

## 시공간검정통계량을 이용한 도시범죄의 핫스팟분석\*

정경석<sup>1</sup> · 문태헌<sup>2\*</sup> · 정재희<sup>3</sup>

## Hotspot Analysis of Urban Crime Using Space-Time Scan Statistics\*

Kyeong-Seok JEONG<sup>1</sup> · Tae-Heon MOON<sup>2\*</sup> · Jae-Hee JEONG<sup>3</sup>

### 요 약

본 연구의 목적은 공간적 분포 특성만을 고려하고 있는 기존의 핫스팟분석에 대한 대안적인 방법으로서 공간상에서 나타나는 사건간의 인과관계를 시간영역으로까지 확장하여 동시적 분석이 가능한 시공간분석 방법을 제안하는 것이다. 분석방법으로는 먼저 지리정보시스템을 이용하여 지방중소도시인 M시의 범죄자료를 데이터화 하였고, Ripley K함수와 시공간검정통계량 분석을 통해 M시의 범죄분포 패턴을 지도화 하였다. 연구결과, 범죄위험도가 유의미하게 높은 지역들이 나타났으며, 이들 시공간적 범죄 집중지역들은 기존의 공간분포만을 고려한 범죄분포 패턴과는 다소 차이가 있음을 발견할 수 있었다. 본 연구결과는 시공간적인 범죄분포 특성에 맞는 맞춤형의 경찰 인력 배치와 배분, 그리고 치안행정 서비스 등의 조정을 위한 참고자료로서, 또한 시공간적인 집중을 보이는 이들 지역을 중심으로 물리적 환경 변화의 유도과 공간이용의 개선 효과를 통해 범죄율을 줄여나가는 범죄예방 활동 및 정책수립을 위한 기초자료로도 유용하게 활용될 수 있을 것으로 기대된다.

주요어 : 시공간군집분석, 시공간스캔통계량, 도시범죄, 범죄다발지역

### ABSTRACT

The aim of this study is to investigate crime hotspot areas using the spatio-temporal cluster analysis which is possible to search simultaneously time range as well as space range as an alternative method of existing hotspot analysis only identifying crime occurrence distribution patterns in urban area. As for research method, first, crime data were collected from criminal registers provided by official police authority in M city,

2010년 5월 20일 접수 Received on May 20, 2010 / 2010년 6월 28일 수정 Revised on June 28, 2010 / 2010년 7월 20일 심사완료 Accepted on July 20, 2010

\* 본 연구는 주저자의 박사학위 논문의 일부를 수정·보완한 것이며 2단계 BK21의 지원을 받아 연구되었음.

1 경남발전연구원 남해안발전연구지원센터 Gyeongnam Development Institute

2 경상대학교 도시공학과, 2단계 BK21 및 공학연구원 2nd BK21 and ERI, Dept. of Urban Engineering, Gyeongsang National University

3 경남발전연구원 Gyeongnam Development Institute

\* 연락처 E-mail : thmoon@gnu.ac.kr

Gyeongnam and crime occurrence patterns were drafted on a map by using Geographic Information Systems(GIS). Second, by utilizing Ripley K-function and Space-Time Scan Statistics analysis, the spatio-temporal distribution of crime was examined. The results showed that the risk of crime was significantly clustered at relatively few places and the spatio-temporal clustered areas of crime were different from those predicted by existing spatial hotspot analysis such as kernel density analysis and k-means clustering analysis. Finally, it is expected that the results of this study can be not only utilized as a valuable reference data for establishing urban planning and crime prevention through environmental design(CPTED), but also made available for the allocation of police resources and the improvement of public security services.

**KEYWORDS** : *Spatio-temporal Cluster Analysis, Space-Time Scan Statistics, Urban Crime, Hotspot*

## 서론

일반적으로 공간적 군집현상과 확산현상에 대한 도시계획가 내지 공간연구자의 주 관심은 크게 두 가지로 나뉘어져 있다. 그 중 하나는 공간적인 사상이 특정한 지역에 집중하여 나타나는 핫스팟(hotspot)지역이나 그 반대의 콜드스팟(coldspot)지역을 찾아내는 일이고, 다른 하나는 어떠한 사회·경제적, 그리고 환경적 변인들이 그러한 공간적 군집과 확산을 불러일으켰는가를 밝혀내는 일이다. 이 중 핫스팟분석기법은 산업집적지를 분석하거나 생태학에서 식생의 분포, 의료분야에서 특정질병의 공간적 분포 등을 판별해 내는데 주로 사용되어 왔다(Kang, 2009). 그러나 대부분의 사례가 공간적 군집현상만을 다루고 있을 뿐, 시간에 의한 집중현상을 간과한 채, 시간이라는 주요한 변수를 적절히 고려하지 못한 한계를 보여주었다(Kullodrrff *et al.*, 2005). 뿐만 아니라, 핫스팟분석은 동일한 공간적 사건 자료를 대상으로 하더라도 연구자의 의도나 분석과정에서의 변수 조정 등의 차이에 의해 핫스팟지역이 달라지기도 하며 분포형태나 공간범위가 서로 다르게 나타나기도

한다(Eck *et al.*, 2005).

이러한 핫스팟분석은 지리적 프로파일링과 범죄지도화 연구의 한 방법으로 소개되어 도시범죄 연구에도 자주 활용되고 있다. 그러나 도시범죄 문제는 공간적 요인뿐만 아니라 시간적 요인과의 관련하여 강한 상호관련성을 가지므로 보다 효과적인 도시범죄 예방활동을 위해서는 이러한 시공간적인 군집 특성을 보이는 범죄다발지역(hotspot)을 우선적으로 찾아내는 일이 무엇보다 중요하다. 이들 지역을 중심으로 범죄와 범죄유발요인간의 인과관계를 분석해 낼 수 있다면, 보다 정확한 범죄발생의 원인과 대책을 마련할 수 있기 때문이다.

이러한 배경에서 본 연구의 목적은 공간적 분포 특성만을 고려하고 있는 기존의 핫스팟분석과 달리, 공간상에서 나타나는 사건 간의 인과관계를 시간영역으로까지 확장하여 동시적 분석이 가능한 시공간분석을 통해 범죄다발지역을 새롭게 찾아내고, 이들 지역이 기존의 공간적 분포 특성만을 고려한 핫스팟분석 결과와는 어떤 차이가 있는지를 밝혀보고자 한다.

## 기존 연구동향 및 이론적 고찰

### 1. 도시범죄의 연구동향

도시범죄연구에 있어 GIS 및 통계적 공간 분석과 결합된 범죄분석 관련 소프트웨어의 개발과 활용방안에 대해 많은 논의들이 있어 왔다. 이러한 정량적인 방법들의 적용가능성은 계속 증가하고 있으며 특히 범죄발생의 분포 특성을 밝혀내는 공간기반의 통계분석기법들은 도시범죄행위 분석에 있어 매우 유용한 방법으로 간주되고 있다(Hirschfield *et al.*, 1995; Weisburd, 1997). 범죄행위가 집중되어 발생하는 지역을 정확하게 찾아내는 작업은 왜 범죄가 특정한 장소에서 발생하는지에 대한 메커니즘을 이해하는데 단초적인 정보를 제공해 준다. 따라서 범죄패턴을 설명할 수 있다는 것은 앞으로의 범죄발생이나 예방책에 대해 좀 더 정확한 예측이 가능함을 의미하기도 한다. 그러므로 공간적으로 범죄다발지역을 식별해 낼 수 있는 범죄지도화 방안은 범죄와 관련된 경찰활동, 범죄예방을 위한 사회적 서비스 등의 의사결정에 지대한 영향을 미칠 수 있다(Goldsmith *et al.*, 2000).

최근에는 기존의 환경범죄학에 대한 이론적인 측면을 공간통계분석모형(Spatial Statistics Analysis Model)과 GIS에 통합하고자 하는 방향에서 많은 연구가 이루어지고 있다. 그동안은 주로 공간계량경제학 등에서 활발하게 논의 되어왔으나, 최근 도시 및 지리학 분야로 점차 확대되고 있는 분석방법인 공간적 자기상관분석과 공간적 자기회귀모형(Anselin, 1992)을 이용한 도시범죄 연구들도 점차 증

가하고 있는 추세이다. 이처럼 최근의 도시범죄연구는 범죄발생 자료의 공간 분포를 가시화하고 범죄다발지역을 지도화 하는 영역(Murray *et al.*, 2001; 황선영과 황철수, 2003)과 통계적 공간분석을 이용하여 범죄발생에 대한 원인과 그 영향요인 등을 규명하려는 영역(Shimada, 2004; Andersen, 2006; 정경석 등, 2009)으로 크게 나뉘어 모형개발이 이루어지고 있으며 이들을 통합한 범죄분석통합모형이 소개되고 있기도 하다(정경석, 2010). 그러나 여전히 범죄발생에 대한 시간적 고려는 매우 미흡한 편이다.

표 1에서 알 수 있듯이, 대부분의 공간군집 분석기법들은 주로 점사상의 위치적, 형상적 특징을 확률분포와 분포의 중심경향 및 분산에 관한 측도 등을 통해 사건들이 공간상에 임의적으로 분포하는가, 규칙적으로 분포하는가, 또는 집중적으로 분포하는가를 판별해 내지만, 국지적 차원에서 그러한 분포 형태가 우연에 의한 결과인지 아닌지에 대한 추가적인 검증이 어렵고, 시간의 흐름에 따른 공간 분포의 특성변화를 예측해 내기가 어려운 한계를 지니고 있다. 또한 기존의 시공간 범죄 분석연구들은 주로 범죄발생시기를 시간별, 요일별, 계절별로 구분하여 범죄분포 특성을 밝혀내고 있으나 시간적 분석과 공간적 분석이 통합되지 못하고 분리되어 종합적으로 분석하지 못하는 한계가 있다. 따라서 본 연구에서는 범죄발생 자료의 공간 분포 특성을 가시화하고 범죄지도화 하는 새로운 방법으로서 시공간 탐색에 기반한 범죄 핫스팟 분석방법을 제안하고자 한다.

TABLE 1. 기존의 공간 핫스팟분석의 주요 특징 및 한계

구분		특징 및 한계
시간 분석	시간별 분석	빈도분석에 의한 시간대별, 요일대별, 월 및 계절별 단순 집계화 자체적인 공간분포 특성 파악이 불가능
	요일별 분석	
월(계절)별 분석		
지역 자료 탐색	Pinmap 중심점과 표준거리타원체 (Standard Deviational Ellipse)	점사상의 위치적·형상적 특징을 확률분포와 분포의 중심경향 및 분산에 관한 측도 등을 통해 검증하게 되며, 전역적 공간 분포 특성의 지도화가 가능
	점사상 밀집표면/밀도분석(Kernel Density) 최근린지수 (Nearest Neighbor Index)	공간 분포의 형태적(임의적 분포, 규칙적 분포, 군집적 분포) 특성만을 판단할 수 있을 뿐, 분포 특성이 우연에 의한 결과인지 아닌지에 대한 검증이 불가능 밀도분석의 경우 초기값(셀의 크기 및 Kernel 범위)의 설정 등에 따라 분포패턴이 달라짐 시간에 따른 공간 분포 특성의 변화 예측 파악이 자체적으로 불가능
면사상	Moran's I, Geary's C	행정구역이나, 인구센서스, 교통존과 같은 면단위의 집계자료를 이용하여 인접한 공간효과에 대한 자기상관성 유무를 검증 시각화된 도면형태가 아닌 scatter graph 형태의 지수값으로 표현 시간에 따른 공간 분포 특성의 변화 예측 파악이 자체적으로 불가능
국지 자료 탐색	점사상 최근린 계층군집(Nearest Neighbor Hierarchical Clustering) K-means 군집 (K-means Clustering)	국지적 차원에서의 핫스팟 지역에 대한 지도화가 가능 공간 분포의 형태적(임의적 분포, 규칙적 분포, 군집적 분포) 특성만을 판단할 수 있을 뿐, 분포 특성이 우연에 의한 결과인지 아닌지에 대한 검증이 불가능 연구자가 임의적으로 설정한 군집에 포함될 사건의 개수와 군집 수에 대한 초기값(seeds)에 의해 핫스팟지역의 수와 범위가 민감하게 반응 시간에 따른 공간 분포 특성의 변화 예측 파악이 자체적으로 불가능
	면사상	Local Moran's I, Local Geary's C

2. 시공간 분석의 이론적 고찰

공간통계학 분야에서 주로 활용되고 있는 시공간 분석기법은 크게 두 가지 유형으로 구분된다. 한 부류는 시공간의 상호작용성에 대한 유무를 측정하고 검증하는 것을 주목적으로 하는 분석기법과 다른 하나는 시공간 군집 패턴을 시각적인 그래프와 도면 형태로 검색하고 이를 통계적으로 검증하는 분석기법으로

나뉠 수 있다. 전자의 대표적인 예로는 녹스 테스트 지수(Knox, 1964)와 맨틀테스트 지수(Mantel, 1967)가 있고, 후자의 분석기법으로는 최근린분석과 K함수를 시공간으로 확대한 Jacquez 테스트(Jacquez, 1996)와 시공간 K함수(Diggle *et al.*, 1995), 그리고 Kulldorff의 시공간검정통계량(Kulldorff, 2001) 분석 등이 있다.

TABLE 2. 주요 시공간 핫스팟분석의 특징

구분	특징	
전역 분석	Knox test Mantel test	녹스테스트는 시간적-공간적으로 가깝고 먼 기준에 대한 최적값(cut-off point)을 찾아내기가 어려워 연구자에 의해 주로 임의적으로 결정(Levine, 2005) 맨틀테스트는 시간거리와 공간거리 간 선형성이 보장되어야 분석이 가능하며, 극단적인 이상치에 의해 상관관계가 왜곡될 수 있고, 일반적으로 상관계수값이 매우 낮게 나오는 단점이 있음(Levine, 2005) 녹스 및 맨틀테스트는 시간과 공간의 상호관련성만을 통계적으로 검증할 뿐 시공간 상에 분포 특성이 구체적으로 어떻게 나타나고 있는지에 대한 추가적인 정보제공 불가능(시각적인 도면화 작업이 불가능)
	Jacques test	시간거리와 공간거리 간 선형성이 존재하지 않는 경우에도 적용 가능 Jacques test는 사건발생간의 시간적, 공간적 분포순서(서수적인 시공간 거리측정)를 통해 시공간집중여부를 판단하기 때문에 정보제공이 제한적
국지 분석	시공간 K함수 (space-time K function)	Ripley(1976)가 제시한 K함수를 시공간집중분석방법으로 확대한 것 시간거리와 공간거리 간 선형성이 존재하지 않는 경우에도 적용 가능 기수적인 시공간거리를 분석대상으로 하여 3차원 그래프를 통해 시간과 공간 속에서 발생하는 사건의 패턴을 손쉽게 도식화 하고 예측이 가능 전역적 차원에서 시공간적 집중지역에 대한 시간거리와 공간거리는 3차원 그래프를 통해 대략적인 추정이 가능하나, 국지적 차원에서의 시공간적 군집 현상에 대한 확률적 검증과 도면화된 시각적 정보제공은 불가능
	시공간검정통계량 (space-time scan statistics)	시간거리와 공간거리 간 선형성이 존재하지 않는 경우에도 적용 가능 지역차원의 시공간집중지역 및 분포 패턴을 탐색해 낼 수 있으며, 점사상 뿐만 아니라 면사상에 대한 분석도 가능 시공간적 군집패턴을 시각적인 그래프와 도면 형태로 표현이 가능 시공간적 군집형태가 우연에 의한 것인지 아닌지에 대한 통계적 검증도 가능 과거에 발생했던 시공간집중의 공간적, 시간적 범위를 파악할 수 있을 뿐만 아니라, 현재 진행형의 시공간집중의 공간적 범위와 시간적 범위를 동시에 탐색해 낼 수 있는 장점이 있음(retrospective vs prospective 분석)

이러한 다양한 시공간분석기법들 가운데 본 연구에서는 시공간검정통계량을 중심으로 고찰해 보고자 한다. Kulldorff의 시공간검정통계량은 기존의 점사상 패턴분석에서 제시하지 못했던 지역차원의 시공간집중지역 및 분포 패턴을 탐색해 낼 수 있는 분석방법으로서 점사상 뿐만 아니라 면사상에 대한 분석도 가능한 장점이 있다. 이러한 시공간검정통계량은 공간검정통계량(spatial scan statistics)분석에 그 기반을 두고 있는데, 공간검정통계량은 대상지역에 대해 인근지역을 포함하는 하나의 원을 투영하고 이를 인근지역으로 옮겨 투영된 원의 외부와 내부에 존재하는 사상의 로그

우도비(log likelihood)를 계산하여 공간집중도를 분석하는 방법이다. 즉 공간검정통계량은 관심지역 내 특정지역의 중심에 하나의 원을 투영하고 투영된 원의 지름을 점차 최대치까지 증가시켜 나가면서 최대치의 우도비(likelihood ratio)를 계산하는 방법이며 최대치에 이른 원은 인근지역으로 이동하여 다시 최대치까지 확장하면서 각각의 원에 대한 개별 가능도비를 산정하면서 비교하게 된다. 이때 각각의 원에서 계산되는 우도비는 원 내에서 관측된 사건의 수와 원 외부에서 관측된 사건의 수를 비교해서 산정하게 되는데, 이렇게 산정된 우도비 중에서 최대우도비

(maximum likelihood)를 이용해 구체적으로 어디에서 공간집중이 일어나는지를 파악할 수 있게 된다(Kulldorff and Nagarwalla, 1995; Kulldorff, 1997; 손학기와 박기호, 2008).

이와 같은 공간검정통계량의 분석 원리를 시공간검정통계량으로 확장하면 특정지역의 중심을 기준점으로 원통(cylinder) 형태를 통해 시공간집중을 판단할 수 있게 된다. 원통의 지름과 높이가 각각 공간거리와 시간거리가 되며 원통의 면적과 높이를 변화해 가며 이웃한 지역의 중심점으로 이동해 가면서 최대우도비를 찾게 되는 것이다. 따라서 최대우도비가 관측된 공간과 시간이 바로 시공간집중이 발생한 지역과 시간대 구간으로 결정된다(Kulldorff, 2001; Kulldorff *et al.*, 2005).

시공간검정통계량은 SaTScan이라는 프로그램을 통해 분석이 가능하며, 과거에 시공간집중이 발생했던 공간과 시간적 범위를 주로 찾는 retrospective 시공간순열(space-time permutation) 방식과 현재의 상황을 근거로 현재 시공간집중이 현저하게 발생하고 있는 지역을 찾는 prospective 시공간 순열방식의 두 가지 탐색방법을 제공하고 있다. 따라서 본 연구에서는 이러한 시공간검정통계량 분석 방법을 중심으로 M시의 시공간범죄 집중지역을 찾아내고자 하였다.

## 자료구축 및 연구방법

분석에 사용된 자료는 M시 관할경찰서의 협조를 통해 정리한 2006년도 범죄신고관리대장을 GIS를 통해 점사상으로 구축한 것으로 범죄신고대장에 기록된 주소를 토대로 1/1,000의 수치지도와 지번도를 참조하여 공간자료화 하였고, 사건발생일자 및 시간, 요일, 동, 발생위치 및 주소, 장소, 범죄유형 등을 속성자료화 한 것이다. M시의 2006년도 총 범죄건수는 5,791건이고, 5대 범죄유형별 수는 강간 78건, 강도 43건, 살인 8건, 폭력

2,700건, 그리고 2,962건의 절도가 보고되었다. 그러나 전체 5,791 건의 범죄신고 자료 중에 주소식별이 가능한 3,640건을 대상으로 최종적으로 분석에 활용하였으며, 범죄 핫스팟 분석을 위해 Crimestat 3.0과 ArcGIS 9.2, 그리고 SaTScan 8.1 프로그램이 사용되었다.

연구방법으로는 우선 기존의 선행연구들에서 자주 활용되어 온 kernel 밀도분석과 k-means 군집분석, 그리고 최근린계층군집 분석에 의한 범죄 핫스팟지역을 지도화 한 다음, 시공간검정통계량 분석에 의한 시공간범죄 집중지역의 분포 특성과 범죄발생의 위험 확률 지도를 서로 비교하였다. 이때, 시공간검정통계량분석에 있어 탐색범위(원형 윈도우의 크기)의 적절한 설정이 매우 중요한데, 일반적으로는 전체면적과 전체시간범위를 기준으로 하여 최장 50%까지 확장하면서 각각의 우도비를 계산하게 되나 본 연구에서는 Ripley K군집 분석을 통해 최대의 범죄군집 경향을 보이는 최장거리를 별도로 산정하여 이 값을 기준으로 원형 윈도우 면적의 최대 반경 한계를 설정하였고, 시간표집단위는 시물레이션의 연산속도를 고려하여 2주일(14일)을 기준으로 삼아 전체기간의 50%까지 확장하여 탐색할 수 있도록 시간범위의 한계를 설정하였다. 시간단위를 1일로 하지 않고 2주단위로 표집한 이유는 면사상이 아닌 다량의 점사상 자료를 분석하게 될 경우, 핫스팟지역을 판별하는 정확성의 향상 비율에 비해, 컴퓨터 연산 능력에 따라 수 시간에서 몇일 단위의 과도한 시물레이션 소요시간이 비효율적으로 요구되기 때문이다. 한편, 시공간범죄 집중지역에 대한 통계적 유의성 검증은 몬테카를로 시물레이션의 반복연산을 1,000회로 설정하여 검증하였으며, Prospective와 Retrospective 시공간순열방식으로 각각의 시공간범죄 집중지역을 찾아낸 다음, 그 결과를 서로 비교하였다.

### M시의 도시공간구조 및 공간핫스팟 분석 결과

시공간분석에 앞서 M시의 일반적인 공간구조와 공간핫스팟 분석에 의한 범죄분포 특성을 요약해 보면 다음과 같다.

M시의 공간구조는 합포동, 오동동, 중앙동, 성호동, 동서동을 포함하고 있는 구도심지역과 북동방향의 구시가지 지역, 그리고 남부의 신시가지지역으로 구분된다. 이 중 구도심지역은 상업, 업무 행정기능과 공공시설물이 밀집되어 있고, 구시가지지역은 재래시장 및 소매업 중심의 유통, 상업기능과 산업단지가 위치해 있으며, 중·저층 중심의 주거시설이 밀집되어 있다. 한편 신시가지지역은 향만, 물류기능과 K대학교 등의 교육시설, 그리고 대규모 공동주택의 주거기능이 밀집되어 있다. M시는 근대적 도시형성 역사가 100년이 넘고 영남지역의 주요도시로서 기능을 수행하였기 때문에 상업지역의 비중이 다른 지역에 비해 높은 편이다(약 15.8%). 그러나 시가지의 면적이 매우 협소하므로 상업지역의 절대적인 규모에 있어서는 그리 큰 비중의 도시는 아니다. 한편 주거용도의 토지이용은 전 도시지역에 걸쳐 고르게 분포되어 있는데 반해, 상업적 토

지이용은 구도심을 중심으로 매우 높은 집적을 보이면서 주요 간선도로를 따라 대상형으로 위치해 있다(그림 1 참조). 전체적으로 협소한 지형적인 요인에 의해 도시의 중심성이 약하고, 도시 확장 또한 쉽지 않은 공간구조를 띠고 있으나, 최근에 인근의 C시와 J시와의 행정구역의 통합 개편을 계기로 광역중심도시로 점차 자리매김 할 것으로 기대되는 지역이다.

이러한 M시의 도시범죄 분포 특성을 일반적인 시간분석과 공간핫스팟 분석을 통해 요약해 보면(보다 자세한 내용은 정경석 등(2009)의 연구를 참조 바람), 첫째, M시의 총범죄 건수 중 약 39.8%가 야간시간대에 집중하여 발생하고 있으며, 심야시간대(24.5%), 오후시간대(20.1%), 오전시간대(15.6%) 순으로 나타났다. 특히 야간시간대의 경우, 도시전역에 걸쳐 범죄가 점차 확산되는 추세를 보이며, 동시에 철도역 부근과 합성동 일대의 상업지역, 노산동, 산호동, 오동동, 중앙동, 월영동 일대의 상업과 주거지역들에서 범죄밀도가 집중된 경향을 보였다. 이들 가운데 범죄밀도가 가장 높은 지역은 철도역 부근과 합성동 일대로 나타났다(그림 2 참조).

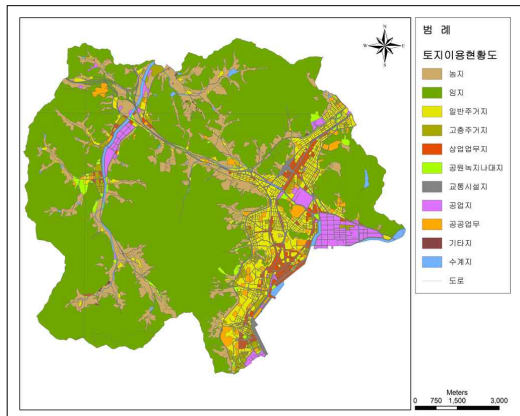


FIGURE 1. M시 토지이용현황도

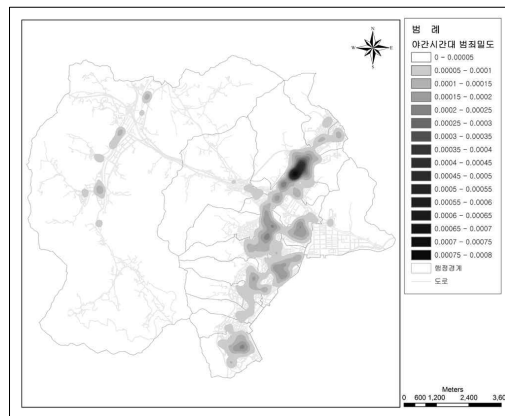


FIGURE 2. M시 야간시간대 총범죄 kernel 밀도분석 결과\*

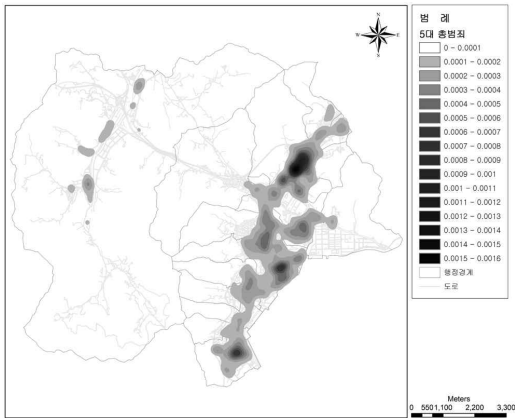


FIGURE 3. 총범죄에 대한 kernel 밀도분석 결과

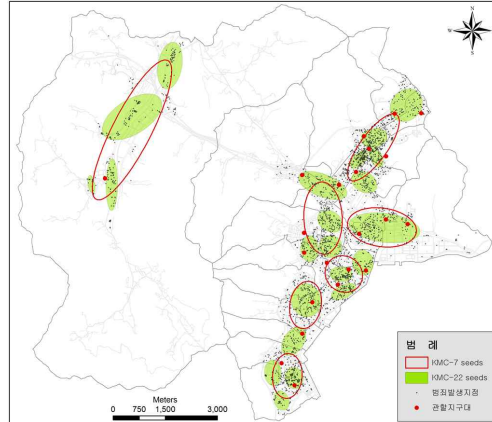


FIGURE 4. 총범죄에 대한 K-means군집분석 결과\*\*

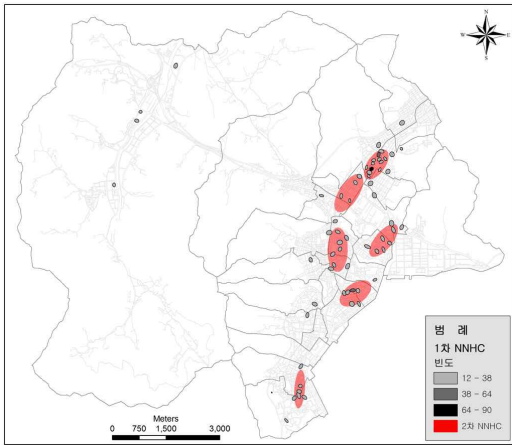


FIGURE 5. 총범죄에 대한 최근린 계층군집 분석결과

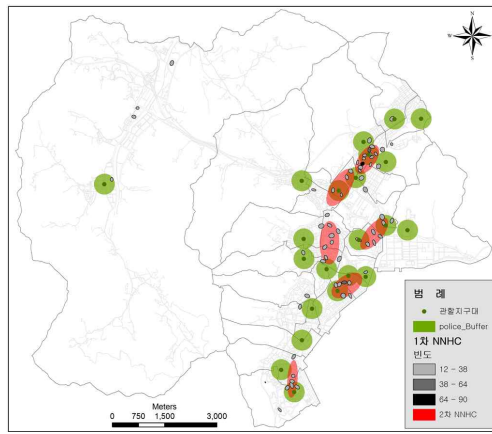


FIGURE 6. 경찰지구대의 버퍼링중첩 분석결과

\*주1) 셀의 크기는 10m×10m로 하였고, 각 범죄발생지점별로 300m 반경 내에 있는 범죄발생건수를 합산하여 km<sup>2</sup>당 범죄발생밀도를 계산한 결과임  
 \*\*주2) 대상지역내 7개 경찰 관할구역과 22개의 경찰관할지구대 수를 각각 초기군집수로 투입하여 분석한 결과임(α=0.01)

둘째, 2006년도 한해의 총범죄건수를 대상으로 범죄밀도분석을 수행한 결과, 앞서 시간대별 분석결과와 매우 유사한 패턴이 나타났다. 철도역 부근과 합성동일대의 상업지역이 범죄밀도가 가장 높게 나타났고, 그 다음으로 월영동과 오동동·동서동 일대, 그리고 산호동과 노산동 일대로 나타났다. 이들 지역은 유동인구가 빈번한 시외버스 터미널과 대형 백

화점 등이 입지한 상가 밀집 지역일 뿐만 아니라, 상업지역 주변으로 제 3종 일반주거지역들이 다수 분포하고 있어 거주밀도가 높고, 용도의 혼재 비율이 상대적으로 매우 높은 지역들이다(그림 3 참조).

셋째, 국지적 차원에서 범죄 핫스팟 지역을 좀 더 정확하게 찾아내기 위해 7개의 경찰관할구역수와 22개의 경찰관할지구대 수를 초기



값으로 각각 투입하여 K-means 군집분석을 수행하였다. 분석결과, 군집수를 7개로 설정하여 분석한 결과에서는 핫스팟지역이 너무 포괄적이어서 육안으로도 쉽게 식별이 가능한 반면, 22개로 설정하여 분석한 결과에서는 내서읍 일대에서 필요 이상의 과도한 군집수가 발견되었다(그림 4 참조).

넷째, 국지적 차원에서의 범죄 핫스팟 지역을 찾아내는 또 다른 방법으로서 군집에 포함될 사건의 수를 12개로 한정하여 최근린 계층군집분석을 수행하였다. 그 결과로서 64개의 1차군집 지역과 6개의 2차 군집 지역이 도출되었다. 1차 군집집단이 가장 빈번하게 출현한 지역은 합성 1·2동과 양덕1동 일대이며, 그 다음으로 회원1동과 노산동 일대, 그리고 산호동과 양덕2동, 오동동과 동서동 일대에서 나타났다. 1차 군집집단 가운데 범죄발생의 빈도가 가장 높게 나타난 지역은 합성2동과 오동동 일대에서 발견되었다. 한편 2차군집 지역은 합성1·2동과 양덕 1동 일대, 석전 1·2동 일대, 회원1동과 노산동 일대, 오동동과 동서동 일대, 산호동과 양덕2동 일대, 그리고 월영동과 문화동 일대에서 나타났다(그림 5 참조).

또한 그림 6은 대상지역에 소재된 치안지구대를 중심으로 반경 300m의 버퍼링 결과를 최근린 계층군집분석에 의해 탐색된 핫스팟지역들과 중첩시켜놓은 것으로서 범죄발생에 대한 방어기제적 성격의 치안지구대가 범죄발생에 어떤 영향을 미치고 있는지를 살펴보기 위함이다. 분석결과, 이들 치안시설물의 입지가 범죄예방차원에서 실제로 효과가 있는지에 대한 명확한 증거를 공간 핫스팟분석에서는 쉽게 발견할 수 없었다. 이들 관할서가 입지한 지역은 비교적 범죄로부터 안전한 콜드스팟(coldspots)지역에 위치한 경우도 있었으나, 오히려 범죄빈도가 매우 높은 핫스팟 지역에

위치하고 있기도 하고, 일부에서는 지구대 주변으로 강한 범죄 핫스팟지역이 나타나고 있었기 때문이다.

이상의 공간핫스팟 분석기법들은 시간적인 요인, 즉 사건간의 발생시간 간격에 대한 고려 없이 단순히 특정한 공간내에 누적되어 온 사건의 위치적 거리만을 고려하고 있기 때문에 범죄발생 가능성이 상대적으로 매우 높은 상업지역들이 주로 범죄핫스팟 지역으로 분류되었다. 그러나 공간적으로 범죄사건이 집중되어 있는 지역이라고 해서 반드시 범죄 발생 시기도 특정 시간대에 집중되어 있는 것은 아니며, 이렇게 찾아진 핫스팟지역이라 하더라도 그것이 단지 우연에 의해 나타난 결과인지, 아니면 실제로 시공간적으로 범죄발생 위험률이 높은 지역인지에 대한 추가적인 확률적 정보를 일반적인 공간핫스팟분석에서는 제공해주지 못하고 있다. 따라서 실제로 공간핫스팟 분석을 통해 탐색된 범죄다발지역이 시간적으로, 그리고 공간적으로 범죄가 집중된 위험지역인지, 그렇지 않으면 단순히 우연의 결과로서 사건들이 집중된 지역인지를 검증해낼 별도의 방법이 필요하며, 이것은 시공간검정통계량 분석을 통해 손쉽게 검증해 낼 수 있다.

## M시의 시공간분석 결과

시공간검정통계량 분석에서 탐색원통은 0에서 전체 대상공간 및 전체 대상기간의 50%까지 확장하면서 탐색하도록 초기값이 설정되어 있다. 본 연구에서는 보다 정확한 분석결과를 유도해 내기 위해 공간거리의 최대치는 Ripley K함수의 적용을 통해 별도로 추정해 내었고, 시간거리는 과도한 시물레이션 소요 시간을 줄이기 위해 2주 간격으로 표집하여 최대 50%까지 확장하여 탐색할 수 있도록 설정하였다.

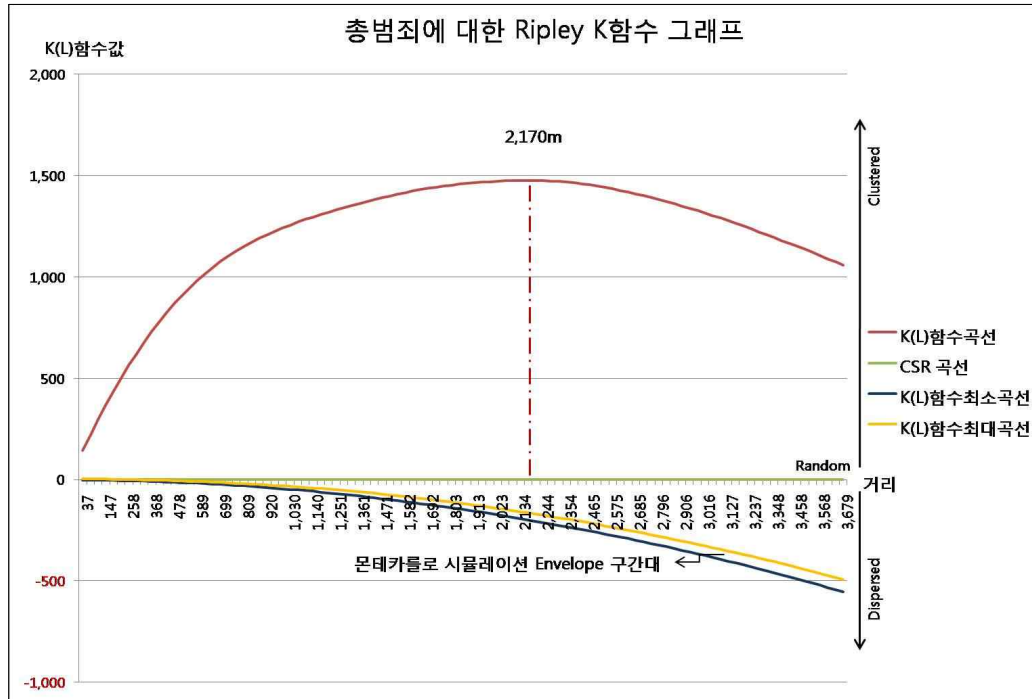


FIGURE 7. M시 총범죄에 대한 Ripley K함수분석 결과

M시의 총범죄를 대상으로 한 Ripley K함수 분석결과가 그림 7과 같다. Y축은 K(L)함수 값을 의미하며 X축은 비교 대상 점사상 간의 거리를 의미한다. K(L)함수의 최대곡선과 최소곡선은 몬테카를로 시뮬레이션을 100회 반복하여 도출된 K(L)함수값의 최대·최소 예측치를 토대로 해서 그려진 것으로 envelope이라 불리는 신뢰구간을 의미한다 ( $\alpha=0.05$ 의 양측검증). CSR(Complete Spatial Randomness) 곡선은 완전한 임의적 분포의 가정하에 생성된 분포곡선을 의미한다. 분석 결과, K(L)함수곡선은 임의적인 분포를 가정한 몬테카를로 시뮬레이션 envelope 구간대를 훨씬 벗어나 그래프 상단에 그려지고 있어 강한 군집적인 패턴을 보인다고 해석할 수 있다. 이러한 군집 경향은 약 2,170m까지는 계속해서 증가하다가 그 이후로 점차 이산적인 분포 경향을 보이면서 K(L)함수 곡선도 하향 추세로 전환되고 있음을 알 수 있다. 따라서

M시의 총범죄에 대한 공간군집성을 판단할 수 있는 최대거리 한계는 약 2,170m라고 볼 수 있다. 따라서 시공간검정통계량 분석에서 공간탐색을 위한 원형크기의 최대 한계를 2,170m로 설정하였다.

표 3은 Prospective와 Retrospective 시공간순열방식으로 시공간검정통계량에 의해 각각 도출된 시공간범죄 집중지역의 분석결과이다. 먼저 Prospective 시공간순열방식으로 탐색된 분석결과를 살펴보면, 가장 집중 가능성이 큰 1차 군집과(the most likely cluster) 3개의 2차 군집집단(second clusters)이 탐색되었다( $p<0.05$ ). 1차 군집지역을 중심으로 표 2를 해석해 보면, 위치ID는 3,640개의 범죄발생지점에 표기된 고유 식별자를 뜻하며 X, Y좌표 및 반경(radian)값은 시공간범죄 집중지역의 중심점과 공간범위를 의미하고, 시점과 종점은 시간범위를 뜻하는 것으로 시공간탐색 결과에 의해 만들어진 탐색원도우인

원통(circular cylinder)의 둘레와 높이를 의미한다. 탐색수는 이 원통에 포함된 범죄발생(위치ID)의 총수를 의미하며, 검정통계량은 원통내의 범죄발생수와 원통 밖의 범죄발생수를 비교하여 계산한 우도비(likelihood ratio) 값으로 x축 162420과 y축 192452의 시작 중심점에서 원의 최대반경을 2,170m까지 확장해 가며 계산한 값들 중에서 반경 841.4141m에서의 우도비 175.9418이 가장 큰 값을 보여주고 있다. 이 최대의 우도비를 갖는 원통크기에 포함된 범죄 사건들 중에서 2006년 10월 23부터 2006년 12월 31일에 걸쳐 발생한 범죄 사건수가 곧 관측수가 된다. 1차 군집의 경우 원통에 포함된 256개의 범죄사건 중 221개가 이 기간에 발생된 것이다. p값은 CSR(Complete Spatial

Randomness)의 가정하에서 몬테카를로 시뮬레이션을 1,000회 반복했을 때 해당 군집결과가 우연에 의한 군집분포일 가능성이 0.001%이라는 것을 의미하며, 예측수는 CSR을 가정한 시뮬레이션의 결과에 의해 도출된 해당지역에서의 범죄발생수에 대한 예측값을 의미한다. ODE는 실제 관측값과 예측값 간의 비율로서 시뮬레이션 예측으로는 약 46건 정도의 범죄사건이 발생할 것으로 예상되는 지역이나, 실제로는 221건의 범죄사건이 발생됨으로써 약 4.8배의 범죄발생 위험률을 보이는 시공간범죄 집중지역임을 의미하고 있다. Retrospective 시공간순열방식으로 탐색된 분석결과에서는 가장 집중 가능성이 큰 1개의 1차 군집과 5개의 2차 군집집단이 각각 탐색되었다( $p < 0.05$ ).

TABLE 3. 최우도값을 갖는 1차 시공간군집과 17개의 2차 시공간군집 결과

탐색 방법	군집 번호	위치 ID	군집 공간범위			군집 시간범위		탐색 수	검정 통계량	p-값	관측 수	예측수	ODE
			X좌표	Y좌표	군집 반경	시점	종점						
prospective S-T permutation*	1(A)	2108	162420	192452	841.4141	2006/10/23	2006/12/31	256	175.9418	0.001	221	46.0659	4.7974
	2(B)	1757	161006	191564	622.1761	2006/08/14	2006/12/31	280	66.9715	0.001	252	112.1538	2.2469
	3(C)	2608	162166	193367	167.0973	2006/11/20	2006/12/31	29	15.5357	0.001	17	2.9956	5.6749
	4(D)	2554	159849	193252	727.9188	2006/07/31	2006/12/31	41	10.8541	0.016	41	17.9769	2.2807
Retrospective S-T permutation**	1(A)	2108	162420	192452	841.4141	2006/10/23	2006/12/31	256	175.9418	0.001	221	46.0659	4.7975
	2(B)	1702	160966	191496	456.4163	2006/08/28	2006/10/22	204	156.5530	0.001	185	36.0923	5.1257
	3(C)	2880	162500	193693	397.4301	2006/02/27	2006/04/23	238	134.2177	0.001	179	39.3615	4.5476
	4(D)	3008	161855	193876	277.4702	2006/05/22	2006/06/04	71	78.6812	0.001	44	2.9077	15.1323
	5(E)	2754	162045	193531	62.3085	2006/01/30	2006/02/12	69	70.1170	0.001	37	2.1800	16.9729
	6(F)	2571	159844	193279	873.5351	2006/07/31	2006/08/13	44	67.3019	0.001	33	1.6681	19.7826

\*주1) 전체 41개의 초기 탐색 핫스팟지역 중 단지 4개 지역만이 유의수준  $\alpha < 0.05$ 를 만족함  
 \*\*주2) 전체 51개의 초기 탐색 핫스팟지역 중 단지 6개 지역만이 유의수준  $\alpha < 0.05$ 를 만족함

그림 8은 Prospective 방식에 의해 도출된 시공간검정통계량 결과를 지도화 한 것으로 총 41개의 군집후보군들 중에서 표 3에 정리된 바와 같이 서로 다른 시공간 탐색범위를

갖는 4개의 시공간범죄 집중지역( $\alpha < 0.05$ )을 도식화 한 것이며, 그림 9는 Retrospective 방식에 의해 도출된 시공간검정통계량 결과로서 총 51개의 후보군들 중 유의수준  $\alpha$

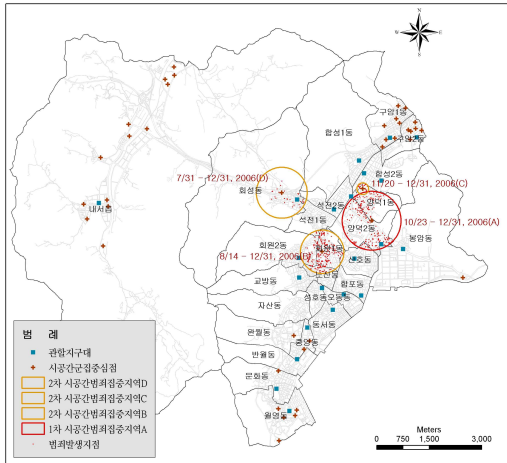


FIGURE 8. 시공간검정통계량에 의한 도시범죄 핫스팟지역(Prospective 분석결과)

<0.05를 만족하는 6개의 시공간범죄 집중지역을 지도상에 표시한 것이다. 그림 2~6의 공간적 분포특성만을 고려했던 일반적인 공간군집 분석결과와 달리 시공간분석에서는 철도역 부근과 합성동 일대가 아닌, 양덕동 일대가 시공간적으로 범죄발생 집중도가 가장 높은 지역으로 나타났으며, 그 다음으로 회원 1동, 합성2동 및 양덕 1동의 경계지역, 그리고 회성동 일대가 범죄발생위험률이 높은 지역으로 나타났다.

Retrospective 방식에 의한 분석결과에서도 Prospective 방식과 유사하게 범죄 집중 가능성이 가장 큰 1차 군집지역이 양덕동 일대에서 발견되고 있으며, 2차 군집지역들도 회원1동과 합성1, 2동, 그리고 회성동 일대에서 나타나고 있다. 그러나 1차 군집지역을 제외한 2차 군집지역들에서는 시공간 순열의 계산방식에 따라 시간적 범위와 공간적 범위구간이 약간씩 서로 다르게 나타나고 있음을 알 수 있다. 특히 Retrospective 방식에 비해, Prospective 방식에서는 대상기간의 종점(2006년 12월)에 더 가까운 쪽으로 시공간 집중지역이 다수 나타나고 있다. 한편, 기존의

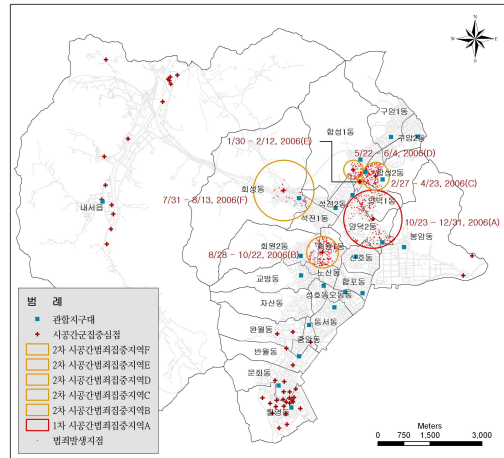


FIGURE 9. 시공간검정통계량에 의한 도시범죄 핫스팟지역(Retrospective 분석결과)

공간핫스팟 분석에서는 주로 상업지역들(합성동, 중앙동, 동서동, 오동동 일대)에서 범죄핫스팟 지역이 집중적으로 나타나고 있는데 반해, 시공간검정통계량 분석에서는 주로 주거지역(양덕동, 회원동, 회성동 일대)을 중심으로 범죄핫스팟 지역이 분포하고 있으며, 일부 구시가지의 합성동 일대의 상업지역을 제외하고는 구도심의 중앙동, 동서동, 오동동 일대의 상업지역과 신시가지의 월영동, 문화동 일대의 상업지역들에서는 통계적으로 유의한 시공간적 집중 현상이 전혀 나타나지 않고 있다. 또한 주거지역에 있어서도 공동주택단지나 단독주택위주의 전용주거지역보다는 다세대 위주의 주거혼재비율이 상대적으로 높은 구도심의 일반주거지역(회원1동과 양덕동)에서 시공간적인 범죄 집중 현상이 높게 나타났으며, 이들 주거지역은 치안지구대 입지의 사각지대라는 점에서 공통적 특징을 보이고 있다(그림 8, 9 참조). 따라서 이 지역에서의 범죄발생 양상은 공간상으로뿐만 아니라 특정기간에 걸쳐서도 범죄가 집중적으로 발생하고 있기 때문에 경찰 인력 및 자원의 효율적인 배분과 순찰 이동경로의 설정, 그리고 셉테드

(CPTED : Crime Privement Through Environmental Design) 적용 등과 같은 다양한 범죄예방활동들이 이들 지역을 중심으로 지속적으로 전개되어야 할 것으로 판단된다.

## 결 론

일반적인 공간핫스팟 분석에서는 그 범죄 집중의 결과가 우연에 의해 나타난 결과인지 실제적으로 범죄발생위험률이 높은 지역으로서 범죄가 빈번하게 발생한 지역인지를 통계적으로 검증해 내기가 쉽지 않은 측면이 있었다. 이러한 관점에서 본 연구는 후속적인 도시범죄 분석을 위한 진단계적 성격의 기초연구로서 시공간적인 범죄다발지역을 통계적인 검증 과정을 거쳐 범죄발생 위험률을 지도화해낸 점과 미시적인 관점에서 후속연구에 필요한 연구대상지역의 선정과 시간 및 공간범위 설정에 대한 보다 객관적인 기준을 제시해 주고 있다는 점에서 기존의 공간거리 위주의 범죄 핫스팟분석 연구들과는 다른 차별성을 띠고 있다.

본 연구를 통해 탐색된 시공간범죄 집중지역은 특정 시간대와 공간에서 집중적으로 발생하는 범죄사건들에 대한 경찰 행정의 신속한 대응과 보다 효율적인 인력배분 및 치안활동의 강화를 위해 최우선적으로 고려되어야 할 지역이라 할 수 있고, 도시계획 관점에서는 셉테드(CPTED)의 적용에 대한 우선순위를 갖는 지역이라 할 수 있으며, 다중이용 시설물과 같은 범죄유발시설들에 대한 계획적 입지와 체계적인 관리가 무엇보다 필요한 지역이라 할 수 있다. 따라서 이러한 범죄분포 특성을 반영한 차별화된 대응전략으로서의 범죄예방 활동이 이루어져야 할 것으로 판단된다. 이러한 의미 있는 연구결과의 도출에도 불구하고 본 연구는 다음과 같은 부분에서의 연구 한계 및 후속적인 연구가 필요하다.

첫째, 본 연구는 특정년도에 한정된 시계열

범죄 자료만이 분석에 활용되었다. 따라서 시공간범죄 집중지역과 범죄발생위험지역을 보다 정확하게 탐색하고 예측해 내기 위해서는 다년간에 걸쳐 누적된 충분한 시계열적 범죄 자료의 추가적인 확보와 그에 따른 분석이 이루어져야 할 것이다.

둘째, 본 연구는 특정 공간과 시기에 걸쳐 범죄율이 높게 나타나고 있는 시공간범죄 집중지역들을 통계적 검증과정을 통해 보다 객관적으로 탐색해 낼 수 있는 시공간에 기반한 새로운 핫스팟 분석방법론을 제안하고 있으나, 이러한 시공간범죄 집중지역이 지역적 특성 가운데 어떤 영향요인과 원인들에 의해 나타나고 있는지에 대해 구체적인 분석을 시도하지는 못하였다. 이것은 또 다른 성격의 연구주제로서 후속적인 연구가 필요한 부분이다. 이를 위해서는 범죄프로파일링 기법뿐만 아니라 좀 더 다양한 사회·경제적, 그리고 물리적인 환경요인들이 고려되어야 하며, 범죄 유형별로 좀 더 세분화시켜 접근할 필요가 있고, 인간 행태에 기반한 환경행태학적 접근에 대한 이해와 적용이 병행되어야 할 것으로 판단된다.

끝으로, 본 연구에서는 군집(cluster)에 대한 탐색범위를 원통으로 정의하여 지도상에 표기하고 있으나, 기존의 공간핫스팟 분석결과에서도 그러하듯이, 그 경계구역이 절대적인 기준이 되는 것은 아니므로, 정책결정의 공간단위로 활용함에 있어서는 세심한 주의가 필요하다. **KAGIS**

## 참고문헌

- 손학기, 박기호. 2008. 부동산 가격변동 핫스팟 탐색을 위한 공간통계기법. *대한지리학회지* 43(3):392-411.
- 정경석, 문태현, 정재희, 허선영. 2009. GIS와 공간통계기법을 이용한 시공간적 도시범죄 패턴 및 범죄발생 영향요인 분석. *한국지리정보학회지*

- 12(1):12-25.
- 정경석. 2010. 공간범죄통합분석모형을 이용한 도시범죄의 시공간적 분포 특성 및 영향요인 분석. 경상대학교 대학원 박사학위논문. 151쪽.
- 황선영, 황철수. 2003. GIS를 활용한 도시 범죄의 공간패턴분석. 국토계획 38(1):53-66.
- Andersen, M.A. 2006. A Spatial Analysis of Crime in Vancouver, British Columbia: A Synthesis of Social Disorganization and Routine Activity Theory. *The Canadian Geographer/Le Géographe canadien* 50(4):487-502.
- Anselin, L. 1992. Spatial data analysis with GIS: An introduction to application in the social sciences. NCGIA Technical Report, pp.92-10.
- Diggle, P.J., R. Haggkvist and S.E. Morris. 1995. Second-order Analysis of Space-Time Clustering. *Statistical Methods in Medical Research* 4:124-136.
- Eck, J.E., S. Chainey, J.G. Cameron, M. Leitner and R.E. Wilson. 2005. Mapping Crime: Understanding Hot Spots. U.S. Department of Justice Office of Justice Programs National Institute of Justice, pp.23-25.
- Goldsmith, V., P. McGuire, J. Mollenkopf and T. Ross. 2000. Analyzing Crime Patterns: Frontiers of Practice. Thousand Oaks, Sage, pp.3-12.
- Hirschfield, A., P. Brown and P. Tod. 1995. GIS and the Analysis of Spatially-referenced Crime Data: Experiences in Merseyside, UK. *International Journal of Geographical Information Systems* 12:191-210.
- Jacquez, G.M. 1996. A k Nearest Neighbour Test For Space-Time in Interaction. *Statistics in Medicine* 15:1935-1949.
- Kang, H.J. 2009. Detecting agglomeration processes using space-time clustering analyses. *The Annals of Regional Science* 45(2):291-311.
- Knox, E.G. 1964. Epidemiology of Childhood Lukaemia in Northumberland and Durham. *British Journal of Preventive and Social Medicine* 18:17-24.
- Kulldorff, M. and N. Nagarwalla. 1995. Spatial disease clusters: Detection and Inference. *Statistics in Medicine* 14:799-810.
- Kulldorff, M. 1997. A Spatial Scan Statistic. *Communications in Statistics-Theory and Methods* 26:1481-1496.
- Kulldorff, M. 2001. Prospective Time-periodic Geographical Disease Surveillance using a Scan Statistic. *Journal of the Royal Statistical Society A* 164:61-72.
- Kulldorff, M., R. Heffernan, J. Hartman, R. Assunção and F. Mostashari. 2005. A Space-Time Permutation Scan Statistic for Disease Outbreak Detection. *PLoS Medicine* 2(3):216-224.
- Levine, N. 2005. CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Location. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington DC. Chapter9. 8pp.
- Mantel, N. 1967. The Detection of Disease Clustering and a Generalised Regression Approach. *Cancer Research* 27:209-220.
- Murray, T.A., I. McGuffog, S.J. Western, and P. Mullins. 2001. Exploratory Spatial Data

- Analysis Techniques for Examining Urban Crime. *British Journal of Criminology* 41(2):309–329.
- Ripley, B.D. 1976. The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability* 13:255–266.
- Shimada, T. 2004. Spatial Diffusion of Residential Burglaries in Tokyo: Using Exploratory Spatial Data Analysis. *Behaviormetrika* 31(2):169–181.
- Weisburd, D. 1997. *Reorienting Crime Prevention Research and Policy: From the Causes of Criminality to the Context of Crime*, National Institute of Justice, Washington, DC. **KAGIS**