

Decision Tree 모델을 이용한 Indoor Traffic 분석

김영준, 유현민, 조영준, 이상연, 홍인기

경희대학교

donjomyo, yhm1620, rkdtmd34, sangyeon, ekhong@khu.ac.kr

Indoor Traffic Analysis Using Decision Tree Model

Young-Jun Kim, Hyun-Min Yoo, Young-Jun Cho, Sang-Yeon Lee, Een-Kee Hong

Kyung Hee Univestiy

요약

모바일 트래픽 사용량이 해가 갈수록 증가하면서 무선통신 네트워크에서 이들을 수용할 수 있도록 효율적인 무선 자원 배분 연구가 활발해지고 있다. 본 논문에서는 모바일 트래픽의 70~80%를 차지하는 실내 환경에서 발생하는 모바일 트래픽을 분석하였다. 실내 환경에서 발생하는 모바일 트래픽이 다른 지역과는 구분되는 특징이 존재함을 증명하고 특징이 존재하는 원인에 대한 분석을 하였다.

I. 서론

이동통신 서비스는 세대 변화를 거쳐가며 다양한 서비스 제공을 통해 가입자와 이동통신 트래픽 사용량을 급증시켰다. 트래픽 사용량 증가의 원인으로서는 스마트폰 장치 기능 향상, 비디오 스트리밍과 같은 데이터 집약적 콘텐츠의 증가, 배포된 네트워크 성능의 지속적인 개선으로 뽑을 수 있다. Ericsson은 글로벌 모바일 데이터 트래픽을 2021년 말까지 월 약 67 Exabytes(EB)에 도달했으며, 2027년에는 약 4.2배 증가하여 월 282 EB에 이를 것으로 예상한다 [1]. 2027년까지의 트래픽 증가에는 AR을 포함한 XR 유형 서비스의 초기 도입이 영향을 미칠 것으로 내다보았다. 그러나 폭발적으로 증가하는 모바일 데이터 트래픽을 수용하기 위해 추가적으로 통신 인프라를 설치하는 데에는 막대한 비용이 든다. 제한된 무선 자원 내에서 모바일 데이터 트래픽을 효율적으로 분배하기 위해선 기지국단위에서 미리 모바일 트래픽 사용량 예측을 기반으로 자원 운용이 되어야 할 것이다.

Ericsson은 모바일 데이터의 70~80%가 실내에서 생성된다고 추정하고 있다 [2]. 추정치로 미루어 보았을 때, 현재 무선 통신의 대부분은 실외의 기지국에서 실내로 서비스되고 있다. 실외 기지국에서 실내로 전파를 송신할 경우, 경로 손실에 의한 전력 감쇠[3]가 심하게 일어나기 때문에 네트워크 운영 측면에서 실내에서 발생하는 모바일 트래픽에 대한 분석이 이루어져야 한다.

본 논문에서는 실내에서 발생하는 모바일 트래픽(Indoor Traffic) 데이터를 Exploratory Data Analysis (EDA)하고 데이터 특징을 분석하였다. Decision Tree [4] 라는 지도학습 모델을 이용해 실내에서 발생하는 모바일 트래픽이 다른 지역의 모바일 트래픽과 분류할 수 있음을 확인하였다. 이는 실내 트래픽이 가지는 고유한 특징에 의해 분류되는 것으로 Decision Tree 모델의 학습 과정에서 특징을 추출하였다.

II. 본론



그림 1. 강남역 반경 1km내 지도

본 논문에서는 실내에서 발생하는 모바일 트래픽의 특징을 분석하기 위해, 그림 1과 같이 실제 강남역 1km 반경내 위치한 셀들에서 발생한 모바일 트래픽의 Uplink 데이터를 사용하였다. 데이터 세트는 2020년 8월 한 달(31일)간 발생한 데이터 트래픽으로 1시간 단위로 집계되었다. 그리고 셀들이 위치한 지역 용도에 따라 데이터를 라벨링하였다. 기지국 셀은 총 136개로 주거지역 46곳, 상업지역 64곳, 기타지역(학교, 공원 등) 18곳, 아파트 지역 8곳에 위치했다. 아파트 지역에서 수집된 모바일 트래픽을 Indoor Traffic으로 가정하였다.

Indoor Traffic이 다른 지역과 차별되는 특징이 존재함을 증명하기 위해 Decision Tree 모델을 이용해 분류하였다. Decision Tree 모델은 여러 가지 규칙을 순차적으로 적용하면서 독립 변수 공간을 분할하는 분류 모형이다. 분류 정확도를 올리기 위해 Decision Tree 모델에서 파생된 모델들로는 Random Forest, Xgboost모델 등이 있다. 파생된 모델들보다는 Decision Tree 모델의 분류 성능은 떨어지나 해석력이 뛰어나다는 장점이 있다. 주로 분류와 회귀 분석에 사용되는 지도학습 모델이다.

Decision Tree 모델의 학습 과정은 그림 2와 같다. 31일(744시간)의 모바일 트래픽 데이터와 장소를 나타내는 인덱스를 하나의 데이터 세트로 학습한다. 장소를 나타내는 인덱스는 True or False값으로 매핑해준다. 테스트 세트의 모바일 트래픽을 보고 장소 인덱스를 맞출 수 있는지 확인한다. 본 논문에서는 장소 인덱스로 주거지역, 상업지역, 아파트 지역을 매핑

하고 분류 정확도 결과를 비교하였다. 그 결과, 아파트 지역 분류 정확도는 85.7%, 주거지역의 분류 정확도 50%, 상업지역 분류 정확도 66.4%가 도출되었다. 분류 시뮬레이션은 각각 지역별로 20번씩 실행한 결과의 평균값이다.

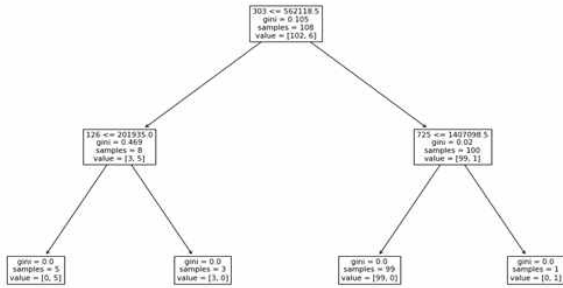


그림 2. Decision Tree 모델의 학습 과정

아파트 지역에서 수집된 모바일 트래픽, 즉, Indoor Traffic이 다른 지역에서 수집된 모바일 트래픽과는 차별되는 특징이 존재함을 확인하였다. Indoor Traffic의 고유한 특징을 분석하기 위해 Decision Tree 모델의 학습 과정에서 분류 기준점이 되는 인덱스를 확인하였다. 학습 데이터 세트를 랜덤하게 바꾸고 Decision Tree 모델의 파라미터를 바꿔가며 자주 나타나는 인덱스들을 분석하였다.

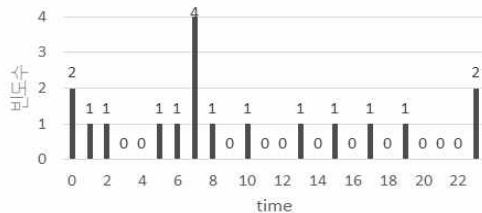


그림 3. 시간대별 기준점 통계

분류 기준점으로 선정되는 빈도가 높은 인덱스는 총 18개 발견하였다. 그림 3은 시간대별 기준점 통계 그래프다. 기준점으로 자주 나타나는 시간대는 23~02시, 05~08시에 나타났다.

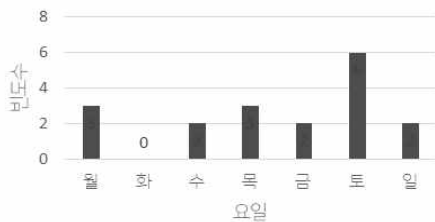


그림 4. 요일별 기준점 통계

그림 4는 요일별로 기준점 통계 그래프를 그렸다. 기준점으로 가장 많이 나타나는 요일은 토요일이었다.

III. 결론

본 논문에서는 Indoor Traffic이 다른 지역의 모바일 트래픽과는 차별되는 특징이 존재함을 증명하고 그 특징에 대해 분석하였다. 아파트 지역에서 수집된 모바일 트래픽이 전부 실내에서 발생한 트래픽이라 가정하고 Indoor Traffic으로 분석하였다. Decision Tree 모델로 지역 용도를 인덱싱하고 분류를 진행한 결과, Indoor Traffic이 다른 지역과는 구분되는 고유한 특징을 가짐을 증명했다.

Decision Tree 모델의 학습 과정을 시각화하여 분류 기준점이 되는 시간대는 23~02시, 05~08시, 요일로는 토요일이 가장 많이 나타났다. 무선 자원 운용 관점에서는 Indoor Traffic의 특성을 이용한 트래픽 예측을 통해

효율적인 무선 자원 관리 방식이 개발되어야 할 것이다. 또한, 실외기지역에서 실내 환경으로 무선 통신 서비스를 제공하면서 발생하는 강한 신호 감쇠 효과를 상쇄할만한 기술 개발도 이루어져야 할 것이다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2018-0-01659, 5G NR 기반 지능형 오픈 스몰셀 기술 개발)

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음" (IITP-2021-0-02046*)

참 고 문 헌

- [1] Ericsson, 'Mobile data traffic outlook', 2021
- [2] Ericsson, "Planning in-building coverage for 5G: from rules of thumb to statistics and AI", 2021
- [3] Graziosi F., Santucci F., Di Claudio V. and Mecella F., 'Radio link performance and traffic capacity of indoor wireless systems integrated with outdoor cellular,' IEEE VTS Fall VTC2000.52nd Vehicular Technology Conference, 2000. pp. 1100-1105 vol.3
- [4] Safavian, S.R., Landgrebe, D., 'A survey of decision tree classifier methodology,' in IEEE Transactions on Systems, Man, and Cybernetics, vol. 21, no. 3, pp. 660-674, May-June 1991