

공공데이터를 이용한 골목길 안전 데이터의 시각화

이창민*, 박상원*

*한국외국어대학교 정보통신공학과

e-mail : sky569865@naver.com

Visualizing of Alley Safety Data making use of public data

Chang-Min Lee*, Sang-Won Park*

*Dept of Information & Communiactions Engineering

, Hankuk University of Foreign Studies

요 약

안전한 도시환경에 대한 관심이 증가하면서 다양한 안전도 측정 모델을 통해 안전도를 측정하여 범죄 예방활동, 재난예측 분야에 적용하고 있다. 안전도 측정 모델을 통해 도로의 안전도 수치를 정량화하여 안전도관련 데이터가 맵핑된 지도와 같은 실제 보행자가 이용할 수 있는 서비스에 대한 필요가 증가하고 있다. 기존의 모델은 적용분야가 구체적이지 않고 모델을 서비스에 적용하기가 어렵고 정량화된 값을 제시하는 모델이 없어 수치가 의미하는 바를 명확히 알 수 없다. 따라서 공공데이터를 활용하여 “도로 안전지수 측정모델”을 개발하여 값을 얻고 이를 구성하는 각각의 공공 데이터를 지도상에 보여주고자 한다.

1. 서론

통계청 범죄율 데이터를 보면 우리나라의 범죄율은 지속적으로 높아지는 추세이다.¹⁾ 특히나 주거지역에서 일어나는 범죄의 경우 해당 지역에 사는 주민들에게 자신이 사는 곳이 안전하지 않다는 인식을 심어주게 된다. 범죄학, 건축학, 도시공학의 분야에서는 범죄율을 낮춰 더 안전한 주거지역을 만들기 위해 다양한 안전도 측정 모델을 통해 안전도를 측정하여 범죄 예방 활동, 재난 예측 분야에 적용하고 있다. 기존의 안전도 측정 모델은 큰 지역이나 도시단위의 범위에서 안전도 분석, 검증이 많이 이루어 졌다. 하지만 지역단위와 같은 광범위한 영역을 검증한 모델은 주거지역에서 도로단위로 일어나는 치안문제를 해결하기 위한 서비스를 제공하기에는 효용성이 떨어진다. 따라서 실질적으로 주거지역에서 필요한 서비스를 제공하기 위해서는 도로 단위에서 측정된 안전도를 도출할 수 있는 모델이 요구된다.

다층 모형 분석 모델[1]을 사용하게 되면 변인 간에 상관관계는 알 수 있지만 독립변인들로부터 체감안전도 값을 정량화된 값으로 제시하지 못한다. 따라서 이 모델이 제시하는 결과만으로는 안전도 모델을 시각화하는 서비스에 적용할 때 제한사항이 많다. 또한 P - S - R 모델을 사용한 경우에는 기존 연구 사례에는 지역방범지수[2]가 있

는데 적용 분야가 구체적이지 않거나 너무 넓은 지역을 분석한 통계라서 실질적인 사용가치가 떨어진다는 단점이 있다. 마지막으로 과거 연구 사례에서 사용된 데이터들은 과거의 통계자료나 지엽적인 데이터를 활용한 경우가 많았다. 이러한 데이터를 사용할 경우 현재의 지표를 바로 반영하기 어렵고 데이터의 변경이 있을 경우 수동으로 일일이 바꾸어 주어야하는 불편함이 있다. 최근 서울시와 몇몇 지방자체 단체의 경우 웹상에 자신의 지역에 해당하는 공공데이터를 파일형태, 오픈API 형태로 제공하고 있는데 이를 제공하는 대표적인 사이트로 공공데이터포털²⁾과 서울 열린 데이터 광장³⁾이 있다. 해당 사이트에서 제공하는 API형태의 데이터를 이용하면 최신 데이터를 쉽게 반영할 수 있고 현재 상황을 잘 반영한 안전도를 제공할 수 있을 것이다. 이것으로 보행자로 하여금 한 눈에 주변상황을 인지할 수 있도록 지도를 이용한 시각화 서비스를 제공할 수 있을 것이고 나아가 안전 경로 추천과 같은 서비스에 적용하여 보행자가 안심하고 골목길을 이용할 수 있을 것으로 기대한다.

2. 관련 연구

기존 연구 사례도 마찬가지로 안전도를 측정하기 위한

1) 통계청, http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=4057

2) 공공데이터포털, <https://www.data.go.kr>

3) 서울 열린데이터 광장, <http://data.seoul.go.kr>

모델을 제시하고 있는데 이를 먼저 비교 분석해 보고자 한다.

2.1 안전도관련 모델 비교

모델명	적용 대상	응용 서비스	적용모델	정량화
지역 방범 지수[2]	행정동의 지역단위	가중치 지도 매핑	P-S-R 모델	O
체감안전도 [1]	지역사회	-	다층자료 분석	X
주민체감안전도[3]	구 단위	-	다층모형 분석	X
생활안전지표[4]	도시	-	-	X
도로 안전도	개별 도로	경로 추천, 매핑	-	O

<표 1> 관련연구 내용비교

지역방범지수[2]는 P-S-R모델을 사용해 정량화된 값을 도출하여 가중치 지도 매핑 서비스에 데이터를 사용하고 있다. 하지만 적용대상이 행정동의 지역단위라서 도시 계획 단위에서 정책 의사 결정 지원에는 도움이 되지만 특정지역의 보행자를 대상으로한 서비스를 제공하기에는 부족하다. 체감안전도[1], 주민 체감안전도[3]의 경우 다층 모형 분석을 통하여 검증을 하였는데 해당 모델은 체감안전도를 구성하는 세 가지 핵심요인을 분류하고 속해 있는 독립변인이 안전도와 더 관련이 있는지에 대한 검증을 하고 있지만 정량화된 값을 제시하지 못하여 모델을 실제로 적용하기가 어렵다는 단점이 있다.

본 논문에서 제안한 “도로 안전도” 모델에서는 개별 도로를 대상으로 하여 정량화 된 가중치를 얻을 수 있다. 따라서 얻은 수치로 안전 경로 추천에 이용할 수 있을 것이다. 생활안전지표[4]는 치안과 관련된 안전도 외에도 보행 안전취약, 노인취약이라는 유형을 두어 AHP분석방법 (Analytic Hierarchy Process)을 통하여 여러 가지 유형에 속한 독립변인들을 비교했다. 비교 후 대안 중요도가 높은 5가지를 유형별로 핵심 변수로 지정을 하였다. 이 분석 방식을 통해 변수의 수를 줄이면서 상대적으로 중요한 변수를 뽑을 수 있기 때문에 변수의 수가 너무 많아서 생기는 오류를 줄여줄 것으로 예상을 할 수 있다. 따라서 본 논문도 대상을 개별 도로로 지정하여 AHP분석방법 (Analytic Hierarchy Process)를 사용해 독립변인을 선정할 것이다. 하지만 생활안전지표[4]에서는 실제 필요한 데이터를 수집하는 단계를 수행하고 검증하는 단계가 없어 실제 효용성이 있는지 확인할 수 가 없다. 따라서 필요한 데이터를 수집하고 값을 도출해 서비스에 이용하는 단계까지 수행할 것이다.

2.2 안전도관련 독립변인 비교

앞에서 제시한 관련 논문에서 공통적으로 가장 많이 사용된 변수 순으로 연령이 네 번으로 제일 많고 성별, 5대 범죄통계, CCTV 설치비율이 세 번 그리고 노후주택비율, 외국인비율, 유동인구 밀도, 가로등 설치 유무, 도보 순찰이 두 모델에서 공통적으로 사용되었다.

본 논문에서는 특정 지역의 안전도를 구하는 것이 아니기 때문에 인구 비율을 의미하는 요소를 제외한 환경적인 독립 변인만을 반영할 것이다. 서울시 열린 광장의 오픈데이터를 활용할 것이기 때문에 도보 순찰을 대체할 수 있는 데이터로 택시 운행량과 버스 운행 유무로 설정하였다. 또한 나머지 모델마다 가지고 있는 개별 독립변인들은 반영하지 않고 적용대상이 도로인 만큼 도로 폭(도로 정보)와 도로 길이를 반영하여 체감안전도 수식을 구성하였다.

3. 도로 안전도 모델

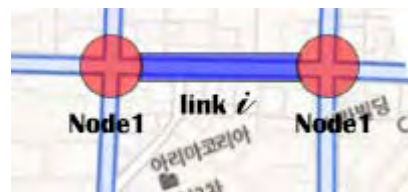
3.1 도로 안전도 지수의 개념 및 구성

각 도로의 특성과 도로의 치안과 관련된 환경적인 요소를 반영해 안전 경로 탐색에 사용할 가중치를 만든다. 이 가중치 값을 “도로 안전도”의 기본 개념으로 한다.

도로의 특성과 도로 안전도 구성하는 환경 요소(범죄 억제요소)를 기본으로 하여 각 요소에 알맞거나 대체할 수 있는 열린 광장의 오픈데이터를 사용하여 구성하였다.

3.2 “도로 안전도” 지수 개발 절차

개발 절차는 크게 평가지표를 선정하고 평가지표를 표준화하는 식을 산정한 후 각 도로의 특성에 맞게 수치를 반영하여 값을 도출하였다. “도로 안전도” 지수는 도로의 특성과 범죄 억제요소에 해당하는 지표를 CPTED(범죄 예방 환경 설계)에 부합하는 요소를 추가하는 방식으로 선정하였다. “도로 안전도” 지수 산정 식은 가법형 함수를 바탕으로 총 가중치의 양을 지정하고 각 변수의 상관관계를 고려하여 산정하였다.



$$\text{도로 안전도}_i = \frac{a_{i1} \cdot r_{i1}}{d_{i1}} + \frac{a_{i2} \cdot r_{i2}}{d_{i2}} + \frac{a_{i3} \cdot r_{i3}}{d_{i3}} + a_{i4} \cdot r_{i4} + a_{i5} \cdot r_{i5} \dots \dots \dots (1)$$

(그림 1) 링크정의

여기서, a_{ij} 는 링크의 평가지표 j 의 도로안전도 가중치
 r_{ij} 은 링크의 평가지표 j 의 개별 값
 d_i 는 링크의 거리 값

그림1은 링크 i 의 정의를 나타내고 있다. 표준노드 링크⁴⁾

4) 표준노드 링크, <https://nodelink.its.go.kr/>

의 정의를 따랐고, 노드는 차량이 도로를 주행함에 있어 속도의 변화가 발생하는 곳을 표현한 곳이고, 링크란 노드와 노드를 연결한 선을 의미한다. 이 정의에서 링크의 길이가 매우 길어지는 경우는 500m 단위로 링크를 나누었다. 식1은 CCTV 개수(r1), 방법등, 가로등(r2), 택시 운행량(r3), 버스 경로 유무(r4), 도로 폭(r5), 도로 길이(d)을 의미한다. 각 변수들은 거리와 상관관계가 있는 변수와 개별 독립변수로 나누어 개별 값을 표준화하여 대입 후 가산한다.

r1은 공공데이터포털에서 제공하는 전국 CCTV표준데이터를 활용하였다. 전국 CCTV표준 데이터는 전국 지역별로 CCTV의 설치위치와 관리기관에 해당하는 정보가 있는데 설치위치에 해당하는 경위도 값을 얻어왔다. r2는 국립재난안전연구원의 생활안전지도⁵⁾ 측에 파일데이터를 공문을 통해 요청하였고 직접 제공받았다. 가로등은 대로변에 있는 일반 가로등을 의미하고 방법등은 경찰청이나 지역 자치단체에서 필요한 위치에 임의로 설치한 가로등을 뜻한다. 택시 운행량(r3) 데이터는 서울시 열린 데이터 광장의 서울시 택시 운행 분석데이터 정보를 활용하였다. 도로(링크)의 택시 승.하차에 해당하는 값을 얻었다. 지금까지 얻은 r1, r2, r3값은 단위 링크당 개수, 횡수에 해당하는 데이터이기 때문에 전체 길이가 아닌 링크의 단위길이 당 데이터로 만들어주어야 한다. 따라서 세 독립변수에 해당하는 데이터는 링크의 총 길이로 나누어 준다. r4는 공공데이터포털의 정류소정보조회 서비스를 활용하였다. 조사하려는 지역의 링크의 버스정류장 유무를 나타내고 0과 1값으로 나타낸다. r5는 각 링크의 도로 폭을 나타내는 것이고 서울시 열린데이터 광장의 서울시 도로명 정보에서 가져왔다. m단위이며 가중치 값을 곱하여 가산한다.

4. 도로 안전도 데이터 시각화



(그림2) 도로 안전도 지도 맵핑

웹에서 구현한 데이터 시각화 서비스는 지도에 데이터를 맵핑하는 방식을 기본으로 구현하였다. 첫 번째로는 앞선 장의 도로 안전도 모델을 사용하여 성동구 지역의 90개 링크에 한해서 링크별로 도로 안전도 수치를 얻었다. 안전도 모델을 활용하여 얻은 값을 최대값으로 나누어 10단계로 나누었다. 그리고 해당 지역의 링크에 10개의 색상 값으로 안전도 레벨을 나타내어 지도에 맵핑했다. 두 번째로는 도로 안전도 모델을 구성하는 변수들에 해당하는 각각의 데이터들의 경위도 좌표를 기준으로 지도에 마커 오버레이로 표시를 하였다. 변수로는 CCTV, 방법등, 가로등에 해당하는 데이터를 표시하였다. 세 번째로는 성동구 지역 90개 링크를 클릭하면 링크 기준으로 10m내외에 있는 안전도 관련 데이터 중 가로등, 방법등 데이터에 한해서 개수를 보여주고 해당 링크와 같은 노드를 공유하는 인접 링크와의 가로등, 방법등 데이터를 비교할 수 있도록 데이터를 그래프로 렌더링하여 보여주었다.

왼쪽 그림2는 도로 안전도 수치를 지도에 맵핑한 것이다. 네이버지도API를 이용하여 링크에 해당하는 데이터를 받아 색상 값으로 바꾸어 표시하였다. 한정된 지역에 대해 표현을 한 것이지만 확장하여 도로를 이용하는 보행자들이 모두 웹에 접속만하면 자신이 있는 위치를 기준으로 어떤 도로가 얼마나 안전한지를 한눈에 인지할 수 있다.



(그림3) 가로등, 방법등 데이터 지도 맵핑

그림 3은 가로등, 방법등에 해당하는 버튼을 누르게 되면 지역에 분포하고 있는 해당 변수의 경위도 좌표를 불러온다. 웹페이지에서 네이버 지도 API⁶⁾에서 제공하고 있는 마킹 오버레이를 활용하여 표시하였다. 사용자 입장에서는 자신의 위치를 기준으로 개별 안전 데이터가 어떻게 분포하고 있는지 한눈에 볼 수 있다. 그리고 각각의 도로 링크를 기준으로 그림3에서 보이는 데이터들이 10m 내외에 있으면 링크가 가지고 있는 도로 안전도 데이터의 개수를 증가시켜 링크마다 각각의 데이터가 몇 개가 들어있는지 테이블화 하였다. 이를 이용해 그림4에서 보이는 바와 같

5) 생활안전지도, <http://www.safemap.go.kr/cmm/main/mainPage.do>

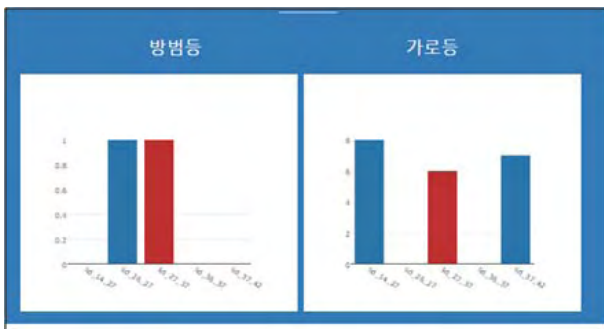
6) 네이버 지도 API, <https://navermaps.github.io/maps.js/>

이 개별 링크를 클릭했을 때 해당 링크가 색이 다르게 표시가 되며 링크가 가지고 있는 데이터를 지도 아래쪽에 표시하여 주는 기능을 가지고 있다.



(그림4) 개별 링크 마우스 이벤트

그림 4화면에서 링크를 클릭하게 되면 아래 그림 5와 같은 그래프가 지도 아래에 표시가 된다. 그래프는 클릭한 링크가 가지고 있는 방법등, 가로등 데이터를 보여주며 링크와 같은 노드를 공유하고 있는 인접 링크가 가지고 있는 데이터도 함께 보여준다.



(그림5) 링크별 가로등,방법등 개수 비교 그래프

그림5에서 빨간색으로 보여지는 막대가 클릭한 링크에 해당하는 데이터의 개수를 의미하며 인접 링크가 가지고 있는 데이터의 개수는 파란색으로 표현을 하였다. 인접 링크의 개수에 맞게 그래프의 x축의 길이를 조정하여 개수를 맞춘다. 그리고 마찬가지로 y축의 데이터의 개수도 최대 값을 최대 길이로하여 나머지 값을 비례하게 표시하도록 그래프를 렌더링 하였다. 이를 확장하여 가지고 있는 모든 링크별 안전도 데이터를 다음과 같이 표현할 수 있을 것이다.

5. 결론 및 활용방안

데이터 수집에 한계가 있었던 과거에는 한정적인 데이터로 넓은 지역에 대한 안전도 모델만을 제시하고 지도에 위험도를 맵핑한 한계가 있었다.

본 논문에서는 적용대상을 도로단위로 좁혀 지역주민이 실제 사용할 수 있는 안전도 모델을 제시하였다. “도로 안전도” 모델을 통해 정량화 된 값을 구하여 안전 경로 추천 서비스를 제공할 수 있도록 하였다. 또한 오픈 데이터를 사용함으로써, 추가적인 데이터를 모델의 변경없이 쉽게 반영할 수 있다. 또한 이러한 공공데이터를 활용하여 웹사이트의 지도에 데이터를 맵핑 함으로써 안전지역에 대한인지와 심리적인 안정감 그리고 범죄 예방 효과를 얻을 수 있을 것으로 예상된다.

참고 문헌

- [1] 최천근 (2012). “다층자료분석을 활용한 체감안전도 결정요인에 관한 연구.” 한국행정학보, 46(1), 103-130.
- [2] 석상목, 권희운, 송기성, 이하경, 황정래 (2017). “범죄예방 정책 의사결정 지원을 위한 지역방법지수(RCPI) 개발.” 한국측량학회 학술대회자료집, 134-137.
- [3] 이병도, 김종길, 유영현 (2015). “주민체감안전도에 영향을 미치는 요인에 관한 다층모형분석.” 사회과학연구, 22(4), 49-70.
- [4] 황선아, 김종구, 강운원 (2015). “안전도시 구축을 위한 생활안전지표 개발.” 대한토목학회 학술대회, 165-166.



이 창 민

2012년~현재 한국외국어대학교 정보통신공학과.
관심분야는 웹, 데이터베이스, 데이터 시각화.