

IoV 데이터와 도로 분할 알고리즘을 이용한 택시 정차위치 파악

임동진^{1,2*} · Athita Onueam^{1,2} · 정한민²

¹과학기술연합대학원대학교 · ²한국과학기술정보연구원

Finding Stop Position of Taxis using IoV data and road segment algorithm

Dong-jin Lim^{1,2*} · Athita Onueam^{1,2} · Han-min Jung²

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

E-mail : glow147@kisti.re.kr / athita@kisti.re.kr / jhm@kisti.re.kr

요 약

손님을 태우기 위해 도로에서 불법으로 정차하는 택시들은 교통체증을 유발하고 때때로 교통사고의 원인이 되기도 한다. 택시들의 정차 위치는 택시 기사들의 오랜 경험에 의해 정해지고 있다. 이번 연구에서는 시간대별 정차 위치를 파악해 택시기사들과 지역을 처음 방문하는 손님들에게 정보를 제공하고자 한다. 이를 위해 택시 40대에 설치된 센서에서 수집되는 Internet of Vehicle(IoV) 데이터를 이용하였다. 기존의 연구들은 택시를 중심으로 군집을 형성하였다. 이 방법은 택시의 위치에 따라 군집의 위치가 변한다. 또한 택시가 범위 내에 일정 대수 이상이 있어야 군집이 형성되기 때문에 실시간으로 정차위치를 파악하기가 어렵다. 이번 연구에서는 이러한 문제를 해결하기 위해 도로 분할 알고리즘을 사용했다. 기존 연구들과 달리 도로를 중심으로 군집이 형성되기 때문에 군집의 위치가 고정적이며 택시 대수에 영향을 받지 않기 때문에 실시간으로 정차 위치를 파악하는 것이 가능하다. 도로 분할은 30M 단위로 이루어져 있으며, 시간대별, 평일, 주말로 분할된 택시 위치 데이터를 가장 가까운 포인트에 매핑하였다. 매핑결과 주말의 경우 운행하는 택시 수가 적어 시간대 별로 큰 차이를 보기 어려웠으나 평일의 경우 출퇴근 시간대와 심야 시간대간의 정차 위치 차이를 확인할 수 있었다. 이 연구결과를 통해 택시 불법 주정차 방지와 택시 승강장 설치위치 기준을 제안할 수 있을 것으로 기대한다.

ABSTRACT

Taxis that are illegally parked on the road to catch customer can cause traffic congestion and sometimes cause traffic accidents. Stop position of taxis is determined by the long term experience of taxi drivers. In this study, We provide information to taxi drivers and customer who visit in first time through finding stop position of taxis by time. To do this, we used the Internet of Vehicle (IoV) data collected from sensors installed in 40 taxis. Previous studies attempted by forming a cluster around a taxi. Since this method is centered on a taxi, the position of the cluster changes depending on the location of the taxi. In this study, we use a road segmentation algorithm to solve these problems. Unlike the previous studies, since the cluster is formed around the road, the position of the cluster is fixed and it is not affected by the number of taxis, so it is possible to grasp the stop position in real time. The road segmentation is made up of 30m units, and map the taxi location data divided into hourly, weekday, and weekend to the nearest point. As a result of the mapping, it was difficult to see a big difference in the time of week because there were few taxis to operate on weekends, but in case of weekdays, the difference of stop position between the commute time zone and the night time zone was confirmed. The results of this study suggest that it will be possible to propose the prevention of taxi illegally driving taxi and the location of the taxi stand.

키워드

Taxi, Internet of Vehicle data, Road segment, Clustering

* corresponding author

I. 서론

현행 운수사업법에 따르면 지자체가 관할 지방경찰청과 협의해 택시 승강장 설치 및 시설 기준을 정하도록 규정하고 있다. 하지만 의견차이로 인한 오락가락한 위치 선정과 영터리 관리로 시민들과 택시기사 모두에게 외면 받는 경우가 많다. 본 논문에서는 데이터 시각화를 통해 택시들이 자주 정차하는 위치를 파악하고 정보를 제공해 이 문제를 해결하고자 한다. 기존 연구에서는 택시의 주요 정차 위치를 파악하기 위해 먼저 지도 위에 정차 데이터를 표시한 후 택시들의 위치를 파악하거나[1] 택시들의 중심 위치를 찾아서 정차 위치를 파악하였다[2]. 하지만 택시 중심의 군집화는 택시의 위치의 따라 군집의 위치가 변하게 된다. 하지만 택시 승강장 설치 위치 제안을 위해서는 고정된 지점의 정차 여부가 중요하다. 이 문제를 해결하기 위해 이번 연구에서는 도로 분할 알고리즘을 이용하였다[3]. 도로 분할 알고리즘은 대구광역시의 주요 도로, 2차선 도로, 자동차 전용 도로, 주거용 도로를 일정 거리로 분할한 알고리즘으로 30M 단위로 분할이 이루어진 데이터를 사용하였다. 이번 연구에서는 그림 1처럼 30M 단위로 분할된 도로 포인트와 정차된 택시의 거리 계산을 하여 가장 가까운 거리의 포인트에 매핑을 했다. 택시 데이터는 시간대별, 평일, 주말로 분류된 데이터를 이용하였다.



그림 1. 대구광역시 도로 분할(30M)

먼저 2장에서는 택시 데이터와 전처리 과정에 대해 설명하고, 매핑 방법에 대해 자세하게 설명한다. 다음으로 3장에서는 시각화 결과에 대해 설명한다. 마지막으로 4장에서는 연구결과를 요약하고 향후 과제에 대해 설명하고자 한다.

II. 데이터

택시 데이터는 대구광역시의 40개 택시에 설치된 센서에서 10초 단위로 수집된다. 센서 데이터는 다음 표 1과 같이 대기질 오염 정보와 함께 GPS, SPEED 데이터가 수집된다. 이번 연구에서는 2017년 06월부터 2018년 3월까지 수집된 약 2437만개의 레코드를 사용하였으며, 그 중 GPS

정보와 SPEED 데이터만을 이용하여 연구를 진행하였다.

표 1. 센서 데이터 정보

Section	Example value	단위	Section	Example value	단위
시간	2017-04-27 07:13:31	반통용 시분초	NO2(이산화질소)	0.049	ppm
경도	128.60176	Degree	SO2(이황산화가스)	0.052	ppm
위도	35.87141	Degree	CO(일산화탄소)	0.35	ppm
온도	27	℃	VOC	1	ppm
습도	98	%	대기압	1013	hPa
PM2.5	27	µg/m³	소음	54	dBA
PM10	98	µg/m³	진동(차이로)	49.78.72	mg(xyz)

연구를 위해서는 정차된 데이터만 필요한데, 센서 데이터가 10초 마다 저장되기 때문에 정차의 기준이 필요하다. 정차 기준은 경찰청에서 제공하는 교통신호기 설치 관리 매뉴얼에 따라 다음 그림 2의 조건으로 신호가 2분 이내에 바뀐다고 가정, SPEED가 0인 구간이 2분 이상일 경우 정차한 것으로 정했다.

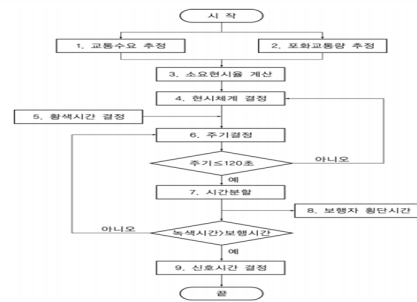


그림 2. 일반적인 신호시간 결정과정

데이터 전처리 결과 24368058개의 레코드 중 132883개의 레코드가 추출되었다. 이 데이터를 이용하여 분할된 도로에 매핑을 진행하였다. 30M로 분할된 도로와 마찬가지로 택시와 분할된 도로의 점 사이의 거리가 30M 이내이며, 다음 그림 3과 같이 겹치는 곳이 존재할 수도 있으므로 거리가 가장 가까운 곳으로 배치가 되도록 하였다.

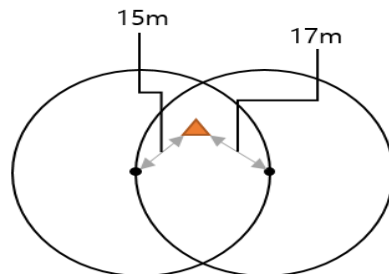


그림 3. 도로 매핑 예시

최종적으로 4873개의 레코드가 서울, 부산 등에 위치하여 대구광역시 도로에 매핑 되지 못하였다.

III. 결과

시각화 결과는 다음 그림 4와 같이 표현된다. 평일과 주말을 선택할 수 있으며, 시간과 최소 표현 택시 대수를 슬라이드로 정할 수 있다.



그림 4. 시각화 결과

다음 그림 5와 그림 6은 출근 시간대의 주요 정차 위치와 심야 시간대의 주요 정차 위치이다.



그림 5. 출근 시간

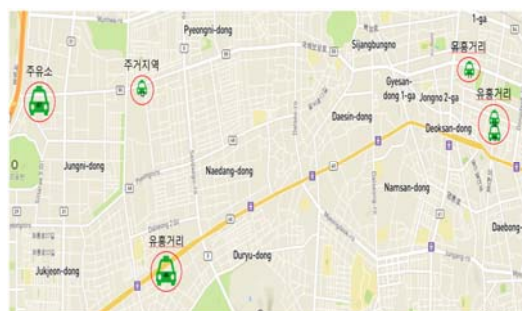


그림 6. 심야 시간

출근 시간대는 주로 지하철역과 버스 정류장

근처에 정차하는 경우가 많았으며, 심야 시간대에는 유흥거리 근처에 정차하는 경우가 대부분 이었다.

IV. 결론

본 논문은 택시의 주요 정차 위치를 파악해 택시 승강장의 설치 위치를 제안하며, 그동안 택시 기사들의 경험적으로 선택되었던 정차 위치를 정량적인 방법을 통해 정보를 제공하고자 했다. 이를 위해 2017년 6월부터 2018년 3월까지 약 9개월 동안 수집된 IoV 데이터를 이용하였으며, 도로 분할 알고리즘을 통해 도로를 중심으로 군집화하였다. 시각화 결과 출퇴근 시간대에는 지하철역과 버스 정류장 근처에 정차하는 경우가 많았으며, 낮 시간대에는 주로 대형마트 등이 있는 곳에 정차하는 경우가 많았다. 심야 시간대에는 유흥거리 근처에 정차하는 경우가 많았다. 이번 연구에서는 센서가 설치된 택시의 대수가 적어 9개월 동안의 데이터를 한 번에 표시해 몇 달에 거쳐 한 대의 택시가 동일 지점에 정차했다면 주요 정차 위치로 파악되는 문제가 있었다. 향후 센서가 설치된 택시의 대수가 늘어난다면 위의 문제를 해결할 수 있을 것으로 기대하며, 또한 향후 연구를 통해 정차 대수뿐만 아니라 평균 대기시간을 제공할 수 있을것으로 기대한다.

References

- [1] Lee, Junghoon. "Analysis on the waiting time of empty taxis for the taxi telematics system." Convergence and Hybrid Information Technology, 2008. ICCIT'08. Third International Conference on. Vol. 1. IEEE, 2008.
- [2] Hu, Xiaowei, Shi An, and Jian Wang. "Exploring urban taxi drivers' activity distribution based on GPS data." Mathematical Problems in Engineering 2014.
- [3] Onueam, Athita, et al. "Adaptive Road Segmentation of OpenStreetMap Data Using Haversine Formula." INTERNATIONAL CONFERENCE ON FUTURE INFORMATION & COMMUNICATION ENGINEERING. Vol. 10. No. 1. 2018.