



69

순서형 프로빗 모형을 이용한 해양 교통사고 심각도 영향요인 분석

김은아 (계명대학교 학석사연계과정)
장혁준 (계명대학교 학사과정)
정재훈 (계명대학교 학사과정)
노유나 (한국해양교통안전공단 차장)
권오훈 (계명대학교 조교수)

한국 해양교통안전공단의 최근 5년 해양 사고정보에 따르면, 해마다 해양 교통사고가 꾸준히 증가하고 있다. 해양 교통사고 건수 및 인명피해가 증가함에 따라 해양 교통사고 원인과 사고 심각도에 영향을 주는 요인 파악이 필요하다. 사고 심각도 분석은 SPSS 26.0 통계 패키지를 사용하여 순서형 프로빗 모형을 적용하였으며, 상관분석을 통해 적절한 독립변수를 설정하였다. 이 모형을 사고 심각도에 영향을 주는 요인을 찾아내고 각 요인의 속성값 변화에 따른 사고 심각도 결정 확률을 추정한다. 이 모형을 사용하여 사고의 심각도가 순서에 따라 분류하여 비교·분석으로 사고에 영향을 주는 요인을 도출하였으며 유의성이 확보되지 않은 변수들은 순차적으로 제거하여 해양 교통사고에 미치는 영향 요인에 대해 설명력이 높은 모형을 구축하였다. 그 결과로, 선박의 용도가 비어선 일 경우 선박의 톤수가 10t 미만일 경우 가장 큰 요인으로 발견되었으며, 사고 발생 시간대에서 0~6시와 12~18시에 사고 심각도가 높아지는 것을 확인할 수 있다. 위와 같은 결과를 통해 해양 교통사고 발생 시 사고 심각도에 영향을 미치는 요인을 발견함으로써 이에 대한 구체적 대안을 마련한다면 사고의 심각도를 낮출 수 있을 것으로 판단된다. 본 연구에서 구축한 순서형 프로빗 모형은 시간적·공간적 관점에서 분석하여 사고 심각도에 영향을 미치는 요인을 찾을 수 있지만, 변수를 많이 넣었을 경우 적합도가 떨어진다는 단점이 있다. 또한 단순히 순서형 프로빗 모형으로 분석을 하기에는 한계가 존재하므로 심각도를 찾는 다양한 분석 모형을 통해 해양 교통사고 심각도를 분석하는 연구가 활발히 이루어져야 한다. 현재 해양 교통사고 관련 연구가 많이 이루어지지 않은 실정에서 본 연구는 해양 교통사고 심각도에 관한 연구를 수행했다는 데 의의를 둘 수 있다. 본 연구의 결과가 해양 교통사고 심각도 감소를 위한 기초자료로 활용되기를 기대한다.

71

대전광역시 버스노선 분석 및 개편에 관한 연구

임효진 (한밭대학교 학부과정)
김명수 (한밭대학교 교수)

사람들이 많이 이용하는 대중교통수단 중 하나인 버스는 모든 도로상을 운행하며 비용이 저렴하고, 환경보호, 교통 체증과 주차 문제를 해소하는 장점이 있지만, 심한 굴곡도, 노선의 과다 중복, 긴 배차간격 등의 문제들도 가지고 있다. 본 연구에서는 대전광역시의 버스 현황을 분석하여 문제점을 파악하고, 국내 다른 도시들의 버스노선 개편 사례를 연구하여 적합한 해결책을 통한 효율적인 버스노선 운영 방안을 제시하고자 한다. 대전광역시 버스의 노선 중복도, 굴곡도, 운행거리, 운행시간 등을 위주로 문제점을 파악하였고, 서울특별시, 부산광역시, 대구광역시, 제주특별자치도 등의 국내 도시들의 버스노선 개편 사례를 통하여 대전광역시의 버스노선 개편 방안을 제시하였다. 이용객 수와 굴곡도에 따른 노선 변경으로 버스노선 개편 계획하였고, 이로 인해 버스이용객, 버스운전자, 버스업체, 대전광역시, 환경적인 측면에서 긍정적인 효과를 기대할 수 있을 것으로 보았다.

70

활동인구 기반의 실시간 동적 접근성 산출과 적용 방법론 연구

박진우 (경기대학교 학부과정)
채오성 (경기대학교 학부과정)
박찬규 (경기대학교 학부과정)
김정화 (경기대학교 조교수)

일반적으로 받아들여지는 교통 접근성은 통행 시간(Travel Time)이라는 직관적인 지표를 사용한다. 기존의 통행시간만을 이용하는 교통 접근성 산출의 경우 시간의 흐름에 따른 접근성의 환경 변화를 고려하지 못해 하루 중 다양한 시간대의 실제 접근성을 대표하는 기능이 떨어진다고 볼 수 있다. 특히 출·퇴근 시간, 심야 시간대에서 기존의 교통 접근성을 대입할 시에 실제 상황과의 괴리가 발생하게 된다. 본 연구에서는 접근성과 공간적 형평성을 조사하는 두 가지 예를 통해 국내에서의 교통 접근성에서 시간적 측면의 영향을 보여주고 정적 접근성과 동적 접근성의 비교 분석을 진행하기 위해 서울시를 425개의 행정동별로 구분하고, 사회적 활동의 장소를 서울시 내의 운영 중인 약국으로 정하였다. 이를 위해 시간대별 약국 데이터와 시간대별, 행정동별로 분류된 서울시 생활인구 데이터를 사용하고 PYTHON기반의 웹 크롤링 통해 출발지와 목적지간의 가장 빠른 경로를 산출하여 분석하였다.

72

시계열 변화에 민감한 대중교통 통행특성 분석: 빈도, 거리, 시간

김나연 (충북대학교 학사)
송태진 (충북대학교 교수)

최근 교통, 에너지, 네트워크 등 다양한 시스템을 통합한 스마트시티가 대두되면서 이를 연결하고 제어하기 위한 빅 데이터의 중요성이 확대되었다. 교통카드 이력데이터는 대중교통을 이용할 시에 기록되는 데이터로 버스, 지하철 등의 대중교통 통행특성을 확인할 수 있고, 교통카드의 이용률이 100%에 가까워지면서 이는 전수데이터로써 활용성이 높다. 10년 동안 교통카드 이력데이터를 활용한 다양한 연구들이 진행되고 있으나 대부분의 선행연구에서는 데이터 접근의 한계로 적게는 하루, 많게는 일주일의 데이터 샘플을 활용했기 때문에 시계열적 특성을 반영하지 못하였다. 대중교통의 이용은 수단 특성상 계절별, 출퇴근의 영향을 받는 요일 및 시간대 등과 같은 시계열적 요소들에 영향을 많이 받는다. 시계열적 특성을 반영한 연구라도 데이터의 평균만을 통해 대중교통 통행을 파악한 정도이다. 본 연구에서는 향후 대중교통 정책 및 운영을 위한 기초연구로 시계열 변화에 민감한 대중교통 통행특성을 빈도, 거리, 시간 등을 분석했다. 이를 위해 수도권 지역의 대중교통 카드이력데이터를 활용하여 Linked Trip을 생성하는 알고리즘을 개발, 이를 개편화했다. 해당 데이터는 수도권지역의 1년 치 데이터로 이를 분석하기 위해서는 슈퍼컴퓨터가 필요한 방대한 데이터이다. 따라서 본 연구는 원천데이터를 최소화하는 데이터 마이닝을 거쳐 유효한 데이터베이스를 추출하여 요일별, 시간별, 월별 시계열 분석을 진행하였다. 분석에 사용된 통행특성은 통행량, 통행시간, 통행거리, 승·하차시간, 환승횟수, 통행수단으로 이들 모두 시계열 특성에 따라 상이한 것으로 나타났다. 그 중 요일과 월에 편차가 큰 구간을 추가적으로 발견하여 시계열 변화에 민감한 통행특성을 파악해냈다. 이후 연구에서 기상 및 기후, 공간 정보등의 이종데이터와 연계한다면 대중교통 통행특성 분석의 정확성을 높일 수 있을 것이다. 이러한 연구 결과를 통해 통행패턴이 상이한 시계열 요인을 고려하여 맞춤형 서비스를 제공하고 향후 스마트시티 구현을 위한 MaaS에 통행량 예측 알고리즘 등에 적용할 수 있다. 또한 버스 및 지하철 노선도, 배차간격 등과 같은 대중교통 운영을 위한 대중교통 수요예측 알고리즘 개발 연구에 초석이 되는 자료로 활용될 것이다.

시계열 변화에 민감한 대중교통 통행특성 분석: 빈도, 거리, 시간

Identification of Sensitive Trip Characteristics by Temporal Variation in Public Transit: Frequency, Distance, and Duration

김나연¹ · 송태진²

¹충북대학교 도시공학과, ²충북대학교 도시공학과

I. 서론

1. 연구의 배경 및 목적

최근 교통, 에너지, 네트워크 등 다양한 시스템을 통합한 스마트시티가 대두되면서 이를 연결하고 제어하기 위한 빅 데이터의 중요성이 확대되었다. 교통카드 이력데이터는 대중교통을 이용할 시에 기록되는 데이터로 버스, 지하철 등의 대중교통 통행 특성을 확인할 수 있고, 교통카드의 이용률이 100%에 가까워지면서 이는 전수데이터로써 활용성이 높다. 10년 동안 교통카드 이력데이터를 활용한 다양한 연구들이 진행되고 있으나 대부분의 선행연구에서는 데이터 접근의 한계로 적게는 하루, 많게는 일주일의 데이터 샘플을 활용했기 때문에 시계열적 특성을 반영하지 못하였다. 대중교통의 이용은 수단 특성상 계절별, 출퇴근의 영향을 받는 요일 및 시간대 등과 같은 시계열적 요소들에 영향을 많이 받는다. 시계열적 특성을 반영한 연구라도 데이터의 평균만을 통해 대중교통 통행을 파악한 정도이다. 본 연구에서는 향후 대중교통 정책 및 운영을 위한 기초연구로 시계열 변화에 민감한 대중교통 통행특성을 빈도, 거리, 시간 등을 분석했다. 이러한 연구 결과를 통해 통행패턴이 상이한 시계열 요인을 고려하여 맞춤형 서비스를 제공하고 향후 스마트시티 구현을 위한 MaaS에 통행량 예측 알고리즘 등에 적용할 수 있다. 또한 버스 및 지하철 노선도, 배차간격 등과 같은 대중교통 운영을 위한 대중교통 수요예측 알고리즘 개발 연구에 초석이 되는 자료로 활용될 것이다.

2. 연구의 범위 및 방법

본 연구에서는 2018 수도권 지역의 1년치 대중교통 이력데이터를 분석하여 월별, 요일별, 시간대별과 같은 시계열적 특성에 따른 대중교통 통행패턴의 차이를 비교·분석하고자 하였다. 특히 대중교통 수단선택 및 개별 통행자의 승·하차시간 및 이용시간/거리, 환승횟수와 같은 통행특성에 영향을 미치는 시계열적 특성을 파악하고자 하였다. 본 연구의 과정은 아래와 같다.

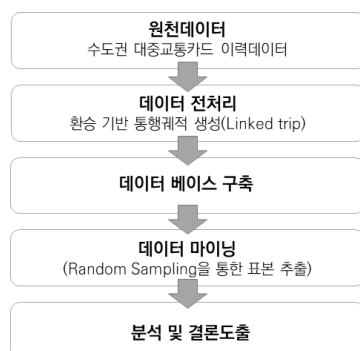


그림 1 연구의 과정

II. 본론

1. 데이터 전처리 과정

본 연구에서는 2018년에 수집된 수도권지역 교통카드 이력데이터를 활용하였다. 수도권 외의 많은 비수도권의 지역에서 수집되는 교통카드 이력데이터는 하차정보가 누락되어 활용에 제약이 있다. 수도권의 경우, 현재 수도권 통합환승할인제도가 시행되고 있어 대부분의 통행자가 승하차 시에 교통카드 태그(tag)를 하기에 하차정보의 누락이 비수도권 데이터에 비해 현저히 적어 활용도 및 신뢰성이 높다. 또한 본 연구에서 활용된 데이터는 선행연구와 같이 일주일 미만의 단기간 동안 수집된 것이 아닌 2018년 1년 동안 수집된 데이터이다. 교통카드 이력데이터는 통행자가 카드를 태그할 시에 기록되는 단순 이력데이터로 전처리 과정이 필요하다. 본 연구에서는 고유 식별화된 개별 ID를 기준으로 최대 5번의 환승을 고려한 통행패턴을 생성하였다. 그 후 각 통행패턴에 기종점 정류장(혹은 역) 정보를 매칭하였고, 이를 통해 생성된 Linked trip에 승차일자 및 승·하차 시간과 같은 시계열 정보를 매칭하였다. 또한 승차일시와 하차일시를 이용해 수단 이용시간을 구했으며, 탑승 및 하차 정류장 정보를 통해 이동거리를 계산해 수단이용거리를 도출하여 해당 데이터에 추가하였다. 데이터 전처리 과정을 통해 개별 식별화된 ID, 탑승 및 하차 일시, 탑승 및 하차 정류장 정보, 환승정보, 노선정보, 수단 이용시간, 수단이용거리 등으로 이루어진 데이터 베이스를 구축하였다.

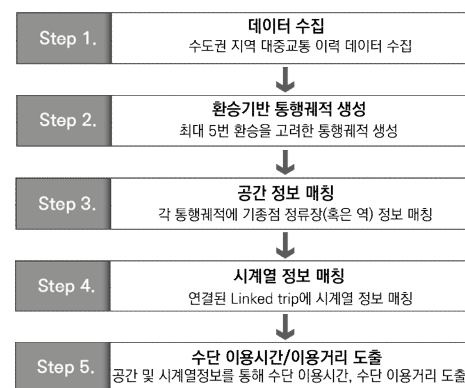


그림 2 데이터 전처리 과정

2. 데이터 마이닝

2018 수도권 교통카드 이력데이터는 2018년 1월 1일부터 2018년 12월 31일까지 수집된 데이터이며 수도권지역의 대중교통을 이용하는 통행자의 전수데이터로 볼 수 있다. 이는 데이터의 크기가 매우 크기에 해당 데이터로 분석을 돌리는 것은 한계가 있다. 따라서 본 연구에서는 random sampling을 통해 데이터 마이닝을 하는 방식을 채택하였다. 365일 데이터의 0.01%를 비복원 random sampling을 하여 표본을 추출하였다. 이에 n번의 동질성 검증을 거쳐 최종 표본을 추출하였다.

3. 분석

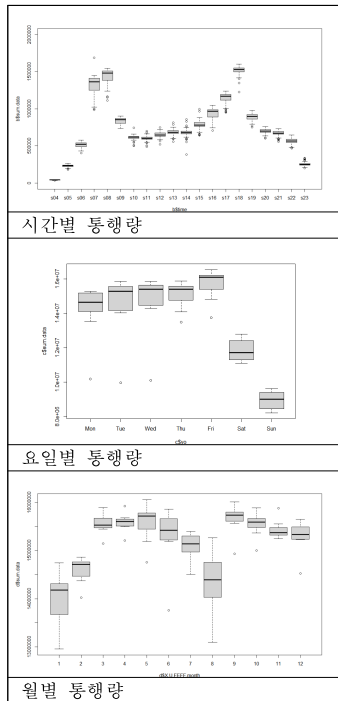


표 1. 시계열적 특성별 통행패턴

시계열적 특성에 따른 대중교통 통행 패턴을 분석하기 위해 월별, 시간별, 요일별로 하루 총 통행량에 대한 박스그래프를 작성하여 패턴을 분석하였다. 시간대별 분석결과 출퇴근 시간대와 일치하게 통행량이 급격히 증가하는 것으로 나타났다. 요일별로는 요일 특성 상 금요일에 통행량이 가장 많은 반면 주말, 특히 일요일의 통행량이 급격히 감소하는 것으로 나타났다. 마지막 월별 통행패턴 분석결과 1월과 8월에 통행량이 현저히 낮아지는 것으로 나타났으며 동시에 표준편차의 범위도 넓은 것으로 나타났다. 이는 기상 및 기후 변화가 상이한 계절별 특성의 영향을 받은 것으로 예상할 수 있다. 따라서 각 시계열적 요소들이 모두 통행량에 유의한 영향을 미치는 것으로 나타났다. 통행량뿐만 아니라 시계열적 특성이 개별 통행자의 버스나 지하철 수단선택에 미치는 영향과 환승횟수, 승·하차 일시, 수단 이용시간 및 이용 거리에 미치는 영향을 살펴보았다. 또한 각 시계열적 특성별, 통행특성별 일원배치 분산분석(ANOVA 분석)을 진행하였다.

data

	제곱합	자유도	평균제곱	F	유의확률
집단-간	2.616E+13	11	2.378E+12	15.473	.000
집단-내	1.353E+13	88	1.537E+11		
전체	3.969E+13	99			

그림 3 월별 ANOVA분석 결과

data

	제곱합	자유도	평균제곱	F	유의확률
집단-간	5.025E+14	6	8.375E+13	94.699	.000
집단-내	1.353E+14	153	8.844E+11		
전체	6.378E+14	159			

그림 4 요일별 ANOVA분석 결과

time.data

	제곱합	자유도	평균제곱	F	유의확률
집단-간	6.582E+13	19	3.464E+12	1287.503	.000
집단-내	1.184E+12	440	2690456337		
전체	6.700E+13	459			

그림 5 시간별 ANOVA분석 결과

III. 결론

4. 결론 및 향후 과제

2018년 1년 동안의 수도권 교통카드 이력데이터를 기반으로 빈도별, 거리별, 시간별로 통행특성을 분석한 결과, 시계열 변화에 민감한 대중교통 특성을 파악할 수 있었다. 통행량은 시간대별로는 출퇴근 시간대와 동일하게 통행량이 증가하였으며, 요일별 특성상 금요일에 가장 많은 통행량이 나타나며 주말에는 통행량이 현저히 낮아졌다. 또한 개별 통행자의 환승횟수, 수단 이용 시간 및 이용 거리도 시계열 변화에 따라 유의하게 나타남을 알 수 있었다. 월별 통행패턴은 1월과 8월에 통행량이 현저히 낮으며 표준편차의 범위도 넓은 것으로 나타났다. 따라서 통행패턴이 상이한 월별, 요일별, 시간별로 해당 시계열적 특성을 고려한 맞춤형 서비스를 제공해야 함을 사료한다. 또한 기상 및 기후, 공간 정보 등 이중 데이터와 연계한다면 대중교통 통행특성의 정확성을 높일 수 있을 것이다. 이러한 연구의 결과를 통해 향후 MaaS와 같은 다 수단 연계 모빌리티 서비스 혹은 맞춤형 서비스를 제공하기 위한 기초연구로 통행량 예측 알고리즘 등에 적용될 수 있으며, 대중교통 수요를 예측해 버스 및 지하철 노선도, 배차간격 등을 조정하는 교통운영을 위한 연구에 초석이 되는 자료로 활용될 것이다.

참고문헌

- Shi, Z et al. Analysis of the Temporal Characteristics of the Elderly Traveling by Bus Using Smart Card Data. ISPRS Int. J. Geo-Inf. 2020, 9, 751.
- Kohei Hirako et al. Estimations of Bus Stop Territories using Reachable Area Analysis Focusing on Travel Behavior of Elderly to Medical Facilities, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 06 (June 2020),
- 김경태, 이인목, 민재홍. "교통카드 자료를 활용한 버스 OD 추정 방법론 연구." 한국철도학회논문집 21.11 (2018): 1155-1161.
- 김경태(Kyoungtae Kim), 이인목(Inmook Lee). 교통카드 데이터를 활용한 하차정류장 추정 방법론 연구. 한국철도학회논문집 20.5 (2017): 692-702.
- 박준환, 김순관, 조종석, 허민욱. 대중교통 OD구축을 위한 대중교통카드 데이터의 오류와 결측 분석 및 보정에 관한 연구. 大韓交通學會誌 26.2 (2008): 109-119.
- 신강원. 통행사슬 구조를 이용한 교통카드 이용자의 대중교통 통행중점 추정. 대한교통학회지, (2016): 437-448
- 유봉석, 추상호. 의사결정 학습 모델 기반 교통카드 데이터 하차 정류장 추정 모델 연구. 한국ITS학회논문지 18.6 (2019): 11-30.
- 이석주, 장동익. (2015). 교통카드 이용실적 자료 기반의 대중교통 이용자 시공간 패턴 분석 및 시뮬레이션 모형 구축. 한국교통연구원 기본연구보고서, 1-113.
- 김순관. 교통카드 데이터를 활용한 OD 추정 및 활용. 서울시 정개발연구원, 2007.