

도시 스케일의 교통 흐름 시뮬레이션을 위한 궤적 데이터 시각화

최남식^{1,2*} · Athita Onuean^{1,2} · 정한민^{1,2}

¹한국과학기술연합대학원대학교 · ²한국과학기술정보연구원

On Visualization of Trajectory Data for Traffic Flow Simulation of Urban-scale

Namshik Choi^{1,2*} · Athita Onuean^{1,2} · Hanmin Jung^{1,2}

¹University of Science and Technology · ²Korea Institute of Science and Technology Information

E-mail : munr3@kisti.re.kr / athita@kisti.re.kr / jhm@kisti.re.kr

요 약

교통량이 증가하고 도로 네트워크가 복잡해짐에 따라 정확한 교통 흐름 파악을 통해 교통의 원활한 흐름을 유도하는 것은 많은 국가의 관심사항이다. 교통 흐름을 효과적으로 알기 위한 다양한 분석 기술 및 연구들이 있어 왔지만 위치(GPS) 데이터를 포함한 데이터 시각화를 통해 먼저 교통 흐름의 패턴을 찾는 것이 필요하다. 본 논문에서는 실제 도시의 교통 궤적을 시뮬레이션한 내용을 도구로 사용함으로써 교통 흐름의 패턴을 시각화하는 것을 목표로 한다. 이에 24시간운행 되어 지고 정해진 경로가 없는 특징을 가진 실제 택시 40대에 센서 모듈을 설치하여 IoV(Internet of Vehicle)데이터를 수집하고 이 데이터를 이용하여 전처리 과정을 거친 후 오픈소스 기반의 데이터 시각화 도구를 우리의 데이터 특성에 적합하도록 개선하였다. 해당 시각화 모델은 시간 흐름에 따른 차량 트랙킹 Dot을 통해 차량 밀집 지역과 이동 경로 패턴 인식이 가능하므로 도시 내에서 또는 도시와 도시간의 교통 흐름 파악을 통해 도시 환경 문제 개선에 기여할 것으로 기대된다.

ABSTRACT

As traffic volume increases and road networks become more complicated, identifying for accurate traffic flow and driving smooth traffic flow are a concern of many countries. There are various analytical techniques and studies which desire to study about effective traffic flow. However, the necessary activity is finding the traffic flow pattern through data visualization including location information. In this paper aim to study a real-world urban traffic trajectory and visualize a pattern of traffic flow with a simulation tool. Our experiment is installing the sensor module in 40 taxis and our dataset is generated along 24 hours and unscheduled routes. After pre-processing data, we improved an open source traffic visualize tools to suitable for our experiment. Then we simulate our vehicle trajectory data with a dots animation over a period of time, which allows clearly view a traffic flow simulation and a understand the direction of movement of the vehicle or route pattern. In addition we further propose some novel timelines to show spatial-temporal features to improve an urban environment due to the traffic flow.

키워드

Traffic flow, Data visualization, IoV(Internet of Vehicle), Urban Environment

1. 서 론

최근 예측에 따르면 2020년까지 1억 2천만 개의 “사물”이 인터넷에 연결될 것이고 이중 차량이 상

당 부분을 구성할 것이라고 한다. 자동차-IT 융합에서는 이를 IoV(internet of vehicles)라 하고 중요한 키워드 중의 하나로 떠오르고 있다. 특히 차량에 설치된 각종 센서와 외부의 인터넷 무선통신을 통해 수집된 데이터는 도시환경(교통 문제, 공기질 문제 등) 문제를 해결하는데 중요한 정보가

* corresponding author

될 것이다.

특히 교통량이 증가하고 도로 네트워크가 복잡해짐에 따라 정확한 교통 흐름 파악을 통해 교통을 제어하여 원활한 흐름을 유도하는 것은 많은 국가의 관심사항이다.

과거부터 현재까지 교통 흐름을 효과적으로 알기 위한 다양한 분석 기술 및 연구들이 진행되어 왔다. 특히 이전의 연구에서는 수학적, 통계적 기법을 통해 교통 흐름을 알고자 하는 연구가 많이 진행되었다[1][2].

도시 도로망의 복잡성을 수학적으로 계산하여 교통 흐름을 분석하거나 도로망에서 루프 탐지 센서에서 수집한 데이터를 기반으로 교통 흐름을 예측하는 연구도 있다[3]. 하지만 교통 흐름은 여러 가지 요소에 영향을 받기 때문에 불규칙하고 무작위로 나타난다. 따라서 모델이 구축되더라도 도시마다의 특성을 반영하지 못한다거나 루프 탐지를 통한 데이터는 여러 예외 상황이 발생할 확률이 높아 흐름 파악에 효과적이지 않다. 이에 본 논문에서는 직접 차량에서 수집한 위치 데이터를 지도위에 시각화함으로써 도시 전체의 교통 흐름을 알고 위치(GPS) 데이터만으로는 알 수 없었던 교통 패턴을 시각화를 통해 찾는 것을 목표로 한다. 또 최근 데이터 시각화의 추세에 따라 인터랙티브하고 시각화 공유를 위한 웹 환경에서의 시각화를 구현하고자 한다.

본 논문의 구성은 다음과 같다. 본문에서는 시각화 과정의 전체적인 구성에 대해 설명하고 실험 및 결과에서는 각 구성의 세부적인 구현에 대한 설명과 시각화 결과에 대해 설명한다. 마지막으로 본 연구의 결론 및 향후 연구 계획을 제시한다.

II. 본 론

그림1은 데이터 시각화의 전체 과정 및 처리 흐름을 나타내며 과정은 크게 데이터 전처리와 시각화 두 단계로 분리할 수 있다. 먼저 데이터베이스에서 특정 필드의 데이터를 추출하는 과정을 거친다. 추출된 데이터는 클리닝 과정을 거치면서 노이즈 데이터를 삭제하게 되고 인코딩 과정을 거쳐 시각화에 필요한 데이터 셋이 준비된다. 마지막으로 이를 이용하여 시각화 모델을 구현하게 된다.

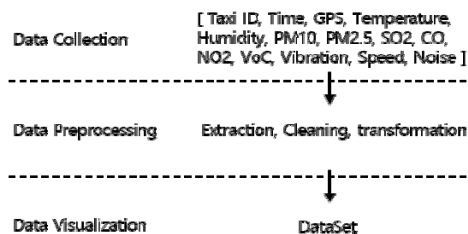


그림 1. Procedure for Visualization

2.1 데이터 전처리

데이터 시각화를 위해서는 그에 적합한 데이터가 필요하며 주어진 원 데이터를 그대로 사용하기보다는 목적에 적합하게 데이터를 가공하는 작업이 이루어지는데 이를 ‘데이터 전처리’라고 한다.

본 논문의 데이터 전처리에는 데이터를 추출(Extraction), 클리닝(Cleaning), 변형(Transformation) 등의 작업이 포함된다.

추출(Extraction) 과정에서는 데이터베이스의 전체 속성에서 시각화에 사용되어지는 필드[택시ID, 시간, GPS, 속도]만 추출되어지고 클리닝(Cleaning) 과정에서는 동적 시각화의 목적에 맞게 일정 시간 이상 정체 되어 있는 위치(GPS) 데이터에 대해서는 노이즈로 판단하여 클리닝 하게 된다. 또 시각화의 타이머 속도를 높이기 위해 10초마다 수집된 데이터의 interval을 1분으로 바꾸기 위해 클리닝이 이루어졌다. 변형(Transformation) 과정에서는 기존의 오픈 소스를 개선하여 시각화를 진행하는 만큼 오픈 소스 알고리즘에 부합하는 위치(GPS) 데이터 인코딩(ASCII) 과정을 거치게 된다.

2.2 데이터 시각화

데이터 시각화는 데이터를 쉽게 이해 할 수 있도록 시각적으로 표현하고 전달되는 과정을 말한다. 특히 위치 데이터를 기반으로 한 시각화의 경우 시각화를 통해 시간에 따른 위치의 이동 패턴을 알 수 있어 중요하다.

본 논문은 기존 연구를 데이터의 특성과 시각화 목적에 맞도록 개선하여 제안된 연구이다. 이전 시각화(Holiday Taxi Traffic Trends-NYC Airports)의 경우 주말 동안 출발지(NYC 공항)에서 목적지까지 택시 흐름에 대한 패턴을 보는 것이 목적이다[4]. 하지만 이 경우 도시 전체의 교통 흐름을 알기에는 위치 데이터 범위가 좁은 한계를 가지며 Google의 Direction API가 Output으로 주는 중간지점 데이터를 시각화하였기 때문에 실제 택시가 해당 경로로 운행되었는지 불확실하다.

이를 보완하기 위해 본 연구에서 사용된 데이터는 차량이 운행되는 모든 시간동안 수집된 데이터를 이용하여 시각화함으로써 넓은 범위의 도시 교통 흐름을 알 수 있는 장점이 있다.

III. 실험 및 결과

3.1 데이터

본 논문에서 사용한 데이터는 2017년 6월부터 ~18년 3월까지 9개월 동안 대구시 택시 40대에서 수집한 데이터이다. 데이터는 택시갯등에 설치된 10개의 센서로 부터 수집된 데이터이고 총 25,149,069개의 레코드가 있다. 데이터의 필드 구조는 다음과 같다. [택시ID, 시간, 위도, 경도, 온습도, PM10, PM2.5, ACC, 속도, 소음, SO2, NO2, VoC]. 데이터 수집 전체 아키텍처는 그림2 이고 수집 interval은 10초 간격으로 차량의 LTE 통신을 통해

서버에 전송된다.

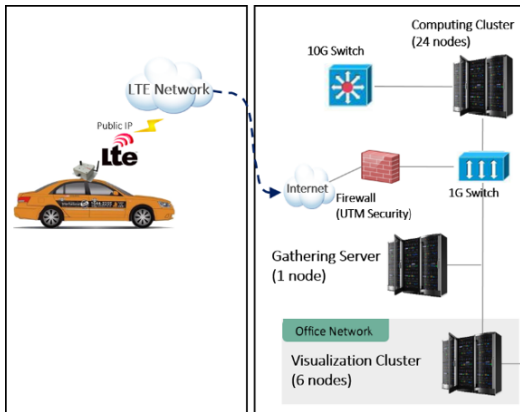


그림 2. Taxi-based vehicle sensor networks architecture

3.2 데이터 전처리

시각화에 유효한 데이터 취합을 위해 데이터 추출, 클린징, 변경의 순으로 전처리 과정을 수행한다. 데이터 추출은 전체 DB에서 각 택시 ID를 기준으로 시각화에 필요한 속성(택시ID, 시간, GPS, 속도)만을 가지는 csv파일을 생성하였다. 다음으로 데이터 클린징은 그림3과 같이 택시가 2분 이상 10m이상의 이동이 없는 경우 정차된 경우로 판단하여 클린징 하였다. 기존의 설정 근거는 신호등의 평균 대기시간이 2분정도 소요되는 것과 위치(GPS) 데이터의 오차율(10~15m)을 감안하여 설정하게 되었다.

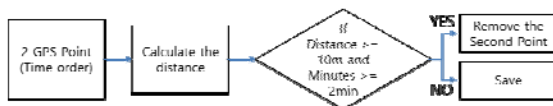


그림 3. Data cleaning Procedure-I

두 번째 데이터 클린징은 그림4와 같이 데이터의 시간 간격을 1분으로 바꾸기 위해 해당 분에서 가장 처음 발생된 데이터만 저장하고 나머지 데이터는 클린징 하였다. 클린징 과정을 통해 총 25,149,069개의 레코드에서 데이터 크기를 1/4 크기로 줄일 수 있었다.

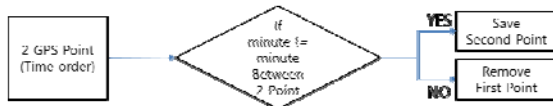


그림 4. Data cleaning Procedure-II

데이터 변경은 기존의 연구에서 사용된 Google Direction API를 사용하기 위해서는 데이터 량이 많은 경우 상당한 시간과 비용이 소모 된다. 이에 본 연구에서는 MapBox라는 OSM(open street map)

오픈 소스 맵핑 서비스에서 제공하는 라이브러리를 바탕으로 위도, 경도를 base64 인코딩 체계로 이진 값을 아스키코드로 변환한 값을 사용하여 시각화 오픈소스의 알고리즘 수정을 최소화 할 수 있었다.

3.3 데이터 시각화

시각화는 JavaScript로 구현된 오픈 소스[4][5]를 기반으로 우리의 데이터 특성에 맞게 수정하여 구현 하였다.

시각화의 기본 바탕이 되는 대구시 지도는 OSM (Open Street Map)기반의 Mapbox에서 제공하는 지도를 사용하였고 그 지도위에 데이터 맵핑을 위해 오픈 소스 JavaScript 라이브러라인 leaflet을 사용한다. 특히 택시의 움직임을 동적으로 표현하기 위해 데이터와 웹페이지 이미지를 바인딩 해주는 D3.js 라이브러리도 사용한다.

택시가 이동한 트래킹 Dot을 나타내기 위해 D3 라이브러리를 사용하여 택시가 지도상에서 움직이는 듯한 애니메이션적인 효과 부분을 추가하였고, 트래킹 Dot이 10초(데이터 상 6시간 정도)가 지난 후 사라지는 효과를 위해 관련 파라미터를 수정하였다.

또 40대 택시 전체가 아닌 개별 택시의 이동 흐름을 시각화하기 위해 특정 택시를 활성화, 비활성 할 수 있는 인터페이스를 추가하였다.

3.4 궤적 데이터의 시각적 패턴

그림5는 데이터 시각화의 순간 스냅 샷을 보여준다. 택시가 지나간 경로에는 트래킹을 위한 노란색 Dot을 남기게 되는데 이 Dot은 10초간 유지된다. 10초는 데이터 시간으로 약 6시간정도 유지하게 된다. 노란색 포인트를 통해 택시의 밀집 운영 지역을 유추 할 수 있는데 출퇴근(AM 05:00~AM 10:00) 시간대에는 대구역과 중구 지역에 많은 택시 Dot들이 밀집되어 있는 것을 알 수 있다. 그리고 PM 10:00 ~ AM 03:00과 같이 늦은 저녁부터 새벽 시간에는 택시가 대구시내로부터 대구시 외곽지역 또는 경산시 같은 거주지 위주의 지역으로 이동했다가 다시 대구 시내로 돌아오는 패턴을 보였다.

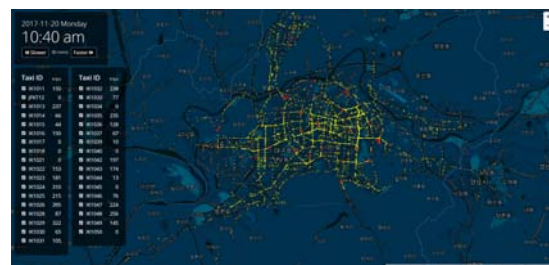


그림 5. Trajectory data Visualization snapshot

IV. 결론 및 향후과제

본 논문에서는 실제 IoV(internet of vehicles) 환경에서 9개월 동안 수집한 대구 택시 40대의 위치(GPS) 데이터를 Time Series에 따라 동적으로 시각화하였다. 그 결과 도시의 교통 흐름을 시각적으로 인지 할 수 있었고 만약 데이터의 범위가 확장되는 경우 도시와 도시간의 교통 흐름도 알 수 있을 것이다. 또 부수적으로 대구시의 경제 거점 지역을 유추 할 수 있고 택시 운행이 빈곤한 지역의 경우 교통 시스템 개선에 도움이 될 것으로 기대된다. 향후 연구로는 본 연구의 교통 흐름 파악을 위한 시각화를 확장하여 교차로의 교통 신호정보 및 날씨 정보 등 교통 정체와 관련 데이터를 추가하여 도심 교통 정체에 대한 데이터 분석 연구가 필요하다.

References

- [1] Huang, L., Zhu, X., Ye, X., Guo, W., & Wang, J. . "Characterizing street hierarchies through network analysis and large-scale taxi traffic flow: a case study of Wuhan, China." *Environment and Planning B: Planning and Design*, 43(2), pp. 276-296, Jan, 2016
- [2] Kong, X., Xu, Z., Shen, G., Wang, J., Yang, Q., & Zhang, B. "Urban traffic congestion estimation and prediction based on floating car trajectory data." *Future Generation Computer Systems*, 61, pp. 97-107, Dec, 2016
- [3] Tang, J., Liu, F., Zhang, W., Zhang, S., & Wang, Y, "Exploring dynamic property of traffic flow time series in multi-states based on complex networks: Phase space reconstruction versus visibility graph." *Physica A: Statistical Mechanics and its Applications*, 450, pp. 635-648, Jan, 2016
- [4] Holiday Taxi Traffic Trends - NYC Airports[Internet]. Available : <http://github.com/dheerosaur/nyc-holiday-taxi-viz>
- [5] NYC Taxis: A Day In the Life : <http://github.com/chriswhong/nyctaxi>