

공간더빈모형을 활용한 범죄 발생 환경요인 분석

박현지¹, 정지영², 박현수³, 유예원⁴, 유제혁⁵^{1,2,3,4}덕성여자대학교 컴퓨터공학전공 학부생⁵덕성여자대학교 데이터사이언스학과 교수

phjgina0314@gmail.com, gyjeong021@gmail.com, dontlikesm@gmail.com

yoovy0106@gmail.com, jhrew@duksung.ac.kr

Analysis of Environmental Factors for Crime Occurrence based on Spatial Durbin Model

Hyeon-Ji Park¹, Ji-Young Jeong², Hyun-Soo Park³,Yewon Yoo⁴, Jehyeok Rew⁵^{1,2,3,4}Dept. of Computer Engineering, Duksung Women's University⁵Dept. of Data Science, Duksung Women's University

요 약

사회에서 발생하는 범죄는 다양한 원인이 복합적으로 연관되어 발생하기 때문에, 유형별 범죄에 중요영향을 주는 요인을 파악하는 것이 시급하다. 본 연구는 대학 자취촌 지역에서 집중적으로 발생하는 범죄 유형을 대상으로 하여, 범죄에 영향을 줄 수 있는 환경요인을 선별하고, 공간더빈모형을 활용하여 주요 발생 요인을 분석했다. 대표적 대학가를 포함하는 강북구, 도봉구, 관악구를 50x50m² 공간 그리드로 분할하고, 범죄, 폭력, 절도 등에 영향을 줄 수 있는 41개의 환경 변수를 다양한 공간 통계모형들을 활용하여 상세 분석했다. 이 중, 공간 더빈 모형이 가장 효과적인 통계적 설명력을 가지는 것을 다양한 비교 실험을 통해 검증했으며, 범죄 주요 요인들에 대한 분석적 고찰을 수행했다.

1. 서론

범죄 발생은 사회적, 경제적, 환경적 요인이 복합적으로 작용하는 현상으로, 최근 도시화 현상이 심화하면서 새로운 유형의 범죄 패턴이 다양하게 발생하고 있다. 특히, 대학가 주변인 자취촌 지역은 지속해서 범죄 취약 지역으로 지목되어 왔으나, 이에 관한 체계적인 연구와 대응 전략은 아직 미흡한 실정이다. 따라서 본 연구는, 공간더빈모형(Spatial Durbin Model, SDM)을 사용하여 각 환경요인이 범죄 발생에 미치는 직·간접적인 영향을 분석함으로써, 효과적인 도시 및 지역 계획을 위한 기본 분석을 수행하는데 중요한 통찰을 제공할 수 있다.

2. 공간 모형 적용을 위한 데이터처리

대표적인 대학 자취촌인 강북구, 도봉구, 관악구 일대를 연구 지역으로 선정하고 이를 50x50m² 그리드로 나누어 총 2,833개의 그리드 데이터를 구성했다. 또한, 5대 범죄 중 발생 빈도가 높은 상위 3개(전체, 폭행,

절도)를 종속 변수로 설정하고, 이에 대응되는 생활안전지도 API의 데이터를 shp 파일 형식으로 수집하여 2,833개의 그리드에 각각 매핑했다. 독립 변수는 V-World API, 서울시 공공 데이터 등을 통해 수집한 환경·사회·공간 데이터를 활용했다. 공시지가·유동 인구 등 인구·경제와 층수·면적 등의 건물 상세 정보, 경찰·음식점·숙박시설과 같은 POI(Point of Interest) 지점 정보, 버스정류장, 지하철역 등 대중교통, CCTV·보안 등·녹지·하천 등의 도시경관 정보, 도로 통과도·전환비 등 가로망 구조, 총 6개 데이터 영역으로 분류하고, 총 41개의 상세 데이터를 요인 분석 대상으로 지정했다.

독립 변수의 데이터 전처리를 위해 각각의 분포를 확인한 후, 왜도가 ±1 초과할 경우, 로그 변환을 진행했다. 또, 변수가 개수 값이라면 Min-Max 정규화를, 면적처럼 특정 값이라면 Standard 표준화를 수행한 후, 다중 공선성 제거를 위해 VIF 값이 10 이상으로 나타난 변수는 요인에서 제거했다.

		전체	폭력	절도
주요 요인	비례	건물 개수(0.43***), 전환비(0.29***) 외 7개	건물 개수(0.34***), 통과도(0.13***) 외 4개	건물 높이(0.20***), 전환비(0.25***) 외 5개
	반비례	건물 연면적(-0.20***) 외 1개	휴게음식점(-1.57**) 외 2개	건물 연면적(-0.18***) 외 1개
직접 효과		건물 높이(0.15) 외 10개	건물 높이(0.13) 외 13개	지하철역 거리(-0.15) 외 8개
간접 효과		사교육 시설(0.93) 외 6개	사교육 시설(0.89) 외 9개	휴게 음식점(1.81) 외 7개

<표 4> 범죄 유형 별 공간더빈모형(SDM) 분석

3. 공간 통계모형 적용과 요인 분석결과

본 연구는 일반 회귀모형이 아닌 공간통계 모형을 활용했다. 특히, 일반회귀 모형과는 달리 공간통계에 활용되는 커널 가중치는 거리 감쇠 효과를 반영할 수 있어, 거리가 멀어질수록 상호 작용의 영향력이 감소하는 현실적인 상황에 적용할 수 있다. SDM은 종속 변수와 독립 변수에 공간적 지연을 도입하여 관측치의 상호 의존성 확인 및 변수의 편향을 해결할 수 있으며, 영향력 분석을 통해 직접 효과, 간접 효과, 그리고 전체 효과를 도출할 수 있어 활용에 적합하다.

Moran's I 계수는 전체 범죄 0.762, 폭력 0.744, 절도 발생 0.666으로 나타나 공간적 자기 상관성이 존재하므로, 공간통계 모형의 적용이 적합하다고 판단했다.

Wald 검정과 우도 비 검정(LR 검정)을 통해 'SDM이 SLM(Spatial Lag Model, SLM), SEM(Spatial Error Model: SEM)로 단순화될 수 있다'라고 설정한 귀무가설은 1% 유의수준에서 기각된다.

		전체범죄	폭력	절도
SLM	Wald	72.315	95.987	53.771
	LR	71.393	94.393	52.860
SEM	Wald	164.018	161.652	179.008
	LR	141.117	147.662	124.035

<표 1> 공간통계 모형 검정

	Adjusted R ²	AIC
OLS	0.553405	11256.860
SLM	0.810364	9205.603
SEM	0.384607	9273.327
SDM	0.815589	9192.210

<표 2> 공간통계 모형 설명력 지표

또, 전체 범죄를 대상으로 Adjusted R²와 Akaike Information Criterion을 비교하면, SDM 모형이 가장 높은 설명력을 보였다.

SDM 분석 결과, 범죄별 Pseudo R²와 Spatial Pseudo R²는 <표 3>과 같이 나오며, 이를 통해 SDM 모델이 공간분석에 적합함을 알 수 있다.

	Pseudo R ²	Spatial PR ²
전체범죄	0.8142	0.6451
폭력	0.7687	0.6026
절도	0.6939	0.5218

<표 3> 공간 자료 모형별 적합성 지표 비교

<표 4>의 분석 결과, 모든 범죄 유형이 건축 환경요인에 영향을 받는 것으로 나타났다. 건축물의 높이, 연령, 개수 수치가 높을수록 범죄 발생이 증가하는 반면, 연면적이 넓을수록 감소했다. 이는 좁고 오래된 주택단지가 밀집된 자취촌의 특성을 반영한다. 또한, 통행 유효도인 통과도와 보도 방향 전환을 의미하는 전환비는 범죄 발생에 양(+)의 영향을 미친다. 이는 자취촌 주변의 거리 특성과 연관된다. 폭력의 경우, 주류 판매가 금지된 휴게 음식점이 많을수록 빈도수가 감소했고, 절도는 주변에 금융 시설 수와 비례하는 경향을 보인다. 직접 효과 분석에서는 건물 높이, 지하철역과의 거리 등이 범죄 발생에 영향을 미쳤으며, 간접 효과 분석에서는 인접 공공시설 수, 사교육 시설 수 등이 영향을 주었다. 이러한 패턴은 폭력에서도 유사하게 나타났으며, 절도 범죄에서는 인접 지역의 공공시설과 휴게 음식점 수, 교육 관련 시설 수 등이 간접 효과를 통해 유의한 영향을 미쳤다.

참고문헌

[1] 김선재, 조월, 이수기 "도시환경 특성과 범죄발생의 연관성 분석: 도시 빅데이터와 공간더빈 모형을 활용하여," 한국도시설계학회지, 제23권 제3호 pp.143-156, 2022.