songpa-gu, b) Mt. Dunjisan in Yongsan-gu, c) Backun Park in Nowon-gu

다. 서울의 경우 세종로가 왕복 16차선으로 폭원이 넓은 도로이며, 하나의 차선이 최소 3.5~5m 정도의 넓이를 가지므로 하나의 교통사고 발생 지점으로부터 가장 가까운 도로 링크를 찾을 때 반경 80m를 중심으로 찾을 수 있다. 경찰청에서 수집한 교통사고 발생 지점의 오차 값을 고려하여 본 연구는 반경 100m 이상일 경우 해당 도로 링크와 상관없는 교통사고 지점이라고 정의하였으며, 이는 향후 연구에서 차선 수 또는 도로의 넓이에 따른 고려가 필요하다고 판단된다.

4. 분석 결과

도로 링크에 교통사고 데이터를 할당한 후 시각화한 결과는 Figure 8, Figure 10과 같다. 교통사고가 집중되는 도로 구간을 보다 명확히 확인할 수 있으며, 교통사고 건수가 가장 많이 할당되어 집계된 곳은 Table 1, Table 2와 같다

첫 번째 방법(Method A)은 도로 링크를 10m 단위로 나눈 후 교통사고 점 데이터를 할당하였고, 집계된 건수(count)를 Getis-Ord G_i^* 통계의 핫스팟 분석을

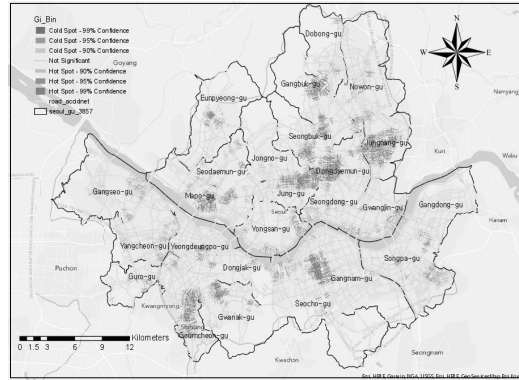


Figure 8. The hotspot links of traffic accidents in Seoul by method A.

수행하였다. 그 결과 Figure 8과 같이 마포구 일부 지역, 중구-동대문구-중랑구로 이어지는 지역, 성북구와 강북구 지역, 서초구-강남구 일부 지역과 금천구-

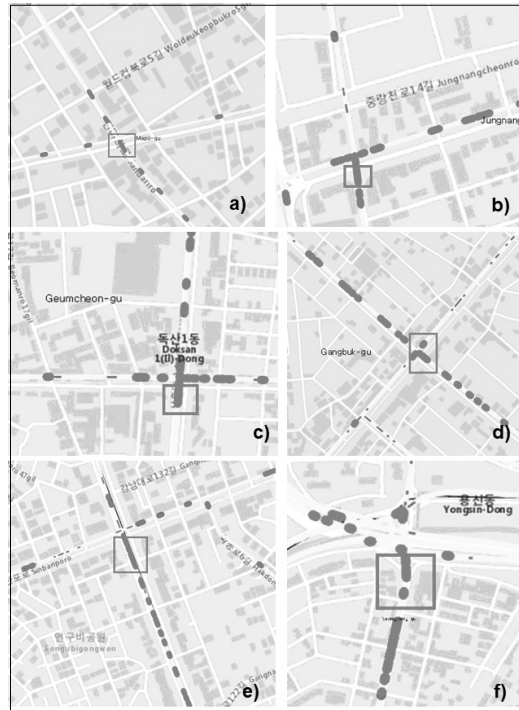


Figure 9. The highest black zones(hotspot links) of traffic accidents by method A.

관악구 일대에 교통사고 핫스팟 구간이 나타나는 것을 확인할 수 있다. Gi*의 통계량이 가장 유의미하게 나타나는 곳은 마포구의 잔다리길, 중랑구의 중랑천길, 면목천길, 금천구의 시흥대로, 성동구의 고산자로, 강북구의 한천로, 서초구와 강남구의 강남대로 일부 구간으로 나타났다. 이들 지역을 분석한 결과, 대부분 교차로 부근에서 높은 사고 건수를 기록하고 있었으며(Figure 9-a-e), 성동구의 고산자로는 내부순환로로 이어지는 인터체인지 부분에서 교통사고가 밀집되어 나타나며 통계적으로 유의미하게 나타났다(Figure 9-f).

Table 1. The black zones(hotspot links) of traffic accidents on top list by method A.

Road_Name	Location of the road	the Number of Accidents
잔다리길 일부구간	마포구 서교동	15
중랑천길 일부구간	중랑구 면목동	7
시흥대로 일부구간	금천구 독산동	7
고산자로 일부구간	성동구 도선동	6
한천로 일부구간	강북구 번동	6
강남대로 일부구간	서초구 서초동	5
...

두 번째 방법(Method B)은 교통사고 데이터를 도로 링크에 할당하고 도로 링크의 길이를 고려하여 미터 당 교통사고를 속성으로 추가하였다. 시각화 결과, 통계적으로 유의한 핫스팟 구간은 중구, 종로구, 동대문구, 중랑구, 강북구, 강남구, 구로구, 관악구, 강동구 등에 속하는 도로로 나타났다(Figure 10). 즉, 교통사고가 밀집된 도로는 구로구의 경인로61길, 강동구의 강동대로, 성북구로 이어지는 창경궁로, 중구와 종로

구에 걸친 율곡로, 종로구의 보문로 등이 통계적으로 가장 유의하게 나타났다(Table 2, Figure 11). 두 번째 방법에 의한 핫스팟은 도로 링크의 길이가 긴 경우에 상대적으로 교통사고 건수가 많을 가능성이 높기 때문에 선형의 클러스터를 보이고 있다. 핫스팟 구간은 대체적으로 서울 도심부에 위치하고 있으며, 외곽에 위치한 강서구, 마포구, 도봉구, 노원구, 강동구에 속한 도로에서는 콜드스팟이 나타나는 것을 확인할 수 있다.

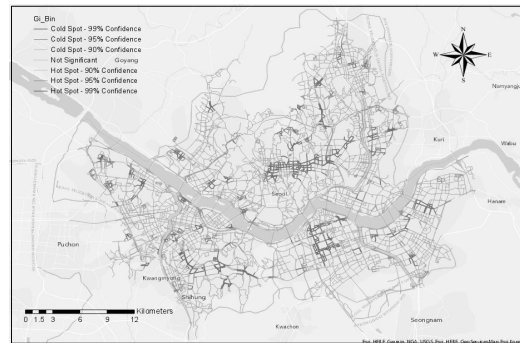


Figure 10. A hotspot analysis result by method B.

Figure 8과 Figure 10을 비교 분석한 결과, 서울시의 교통사고는 도심의 종로구-동대문구-중랑구, 그리고 성북구와 강북구로 이어지는 창경궁로, 율곡로, 보문로, 한천로 부분과 서초구와 강남구의 강남대로와 봉은사로, 금천구의 경인로와 남부순환로 일부 구간에 집중되어 나타남을 알 수 있다. 그러나 상세히 살펴보면 Figure 8에서는 교통사고가 집중되는 곳이 특정 도로 구간으로 명확하게 밀집되어 나타나는 반면, Figure 10에서는 도로 링크의 특성에 따라 교통사고 집중지역이 길게 나타나는 특징을 갖는다. 이는 표를 통해서도 확인할 수 있는데, Table 1에서는 교통사고 집중지역으로 마포구 서교동의 잔다리길이 가장 높게 나타났고 일부 10m 구간에서 2013년에 15건의 교통사고가 발생했다. Table 2에서는 교통사고가 가장 많이 발생하는 도로 구간은 경인로이며, 10미터 당 평균

2.5건의 교통사고가 발생했음을 확인할 수 있다. 즉, Table 1에서는 교통사고가 많이 발생하는 특정구간을 구체적으로 알 수 있는 반면, Table 2에서는 교통사고가 많은 도로를 알 수 있어, 교통사고 다발구간을 나타내는 방식이 다를 수 있다. 만약 교통사고 다발구간을 밝히고 교통경찰을 배치하거나 해당지역의 교통환경을 분석하고 환경을 개선하는 목적이라면 첫 번째 방법처럼 명확한 구간을 파악하는 것이 훨씬 유리하다고 볼 수 있다.

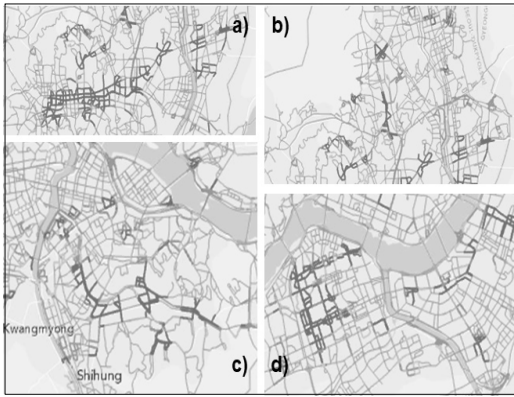


Figure 11. Zoom in to hotspot links. a) a cluster of Jung-gu, Dongdaemun-gu, and Jungnang-gu, b) Gangbuk-gu area, c) Geumcheon-gu area, d) Gangnam-gu and Seocho-gu area.

Table 2. The black zones(hotspot links) of traffic accidents on top list by method B.

Road_Name	Total length (m)	the average accidents values per 10meters
경인로	2,719.6	2.5
강동대로	13,484.5	1.8
창경궁로	7,461.9	1.8
울곡로	8,067.6	1.8
보문로	5,233.0	1.8
...

방법 A와 방법 B 두 가지를 좀 더 확대하여 특정 구간별 비교분석하면 Figure 12와 같다. Figure 12는 이 두 방법의 핫스팟 구간의 차이가 뚜렷하게 나타나는 강남구의 강남대로(Area A), 마포구의 잔다리길(Area B), 도봉구의 도봉로와 우이천로(Area C), 종로구의 율곡로와 창경궁로(Area D) 일대를 보여주고 있다.

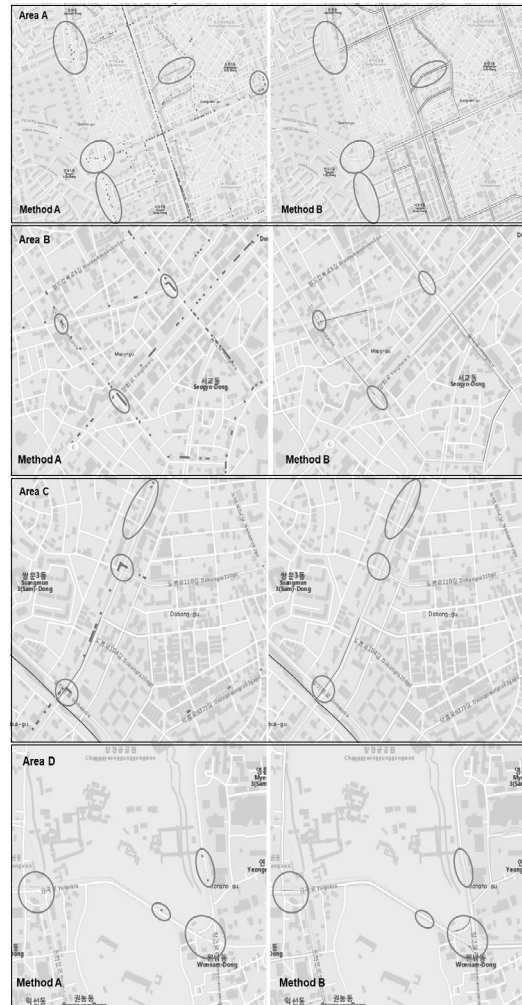


Figure 12. Compare the results of hotspot links by method A and B.

Figure 12의 왼쪽은 방법 A, 오른쪽은 방법 B에 의

한 핫스팟 구간이다. 주요 핫스팟 구간인 강남대로와 봉은사로는 두 방법 모두 공통적으로 교통사고가 밀집된 구간으로 나타났지만, 인터체인지로 진입하는 구간이나 좁은 길 단위에서 도로가 교차하는 부분에서 사고가 많이 발생함을 방법 A는 시각화하여 표현되고 있지만, 방법 B는 이를 표현하기에는 도로 구간의 길이가 긴 것을 알 수 있다. 마포구의 잔다리길 주변도 방법 A에 의해서는 교차로 부근에서 높은 사고 건수가 나타남을 보여주며, 이와 달리 방법 B는 전체 도로 링크를 중심으로 평균 사고 건수가 높은 구간을 찾아 준다. Area 3 일대도 방법 A는 도봉로의 교차로 지점이나 도봉로의 일부 구간들(10m)에 대한 핫스팟 구간을 확인할 수 있지만, 방법 B에 의하면 우이천로와 만나는 교차지점이나 도봉로 110길이 만나는 교차로 지점은 확인하기 어렵다. 앞선 Area A-C와 달리 Area D일대는 율곡로와 창경궁로가 만나는 지점으로 방법 A에서 찾지 못한 핫스팟 구간을 방법 B에서는 교차로 지점과 다른 도로로 이어지는 구간에 교통사고 핫스팟 구간으로 분석되어 시각화된 것을 알 수 있다.

5. 결론

본 연구는 2013년 서울시 교통사고 데이터를 서울시 도로 링크에 할당하여 교통사고가 어떤 도로 구간 상에서 집중적으로 발생하였는지를 시각화하고 통계적 유의성을 확인하였다. 기존의 교통사고 공간적 특성 분석 및 시각화는 네트워크를 고려하지 않고 밀도 추정법에 의한 공간적 밀집 지역을 찾는 연구가 다수 존재하였다. 그러나 교통사고의 경우 도로를 따라 발생하기 때문에 보다 정확한 분석을 위해서는 도로 링크를 따라 사고 데이터를 할당하여 분석하는 것이 필요하다 판단하였다.

따라서 본 연구에서는 두 가지 방법으로 교통사고 점 데이터를 도로 네트워크에 할당하였다. 첫 번째는

도로 링크 데이터를 10m 간격으로 분할하고, 교통사고 지점으로부터 가장 가까운 도로 데이터에 사고 건수를 집계하여 할당하였고, 이를 핫스팟 분석을 수행함으로써 통계적으로 유의한 교통사고 밀집 구간을 추출하였다. 두 번째는 도로 링크에 교통사고 데이터를 할당하고 미터 당 평균 사고 값을 도출하여 핫스팟 분석을 수행하였다. 기존의 교통사고 밀집 지역을 찾기 위한 커널 분석과 비교하였을 때 도로 네트워크를 따라 교통사고 다발 구간을 분석하여 시각화한 결과가 보다 더 정확하며, 교통사고 다발 구간의 특성을 파악하기에 용이함을 확인할 수 있었다.

뿐만 아니라 두 가지 방법에 의한 사고 다발 구간을 분석한 결과, 두 번째 방법에 의한 도로 링크 데이터에 평균 사고 값을 구하여 핫스팟을 시각화 하였을 때는 도로 구간 전체의 평균값으로 추출되는데 비해, 첫 번째 방법인 10m 단위의 일정한 도로 링크별 사고 건수를 추출한 방법은 주요 도로 구간 내에서 특히 교차로나 다른 도로로 진입하는 구간에 주로 사고가 많이 발생하는 것을 구체적으로 나타낼 수 있었다.

본 연구는 도로 네트워크 상에서 발생한 사고 다발 구간을 찾는 연구들에 새로운 시각에서의 시사점을 줄 수 있을 것이라 기대한다. 또한 라인을 따라 발생하는 데이터의 분석 및 시각화에도 적용될 수 있을 것으로 판단된다.

주1. 경기 자동차 등록대수(5,018,564), 경기 도로 10km당 교통사고 발생 건수(38.63), 서울 자동차 등록대수(3,073,263), 서울 도로 10km당 교통사고 발생 건수(48.44): 도로교통공단, 2016.

주2. 가변적 공간 단위 문제: 공간 데이터가 안고 있는 심각한 문제 중 하나로써 공간 패턴을 분석하는데 있어서 속성 데이터가 집계되는 공간 단위가 실제 나타나고 있는 특정한 현상의 공간 패턴과 불일치하는 경우에 문제가 된다. 공간 단위의 가변성을 고려하여 가장 최소단위에서 데이터를 집계하여 집계되는 공간 단위의 유연성을 높이는 것이 중요하다.

참고문헌

References

강영욱, 손세린, 조나혜, 2017. 의사결정나무와 시공간 시각화를 통한 서울시 교통사고 심각도 요인 분석. *지적과 국토정보*. 47(2): 233-254.

Kang YO, Son SR, Cho NH. 2017. Analysis of Traffic Accidents Injury Severity in Seoul using Decision Trees and Spatiotemporal Data Visualization. *Journal of Cadastre & Land InformatiX* 47(2): 233-254.

성병준, 배규한, 유환희, 2015. 진주시 교통사고의 시계열적 공간분포특성 분석. *한국지형공간정보학회지*. 23(2): 3-9.

Sung BJ, Bae GH, Yoo HH. 2015. Analysis of Temporal and Spatial Distribution of Traffic Accidents in Jinju. *The Journal of Korean Society for Geospatial Information Science*. 23(2): 3-9.

손세린, 강영욱. 2017. 서울시 여성운전자 교통사고의 시공간 특성 분석. *한국지도학회지*. 17(2): 89-98.

Son SR, Kang YO. 2017. Spatio-temporal Pattern of Traffic Accident of Female Drivers in Seoul. *Journal of the Korean Cartographic Association*, 17(2): 89-98.

이수진, 조형식, 송원호, 손홍규. 2015. 서울시 교통사고 발생의 공간적 특성과 영향 요인 분석에 관한 연구. *한국지형공간정보학회 추계학술대회*. 132-133.

Lee SJ, Cho HS, Song WH, Sohn HG. 2015. A Study on Spatial Characteristic and Influence Factor of Traffic Accident in Seoul. *Korean Society for Geospatial Information Science*. 132-133.

홍상기. 1998. 시, 공 복합 선형점 자료의 분석을 위한 시각화시스템의 개발. *한국도시지리학회지*, 1(1): 85-100.

Hong SK. 1998. Developing a Visualization System for Spatio - Temporal Linear Point Data. *The Journal of Korean Urban Geographical Society*. 1(1): 85-100.

Anderson TK. 2009. Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*. 41(3): 359-364.

Ceder A, Livneh M. 1978. Further evaluation of the relationships between road accidents and average daily traffic. *Accident Analysis & Prevention*, 10(2): 95-109.

Chen C, Zhang G, Liu XC, Ci Y, Huang H, Ma J, Chen Y, Guan H. 2016. Driver injury severity outcome analysis in rural interstate highway crashes: a two-level Bayesian logistic regression interpretation. *Accid Anal Prev*. 97: 69-78.

Elvik R. 2013. Risk of road accident associated with the use of drugs: a systematic review and meta-analysis of evidence from epidemiological studies. *Accid Anal Prev*. 60: 254-267.

Erdogan S, Yilmaz I, Baybura T, Gullu M. 2008. Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar. *Accident Analysis & Prevention*, 40(1): 174-181.

Kang YO, Cho NH, Son SR. 2018. Spatiotemporal characteristics of elderly population's traffic accidents in Seoul using space-time cube and space-time kernel density estimation. *PLoS one*, 13(5): e0196845.

Levine N, Kim KE, Nitz LH. 1995. Spatial analysis

- of Honolulu motor vehicle crashes: I. Spatial patterns. *Accident Analysis & Prevention*, 27(5): 663-674.
- McSwiggan G, Baddeley A, Nair G. 2017. Kernel density estimation on a linear network. *Scandinavian Journal of Statistics*, 44(2): 324-345.
- Ng JC, Hauer E. 1989. Accidents on rural two-lane roads: differences between seven states (with discussion and closure) (No. 1238).
- Okabe A, Okunuki KI, Shiode S. 2006. SANET: a toolbox for spatial analysis on a network. *Geographical analysis*, 38(1): 57-66.
- Okabe A, Sugihara K. 2012. *Spatial analysis along networks: statistical and computational methods*. John Wiley & Sons.
- Openshaw S. 1984. Ecological fallacies and the analysis of areal census data. *Environment and planning A*, 16(1): 17-31.
- Prasannakumar V, Vijith H, Charutha R, Geetha N. 2011. Spatio-temporal clustering of road accidents: GIS based analysis and assessment. *Procedia Soc Behav Sci*. 21: 317-325.
- Romano B, Jiang Z. 2017. Visualizing Traffic Accident Hotspots Based on Spatial-Temporal Network Kernel Density Estimation. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (p. 98). 2017, November. ACM.
- Xie Z, Yan J. 2008. Kernel density estimation of traffic accidents in a network space. *Computers, Environment and Urban Systems*, 32(5): 396-406.
- Yamada I, Thill JC. 2004. Comparison of planar and network K-functions in traffic accident analysis. *Journal of Transport Geography*, 12(2): 149-158.
- 도로교통공단, <http://www.index.go.kr>

2018년 5월 04일 원고접수(Received)

2018년 6월 01일 1차심사(1st Reviewed)

2018년 6월 15일 2차심사(2st Reviewed)

2018년 6월 27일 게재확정(Accepted)

초 록

최근 우리나라의 경우 교통사고 예방활동으로 자동차 보유에 따른 교통사고 발생건수는 지속적으로 감소하고 있지만, 서울의 경우 다른 지역에 비해 자동차 1만대 대비 사고 건수는 전국에서 광주와 함께 가장 높게 나타나고 있다. 인적 재난인 교통사고를 예방하기 위한 다양한 연구들이 진행되어 왔다. 특히 교통사고에 대한 공간적 분석을 연구한 초기 연구들은 교통사고 클러스터 지역을 확인하기 위해 행정구역 별 교통사고 건수를 집계하거나, 커널밀도 방법을 통해 밀도를 추정하여 분석하는 경우가 다수를 이루었다. 그러나 교통사고는 도로를 따라 발생하는 사건이기 때문에 도로상에서 교통사고 다발구간을 찾는 것이 더 의미가 있을 수 있다. 따라서 본 연구는 도로 네트워크를 따라 교통사고 집중 지역을 찾고자 하였다. 본 연구에서는 2가지 방법으로 교통사고를 가장 가까운 도로 네트워크에 할당한 뒤, Getis-Ord G_i^* 에 의한 핫스팟 분석을 통해 교통사고 다발구간을 분석하였다. 하나는 10m 단위의 일정한 도로 링크를 중심으로 분석을 수행하였으며, 다른 하나는 도로구간별 단위 길이당 평균 교통사고를 계산하여 교통사고 밀집구간을 분석하였다. 첫 번째 방법에 의한 분석 결과 교통사고가 집중되는 특정 도로 구간을 명확하게 확인할 수 있는 반면, 두 번째 방법에 의한 분석 결과 도로링크의 특성에 따라 교통사고 집중지역이 길게 나타나는 특징을 확인할 수 있었다. 두 방법에 의한 교통사고 다발구간이 다르게 나타나는 것을 알 수 있으며, 향후 해당 지역의 교통환경을 분석하고 개선하기 위해서는 보다 명확한 구간을 파악하는 것이 유의미할 수 있다.

주요어 : 교통사고, 도로 네트워크, 핫스팟 분석, 시각화, 교통사고 다발구간